MIT Open Access Articles

Proof of the satisfiability conjecture for large k

PROOF OF THE SATISFIABILITY CONJECTURE FOR LARGE k

JIAN DING[*], ALLAN SLY[†], AND NIKE SUN

University of Chicago;
University of California–Berkeley and Australian National University;
Microsoft Research and Massachusetts Institute of Technology

ABSTRACT. We establish the satisfiability threshold for random k-SAT for all $k \geq k_0$, with k_0 an absolute constant. That is, there exists a limiting density $\alpha_{\text{sat}}(k)$ such that a random k-SAT formula of clause density α is with high probability satisfiable for $\alpha < \alpha_{\text{sat}}$, and unsatisfiable for $\alpha > \alpha_{\text{sat}}$. We show that the threshold $\alpha_{\text{sat}}(k)$ is given explicitly by the one-step replica symmetry breaking prediction from statistical physics. The proof develops a new analytic method for moment calculations on random graphs, mapping a high-dimensional optimization problem to a more tractable problem of analyzing tree recursions. We believe that our method may apply to a range of random CSPs in the 1-RSB universality class.

1. INTRODUCTION

A **constraint satisfaction problem** (CSP) consists of variables $x_1, \ldots, x_n$ subject to constraints $a_1, \ldots, a_m$. This general framework encompasses several fundamental problems in computer science, the most classic example being **boolean satisfiability** (SAT). Other examples include various natural problems in graph combinatorics; such as (proper) coloring, independent set, and cut or bisection problems. In each of these CSPs, the variables x_i take values in some fixed alphabet $\mathcal{X}$, and it is of interest to understand properties of the subset SOL $\subseteq \mathcal{X}^n$ of valid assignments: its total size, say, or the maximum value of some objective function.

In many cases, even deciding if SOL is nonempty — which requires, *a priori*, exhaustive search over $\mathcal{X}^n$ — is NP-complete [Kar72], and thus believed to require super-polynomial time in worst-case instances. This worst-case intractability of CSPs was one of the early motivations to develop some "average-case" theory for CSPs. For instance, one approach [Lev86] is to study the typical runtime of algorithms in a **random** CSP: formally, a sequence $(\mathbb{P}^n)_{n\geq 1}$ where each $\mathbb{P}^n$ is a probability measure on CSPs of n variables; the interest is in asymptotic behaviors as $n \to \infty$.

Since their introduction into the computer science literature, random CSPs have become a subject of interest among physicists and mathematicians as well, inspired in part by early numerical experiments [CKT91, MSL92] suggesting phase transition phenomena. On the basis of heuristic analytic methods, physicists predict that they exhibit a rich array of phenomena. However, the rigorous analysis of random CSPs poses substantial mathematical difficulties, and many of the physics predictions remain challenging open problems. This paper considers one of these predictions which has been especially well-studied, the satisfiability threshold conjecture.

1.1. **Main result.** The **random k-SAT model** is as follows: variables $x_1, \ldots, x_n$ take values TRUE $\equiv$ + or FALSE $\equiv$ -. They are subject to constraints $a_1, \ldots, a_M$ where M is a Pois$(n\alpha)$ random variable. Conditioned on M, the constraints are independent. Each constraint is a random disjunctive clause: it is the boolean OR of k independent literals, with each literal sampled uniformly at random from $\{+x_1, -x_1, \ldots, +x_n, -x_n\}$. The clause is satisfied if at least one of its k literals evaluates to +. The entire instance is satisfied if every clause is satisfied. This defines a probability measure $\mathbb{P}^{n,\alpha}$ over k-SAT problem instances;[1] and the sequence $(\mathbb{P}^{n,\alpha})_{n\geq 1}$ is what is commonly termed the **random k-SAT model at density α**.

A k-SAT problem instance is naturally encoded by a bipartite graph $\mathcal{G} = (V, F, E)$ where V is the set of variables, F is the set of clauses, and E is the set of edges. The presence of an edge $e = (av) \in E$ indicates that variable v participates in clause a. The edge always comes with a sign L_e which is + or - depending on whether $+x_v$ or $-x_v$ appears in clause a. Thus $\mathbb{P}^{n,\alpha}$ can be regarded as the law of a random bipartite graph with signed edges. This is a bipartite analogue of the standard Erdős–Rényi random graph; and for this reason we sometimes refer to this model as "random Erdős–Rényi k-SAT."

---

[1]Although $\mathbb{P}^{n,\alpha}$ depends on k as well as α, one typically considers the problem for fixed k, so we suppress this dependence from the notation.

It has been notoriously challenging to characterize a most basic property of this model: what fraction of randomly sampled instances are satisfiable? Based on numerical simulations and non-rigorous arguments, it is proposed that for each fixed $k \geq 2$, there is a critical value α_{sat} — depending on k but not on n — such that for all $\epsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P}^{n, \alpha_{\mathrm{sat}} - \epsilon}(\text{satisfiable}) = 1 = \lim_{n \to \infty} \mathbb{P}^{n, \alpha_{\mathrm{sat}} + \epsilon}(\text{unsatisfiable}).$$

In words, the model has a sharp transition from satisfiable to unsatisfiable, with high probability.[2] This is known as the **satisfiability threshold conjecture**. For $k = 2$, it is known to be true with threshold $\alpha_{\mathrm{sat}} = 1$ [CR92, Goe92]. It has been a long-standing open problem to establish a satisfiability threshold for any $k \geq 3$. Our main result resolves this conjecture for large k:

Theorem 1 (main theorem). *For $k \geq k_0$, random k-SAT has a sharp satisfiability threshold α_{sat}, with explicit characterization $\alpha_{\mathrm{sat}} = \alpha_\star$ given by Proposition 1.2 below.*

The study of the random k-SAT model has seen important contributions by researchers from several different communities — probability theory, combinatorics, computer science, and statistical physics. In particular, the explicit characterization $\alpha_{\mathrm{sat}} = \alpha_\star$ emerged from the physics literature [MPZ02, MMZ06], via the so-called "one-step replica symmetry breaking" (1-RSB) framework. Subsequent works [KMR+07, MRS08] detailed the implications of 1-RSB for the geometry of the solution space SOL, and the resultant obstacles to locating the threshold. At the same time, a quite separate challenge posed by this model concerns the fluctuating local geometry of the underlying random (bipartite Erdős–Rényi) k-SAT graph. This issue has been most notably considered within the probability and computer science communities [AM02, AP03, CP12, CP16].

The current paper is heavily guided by insights from the aforementioned works. We describe these connections in the remainder of this introductory section, which is organized as follows. In §1.2 we survey the prior rigorous literature on random k-SAT. We then turn to the statistical physics work on this problem (§1.3), and describe the general notion of replica symmetry breaking (RSB) (§1.4). In the specific context of random k-SAT, we explain (§1.5) how this manifests as **one-step** RSB, leading to an explicit threshold prediction (§1.6). Lastly we explain how the underlying graph geometry poses further challenges, and outline our proof strategy to deal with these issues (§1.7).

1.2. **Prior rigorous results.** Exact satisfiability thresholds have been rigorously shown in only a few models, including k-XOR-SAT [MRZ03, PS16] and random 1-in-k-SAT [ACIM01]. Also, as we remarked above, it has been proven for random 2-SAT [CR92, Goe92], along with even finer results characterizing the scaling window [BBC+01]. Compared with all these, however, random k-SAT for $k \geq 3$ is believed to undergo a very different type of transition, as we explain below (§1.4).

For random k-SAT, even the existence of α_{sat} was not known for any $k \geq 3$. To date, the strongest result that applies for *every* $k \geq 2$ is Friedgut's theorem [Fri99]. It states that for each fixed $k \geq 2$, there is a sharp threshold *sequence* $\alpha_{\mathrm{sat}}(n)$ such that for all $\epsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P}^{n, \alpha_{\mathrm{sat}}(n) - \epsilon}(\text{satisfiable}) = 1 = \lim_{n \to \infty} \mathbb{P}^{n, \alpha_{\mathrm{sat}}(n) + \epsilon}(\text{unsatisfiable}). \tag{1.1}$$

The theorem does not imply that $\alpha_{\mathrm{sat}}(n)$ converges to a unique limit, as the conjecture requires. It also gives no quantitative information on $\alpha_{\mathrm{sat}}(n)$.

Complementing Friedgut's theorem, there have been many results giving quantitative bounds on $\alpha_{\mathrm{sat}}(n)$, usually in the limit of large k. An easy calculation of the first moment of assignments gives a fairly accurate upper bound [FP83]. Truncating the first moment to "locally maximal" solutions gives an even more precise bound

$$\limsup_{n \to \infty} \alpha_{\mathrm{sat}}(n) \leq 2^k \ln 2 - \frac{1}{2}(1 + \ln 2) + \epsilon_k \quad [\text{KKS98}].$$

In the above and throughout what follows, ϵ_k denotes any error term that tends to zero as $k \to \infty$. This upper bound is already correct in the second-order term. In contrast, all early *lower* bounds for the k-SAT threshold, which were generally algorithmic in nature, missed the true threshold by a large multiplicative factor — the current best algorithmic result [CO09] gives a lower bound of order $2^k (\ln k)/k$ while the threshold is of order 2^k.

More recent advances in lower bounding α_{sat} have all taken a non-algorithmic route — via the second moment method, in combination with Friedgut's theorem. This route, initiated by [AM02], faces two major challenges in the

---

[2]An event occurs **with high probability** if its probability tends to one in the limit $n \to \infty$.

random k-SAT model. In brief, the first concerns the geometry of the solution space SOL $\subseteq \{+, -\}^n$, while the second concerns the geometry of the underlying bipartite graph — we explain these further below. Important advances on the second issue led to a series of improvements in the lower bound:

$$\liminf_{n \to \infty} \alpha_{\text{sat}}(n) \geq \begin{cases} 2^{k-1} \ln 2 - O(1) & \text{[AM02]}; \\ 2^k \ln 2 - O(k) & \text{[AP03]}; \\ 2^k \ln 2 - \frac{3}{2} \ln 2 + \epsilon_k & \text{[CP12]}. \end{cases} \tag{1.2}$$

These works did not address the first issue (the solution space geometry), which was first discussed in the physics literature. Coja-Oghlan and Panagiotou were the first to address both issues simultaneously: they prove

$$\liminf_{n \to \infty} \alpha_{\text{sat}}(n) \geq 2^k \ln 2 - \frac{1}{2}(1 + \ln 2) - \epsilon_k \quad \text{[CP16]}, \tag{1.3}$$

matching the upper bound of [KKKS98] up to ϵ_k. This gives the best estimate of the k-SAT threshold prior to the current work, which closes the ϵ_k gap for large k.

To explain the difficulties in pinning down an exact threshold, we turn next to a survey of the statistical physics heuristics for this model, leading to the explicit characterization of α_{sat}. Having done this, we can then give a more detailed account of the earlier advances in rigorous lower bounds, as well as the obstacles that remain. With this context, we give an overview of our proof approach at the conclusion of this section.

1.3. **Statistical physics.** Statistical physicists became interested in random CSPs as examples of **spin glasses**, which are models of disordered systems (see e.g. [MP85, MPV87]). Perhaps the most extensively studied such model is the **Sherrington–Kirkpatrick** (SK) spin glass [SK75]: let $(g_{ij})_{i,j\geq 1}$ be an array of i.i.d. gaussian random variables with mean zero and variance $2/n$. The SK spin glass is defined as the probability measure on $\underline{x} \in \{+, -\}^n$ given by

$$\mu(\underline{x}) = \frac{1}{Z} \prod_{1 \leq i < j \leq n} \exp\{\beta g_{ij} x_i x_j\},$$

where Z is the partition function (normalizing constant). The measure μ of course depends on the g_{ij}, so it is a random measure supported on $\{+, -\}^n$. Parisi conjectured in a series of seminal papers [Par79, Par80a, Par80b, Par83] that the SK measure has intricate asymptotics, characterized by an infinitely nested hierarchy. Key aspects of this prediction have been rigorously proved in celebrated works [Gue03, Tal06, Pan13a].

The analogue of the SK measure in random k-SAT is the uniform measure ν over the solution space SOL, which can be regarded as a random measure on $\{+, -\}^n$:

$$\nu(\underline{x}) = \frac{1}{Z} \prod_{1 \leq j \leq M} \mathbf{1}\left\{ \begin{matrix} \underline{x} \text{ satisfies} \\ \text{clause } a_j \end{matrix} \right\} = \frac{\mathbf{1}\{\underline{x} \in \text{SOL}\}}{Z} \tag{1.4}$$

where in this context $Z = |\text{SOL}|$. In contrast with SK, however, it turns out that k-SAT exhibits the most interesting behavior when the number of constraints scales proportionally to the number of variables — in other words, when the graph of interactions $\mathcal{G}$ is **sparse**. This is a central distinction from the SK model, where the graph of interactions is the complete graph on n vertices.

An extensive statistical physics literature demonstrates how heuristics for the SK model (and a larger family of p-**spin models**) can be adapted to the analysis of sparse random CSPs such as random k-SAT. In one sense, the sparsity of interactions makes these models more challenging to analyze. In the SK model, because each vertex has a large number of neighbors, there is a self-averaging effect which is crucial to the analysis. The effect does not occur on sparse graphs, and this turns out to pose major difficulties in the mathematical study of random k-SAT and other sparse models. For example, sparse versions of the SK and p-spin models have been studied [Pan13c, Pan14, Pan15], but remain not nearly as well understood as the complete graph versions (a very incomplete list of references includes, e.g., [Pan13b, AC15, AC16, ACZ20, JT16, Sub21, Mon19, CPS21]).

In spite of this, random k-SAT is expected to exhibit behaviors which are very similar to those of SK, and in certain aspects significantly simpler. In particular, while the SK model is described by an infinitely nested hierarchy (∞-RSB or full-RSB), many random CSPs — k-SAT included — are conjecture to be described by a depth-one hierarchy (1-RSB). This is central to our understanding of this problem, and we describe this next.

1.4. **Replica symmetry and cavity methods.** Let ν be a random measure on $\{+,-\}^n$ (such as in (1.4)). For any function $f : (\{+,-\}^n)^\ell \to \mathbb{R}$, we let $\langle f \rangle_\nu$ denote the expected value of f if its ℓ arguments are independent samples of ν: explicitly,

$$\langle f \rangle_\nu \equiv \sum_{\underline{x}^1,\dots,\underline{x}^\ell} f(\underline{x}^1,\dots,\underline{x}^\ell) \prod_{j \le \ell} \nu(\underline{x}^j).$$

In the physics terminology, the $\underline{x}^j$ are *replicas* of system ν. Let R be the *overlap* (normalized inner product) between two replicas, $R = n^{-1}(\underline{x}^1 \cdot \underline{x}^2)$. The measure ν is termed **replica symmetric** (RS) if this overlap is well-concentrated, in the sense of

$$v_n = \mathbb{E}\left\{ \langle R^2 \rangle_\nu - (\langle R \rangle_\nu)^2 \right\} = o_n(1). \tag{1.5}$$

Otherwise ν is said to be **replica symmetry breaking** (RSB). Note that one can rewrite v_n as the average, over all pairs $i,j \in [n]$, of the expected correlation between $x_i^1 x_i^2$ and $x_j^1 x_j^2$,

$$\mathrm{corr}_{i,j} \equiv \mathbb{E}\left\{ \langle x_i^1 x_j^1 \rangle_\nu \langle x_i^2 x_j^2 \rangle_\nu - \langle x_i^1 \rangle_\nu \langle x_j^1 \rangle_\nu \langle x_i^2 \rangle_\nu \langle x_j^2 \rangle_\nu \right\}.$$

The RS condition (1.5) says that the average correlation $\mathrm{corr}_{i,j}$ is small. Non-concentration of the overlap (RSB) indicates the presence of **long-range correlations**.

If $[n] \equiv \{1,\dots,n\}$ is the vertex set of a **sparse** graph, then most of the contribution to (1.5) comes from vertices i,j which are far apart in the graph. Thus, in the sparse setting, RS is regarded by physicists as being equivalent to **correlation decay**: if $\underline{x}$ is a sample from ν, then x_i, x_j are roughly independent if i,j are far apart in the graph. In other words, *in an RS model, the behavior around a vertex $i \in [n]$ depends only on its local neighborhood.* The commonly studied sparse random graph models are locally tree-like — for example, the random k-SAT graph converges locally in law to a certain (multi-type) Galton–Watson tree. It is expected that sparse models in the RS regime can be accurately analyzed by a certain set of tree approximations, which generally go under the name of **belief propagation** (BP), or **replica symmetric cavity methods**. In this viewpoint, roughly speaking, the stochastic process on the finite graph is approximated by a stochastic process on the limiting tree.

We shall not go into many more details on the RS cavity method, pointing instead to the literature ([MM09, Ch. 14] and refs. therein) for details. We only note here that a key step in the method is to compare graphs $\mathcal{G}, \mathcal{G}'$ where $\mathcal{G}'$ is $\mathcal{G}$ with a random clause $a \in F$ removed. Let ∂a denote the variables incident to a in graph $\mathcal{G}$: these variables are most likely well-separated in $\mathcal{G}'$. Thus, in the RS (correlation decay) regime, one can treat these variables as *(approximately) independent, with laws depending only on their local neighborhoods in $\mathcal{G}'$.* This is a key simplification, leading to explicit tree recursions which can be analyzed. In RS models this is a powerful analytic tool. It leads further to an explicit prediction for the free energy of the model, expressed in terms of a fixed point of the tree recursions — the **RS free energy** or **Bethe free energy** (see [YFW05]). In RSB models, however, it is expected that this method yields false predictions, as the lack of correlation decay invalidates the independence assumption. We discuss this next in the context of SAT.

1.5. **Condensation and one-step replica symmetry breaking.** In a broad class of models, it is believed that RSB arises due to the formation of **clusters**, which are loosely defined as dense regions of the measure. Specifically, for $t \ge 1$ integer, t-**step replica symmetry breaking** (t-RSB) is the special case of RSB in which the overlap R concentrates on exactly $t + 1$ values. The way this can occur is that in the cube $\{+,-\}^n$, there is a hierarchy of scales $\delta_0 \gg \delta_1 \gg \dots \gg \delta_t$ such that there are many clusters of mass δ_i nested within each cluster of mass δ_{i-1}. The maximal scale δ_0 refers to the entire space $\{+,-\}^n$. This scenario is often summarized by a depth-t tree, where the vertices at depth j correspond to the clusters at scale δ_j. It was proposed by Parisi (refs. cited above) that the $t \to \infty$ limit describes symmetry breaking in the SK model.

By contrast, later works ([MPZ02, KMR+07] and refs. therein) indicated that random k-SAT and several other sparse CSPs of interest exhibit 1-RSB, corresponding to a single-depth hierarchy $\delta_0 \gg \delta_1$. In fact, random k-SAT is believed to have a rich phase diagram (Figure 1) which includes both RS and 1-RSB regimes. As we next describe, physicists predict a **condensation threshold** $\alpha_{\mathrm{cond}} \in (0, \alpha_{\mathrm{sat}})$, which marks the onset of 1-RSB.

This conjectural phase diagram was derived [ZK07, MRS08, KMR+07] in the following manner. It starts from the hypothesis that the model is at most 1-RSB. This means that there is at most one hierarchy of clustering, or equivalently that *clusters are replica symmetric.* This means that one can successfully apply RS inference methods,
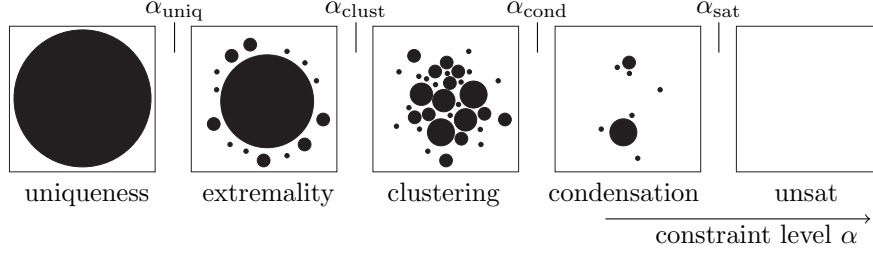
FIGURE 1. *Figure adapted from* [KMR+07]. Conjectural phase diagram of random k-SAT: each panel depicts the typical geometry of the solution space SOL $\subseteq \{+, -\}^n$ in a different regime of the constraint level α. The **satisfiability threshold** α_{sat} is the point beyond which SOL is empty with high probability. The **condensation threshold** marks the onset of symmetry breaking: the model is RS for $\alpha < \alpha_{\text{cond}}$, and is 1-RSB for $\alpha_{\text{cond}} < \alpha < \alpha_{\text{sat}}$.

but at the level of clusters rather than individual solutions. Let $\boldsymbol{\Omega}$ count the total number of clusters in SOL. For $0 \le s \le \ln 2$, let $\boldsymbol{\Omega}_s$ count only those of size approximately $\exp\{ns\}$. By the RS cavity method applied at the level of clusters, it is possible to calculate an explicit function $\Sigma(s)$ such that $\boldsymbol{\Omega}_s$ concentrates around $\exp\{n\Sigma(s)\}$. Note the implicit dependence on α; we write also $\Sigma(s) \equiv \Sigma(s; \alpha)$.

This calculation of $\Sigma(s)$, combined with other insights from the literature, lead physicists to suggest the phase diagram shown in Figure 1 [KMR+07]. As soon as α crosses a **clustering threshold** α_{clust}, the curve $\Sigma(s)$ becomes positive for some interval of s-values. There is some rigorous evidence for the clustering and related phenomena [AC08, GS14]. Note that if $\Sigma(s) \ge 0$, the clusters of size $\exp\{ns\}$ contribute roughly $\exp\{n[s + \Sigma(s)]\}$ to the total number of solutions. If $\Sigma(s) < 0$, clusters of size $\exp\{ns\}$ typically do not occur. One characterization of the condensation threshold is

$$\alpha_{\text{cond}} = \inf\{\alpha : s_1 \ne s_\star\}$$

where s_1 and $s_\star$ are both functions of α, defined by

$$s_1 = \arg\max_s\{s + \Sigma(s) : 0 \le s \le \ln 2\},$$
$$s_\star = \arg\max_s\{s + \Sigma(s) : 0 \le s \le \ln 2 \text{ and } \Sigma(s) \ge 0\}.$$

For $\alpha < \alpha_{\text{cond}}$, the solution space SOL is dominated by clusters of size $\exp\{ns_1\}$, of which there are exponentially many $(\exp\{n\Sigma(s_1)\})$. Since each cluster carries a negligible fraction of the total mass, the asymptotic $(n \to \infty)$ measure is understood as having no clusters — this regime is therefore considered replica symmetric. By contrast, for $\alpha > \alpha_{\text{cond}}$, the solution space SOL is dominated by clusters of size $\exp\{ns_\star\}$, of which there are only a *bounded* number because $\Sigma(s_\star) = 0$. Thus there are clusters carrying a non-vanishing fraction of the total mass, so the asymptotic measure has non-trivial clusters and is considered to be genuinely replica symmetry breaking. This scenario persists up to

$$\alpha_\star = \sup\left\{\alpha : \max_s \Sigma(s) \ge 0\right\}, \tag{1.6}$$

which is the 1-RSB prediction for the satisfiability threshold.

We emphasize that the above derivation is highly non-rigorous, relying on unjustified assumptions regarding the measure ν. Nevertheless we have included the above discussion in order to highlight some of the physics intuition. As we discuss in the next section, several of these ideas have had an important role in recent progress on the rigorous study of random CSPs. We note also that key aspects of the condensation phenomenon have been rigorously verified in random graph coloring [BCH+14] and random regular NAE-SAT [BCO16, SSZ16].

1.6. **Explicit threshold, and sharp upper bound.** In the physics perspective, since Σ is an explicit function, $\alpha_\star$ is already explicitly characterized by (1.6). We now spell this out by making an explicit definition of a function $\Phi(\alpha)$ which corresponds to the physics prediction for $\max_s \Sigma(s; \alpha)$. Throughout what follows, we always assume that k exceeds a large enough absolute constant k_0. Further, in view of known bounds (§1.2) on α_{sat}, we restrict attention to the regime

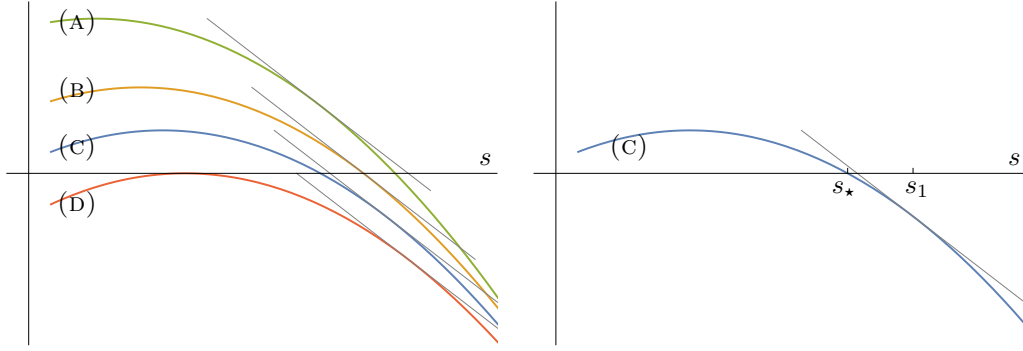$$2^k \ln 2 - 2 \equiv \alpha_{\text{lbd}} \le \alpha \le \alpha_{\text{ubd}} \equiv 2^k \ln 2. \tag{1.7}$$

FIGURE 2. The number of clusters of size roughly $\exp\{ns\}$ concentrates around its mean value $\exp\{n\Sigma(s)\}$. The left panel shows $\Sigma(s) \equiv \Sigma(s; \alpha)$ as a function of s for four different values of α, together with the tangent lines of slope -1. In increasing order of α, the curves indicate (A) $\alpha_{\text{clust}} < \alpha < \alpha_{\text{cond}}$, (B) $\alpha = \alpha_{\text{cond}}$, (C) $\alpha_{\text{cond}} < \alpha < \alpha_{\text{sat}}$, and (D) $\alpha = \alpha_{\text{sat}}$. The right panel shows curve (C) only and indicates the locations of $s_\star$ and s_1.

These assumptions will be made throughout the paper even when not explicitly stated.

Let $\mathscr{P}$ denote the space of probability measures on the half-open interval $[0, 1)$ — this means, in particular, that any $\mu \in \mathscr{P}$ gives zero measure to the event $\{\eta = 1\}$. The interpretation of the measure μ can be explained roughly as follows (it will be formalized in §2.5 below). For an edge $e = (av)$ with literal $\mathtt{L}_{av}$, we can consider the law of x_v "in absence of a," i.e., ignoring the constraint imposed by clause a. A random variable η sampled from the law μ (hereafter denoted "$\eta \sim \mu$") represents the probability "in absence of a" that we have $x_v = \text{-}\mathtt{L}_{av}$. The randomness of η results from the randomness in the neighborhood structure of v.

The above interpretation leads naturally to recursive equations for the random variables η, which are termed "survey propagation" equations in the literature (see [MMW07] and references therein). We express this as a mapping $R \equiv R^\alpha : \mathscr{P} \to \mathscr{P}$ as follows. Given $\mu \in \mathscr{P}$, generate an array of i.i.d. samples from μ,

$$\underline{\eta} \equiv \left((\eta_{ij}^+, \eta_{ij}^-)_{i,j \geq 1} \right). \tag{1.8}$$

Let d^+, d^- be $\text{Pois}(\alpha k / 2)$ random variables, independent of $\underline{\eta}$ and of one another: d^+ and d^+ represent the cardinalities of $\partial v(+a)$ and $\partial v(\text{-}a)$ respectively, where

$$\partial v(+a) \equiv \{b \in \partial v \setminus a : \mathtt{L}_{bv} = \mathtt{L}_{av}\},$$
$$\partial v(\text{-}a) \equiv \{b \in \partial v \setminus a : \mathtt{L}_{bv} = \text{-}\mathtt{L}_{av}\}.$$

If b is the i-th clause in $\partial v(+a)$, the chance "in absence of $\partial v \setminus b$" that x_v is forced to equal $\mathtt{L}_{bv}$ is represented by

$$\hat{u}_i^+ \equiv \prod_{j=1}^{k-1} \eta_{ij}^+$$

— this corresponds to the chance "in absence of b" that $x_u = \text{-}\mathtt{L}_{bu}$ for every $u \in \partial b \setminus v$. We define analogously $\hat{u}_i^-$ to refer to the i-th clause in $\partial v(\text{-}a)$. The chance "in absence of a" that none of the clauses in $\partial v(+a)$ are forcing to x_v (meaning that the value $x_v = \text{-}\mathtt{L}_{av}$ is permitted) is

$$\Pi^+ \equiv \prod_{i=1}^{d^+} \left(1 - \hat{u}_i^+ \right) = \prod_{i=1}^{d^+} \left(1 - \prod_{j=1}^{k-1} \eta_{ij}^+ \right). \tag{1.9}$$

We can define analogously Π^- which corresponds to the chance "in absence of a" that none of the clauses in $\partial v(\text{-}a)$ are forcing to x_v, meaning that the value $x_v = \mathtt{L}_{av}$ is permitted. We sometimes write $\Pi^\pm \equiv \Pi^\pm(\underline{d}, \underline{\eta})$ to emphasize the dependence of $\Pi^\pm$ on the random variables $\underline{d} \equiv (d^+, d^-)$ and $\underline{\eta}$ (from (1.8)). Notice that if all the η's belong to $[0, 1)$, then $\Pi^\pm \in (0, 1]$. The chance "in absence of a" that x_v is not forced by $\partial v(+a)$, but is forced by some clause in $\partial v(\text{-}a)$, is given by $\Pi^+(1 - \Pi^-)$. However, if x_v is simultaneously forced by both $\partial v(+a)$ and $\partial v(\text{-}a)$, this would invalidate the configuration, since it means that not all the clauses in ∂v can be simultaneously satisfied. The chance

"in absence of a" that x_v is not simultaneously forced by both $\partial v(+a)$ and $\partial v(-a)$ is given by $\Pi^+ + \Pi^- - \Pi^+\Pi^-$. We condition on being in a valid configuration simply by taking the ratio of the last two expressions, resulting in

$$R(\underline{d}, \underline{\eta}) \equiv \frac{\Pi^+(1 - \Pi^-)}{\Pi^+ + \Pi^- - \Pi^+\Pi^-} \ . \tag{1.10}$$

We define $\boldsymbol{R}\mu$ to be the law of $R(\underline{d}, \underline{\eta})$. Since $\Pi^{\pm} \in (0, 1]$, it follows that $R(\underline{d}, \underline{\eta}) \in [0, 1)$, so $\boldsymbol{R}\mu$ is indeed also an element of $\mathscr{P}$.

Proposition 1.1 (proved in §4.4: solution of 1-RSB distributional recursion). *Fix k, α and let $\boldsymbol{R} \equiv \boldsymbol{R}^\alpha$ as above. Let $\mu^\ell \equiv \mu^\ell(\alpha) \in \mathscr{P}$ ($\ell \geq 0$) be the sequence of probability measures defined by $\mu^0 = \delta_{1/2}$, and $\mu^\ell = \boldsymbol{R}\mu^{\ell-1}$ for all $\ell \geq 1$. For $k \geq k_0$ and $\alpha_{\mathrm{lbd}} \leq \alpha \leq \alpha_{\mathrm{ubd}}$, this sequence converges weakly as $\ell \to \infty$ to a limit $\mu = \mu^{\infty,\alpha} \in \mathscr{P}$, satisfying $\boldsymbol{R}\mu = \mu$.*

The following is the formal characterization of the 1-RSB prediction $\alpha_\star$ for the k-SAT threshold:

Proposition 1.2 (proved in §10: characterization of 1-RSB threshold prediction). *Given k, α, let $\mu = \mu^{\infty,\alpha}$ be the fixed point of Proposition 1.1. Let $\underline{d}$ and $\underline{\eta}$ be as above, and let $\underline{\eta}' \equiv (\eta_j)_{j\geq1}$ another sequence of i.i.d. samples from μ (independent of $\underline{d}, \underline{\eta}$). Let*

$$\Phi(\alpha) = \mathbb{E}\left[\ln \frac{\Pi^+ + \Pi^- - \Pi^+\Pi^-}{(1 - \prod_{j=1}^k \eta_j)^{\alpha(k-1)}} \right], \tag{1.11}$$

where $\mathbb{E}$ indicates the expectation over $(\underline{d}, \underline{\eta}, \underline{\eta}')$. For $k \geq k_0$, the function Φ is well-defined and strictly decreasing on the interval $\alpha_{\mathrm{lbd}} \leq \alpha \leq \alpha_{\mathrm{ubd}}$, with a unique zero $\alpha_\star \equiv \alpha_\star(k)$.

Recall that a cluster means, generally, a dense region of the measure, where for us the measure of interest is the uniform measure (1.4) over the k-SAT solution space $\mathrm{SOL} \subseteq \{+, -\}^n$. In the regime that we study, with $k \geq k_0$ and $\alpha_{\mathrm{lbd}} \leq \alpha \leq \alpha_{\mathrm{ubd}}$, it turns out that the clusters are well-separated, so that we can simply *define* a cluster to be a connected component of SOL. (Two assignments $\underline{x}, \underline{x}' \in \mathrm{SOL}$ are connected if they differ by a single bit.) With this definition, the above propositions describe the replica symmetric calculation for the **uniform measure on clusters**. Recall from §1.4 that a key step of this calculation is to compute the distribution of a variable incident to a removed clause. In the above, $\eta \in [0, 1)$ represents the probability for such a variable to be frozen to the – value. The randomness in η, as described by μ, reflects the random structure of the local neighborhood of this variable.

The normalization for the uniform measure on clusters is the total number $\boldsymbol{\Omega}$ of clusters, and $\Phi(\alpha)$ is the associated Bethe free energy. As discussed in §1.4, the measure μ should satisfy a tree recursion (the map $\boldsymbol{R}$ of Proposition 1.1), and Φ is expressed in terms of a fixed point for this recursion. The 1-RSB conjecture for random k-SAT says that clusters are RS and so $\Phi(\alpha)$ correctly predicts the free energy, which would mean that $\boldsymbol{\Omega}$ concentrates about

$$\exp\left\{ n\Phi(\alpha) \right\} = \exp\left\{ n \max_s \Sigma(s; \alpha) \right\}$$

(with high probability). This explains why $\alpha_\star$ is defined as the root of $\Phi(\alpha)$.

We emphasize again that the above characterization of $\alpha_\star$ already appears in the physics literature [MMZ06]. To the best of our knowledge, however, it has not been formally proved to be well-defined. Indeed, the proofs of Propositions 1.1 and 1.2 are based on a detailed recursive analysis, which we could not extend to all $k \geq 3$. Nevertheless, these propositions do show that $\alpha_\star$ is at least well-defined for k large enough. Having verified this, it is relatively straightforward to deduce the sharp satisfiability upper bound:

Proposition 1.3 (proved in §4.5: 1-RSB upper bound). *For $k \geq k_0$, random k-SAT at $\alpha > \alpha_\star(k)$ is with high probability unsatisfiable.*

Proposition 1.3 is proved via previously known bounds [FL03, PT04] for the positive-temperature version of the k-SAT model — that is to say, the measure

$$\nu(\underline{x}) = \frac{\exp\{-\beta H(\underline{x})\}}{Z(\beta)}$$

where $H(\underline{x})$ is the number of clauses violated by $\underline{x}$. For any $\beta \in [0, \infty)$, it is proved [PT04] that

$$\frac{\mathbb{E}[\ln Z(\beta)]}{n} \leq \inf_{\zeta, m} \Phi_{1\text{-RSB}}(\beta, \zeta, m) \tag{1.12}$$

where ζ runs over the space of probability measures over $\mathbb{R}$, m is the 1-RSB "Parisi parameter" which goes over $[0, 1]$, and the functional $\Phi_{\text{1-RSB}}$ is given explicitly in the statement of Theorem 4.30 below. The proof of the bound (1.12) in [PT04] is based on a certain Hamiltonian interpolation scheme, inspired by related results for the SK spin glass [GT02, Gue03, FL03].

It remains for us to choose ζ and m to obtain a good upper bound on (1.12). The 1-RSB heuristic suggests to choose ζ in a particular way, such that it is approximately a reparametrization of the measure μ (from Proposition 1.1). With this choice, we show that as soon as α exceeds $\alpha_\star$, we have (see (4.60) below)

$$\lim_{\beta \to \infty} \Phi_{\text{1-RSB}}\left(\beta, \zeta, \frac{1}{\beta^{1/2}}\right) < 0\,.$$

This will imply there are no satisfying assignments with high probability, yielding Proposition 1.3.

1.7. **Sharp lower bound.** The main content of this paper is to prove the matching lower bound to Proposition 1.3. As noted in §1.2, all recent satisfiability lower bounds, including our current result, are proved by the second moment method together with Friedgut's theorem [Fri99]. We now briefly describe the main obstacles to this method, and how they are overcome in our analysis. A more extensive discussion is given in §2.1.

As before, let Z be the total number of k-SAT solutions, and let $\mathbb{E}$ denote expectation with respect to $\mathbb{P} = \mathbb{P}^{n,\alpha}$. The most basic version of the second moment method would be to prove

$$\limsup_{n \to \infty} \frac{(\mathbb{E}Z)^2}{\mathbb{E}[Z^2]} < \infty\,.$$

The Cauchy–Schwarz inequality then gives

$$\liminf_{n \to \infty} \mathbb{P}(Z > 0) > 0$$

at this value of α; and Friedgut's theorem immediately implies satisfiability with high probability at any $\alpha' < \alpha$.

In fact, this basic version of the second moment method fails on random k-SAT at **any** positive clause density — the ratio $\mathbb{E}[Z^2]/(\mathbb{E}Z)^2$ diverges with n for any positive α, including throughout the RS regime. The problem does not go away if we condition on the number of clauses, or make other minor modifications, as we discuss in more detail in §2.1 below. At a high level, the reason is roughly as follows. Recall (§1.4) that in the RS regime, variables far apart in the graph $\mathcal{G}$ are nearly independent, and the behavior of each variable depends only on its local neighborhood. In some "locally homogeneous" models, either all variables have the same local neighborhood, or there is a variety of local neighborhoods but they all give rise to the same variable behavior.[3] This homogeneity does **not** hold for random k-SAT — e.g., some variables are incident to more positive literals, and so are more likely to be TRUE. In the RS regime, there is correlation decay **conditional** on the graph structure — but the moment calculation averages over the graph structure, and as a result non-negligible correlations arise. These "local neighborhood correlations" cause the second moment method to fail, and are a central difficulty of random k-SAT.

In spite of this, the second moment method has been successfully applied to lower bound the number of k-SAT solutions in the RS regime. In all such results ([AM02, AP03, CP13], see (1.2)), a key step is to make some truncation, or reweighting, such that the resulting model becomes locally homogeneous. The result of [CP13] is notable in that it also conditions on the degree profile of the k-SAT instance, an idea which had previously been applied in a simpler model [CP12]. This decreases the effect of local neighborhood correlations and gives an improved lower bound.

The best lower bound prior to this work is due to Coja-Oghlan and Panagiotou ([CP16], see (1.3)). This advance was especially significant in moving the lower bound past the conjectural condensation threshold of random k-SAT. Inspired by the 1-RSB heuristic, the proof of [CP16] applies second moment method to the number of solution **clusters**, rather than the number of individual solutions. This strategy had previously been applied to improve the lower bound for random NAE-SAT [CP12], and to obtain sharp satisfiability thresholds in some locally homogeneous models [DSS13, DSS16b, DSS16a]. The result of [CP16] further incorporates techniques developed in [CP12, CP13] for conditioning on the degree profile.

The result of [CP16] demonstrates that applying the second moment method to the number of solution clusters, and conditioning on the degree profile, can give very good lower bounds. It became clear, however, that in order to

---

[3]See the discussion of "symmetric" models in [CP13, Appx. A]. We use the phrase "locally homogeneous" rather than "symmetric" to avoid confusion with the (separate) issue of replica symmetry.

achieve a sharp lower bound, it would be necessary to condition not only on the degree profile, but on the profile of local neighborhood structures to **arbitrarily large (constant) depth** R. The main work of this paper is to carry out this approach: we establish a satisfiability lower bound $\alpha_{\text{lbd}}(R)$ for each R; and show that $\alpha_{\text{lbd}}(R) \to \alpha_\star$ in the limit $R \to \infty$.

Let us briefly indicate the main difficulties in implementing this strategy. The second moment computation reduces to an optimization problem over a vector ω of empirical marginals, broken down according to the R-neighborhood type — the dimension of this problem diverges with R. The proof of [CP16] solves a version of this problem for marginals ω broken down according to the variable degree. Their analysis relies on an important pre-processing step — for k large, removing $n\epsilon_k$ variables with atypical degree leaves behind an nearly regular graph. This allows for very explicit analysis of the second moment, but costs ϵ_k in the satisfiability lower bound.

To achieve a sharp lower bound, we can only afford to remove $n\epsilon_{k,R}$ variables with $\epsilon_{k,R} \to 0$ in the limit $R \to \infty$. Thus we cannot hope to avoid including increasingly pathological vertices as R grows. Instead, we devise a slightly elaborate preprocessing scheme which ensures that bad vertices are surrounded by large buffers of nice vertices. One portion of the paper is occupied with proving that this scheme indeed removes a vanishing fraction $\epsilon_{k,R}$ of variables.

It remains to solve the second moment optimization problem, where as input we have only rather rough *a priori* estimates on ω that are guaranteed by the preprocessing step. The central new idea in this paper is to update ω in **blocks** corresponding to trees inside the graph. By keeping the rest of ω fixed, we can reduce a non-convex optimization problem on large finite graphs to a convex optimization problem on finite trees of bounded (though diverging with R) depth, with some fixed boundary conditions. For the tree optimization we make a system of weights that act as Lagrange multipliers for the boundary conditions. The weights are set by an inductive construction, where the preprocessing step was specifically designed to ensure that the weights contract in the desired way. Once these weights are set, it becomes relatively easy to read off the desired second moment bound. This analysis is the main technical contribution of this paper, and may be appliable in the analysis of other models which are not locally homogeneous. We refer to Section 3 for a more detailed proof outline.

Acknowledgements. We thank Amir Dembo, Ahmed El Alaoui, Elchanan Mossel, Andrea Montanari, and Lenka Zdeborová for many helpful conversations. We also wish to acknowledge the hospitality of the Theory Group at Microsoft Research Redmond, where part of this work was done. Ahmed El Alaoui and Andrea Montanari reviewed with us a draft of this paper and made innumerable valuable comments, and we especially thank them for their generosity. Finally, we are extremely grateful to the anonymous referee for their careful reading and detailed comments on the paper.

2. Moment method, cluster encodings, and tree recursions

In this section we introduce some of the preliminary formalisms that will be required for the proof. The section is organized as follows:

– In §2.1 we review the standard first and second moment calculations for random k-SAT, and show that the second moment method fails in this model.
– In §2.2 we elaborate on two reasons for the failure of the second moment method: the local inhomogeneity of the random k-SAT graph, and the phenomenon of large clusters (replica symmetry breaking) in the solution space.
– In §2.3 we introduce a combinatorial encoding of clusters, the so-called "frozen model," which will be the basis of our modified moment method approach.
– In §2.4 we introduce two more combinatorial encodings, the warning propagation model and the color model. They are equivalent to the frozen model, but each model has its own drawbacks and advantages.
– In §2.5 we introduce tree recursions for the warning propagation model.
– In §2.6 we introduce weighted versions of the color model, and review the belief propagation (BP) equations in this context. We show that the tree recursions of §2.4 can be recovered as a special case.

Before proceeding further, we first review the formal definition of the model.

Definition 2.1 (bipartite factor graph). A **bipartite factor graph** is a graph $\mathscr{G} = (V, F, E)$ whose vertex set $V \cup F$ is partitioned into **variables** V and **clauses** F, with undirected edges E joining variables to clauses. We generically denote variables u, v, w, clauses a, b, c, and edges e. For each edge e, we write $a(e)$ for the incident clause and $v(e)$ for the incident variable. Each e comes with a **sign** L_e, indicating whether the inclusion of variable $v(e)$ in clause

$a(e)$ is positive ($\mathtt{L}_e = +$) or negative ($\mathtt{L}_e = -$). We allow for multi-edges, so $\mathscr{G}$ might have two edges e, e' joining a to v (possibly with $\mathtt{L}_e \neq \mathtt{L}_{e'}$). If there is a unique edge e joining clause a to variable v, we will sometimes denote it as $e = (av) = (va)$, and write $\mathtt{L}_e \equiv \mathtt{L}_{av}$. For any vertex $x \in V \cup F$, we write ∂x for its neighboring vertices. Similarly we write δx for the edges incident to x. For each clause $a \in F$ we regard ∂a and δa as **ordered tuples**. For each edge $e \in E$ we write $j(e)$ to indicate the position of e in $\delta a(e)$, so $j(e) \in [k]$. For each variable $v \in V$ we partition its neighbors and incident edges according to edge sign:

$$\begin{aligned}
\delta v(+) &\equiv \{e \in \delta v : \mathtt{L}_e = +\}, \quad \partial v(+) \equiv \{a(e) : e \in \delta v(+)\}, \\
\delta v(-) &\equiv \{e \in \delta v : \mathtt{L}_e = -\}, \quad \partial v(-) \equiv \{a(e) : e \in \delta v(-)\}.
\end{aligned} \tag{2.1}$$

(In scenarios with multi-edges, the sets $\partial v(+)$ and $\partial v(-)$ can intersect, and are not in one-to-one correspondence with the sets $\delta v(\pm)$. For this reason, we always work with $\delta v(\pm)$ to avoid ambiguity, except in cases where multi-edges are expressly prohibited.) We will also refer to $\mathscr{G}$ as a **SAT problem instance**. Furthermore we call $\mathscr{G}$ a k-**SAT problem instance** if each clause $a \in F$ has width $|\delta a| = k$.

Definition 2.2 (satisfying assignments). If $\mathscr{G} = (V, F, E)$ is a bipartite factor graph as in Definition 2.1, it defines a mapping $\mathscr{G} : \{+, -\}^V \to \{0, 1\}$ as follows: for any **variable assignment** $\underline{x} \in \{+, -\}^V$,

$$\mathscr{G}(\underline{x}) \equiv \prod_{a \in F} \left\{ 1 - \prod_{e \in \delta a} \frac{1 - \mathtt{L}_e x_{v(e)}}{2} \right\}.$$

The instance $\mathscr{G}$ is called **satisfiable** if and only if its set

$$\mathrm{SOL}(\mathscr{G}) \equiv \mathscr{G}^{-1}(1) \equiv \left\{ \underline{x} \in \{+, -\}^V : \mathscr{G}(\underline{x}) = 1 \right\} \subseteq \{+, -\}^V$$

of **satisfying assignments** is nonempty.

Definition 2.3 (random k-SAT). To generate an instance of the **random k-SAT model at clause density** α, start with n labelled variables $V = [n] \equiv \{1, \dots, n\}$, as well as M labelled clauses $F = [M] \equiv \{1, \dots, M\}$, where M is sampled from the $\mathrm{Pois}(n\alpha)$ distribution. Independently for each clause $a \in F$, sample its k-tuple of variables ∂a uniformly at random from $V^k = [n]^k$, then sample the k-tuple of signs $(\mathtt{L}_e)_{e \in \delta a}$ uniformly at random from $\{+, -\}^k$. This defines a family of probability measures $\mathbb{P} \equiv \mathbb{P}^{n,\alpha}$ over k-SAT instances, indexed by n and parametrized by the expected clause density α. Writing $\mathbb{P}_{n,m}$ for the measure $\mathbb{P}$ conditioned on $M = m$, we can decompose

$$\mathbb{P} \equiv \mathbb{P}^{n,\alpha} = \sum_{m \geq 0} \mathrm{pois}_{n\alpha}(m) \mathbb{P}_{n,m}, \quad \mathrm{pois}_\lambda(m) \equiv \frac{e^{-\lambda} \lambda^m}{m!}. \tag{2.2}$$

Note that the conditional measure $\mathbb{P}_{n,m}$ does not depend on α.

Remark 2.4. A SAT problem instance can be equivalently viewed as a hypergraph, with vertices and hyperedges corresponding to variables and clauses respectively. Each hyperedge should be viewed as an ordered tuple $(v_1, \dots, v_k)$ of variables, with corresponding signs $(\mathtt{L}_1, \dots, \mathtt{L}_k)$. A k-SAT instance thus corresponds to a k-uniform hypergraph. If $\mathscr{G} = (V, F, E)$ is the bipartite factor graph representation of a SAT instance, we assign length $1/2$ to all its edges, so that graph distances in $\mathscr{G}$ will be consistent with the standard graph distances in the hypergraph representation. For any vertex x in $\mathscr{G}$ we will write $\partial_s x$ for the set of vertices at distance s from x. We will write $N(x) \equiv \partial_1(x)$. If v is a variable then $N(v)$ is the set of variables sharing a clause with v, and we often refer to these as the "neighboring variables of v."

2.1. Moments of satisfying assignments.

We now review the standard first and second moment calculations for random k-SAT. The purpose of this discussion is to illustrate the main obstructions to proving sharp bounds on the satisfiability threshold. These issues were known prior to our work, and we refer to further detailed discussions in the existing literature [AM02, AP03].

For comparison, we will present the moment calculations for random k-SAT as well as a closely related model, **random k-NAE-SAT**, which has also been extensively studied (notably by [AM02, CP12]). On a bipartite factor graph $\mathscr{G} = (V, F, E)$, a variable assignment $\underline{x} \in \{+, -\}^V$ is called an **not-all-equal-SAT (NAE-SAT) assignment** if both $\underline{x}$

and $-\underline{x}$ are valid SAT assignments. Thus, while the SAT assignments of $\mathcal{G}$ are given by $\text{SOL}(\mathcal{G}) \equiv \mathcal{G}^{-1}(1)$, the NAE-SAT assignments are given by $\text{NAE}(\mathcal{G}) \equiv \text{SOL}(\mathcal{G}) \cap [-\text{SOL}(\mathcal{G})]$. We define the corresponding **partition functions**

$$Z \equiv Z(\mathcal{G}) \equiv \left| \text{SOL}(\mathcal{G}) \right| \equiv \left\| \left\{ \text{SAT assignments of } \mathcal{G} \right\} \right\|,$$

$$Z_{\text{NAE}} \equiv Z_{\text{NAE}}(\mathcal{G}) \equiv \left| \text{NAE}(\mathcal{G}) \right| \equiv \left\| \left\{ \text{NAE-SAT assignments of } \mathcal{G} \right\} \right\|.$$

Clearly, $Z_{\text{NAE}}(\mathcal{G}) \leq Z(\mathcal{G})$ for any instance $\mathcal{G}$.

For $M \sim \text{Pois}(n\alpha)$, we have $|M - n\alpha| \leq n^{1/2} \ln n$ with high probability. We can then see from (2.2) that in order to show $\mathbb{P}^{n,\alpha}(E_n) = o_n(1)$ for some event E_n, it is sufficient to show $\mathbb{P}_{n,m}(E_n) = o_n(1)$ uniformly over all m satisfying $|m - n\alpha| \leq n^{1/2} \ln n$. We will compute first and second moments of $Z(\mathcal{G})$ and $Z_{\text{NAE}}(\mathcal{G})$ for $\mathcal{G}$ distributed according to the conditional measure $\mathbb{P}_{n,m}$. The reason to work with $\mathbb{P}_{n,m}$ rather than $\mathbb{P}^{n,\alpha}$ is that the fluctuations in M alone are already enough to make the second moment method fail under $\mathbb{P}^{n,\alpha}$. Fixing the number of clauses is an easy way to remove some variance from the second moment calculation, and will allow us to see the more difficult sources of variance that remain after m is fixed. The second moment also fails under $\mathbb{P}_{n,m}$ for reasons that are more subtle and that determine the proof strategy.

For ease of exposition, for the current discussion we will assume that assume $n\alpha$ is an integer, and consider $\mathbb{P}_{n,m}$ only for $m = n\alpha$.[4] Let $\mathbb{E}_{n,n\alpha}$ denote expectation with respect to $\mathbb{P}_{n,n\alpha}$. We then have

$$\mathbb{E}_{n,n\alpha} Z = 2^n \left(1 - \frac{1}{2^k} \right)^m = \exp \left\{ n \left[\ln 2 + \alpha \ln \left(1 - \frac{1}{2^k} \right) \right] \right\} \equiv \exp\{n f_{\text{SAT}}(\alpha)\}, \tag{2.3}$$

$$\mathbb{E}_{n,n\alpha} Z_{\text{NAE}} = 2^n \left(1 - \frac{2}{2^k} \right)^m = \exp \left\{ n \left[\ln 2 + \alpha \ln \left(1 - \frac{2}{2^k} \right) \right] \right\} \equiv \exp\{n f_{\text{NAE}}(\alpha)\}.$$

Write α_1 for the solution of $f_{\text{SAT}}(\alpha) = 0$; this is the **first moment threshold** for random k-SAT. For $\alpha > \alpha_1$ the expected value of Z is exponentially small in n, so (by Markov's inequality) it holds with high probability that Z is zero, meaning that the instance is unsatisfiable. Thus the satisfiability transition for random k-SAT is upper bounded by the first moment threshold, which occurs just below $2^k \ln 2$. The analogous statement holds for random k-NAE-SAT: the satisfiability transition is upper bounded by the solution of $f_{\text{NAE}}(\alpha) = 0$, which occurs just before $2^{k-1} \ln 2$.

To lower bound the satisfiability threshold, one approach is to apply the **second moment method**, based on the following consequence of the Cauchy–Schwarz inequality:

$$\frac{(\mathbb{E}Z)^2}{\mathbb{E}[Z^2]} = \frac{(\mathbb{E}[Z \mathbf{1}\{Z > 0\}])^2}{\mathbb{E}[Z^2]} \leq \mathbb{P}(Z > 0). \tag{2.4}$$

In the most naive application, one could take $\mathbb{E}$ to be expectation with respect to the overall measure $\mathbb{P}^{n,\alpha}$, and try to show a bound of the form

$$\limsup_{n \to \infty} \frac{\mathbb{E}(Z^2)}{(\mathbb{E}Z)^2} \leq C(k,\alpha) < \infty, \tag{2.5}$$

for any positive α. In fact, as we already alluded to, the bound (2.5) is false for all positive α: indeed, one can use the above calculation (2.3) to see that for all $|n(\alpha' - \alpha)| \leq n^{1/2} \ln n$, we have

$$\mathbb{E}_{n,n\alpha'} Z = \exp\{n f_{\text{SAT}}(\alpha')\} \ll \sum_{m \geq 0} \text{pois}_{n\alpha}(m) \mathbb{E}_{n,m} Z = \mathbb{E}Z, \tag{2.6}$$

with $\text{pois}_{n\alpha}$ as defined by (2.2). It follows by Markov's inequality that $Z \ll \mathbb{E}Z$ with high probability. This implies that (2.5) must be false, otherwise we would have a contradiction to (2.4). By a similar calculation, the bound (2.5) also fails with Z_{NAE} in place of Z.

A more promising approach is to try to establish a conditional second moment bound, of the form

$$\limsup_{n \to \infty} \left[\sup \left\{ \frac{\mathbb{E}_{n,m}(Z^2)}{(\mathbb{E}_{n,m}Z)^2} : |m - n\alpha| \leq n^{1/2} \ln n \right\} \right] \leq C(k,\alpha) < \infty. \tag{2.7}$$

---

[4]Strictly speaking, we should fix α and consider $m = n\alpha'$ for all $|n(\alpha' - \alpha)| \leq n^{1/2} \ln n$. However the calculations that follow are not very sensitive to slight perturbations in α: it the moment method gives satisfiability (or unsatisfiability) with high probability under $\mathbb{P}_{n,n\alpha}$, then it also gives the same result under $\mathbb{P}_{n,n\alpha'}$ for $|n(\alpha' - \alpha)| \leq n^{1/2} \ln n$, unless α is exactly at a threshold. For this reason we prefer to keep the notation simple in this introductory discussion, and consider only $\alpha = \alpha'$.

If (2.7) were to hold at some positive clause density α, then substituting it into (2.4) would show that satisfiability occurs with asymptotically positive probability under $\mathbb{P}^{n,\alpha}$:

$$\liminf_{n\to\infty} \mathbb{P}^{n,\alpha}(\text{SAT}) \geq \liminf_{n\to\infty}\left[\inf\left\{\mathbb{P}_{n,m}(\text{SAT}) : |m - n\alpha| \leq n^{1/2}\ln n\right\}\right] \geq \frac{1}{C(k,\alpha)} > 0.$$

Then Friedgut's theorem (1.1) immediately gives satisfiability with high probability at any $\alpha - \epsilon < \alpha$, which would imply a satisfiability lower bound, $\alpha_{\text{sat}} \geq \alpha$.

In fact, we will see that the bound (2.7) also fails for random k-SAT at all positive α, although it gives a non-trivial lower bound for the random k-NAE-SAT model ([AM02], and reviewed below). To see this, we decompose

$$Z^2 = \sum_z Z^2[z], \quad Z^2[z] \equiv \left|\left\{\begin{array}{c}\text{pairs }(\underline{x}^1, \underline{x}^2) \in \text{SOL}(\mathcal{G}) \times \text{SOL}(\mathcal{G}) \\ \text{with }|\{v \in V : x_v^1 = x_v^2\}| = nz\end{array}\right\}\right|, \tag{2.8}$$

where the sum is over $z = j/n$ for integer $0 \leq j \leq n$. For example,

(i) The value $z = 1$ corresponds to identical configurations $\underline{x}^1 = \underline{x}^2$, and $Z^2[1] = Z$;

(ii) The value $z = 0$ corresponds to antipodal configurations $\underline{x}^1 = -\underline{x}^2$, and $Z^2[0] = Z_{\text{NAE}}$;

(iii) The value $z = 1/2$ corresponds to configurations $\underline{x}^1, \underline{x}^2$ that "look independent" in the sense that if v is a uniformly random variable, knowing x_v^1 does not give any information about x_v^2.

For each z we calculate the corresponding second moment contribution to be

$$\mathbb{E}_{n,n\alpha}[Z^2[z]] = 2^n\binom{n}{nz}\left(1 - \frac{2}{2^k} + \left(\frac{z}{2}\right)^k\right)^{n\alpha} = \frac{\exp\{n\mathsf{f}_{\text{SAT},2}(z;\alpha)\}}{n^{O(1)}},$$

where the exponent $\mathsf{f}_{\text{SAT},2}(z;\alpha)$ can be derived using Stirling's formula, and does not depend on n:

$$\mathsf{f}_{\text{SAT},2}(z;\alpha) \equiv \underbrace{\left[\ln 2 + \mathcal{H}(z)\right]}_{\text{entropy term}} + \underbrace{\alpha\ln\left(1 - \frac{2}{2^k} + \left(\frac{z}{2}\right)^k\right)}_{\text{probability term}} \tag{2.9}$$

where $\mathcal{H}(z) = -z\ln z - (1-z)\ln(1-z)$ denotes the standard entropy function. For comparison,

$$\mathsf{f}_{\text{NAE},2}(z;\alpha) \equiv \underbrace{\left[\ln 2 + \mathcal{H}(z)\right]}_{\text{entropy term}} + \underbrace{\alpha\ln\left(1 - \frac{4}{2^k} + \frac{(z^k + (1-z)^k)2}{2^k}\right)}_{\text{probability term}}$$

gives the corresponding exponent for $\mathbb{E}[(Z_{\text{NAE}})^2[z]]$.

By comparing (2.3) with (2.9), we find that $\mathsf{f}_{\text{SAT}}(1/2;\alpha) - 2\mathsf{f}_{\text{SAT}}(\alpha)$ is exactly zero. However, one can notice in (2.9) that the entropy term is maximized at $z = 1/2$, but for any positive α the probability term is strictly increasing with z. This will mean that $\mathsf{f}_{\text{SAT},2}(z;\alpha)$ has strictly positive derivative with respect to z at $z = 1/2$, i.e., $z = 1/2$ is not the maximizer for $\mathsf{f}_{\text{SAT},2}(z;\alpha)$. This will imply that the ratio

$$\frac{\mathbb{E}_{n,n\alpha}[Z^2]}{(\mathbb{E}_{n,n\alpha}Z)^2} = n^{O(1)}\exp\left\{n\sup\left\{\mathsf{f}_{\text{SAT},2}(z;\alpha) - 2\mathsf{f}_{\text{SAT}}(\alpha) : 0 \leq z \leq 1\right\}\right\}$$

grows exponentially with n, which is in contradiction to (2.7). By contrast, the function $\mathsf{f}_{\text{NAE},2}$ is *stationary* at $z = 1/2$ at any fixed α, in fact attains its global maximum at $z = 1/2$ for a non-trivial range of α; see Figure 3. This observation was used in previous work ([AM02], see (1.2)) to lower bound the satisfiability threshold for random k-NAE-SAT.

2.2. Local inhomogeneity and replica symmetry breaking.

The calculation of §2.1 demonstrates two distinct (though entangled) issues, which we already mentioned in Section 1 — (I) lack of "local homogeneity" and (II) large solution clusters and "replica symmetry breaking." These issues manifest themselves in the above calculation roughly as follows:

(I) For all positive α, the point $z = 1/2$ is not a local maximizer of $\mathsf{f}_{\text{SAT},2}(z;\alpha)$;

(II) For all $\alpha \in (\alpha_2, \alpha_1)$ (where $\alpha_2 < \alpha_{\text{sat}} < \alpha_1$), the function $\mathsf{f}_{\text{SAT},2}(z;\alpha)$ has another local maximizer z' slightly below 1, such that $\mathsf{f}_{\text{SAT},2}(z';\alpha) > 2\mathsf{f}_{\text{SAT}}(\alpha)$.

(A) $z \mapsto f_{\text{NAE},2}(z; \alpha) - 2f_{\text{NAE}}(\alpha))$

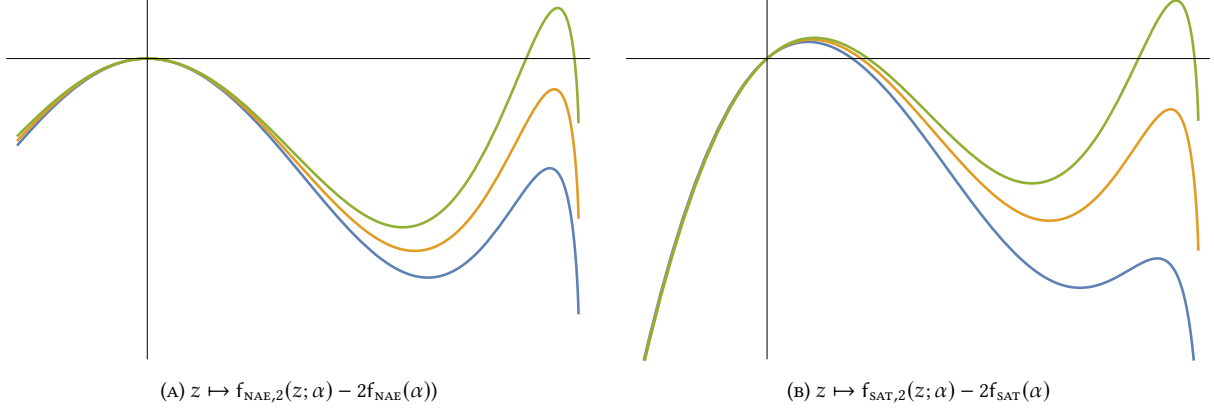(B) $z \mapsto f_{\text{SAT},2}(z; \alpha) - 2f_{\text{SAT}}(\alpha)$

FIGURE 3. Comparison of second moment with first moment squared for random NAE-SAT and SAT
for $k = 6$ (with same qualitative phenomena occuring for all k).
In each figure, the horizontal axis is at zero while the vertical axis is at $z = 1/2$.
Each curve corresponds to a different value of α, with the uppermost curve in the
right panel corresponding to the numerically computed value of $\alpha_{\star}(k)$ [MMZ06].

Of course, the desired second moment bound (2.7) cannot succeed in the presence of either (I) or (II). Point (II) reflects the fact that the second moment can be dominated by an exponentially rare event where there is an unusually large number of pairs of nearby solutions (z near one), i.e., there is an **atypically large cluster** of solutions. For both random k-SAT and random k-NAE-SAT, the α_2 of point (II) occurs strictly below the satisfiability threshold; in fact, it occurs just below the condensation threshold α_{cond} that marks the onset of **replica symmetry breaking** (see Figure 1). For an extensive discussion of this issue, we refer the reader to two works [DSS13, SSZ16] on random **regular** k-NAE-SAT: this is a simplified model where the first issue (I) does not arise at all, and as a result there is a very precise correspondence between replica symmetry breaking and problems in the moment method.

We next turn our attention to point (I). The reflects that the moment calculation favors **graphs** that are slightly rare, for which the solution set is unusually large. The phenomenon results from the inherent asymmetry between + and – in the SAT predicate, which is absent from NAE-SAT. For instance, it is reasonable to expect that to have more SAT assignments, it is favorable to have atypically many edges all take the same sign. To be more explicit, let $\mathscr{D}$ be the empirical degree profile of $\mathscr{G}$,

$$\mathscr{D}(d^+, d^-) = \frac{1}{n} \left| \left\{ v \in V : |\delta v(+)| = d^+ \text{ and } |\delta v(-)| = d^- \right\} \right|. \tag{2.10}$$

Under $\mathbb{P}_{n,n\alpha}$, we expect that the random profile $\mathscr{D}$ is concentrated around the typical profile $\mathscr{D}_{\star} = (\text{pois}_{\alpha k/2})^{\otimes 2}$ (using the notation of (2.2)), with gaussian fluctuations: that is to say, we expect

$$\frac{\mathbb{P}_{n,n\alpha}(\mathscr{D})}{\mathbb{P}_{n,n\alpha}(\mathscr{D}_{\star})} \le \exp \left\{ -n \left(\mathscr{D} - \mathscr{D}_{\star} \right)^{\text{t}} \boldsymbol{K} \left(\mathscr{D} - \mathscr{D}_{\star} \right) \right\} \tag{2.11}$$

for some fixed positive semi-definite $\boldsymbol{K}$. On the other hand, we also expect

$$\frac{\mathbb{E}_{n,n\alpha}(Z \mid \mathscr{D})}{\mathbb{E}_{n,n\alpha}(Z \mid \mathscr{D}_{\star})} \ge \exp \left\{ n \left(\boldsymbol{u}, \mathscr{D} - \mathscr{D}_{\star} \right) \right\} \tag{2.12}$$

for some fixed vector $\boldsymbol{u} \ne 0$. In particular, $(\boldsymbol{u}, \mathscr{D})$ might count the fraction of + edges in $\mathscr{D}$: the typical fraction is $(\boldsymbol{u}, \mathscr{D}_{\star}) = 1/2$, but we would expect that $\mathbb{E}_{n,n\alpha}(Z \mid \mathscr{D})$ could be made larger by taking $(\boldsymbol{u}, \mathscr{D}) = 1/2 + \delta$ for small positive δ. This is to say that $\mathscr{D} = \mathscr{D}_{\star}$ is optimal for (2.11) but not for (2.12), so it is not optimal for the product of (2.11) and (2.12). It follows that for all $\mathscr{D}'$ sufficiently close to $\mathscr{D}_{\star}$ we have

$$\mathbb{E}_{n,n\alpha}(Z \mid \mathscr{D}') \ll \sum_{\mathscr{D}} \mathbb{P}_{n,n\alpha}(\mathscr{D}) \, \mathbb{E}_{n,n\alpha}(Z \mid \mathscr{D}) = \mathbb{E}_{n,n\alpha} Z. \tag{2.13}$$

This can be viewed as a (slightly more complicated) analogue of (2.6). It implies $Z \ll \mathbb{E}_{n,n\alpha} Z$ with high probability. As a result (2.7) must fail, otherwise (2.4) would be contradicted. On the other hand, for random k-NAE-SAT, (I) does

not occur, and the second moment method succeeds for some range of positive α (until (II) arises). From this we can conclude that $\mathscr{D}_\star$ is optimal for $\mathbb{E}_{n,n\alpha}(Z_{\text{NAE}} \mid \mathscr{D}_\star)$.

Another point of view (which is really another side of the same coin) is that (I) reflects the "local inhomogeneity" of SAT solutions. Conditional on any instance $\mathscr{G}$, let $\mu_{\mathscr{G}}$ be the uniform measure over the set of satisfying assignments SOL$(\mathscr{G})$ (assuming that it is nonempty). We emphasize that $\mu_{\mathscr{G}}$ is a **random measure**, since SOL$(\mathscr{G})$ is a random set. We use $\langle \cdot \rangle_{\mathscr{G}}$ to denote averaging with respect to $\mu_{\mathscr{G}}$:

$$\langle f \rangle_{\mathscr{G}} \equiv \sum_{\underline{x} \in \{+,-\}^n} \mu_{\mathscr{G}}(\underline{x}) f(\underline{x}) = \frac{1}{|\text{SOL}(\mathscr{G})|} \sum_{\underline{x} \in \text{SOL}(\mathscr{G})} f(\underline{x}) \,.$$

We will write $[\cdot]_{\mathscr{G}}$ for the average with respect to the uniform measure over the NAE-SAT solutions NAE$(\mathscr{G})$. We say that SAT solutions are **locally inhomogeneous** because on general instances $\mathscr{G}$, the variable mean $\langle x_v \rangle_{\mathscr{G}}$ is not constant over $v \in V$, since we expect it to depend for instance on $|\delta v(+)| - |\delta v(-)|$. This is in contrast with the NAE-SAT model, where the symmetry NAE$(\mathscr{G}) = -$NAE$(\mathscr{G})$ implies $[x_v]_{\mathscr{G}} = 0$ for all $v \in V$, so that the model enjoys **local homogeneity**. This property affects the moment calculation in the following way. Given $\mathscr{G}$, form a new graph $\mathscr{G}'$ as follows: add a new clause a, connect a to a uniformly random k-tuple of existing variables in $\mathscr{G}$, and sample uniformly random signs $(L_e)_{e \in \delta a}$. Then SOL$(\mathscr{G}')$ is a subset of SOL$(\mathscr{G})$, and

$$\frac{Z(\mathscr{G}')}{Z(\mathscr{G})} = \sum_{\underline{x} \in \text{SOL}(\mathscr{G})} \frac{\mathbf{1}\{\underline{x} \text{ satisfies } a\}}{|\text{SOL}(\mathscr{G})|} = 1 - \mu_{\mathscr{G}}\left(L_e x_{v(e)} = - \text{ for all } e \in \delta a\right).$$

The variables $v(e)$ ($e \in \delta a$) are typically far apart from one another in the original graph $\mathscr{G}$. We will now make the simplifying assumption that they are **approximately independent** under $\mu_{\mathscr{G}}$. This assumption is not rigorous, but for $\alpha \leq \alpha_{\text{cond}}$ it is not unreasonable, since it matches the physics prediction that $\mu_{\mathscr{G}}$ exhibits some form of **correlation decay**. The assumption allows us to simplify the above as

$$\frac{Z(\mathscr{G}')}{Z(\mathscr{G})} \doteq 1 - \prod_{e \in \delta a} \mu_{\mathscr{G}}(L_e x_{v(e)} = -) = 1 - \prod_{e \in \delta a} \frac{1 - L_e \langle x_{v(e)} \rangle_{\mathscr{G}}}{2} \,. \tag{2.14}$$

The key point is that, conditional on $\mathscr{G}$, the ratio (2.14) is a **nondegenerate random variable**, due to the randomness in the literals L_e and in the choice of the variables $v(e)$ ($e \in \delta a$). By contrast, under the same assumptions,

$$\frac{Z_{\text{NAE}}(\mathscr{G}')}{Z_{\text{NAE}}(\mathscr{G})} = 1 - \prod_{e \in \delta a} \frac{1 - L_e [x_{v(e)}]_{\mathscr{G}}}{2} - \prod_{e \in \delta a} \frac{1 + L_e [x_{v(e)}]_{\mathscr{G}}}{2} = 1 - \frac{2}{2^k} \,,$$

a **deterministic constant**, simply because $[x_v]_{\mathscr{G}}$ is constant. However, one can imagine building the entire graph $\mathscr{G} = \mathscr{G}_n \sim \mathbb{P}_{n,n\alpha}$ by a sequence $(\mathscr{G}_0, \dots, \mathscr{G}_n)$ where each $\mathscr{G}_i$ is roughly distributed according to $\mathbb{P}_{i,i\alpha}$, and is formed from $\mathscr{G}_{i-1}$ by a small number of random local changes as above: adding a variable, and adding or deleting a small number of clauses. We can then represent $Z(\mathscr{G})$ as a telescoping product

$$Z(\mathscr{G}) \doteq Z(\mathscr{G}_0) \prod_{i=1}^n \frac{Z(\mathscr{G}_i)}{Z(\mathscr{G}_{i-1})} \,,$$

where we expect the n terms in the product to be roughly independent from one another, and we think of each term as being analogous to (2.14). This would suggest that, under $\mathbb{P}_{n,\alpha}$, the variance of $\ln Z(\mathscr{G})$ is of order n, while the variance of $\ln Z_{\text{NAE}}(\mathscr{G})$ is small. This is consistent with the fact that $Z \ll \mathbb{E}_{n,n\alpha} Z$ with high probability for all positive α, while Z_{NAE} concentrates around $\mathbb{E}_{n,n\alpha} Z_{\text{NAE}}$ if α is not too large.

To prove an exact satisfiability lower bound by the second moment method, it is necessary to address both issues of local inhomogeneity (I) and replica symmetry breaking (II). For (II), the idea is to count solution clusters instead of individual solutions. On a given instance $\mathscr{G}$, recall that $\mathbf{\Omega} \equiv \mathbf{\Omega}(\mathscr{G})$ denotes the number of SAT solution clusters (connected components of SOL$(\mathscr{G})$), and let $\mathbf{\Omega}_{\text{NAE}} \equiv \mathbf{\Omega}_{\text{NAE}}(\mathscr{G})$ denote the number of NAE-SAT solution clusters (connected components of NAE$(\mathscr{G})$). Since the random variables $\mathbf{\Omega}$ and $\mathbf{\Omega}_{\text{NAE}}$ give unit weight to each cluster regardless of cluster size, their moments are not affected by atypically large clusters. This approach is implemented in several works on models that are locally homogeneous [DSS13, DSS16a, SSZ16]. An important technical ingredient is a combinatorial representation of clusters as elements $\underline{x} \in \{+, -, f\}^n$, which will be explained in the remainder of this section. We let CL$(\mathscr{G})$ be the set of clusters of $\mathscr{G}$, regarded as a subset of $\{+, -, f\}^n$.

We now define $\nu_{\mathscr{G}}$ to be the uniform measure on $\text{CL}(\mathscr{G})$. Thus $\nu_{\mathscr{G}}$ is a random measure over $\{+,-,\mathtt{f}\}^n$. Problem (I) in this context is that SAT clusters are not locally homogeneous, in that ν_v (the marginal law of x_v under $\nu_{\mathscr{G}}$) is not constant over $v \in V$. In fact, although we saw that NAE-SAT solutions are locally homogeneous, it turns out that NAE-SAT clusters are not: in the $\{+,-,\mathtt{f}\}^n$ representation, a variable is more likely to be $\mathtt{f}$ if it has low degree. (This problem goes away in the random regular k-NAE-SAT model where all variables have the same degree.) To address this problem, it is natural to consider conditioning on a degree profile $\mathscr{D}$ that is close to the typical one $\mathscr{D}_{\star}$. This takes care of (2.13). It also makes the model more locally homogeneous in the sense that conditioning on $\mathscr{D}$ partitions the variables v into different classes according to their $\pm$ degrees $(|\delta v(+)|, |\delta v(-)|)$; and the fluctuations of ν_v **within each class** are smaller than the fluctuations of ν_v over all $v \in V$.

However, as we already suggested in §1.7, conditioning on $\mathscr{D}$ alone does not resolve the problem, because ν_v depends on much more than the $\pm$ degree of v: we expect it to depend on the entire **local neighborhood structure** of $\mathscr{G}$ near v. This motivates the following definitions Let $B_R(v)$ be the R-neighborhood of v, which we regard as a graph rooted at v. Let $\mathscr{D}_R \equiv \mathscr{D}_R(\mathscr{G})$ be the probability measure on rooted graphs defined by

$$\mathscr{D}_R(T) = \frac{|\{v \in V : B_R(v) \cong T\}|}{|V|},$$

where $\cong$ denotes rooted graph isomorphism. We regard this $\mathscr{D}_R$ as a **generalized degree profile**, and note that the ordinary degree profile $\mathscr{D}$ of (2.10) coincides with $\mathscr{D}_R$ for $R = 1/2$ (recalling Remark 2.4). With high probability under $\mathbb{P} = \mathbb{P}_{n,n\alpha}$, the measure $\mathscr{D}_R$ lies within $o_n(1)$ total variation distance of a measure $\mathscr{D}_{\star,R}$, which is the law of the first R levels of a certain Poisson Galton–Watson tree (§4.1). Conditioning on $\mathscr{D}_R$ partitions the variables v into different classes according to the structures of their local neighborhoods $B_R(v)$. The expectation is that ν_v becomes constant within each class in the limit $R \to \infty$. However, conditional on $\mathscr{D}_R$ for large fixed R, typically the SAT solution clusters will still be locally inhomogeneous, because ν_v will fluctuate slightly within each $B_R(v)$ class. It will still be the case that $\Omega \ll \mathbb{E}(\Omega \,|\, \mathscr{D}_R)$ with high probability.

In prior works that have used the second moment method to lower bound the satisfiability threshold in random k-SAT, the proof strategies follow the same basic conceptual outline: start from $X = \text{SOL}(\mathscr{G})$ or $X = \text{CL}(\mathscr{G})$, and fix a radius R. Then devise some $X_R \subseteq X$ which is locally homogeneous, and perform the second moment method on $|X_R|$ given $\mathscr{D}_R$ (near $\mathscr{D}_{\star,R}$). The choices of R and X have been, essentially, as follows:

	R	X	X_R	resulting lower bound on α_{sat}
[AM02]	0	$\text{SOL}(\mathscr{G})$	NAE-SAT assignments	$2^{k-1} \ln 2 - O(1)$
[AP03]	0	$\text{SOL}(\mathscr{G})$	balanced SAT assignments	$2^k \ln 2 - O(k)$
[CP13]	1/2	$\text{SOL}(\mathscr{G})$	"judicious" SAT assignments	$2^k \ln 2 - \frac{3}{2} \ln 2 + \epsilon_k$
[CP16]	1/2	$\text{CL}(\mathscr{G})$	"judicious" SAT clusters	$2^k \ln 2 - \frac{1}{2}(1 + \ln 2) - \epsilon_k$.

We use a similar notion of "judicious" clusters, and defer the exact definition to Section 3 (Definition 3.26); the sole purpose of the condition is to enforce local homogeneity.

Of course, the above approach can only succeed when $X_R \ne \varnothing$ (with high probability). However, the set X is in reality **not** locally homogeneous given $\mathscr{D}_R$ for any fixed R, so we expect

$$\frac{|X|}{|X_R|} \le \frac{1}{\exp(n\epsilon_R)}$$

for ϵ_R positive but vanishing in the limit $R \to \infty$. This suggests a regime $(\alpha - \delta_R, \alpha_{\text{sat}})$ in which $X_R = \varnothing$ with high probability, where δ_R is positive but vanishing as $R \to \infty$. Thus, to achieve the exact satisfiability threshold under this scheme, it is necessary to take $R \to \infty$. This is precisely the strategy of this paper: we take $X = \text{CL}(\mathscr{G})$, and perform the second moment computation on a subset of "judicious" configurations $X_R \subseteq X$, conditional on $\mathscr{D}_R$. In this way we prove $\alpha_{\text{sat}} \ge \alpha_{\star} - o_R(1)$ where $\alpha_{\star}$ is the predicted threshold of Proposition 1.2, and the result follows by taking $R \to \infty$.

2.3. **Combinatorial encoding of clusters.**

The remainder of the current section is dedicated to the combinatorial representation of $X = \text{CL}(\mathscr{G})$, where we follow [CP16]. For further background, we refer to [Par05, MMW07, MM09] and the references therein.

Recall that a variable v has incident edges δv, which connect to its neighboring clauses ∂v, where we regard both δv and ∂v as unordered multisets. A k-SAT solution is given by a configuration $\underline{x} \in \{+,-\}^V$ such that every clause

$a \in F$ is satisfied, meaning that the k-tuple $(L_{av}x_v)_{v \in \partial a}$ is not identically -. We now introduce a new spin $\mathtt{f} \equiv \mathtt{free}$, and use it to define the combinatorial model of k-SAT solution clusters.

Definition 2.5 (frozen model). Throughout this paper we take the convention that if $L \in \{\text{-}, \text{+}\}$ and $x = \mathtt{f}$, then $Lx \equiv \mathtt{f}$. On a k-SAT instance $\mathscr{G} = (V, F, E)$, a **frozen configuration** is a vector $\underline{x} \in \{\text{+}, \text{-}, \mathtt{f}\}^V$ such that

  (i) Each clause $a \in F$ is **satisfied**, meaning that for at least one $e \in \delta a$ we have $L_e x_{v(e)} \in \{\text{+}, \mathtt{f}\}$;

  (ii) A variable $v \in V$ takes value $x_v \neq \mathtt{f}$ if and only if it is **forced** to do so, meaning that for at least one $e \in \delta v$ we have $L_{e'} x_{v(e')} = \text{-} L_e x_{v(e)}$ for all $e' \in \delta a(e) \setminus e$.

(The definition makes sense even in the presence of multi-edges.)

Remark 2.6. Place a graph structure on the set $\text{SOL}(\mathscr{G}) \equiv \mathscr{G}^{-1}(1)$ of satisfying assignments by putting an edge between any pair of assignments at Hamming distance one, and define a **cluster** of solutions to be a (maximal) connected component of the graph $\mathscr{G}^{-1}(1)$. As has been explained in the literature (see e.g. [Par05, MMW07, MM09]), frozen configurations encode clusters in the sense that there is a natural mapping

$$\text{coarsen} : \mathscr{G}^{-1}(1) \longrightarrow \left\{ \text{frozen configurations on } \mathscr{G} \right\}. \tag{2.15}$$

If $\mathscr{G}$ has no multi-edges, then each cluster $\mathscr{C} \subseteq \mathscr{G}^{-1}(1)$ maps to a single frozen configuration $\text{coarsen}(\mathscr{C})$. To define the map, given a configuration $\underline{x} \in \{\text{+}, \text{-}, \mathtt{f}\}^V$ that does not violate any clauses (in the sense of Definition 2.5(i)), let us say that a variable $v \in V$ is **blocked** with respect to $\underline{x}$ if there is some $e \in \delta v$ such that $L_{e'} x_{v(e')} = \text{-} L_e x_{v(e)}$ for all $e' \in \delta a(e) \setminus e$ (cf. Definition 2.5(ii)). Define $\text{co}(\underline{x}) = \underline{y}$ where

$$y_v = \begin{cases} x_v & \text{if } v \text{ is blocked with respect to } \underline{x}; \\ \mathtt{f} & \text{otherwise.} \end{cases}$$

The map of (2.15) is defined by iterating co until the configuration stabilizes, i.e., $\text{coarsen} \equiv \text{co}^\infty$. Note that termination happens in finite time, since if $\text{co}(\underline{x}) \neq \underline{x}$ then $\text{co}(\underline{x})$ has strictly more $\mathtt{f}$ variables than $\underline{x}$. If $\mathscr{G}$ has no multi-edges, then the value of $\text{co}(\underline{x})$ at any variable v is a function of the neighboring values $x_u, u \in N(v)$. Consequently, if $\underline{x}$ and $\underline{x}'$ differ in a single coordinate, then $\text{co}(\underline{x}) = \text{co}(\underline{x}')$. Iterating this gives $\text{coarsen}(\underline{x}) = \text{coarsen}(\underline{x}')$ as long as $\underline{x}, \underline{x}'$ lie in the same cluster. This shows that the map (2.15) takes each cluster $\mathscr{C}$ to a single frozen configuration.

Remark 2.7. Another mapping one might consider is $\text{cube} : \mathscr{G}^{-1}(1) \to \{\text{+}, \text{-}, \mathtt{f}\}^V$, where $\text{cube}(\underline{x}) = \underline{y}$ is defined by

$$y_v = \begin{cases} x_v & \text{if } v \text{ only takes value } x_v \text{ in the cluster containing } \underline{x}; \\ \mathtt{f} & \text{otherwise.} \end{cases}$$

This also has the property that each cluster is mapped to a single point in $\{\text{+}, \text{-}, \mathtt{f}\}^V$, which in this case simply encodes the minimal Hamming subcube containing that cluster. In many situations the maps coarsen and cube are identical, but one can construct cases where they differ; see Figure 4. The frozen model is preferable precisely because it is defined only by **local** constraints.

2.4. Warning propagation and color model.
We now introduce two more combinatorial models, which are both equivalent to the frozen model, but will be important for analytical purposes. The first is the "warning propagation" (WP) model which has appeared widely in the physics literature (see [MPZ02, BMZ05, MMZ06, MMW07, MM09]). In this model, a solution cluster is represented by a "warning configuration" $\underline{w} \equiv (w_e)_{e \in E}$, where each $w_e \equiv (\dot{w}_e, \hat{w}_e)$ represents a pair of "warnings" sent across e in either direction:

$$\dot{w}_e \text{ represents the warning across } e \text{ from } v(e) \text{ to } a(e);$$
$$\hat{w}_e \text{ represents the warning across } e \text{ from } a(e) \text{ to } v(e).$$

Each warning concerns the evaluation of $L_e x_{v(e)}$ within the cluster; the warning from each endpoint of e represents the state of $L_e x_{v(e)}$ "in absence of" the opposite endpoint. The possible warnings are $\text{+}, \text{-}, \mathtt{f}$, where $\pm$ indicates a warning that $L_e x_{v(e)}$ must be $\pm$ (within the cluster), while $\mathtt{f}$ indicates no warning. In the warning propagation model, the clause $a(e)$ can only force the variable $v(e)$ to agree with L_e, so all clause-to-variable warnings $\hat{w}_e$ will be either + or $\mathtt{f}$. The formal definition is as follows:
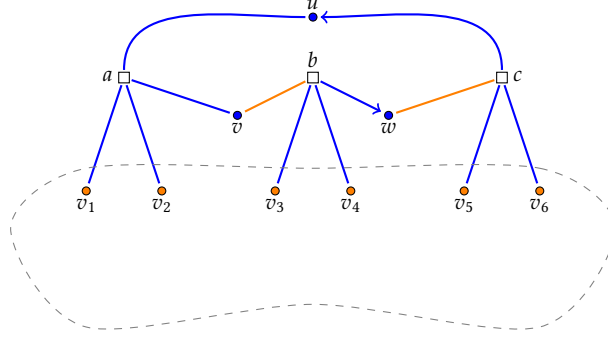
FIGURE 4. An example where the maps coarsen and cube differ (Remarks 2.6 and 2.7). First ignore the portion of the figure above the variables v_i ($1 \le i \le 6$): the dashed gray curve represents a k-SAT instance $\mathcal{G}$, such that SOL($\mathcal{G}$) has a cluster γ in which the variables v_i ($1 \le i \le 6$) always take value $x_{v_i} = -$. Suppose that in $\mathcal{G}$ we have coarsen($\underline{x}$) = cube($\underline{x}$) = $\underline{z}$. Given $\mathcal{G}$, form a new instance $\mathcal{G}' \supset \mathcal{G}$ by adding the rest of the figure, so $\mathcal{G}' \setminus \mathcal{G}$ contains the variables u, v, w as well as the clauses a, b, c. Each edge $e \in \mathcal{G}' \setminus \mathcal{G}$ is colored blue if $\mathsf{L}_e = +$, orange if $\mathsf{L}_e = -$. For $\underline{x} \in \gamma$, we consider how to extend $\underline{x}$ to a satisfying assignment $\underline{y}$ of $\mathcal{G}'$. If $y_u = -$, then clause a forces $y_v = +$, and then clause b forces $y_w = +$, and then clause c is violated. Consequently we must have $y_u = +$, which ensures that clauses a and c are satisfied. For clause b to be satisfied, we must have $(y_v, y_w) \ne (+, -)$. Thus a valid extension of $\underline{x} \in \gamma$ is given by $\underline{y} = (\underline{x}, +, +, +)$ where $(+, +, +)$ indicates the values on (u, v, w). From the above discussion, in $\mathcal{G}'$ we have cube($\underline{y}$) = $(\underline{z}, +, \mathtt{f}, \mathtt{f})$. On the other hand, the arrows in the figure indicate that in the initial configuration $\underline{y} = (\underline{x}, +, +, +)$, variable w is blocked by clause b, and variable u is blocked by clause c. Thus variable v is unblocked so the coarsen map first changes y_v to $\mathtt{f}$. This makes w unblocked, so next y_w is changed to $\mathtt{f}$, which makes u unblocked. As a result coarsen($\underline{y}$) = $(\underline{z}, \mathtt{f}, \mathtt{f}, \mathtt{f}) \ne$ cube($\underline{y}$). We emphasize that this example required the introduction of a short cycle in $\mathcal{G}' \setminus \mathcal{G}$, and so we expect such occurrences to be rare in the random k-SAT model.

Definition 2.8 (warning propagation model). For integers $K \ge 1$ and $D \ge 0$, define the mappings

$$\hat{W}(\dot{\mathtt{w}}_1, \ldots, \dot{\mathtt{w}}_K) = \begin{cases} + & \text{if } \dot{\mathtt{w}}_i = - \text{ for all } 1 \le i \le K, \\ \mathtt{f} & \text{otherwise;} \end{cases}$$

$$\dot{W}(\hat{\mathtt{w}}_1, \ldots, \hat{\mathtt{w}}_D) = \begin{cases} + & \text{if } + \in \{\hat{\mathtt{w}}_1, \ldots, \hat{\mathtt{w}}_D\} \subseteq \{+, \mathtt{f}\}, \\ - & \text{if } + \in \{\hat{\mathtt{w}}_1, \ldots, \hat{\mathtt{w}}_D\} \subseteq \{-, \mathtt{f}\}, \\ \mathtt{f} & \text{if } \{\hat{\mathtt{w}}_1, \ldots, \hat{\mathtt{w}}_D\} \subseteq \{\mathtt{f}\}, \\ \varnothing & \text{otherwise.} \end{cases}$$

In the case $D = 0$, the input to the function $\dot{W}$ is empty, and the output is $\mathtt{f}$. On a k-SAT instance $\mathcal{G} \equiv (V, F, E)$, a **warning configuration** is a tuple $\underline{\mathtt{w}}$ of spins $\mathtt{w}_e \equiv (\dot{\mathtt{w}}_e, \hat{\mathtt{w}}_e) \in \{+, -, \mathtt{f}\} \times \{+, \mathtt{f}\}$, indexed by edges $e \in E$, such that the **warning propagation equations** are satisfied:

$$\hat{\mathtt{w}}_e = \widehat{\mathrm{WP}}_e[\underline{\mathtt{w}}] \equiv \hat{W}_{k-1}\left(\dot{\mathtt{w}}_{e'} : e' \in \delta a(e) \setminus e\right),$$

$$\dot{\mathtt{w}}_e = \dot{\mathrm{WP}}_e[\underline{\hat{\mathtt{w}}}] \equiv \dot{W}\left(\mathsf{L}_{e'}\mathsf{L}_e\hat{\mathtt{w}}_{e'} : e' \in \delta v(e) \setminus e\right). \tag{2.16}$$

for all edges $e \in E$. (Note that for $\underline{\mathtt{w}}$ to be a valid warning configuration, we require each $\dot{\mathtt{w}}_e$ to be an element of $\{+, -, \mathtt{f}\}$, which means $\dot{\mathrm{WP}}_e$ must not output $\varnothing$ for any $e \in E$. We remark also that taking $\dot{\mathtt{w}}_e = \hat{\mathtt{w}}_e = \mathtt{f}$ for all $e \in E$ gives rise always to a valid warning configuration.)

Remark 2.9. Warning configurations are in bijective correspondence with frozen configurations of $\mathcal{G}$. If $\underline{\mathtt{w}}$ is a valid warning configuration of $\mathcal{G}$, we can obtain a valid frozen configuration $\underline{x}$ of $\mathcal{G}$ simply by taking

$$x_v = \dot{W}\left(\mathsf{L}_e\hat{\mathtt{w}}_e : e \in \delta v\right)$$

for all $v \in V$. Conversely, if $\underline{x}$ is a valid frozen configuration of $\mathcal{G}$, we can obtain a valid warning configuration $\underline{\mathtt{w}}$ of $\mathcal{G}$ by to the following procedure:

(i) If $e \in \delta v$ with $x_v = \mathtt{f}$, then we must have $\mathtt{w}_e = (\mathtt{f}, \mathtt{f})$.

(ii) If $e \in \delta v$ with $x_v \in \{+, -\}$ and $L_e x_v = -$, then we must have $w_e = (-, f)$.

(iii) If $e \in \delta v$ with $x_v \in \{+, -\}$ and $L_e x_v = +$, then

$$\hat{w}_e = \begin{cases} + & \text{if } w_{e'} = (-, f) \text{ for all } e' \in \delta a(e) \setminus e \\ f & \text{otherwise.} \end{cases}$$

(iv) The above determines of $\hat{w}_e$ for all e. We can then determine $\dot{w}_e$ for all e by applying the WP rules (2.16).

We leave the reader to verify that these mappings are inverses of one another.

In fact, we will rarely use the full warning propagation model, and instead will work primarily with the following especially concise simplification of WP, introduced by [CP16]. Note that, under the rules of Definition 2.8, the pair $w_e \equiv (\dot{w}_e, \hat{w}_e)$ can take any value in $\{+, -, f\} \times \{+, f\}$ except for $(-, +)$, which represents a pair of conflicting warnings that will invalidate the entire configuration (since in this case $\dot{\text{WP}}_{e'}$ will output $\varnothing$ for some $e' \in \delta v(e) \setminus e$.

Definition 2.10 (color model). Let proj be the mapping which sends warnings $w \equiv (\dot{w}, \hat{w})$ to **colors** $\sigma \in \{r, y, g, b\}$, according to the following rules:

$$\text{proj}(\dot{w}, \hat{w}) = \begin{cases} r \equiv \text{red} & \text{if } \hat{w} = + \text{ and } \dot{w} \in \{+, -\} & \text{(clause forces this edge to be satisfied),} \\ y \equiv \text{yellow} & \text{if } (\dot{w}, \hat{w}) = (-, f) & \text{(variable forced to negate edge; clause not forcing on edge),} \\ g \equiv \text{green} & \text{if } (\dot{w}, \hat{w}) = (f, f) & \text{(no forcing warnings across this edge),} \\ b \equiv \text{blue} & \text{if } (\dot{w}, \hat{w}) = (f, +) & \text{(variable forced to affirm edge; clause not forcing on edge).} \end{cases}$$

A valid **coloring** of $\mathcal{G} = (V, F, E)$ is any configuration $\underline{\sigma} \in \{r, y, g, b\}^E$ that can be obtained by taking a valid warning configuration $\underline{w}$ and applying proj on each edge.

Remark 2.11. Colorings are in bijective correspondence with frozen configurations of $\mathcal{G}$. To see this, let $\underline{\sigma}_{\delta v}$ denote the colors on the incident edges of variable v, recall the notation (2.1), and define

$$x_v = \text{eval}_v(\underline{\sigma}_{\delta v}) \equiv \begin{cases} + & \text{if } \underline{\sigma}_{\delta v(-)} \text{ is all } y, \text{ and } \underline{\sigma}_{\delta v(+)} \text{ is all } \{r, b\} \text{ with at least one } r \\ - & \text{if } \underline{\sigma}_{\delta v(+)} \text{ is all } y, \text{ and } \underline{\sigma}_{\delta v(-)} \text{ is all } \{r, b\} \text{ with at least one } r \\ f & \text{if } \underline{\sigma}_{\delta v} \text{ is all } g \\ \varnothing & \text{otherwise.} \end{cases} \tag{2.17}$$

If $\underline{\sigma}$ is a valid coloring of $\mathcal{G}$, then taking $x_v = \text{eval}_v(\underline{\sigma}_{\delta v})$ for all $v \in V$ defines a valid frozen configuration $\underline{x}$ of $\mathcal{G}$. By Remark 2.9, $\underline{x}$ corresponds to a unique warning configuration $\underline{w}$. We leave the reader to verify that proj maps $\underline{w}$ back to the starting $\underline{\sigma}$, which completes the correspondence.

It is essential to us that all the combinatorial models that we consider (Definitions 2.5, 2.8, and 2.10) are defined only by **local** constraints. This implies that they are all **factor models** — here, it simply means that the counting measure on valid configurations can be expressed as a product of **local factors**, where each factor is an indicator function that checks one of the local constraints. For factor models in much greater generality, there is a rich physics formalism (see e.g. [MM09, Ch. 9]), as well as a natural way to turn the moment calculation into an analytic optimization problem problem (see §3.5).

Let $\mathcal{G} = (V, F, E)$ be a random k-SAT instance. We will give the explicit factors for the WP and color models. The counting measure on warning configurations of $\mathcal{G}$ is given simply by

$$\mathbf{1}\left\{ \begin{array}{l} \underline{w} \text{ is a valid warning} \\ \text{configuration on } \mathcal{G} \end{array} \right\} = \prod_{v \in V} \varphi_v(\underline{w}_{\delta v}) \prod_{a \in F} \hat{\varphi}_a(\underline{w}_{\delta a})$$

where each factor φ_x simply checks that the warnings leaving x are indeed obtained by applying the appropriate warning propagation maps on the incoming warnings:

$$\varphi_v(\underline{w}_{\delta v}) \equiv \prod_{e \in \delta v} \mathbf{1}\{\dot{w}_e = \dot{\text{WP}}_e[\hat{w}]\},$$

$$\hat{\varphi}_a(\underline{w}_{\delta a}) \equiv \prod_{e \in \delta a} \mathbf{1}\{\hat{w}_e = \widehat{\text{WP}}_e[\dot{w}]\}.$$

We will abuse notation slightly and also use $\varphi, \hat{\varphi}$ for the factors of the coloring model. The counting measure on valid colorings is given by

$$\mathbf{1}\left\{\begin{array}{l} \underline{\sigma} \text{ is a valid} \\ \text{coloring of } \mathscr{G} \end{array}\right\} = \prod_{v \in V} \varphi_v(\underline{\sigma}_{\delta v}) \prod_{a \in F} \hat{\varphi}_a(\underline{\sigma}_{\delta a}), \tag{2.18}$$

where the variable factor $\varphi_v(\underline{\sigma}_{\delta v})$ is simply the indicator that $\underline{\sigma}_{\delta v}$ can be obtained by applying proj to a configuration $\underline{w}_{\delta v}$ for which $\varphi_v(\underline{w}_{\delta v}) = 1$; and the clause factor $\hat{\varphi}_a$ is analogously defined. The color model factors can be described much more explicitly, as follows: the clause factor is

$$\hat{\varphi}_a(\underline{\sigma}_{\delta a}) = \begin{cases} 1 & \text{if } \underline{\sigma}_{\delta a} \text{ has exactly one spin } \mathtt{r}, \text{ all other spins } \mathtt{y} & \text{(forcing clause);} \\ 1 & \text{if } \underline{\sigma}_{\delta a} \text{ has no spins } \mathtt{r} \text{ and at least two spins } \{\mathtt{g}, \mathtt{b}\} & \text{(non-forcing clause);} \\ 0 & \text{otherwise.} \end{cases} \tag{2.19}$$

(As far as clauses are concerned, the colors $\mathtt{b}$ and $\mathtt{g}$ are interchangeable.) The variable factor is given by

$$\varphi_v(\underline{\sigma}_{\delta v}) = \mathbf{1}\left\{\operatorname{eval}_v(\underline{\sigma}_{\delta v}) \neq \varnothing\right\}, \tag{2.20}$$

for eval_v defined by (2.17). Note that the clause factor does not depend on the incident edge signs, but the variable factor does, since the edge signs enter into the definition of eval_v.

Remark 2.12. We will sometimes refer to the color model above as the **single-copy** color model, to distinguish it from the **pair** color model which appears in the second moment: the latter is supported on pairs $(\underline{\sigma}^1, \underline{\sigma}^2)$ with each $\underline{\sigma}^j$ a valid coloring of the same graph $\mathscr{G}$. If we extend $\varphi_v, \hat{\varphi}_a$ to pair inputs by setting

$$\varphi_v(\underline{\sigma}^1_{\delta v}, \underline{\sigma}^2_{\delta v}) \equiv \prod_{j=1,2} \varphi_v(\underline{\sigma}^j_{\delta v}),$$

$$\hat{\varphi}_a(\underline{\sigma}^1_{\delta a}, \underline{\sigma}^2_{\delta a}) \equiv \prod_{j=1,2} \hat{\varphi}_a(\underline{\sigma}^j_{\delta a}),$$

then the counting measure for the pair model is also expressed by (2.18), provided that for each vertex $x \in V \cup F$ we interpret $\underline{\sigma}_{\delta x}$ as $\equiv (\underline{\sigma}^1_{\delta x}, \underline{\sigma}^2_{\delta x})$.

To summarize what we have discussed in §2.3–2.4: on a given k-SAT problem instance $\mathscr{G} = (V, F, E)$, assuming there are no multi-edges, there is a map coarsen which sends solution clusters to frozen configurations (Remark 2.6). We then have bijections (Remarks 2.9 and 2.11)

$$\left\{\begin{array}{c} \text{frozen configurations} \\ \underline{x} \in \{\text{+}, \text{-}, \mathtt{f}\}^V \end{array}\right\} \longleftrightarrow \left\{\begin{array}{c} \text{warning configurations} \\ \underline{w} \in \{\text{++}, \text{+f}, \text{-f}, \text{f+}, \text{ff}\}^E \end{array}\right\} \longleftrightarrow \left\{\begin{array}{c} \text{valid colorings} \\ \underline{\sigma} \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^E \end{array}\right\}. \tag{2.21}$$

The models are all defined by local constraints, and so can be written as factor models. For the color model the factors can be written in an especially explicit way, given above by (2.17), (2.19), and (2.20). In the remainder of this section we introduce **tree recursions** for these models.

2.5. **Tree recursions for warnings.** To motivate what comes next, we note that the central analytical challenge of this paper is to prove a lower bound for colorings (Definition 2.10) on the random k-SAT graph $\mathscr{G}$ (Definition 2.3). In view of the bijection (2.21), this translates to a lower bound on frozen configurations (Definition 2.5), which we then show can be "completed" to satisfying assignments of $\mathscr{G}$. This short synopsis hides many technicalities which will appear later. For now, however, we will focus on the fundamental issue of local homogeneity, and consider a basic question: if γ is a uniformly random solution cluster and $\underline{x}$ is the corresponding frozen configuration, how should the marginal law of x_v depend on the local neighborhood structure around v? We do not directly answer this question, but in the rest of this section we describe the physics prediction in mathematical terms. In later sections we use the prediction to obtain the rigorous result.

Let $\mathscr{G} = (V, F, E)$ be any SAT instance. Recall the notation (2.1). If v does not participate in any multi-edges, then for each clause $a \in \partial v$ we shall denote

$$\begin{array}{ll} \partial v(\text{+}a) \equiv \{b \in \partial v \setminus a : \mathtt{L}_{bv} = \text{+}\mathtt{L}_{av}\}, & \delta v(\text{+}a) \equiv \{(bv) : b \in \partial v(\text{+}a)\}, \\ \partial v(\text{-}a) \equiv \{b \in \partial v \setminus a : \mathtt{L}_{bv} = \text{-}\mathtt{L}_{av}\}, & \delta v(\text{-}a) \equiv \{(bv) : b \in \partial v(\text{-}a)\}. \end{array} \tag{2.22}$$

For any vertex $x \in V \cup F$ we can consider its depth-ℓ neighborhood $B_\ell(x)$, which we regard as being rooted at x. If $x \in V$ we assume that ℓ is a nonnegative integer; if $x \in F$ we assume that $\ell - 1/2$ is a nonnegative integer. If ℓ is any constant and $\mathcal{G}$ is an instance of **random** k-SAT, then $B_\ell(x)$ will be acyclic for most vertices of $\mathcal{G}$, with high probability. Since this is the predominant case, we will assume it for the remainder of the section. If $B_\ell(x)$ is acylic then it is a **bipartite factor tree**, by which we mean a bipartite factor graph that is also a tree. We will write $B_\ell(x) \equiv T \equiv (V_T, F_T, E_T)$.

Our next goal is to define a frozen model on $T = B_\ell(x)$ that will approximate, in some sense, what we expect to see in the frozen model in the full random graph $\mathcal{G}$. To do this, we will impose certain **boundary conditions** on T, which we now explain. Let $\mathcal{L}T$ denote the leaf vertices of T; they are all variables by the conditions on ℓ. The set $\mathcal{L}T$ contains (but may be strictly larger than) the set $\partial_\ell x$ of variables at distance exactly ℓ from x. We now take advantage of the correspondence (2.21) and think in terms of warning configurations. Let $\dot{\underline{w}}_\partial$ denote a vector of boundary input warnings $\dot{w}_{ua(u)}$, where u runs over $\mathcal{L}T$ and $a(u)$ denotes the unique clause neighboring u. Any $\dot{\underline{w}}_\partial$ has at most one completion to a valid warning configuration $\underline{w}$ on T. One way to see this is to apply the WP maps (2.16) of Definition 2.8 recursively, started from the boundary inputs $\dot{\underline{w}}_\partial$: either at some point $\dot{\text{WP}}$ outputs $\varnothing$ and there is no valid completion of $\dot{\underline{w}}_\partial$, or the process eventually terminates at the unique valid completion $\underline{w}$ of $\dot{\underline{w}}_\partial$. For example, if $\text{L}_e = \text{+}$ for all edges e in T, and all entries of $\dot{\underline{w}}_\partial$ are in $\{\text{+}, \text{f}\}$, then it is easy to see that $\dot{\underline{w}}_\partial$ can be completed to a valid warning configuration $\underline{w}$ on T, with $\dot{w}_e = \hat{w}_e = \text{f}$ on every internal edge e of T.

Definition 2.13. On the tree $T = B_\ell(x)$, let $\dot{\underline{w}}_\partial$ be a tuple of random boundary input warnings, where each $\dot{w}_{ua(u)}$ is sampled independently from $\text{unif}(\{\text{+}, \text{-}\})$ if $u \in \partial_\ell x$, and $\dot{w}_{ua(u)} = \text{f}$ for $u \in \mathcal{L}T \setminus \partial_\ell x$. Explicitly,

$$\mu_\partial(\dot{\underline{w}}_\partial) = \left\{ \prod_{u \in \partial_\ell x} \frac{\mathbf{1}\{\dot{w}_{ua(u)} \in \{\text{+}, \text{-}\}\}}{2} \right\} \left\{ \prod_{w \in \mathcal{L}T \setminus \partial_\ell x} \mathbf{1}\{\dot{w}_{wa(w)} = \text{f}\} \right\}. \tag{2.23}$$

Let ν_T be the law of the completion $\underline{w}$ of $\dot{\underline{w}}_\partial$, conditioned on the event that the completion exists:

$$\nu_T(\underline{w}) \cong \sum_{\dot{\underline{w}}_\partial} \mathbf{1}\left\{\underline{w} \text{ is the completion of } \dot{\underline{w}}_\partial \text{ on } T\right\} \mu_\partial(\dot{\underline{w}}_\partial)$$

where $\cong$ indicates proportionality up to the normalization that makes ν_T a probability measure. The measure ν_T is well-defined. Indeed, in the special case that $\text{L}_e = \text{+}$ for all edges e of T, we noted above that $\dot{\underline{w}}_\partial$ has a valid completion as long as all its entries are in $\{\text{+}, \text{f}\}$, which guarantees that ν_T is well-defined. Similarly, for general T, there is always at least one choice of $\dot{\underline{w}}_\partial$ that has a valid completion on T and has $\mu_\partial(\dot{\underline{w}}_\partial) > 0$, so ν_T is well-defined in general. By passing through the bijection (2.21), ν_T induces a probability measure on valid frozen configurations of T, which we term the **frozen model on T with rigid boundary**.[5]

This tree frozen model is characterized by a set of recursions, as follows. For any variable-clause edge (yz) in T (where either y or z is the variable), let T_{yz} be the component of $T \setminus z$ containing y — including the edge (yz), but not including z itself. If y is a variable then we call T_{yz} a **variable-to-clause tree**. If y is a clause then we call T_{yz} a **clause-to-variable tree**. In either case, we write $\mathcal{L}T_{yz}$ as shorthand for $T_{yz} \cap \mathcal{L}T$. Given inputs $\dot{w}_{va}$ for all $v \in \mathcal{L}T_{yz}$, we can apply the WP maps ((2.16) from Definition 2.8) recursively started from $\mathcal{L}T_{yz}$. As long as $\dot{\text{WP}}$ never outputs $\varnothing$, this produces all the warnings on T_{yz} in the direction of z, which we term a **completion on T_{yz}**, and denote by $\underline{w}_{T,yz}$. Let $\mu_{T,yz}$ denote the marginal of $\mu_{\mathcal{L}T}$ (as defined by (2.23)) on $\mathcal{L}T_{yz}$, and let

$$\nu_{T,yz}(\underline{w}_{T,yz}) \cong \sum_{\dot{\underline{w}}_{\mathcal{L}T_{yz}}} \mathbf{1}\left\{\underline{w}_{T,yz} \text{ is the completion of } \dot{\underline{w}}_{\mathcal{L}T_{yz}} \text{ on } T_{yz}\right\} \mu_{T,yz}(\dot{\underline{w}}_{\mathcal{L}T_{yz}}), \tag{2.24}$$

where again $\cong$ refers to the normalization. Let $\mathbf{F}_{T,yz}$ be the marginal law of the $y \to z$ warning under $\nu_{T,yz}$. We will also designate some special notation for marginal laws of warnings around the root of the tree, as follows. For $T = B_\ell(x)$ and $y \in \partial x$:

– The law $\mathbf{F}_{T,xy}$ depends only on the subtree T_{xy}, so we write $\mathbf{F}_{T,xy} \equiv \mathbf{F}_\ell(T_{xy})$, where the subscript ℓ indicates that the vertices at distance exactly ℓ from (xy) have a special role in (2.23).

---

[5]We chose the "rigid boundary" terminology because it is reasonably succinct, but note that it is not fully descriptive since the variables in $\mathcal{L}T \setminus \partial_r v$ are free rather than rigid.

– The law $\mathbf{F}_{T,yx}$ depends only on the subtree T_{yx}, so we write $\mathbf{F}_{T,yx} \equiv \mathbf{F}_{\ell-1/2}(T_{yx})$, where the subscript $\ell - 1/2$ indicates that the vertices at distance exactly $\ell - 1/2$ from (xy) have a special role in (2.23).

Finally, if U_{xy} is any tree that agrees with T_{xy} up to distance ℓ around (xy), then we define $\mathbf{F}_\ell(U_{xy}) \equiv \mathbf{F}_\ell(T_{xy})$. This notation will appear again in the following sections.

For $T = B_r(v)$, for each edge $(au) \in E_T$ where a is the clause and u is the variable, we will abbreviate $\hat{u}_{au} \equiv \mathbf{F}_{T,au}$ and $\eta_{ua} \equiv \mathbf{F}_{T,ua}$ for the marginal laws of $\hat{\mathtt{w}}_{au}$ and $\dot{\mathtt{w}}_{ua}$ induced by (2.24). These are both probability measures on $\{+, -, \mathtt{f}\}$. We also denote the scalar values $\eta_{ua} \equiv \boldsymbol{\eta}_{ua}(-)$ and $\hat{u}_{au} \equiv \hat{\boldsymbol{u}}_{au}(+)$. If a variable u lies at depth r in T (i.e., $u \in \partial_r v$), then it neighbors exactly one clause $a = a(u) \in T$, and the definition gives $\boldsymbol{\eta}_{ua} = \mathrm{unif}(\{+, -\})$ and consequently $\eta_{ua} = 1/2$. If instead the variable u is a leaf at distance less than r from v (i.e., $u \in \mathcal{L}T \setminus \partial_r v$), then the definition makes $\boldsymbol{\eta}_{ua}$ fully supported on $\{\mathtt{f}\}$, and consequently $\eta_{ua} = 0$. This defines η on every leaf edge of T. The value of η on every other edge can be computed recursively:

– For an edge $(aw) \in E_T$ where η_{ua} was already determined for all $u \in \partial a \setminus w$,

$$\hat{\boldsymbol{u}} \equiv \left(\hat{\boldsymbol{u}}_{aw}(+), \hat{\boldsymbol{u}}_{aw}(\mathtt{f})\right) = \left(\prod_{u \in \partial a \setminus w} \eta_{ua}, 1 - \prod_{u \in \partial a \setminus w} \eta_{ua}\right). \tag{2.25}$$

In particular this determines the value of $\hat{u}_{aw} \equiv \hat{\boldsymbol{u}}_{aw}(+)$.
– For an edge $(aw) \in E_T$ where $\hat{u}_{bw}$ was already determined for all $b \in \partial w \setminus a$, let

$$\Pi^+_{wa} = \prod_{b \in \partial w(+a)} \left(1 - \hat{u}_{bw}\right), \quad \Pi^-_{wa} = \prod_{b \in \partial w(-a)} \left(1 - \hat{u}_{bw}\right),$$

where $\partial w(+a)$ and $\partial w(-a)$ are defined by (2.22). Then

$$\left(\boldsymbol{\eta}_{wa}(+), \boldsymbol{\eta}_{wa}(-), \boldsymbol{\eta}_{wa}(\mathtt{f})\right) = \frac{\left(\Pi^-_{wa}(1 - \Pi^+_{wa}), \Pi^+_{wa}(1 - \Pi^-_{wa}), \Pi^+_{wa}\Pi^-_{wa}\right)}{\Pi^+_{wa} + \Pi^-_{wa} - \Pi^+_{wa}\Pi^-_{wa}}. \tag{2.26}$$

In particular this determines the value of $\eta_{wa} \equiv \boldsymbol{\eta}_{wa}(-)$.

For any non-leaf variable $w \in T \setminus \mathcal{L}T$, combining the steps above shows that η_{wa} can be expressed in terms of the η_{ub} (for $b \in \partial w \setminus a$ and $u \in \partial b \setminus w$) by the recursive relation

$$\eta_{wa} = R_{wa}(-) = \frac{\Pi^+_{wa}(1 - \Pi^-_{wa})}{\Pi^+_{wa} + \Pi^-_{wa} - \Pi^+_{wa}\Pi^-_{wa}}. \tag{2.27}$$

(Note the clear similarity between (2.27) and the "survey propagation" equations introduced in (1.10).) The probability measures $\boldsymbol{\eta}_{wa}$ and $\hat{\boldsymbol{u}}_{aw}$ can also be computed from this recursion, as indicated by (2.25) and (2.26). Finally, note that since we start from boundary input $\eta_{ua(u)} \in [0, 1)$ for all $u \in \mathcal{L}T$, we have for all other $(aw) \in E_T$ that $\Pi^\pm_{wa} \in (0, 1]$, and so $\eta_{wa} \in [0, 1)$ for all $(aw) \in E_T$. The empty product is understood to be one, so if $w \notin \mathcal{L}T$ with $\partial w(-a) = \varnothing$ then $\Pi^-_{wa} = 1$, implying $\eta_{wa} = 0$.

2.6. **Weighted models and belief propagation.** We will show at the end of this section that the frozen model recursions of §2.5 can be retrieved as a special case of the belief propagation (BP) equations for the color model. In preparation, we will briefly review BP in the slightly generalized setting of **weighted** color models, which will be used throughout our proof, and which we now introduce. We limit our discussion here to the single-copy color model; the definitions and notations generalize to the pair model (Remark 2.12) in the obvious manner.

A weighted color model on $\mathcal{G} = (V, F, E)$ is defined by multiplying (2.18) with edge weights $\gamma_e : \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\} \to (0, \infty)$ for all $e \in E$. For our purposes it will often be convenient to consider the weight γ_e on edge $e = (av)$ as "belonging" to the incident clause a: that is to say, we replace $\hat{\varphi}_a$ in (2.18) by the weighted clause factor

$$\hat{\varphi}_a(\underline{\sigma}_{\delta a}; \Gamma_a) \equiv \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{e \in \delta a} \gamma_e(\sigma_e) \quad \text{where } \Gamma_a \text{ denotes the tuple } (\gamma_e)_{e \in \delta a}.$$

At other times it is more convenient to consider the edge weight as "belonging" to the incident variable v. If v is an internal variable of $\mathcal{G}$ ($|\delta v| \geq 2$), it will be natural to parametrize the weights in a slightly different way: namely, we

set a weight λ_v for the variable's frozen model spin $x_v = \text{eval}_v(\underline{\sigma}_{\delta v}) \in \{\text{+}, \text{-}, \text{f}\}$, then set weights λ_e for the incident edge colors σ_e, so that φ_v is replaced by the weighted variable factor

$$\varphi_v(\underline{\sigma}_{\delta v}; \Lambda_v) = \varphi_v(\underline{\sigma}_{\delta v})\lambda_v(x_v) \prod_{e \in \delta v} \lambda_e(\sigma_e) \quad \text{where } \Lambda_v \text{ denotes } (\lambda_v, (\lambda_e)_{e \in \delta v}). \tag{2.28}$$

If v is a leaf variable of $\mathscr{G}$ ($|\delta v| = 1$), we will denote $\Lambda_v \equiv \lambda_v \equiv \lambda_e$ for the weight on the unique edge e incident to v. In this case we have simply $\varphi_v(\underline{\sigma}_{\delta v}; \Lambda_v) = \lambda_e(\sigma_e)$. Inserting the weighted factors into (2.18) defines a weighted measure on valid colorings of $\mathscr{G}$, which we refer to as the **weighted color model**:

$$\prod_{v \in V} \varphi_v(\underline{\sigma}_{\delta v}; \Lambda_v) \prod_{a \in F} \hat{\varphi}_a(\underline{\sigma}_{\delta a}; \Gamma_a). \tag{2.29}$$

Clearly, scaling any Γ_a or Λ_v by a positive constant has no effect other than to scale the entire measure, so we will always anchor the clause weights by fixing the convention $\gamma_e(\text{y}) = 1$. Likewise we anchor the variable weights by fixing $\lambda_v(\text{+}) = 1$ and $\lambda_e(\sigma) = 1$ for all $\sigma \in \{\text{y}, \text{g}, \text{b}\}$. We will often denote $\lambda_e \equiv \lambda_e(\text{r})$.

Let us remark that clause weights can be re-interpreted as variable weights, and vice versa, simply by shifting the factors: for example, a single-copy model consisting of Γ-weighted clauses surrounded by unweighted variables can be transformed into a model with Λ-weighted variables and unweighted clauses by setting

$$\lambda_v(\text{-}) = \prod_{e \in \delta v(\text{+})} \frac{\gamma_e(\text{y})}{\gamma_e(\text{b})} \prod_{e \in \delta v(\text{-})} \frac{\gamma_e(\text{b})}{\gamma_e(\text{y})},$$

$$\lambda_v(\text{f}) = \prod_{e \in \delta v(\text{+})} \frac{\gamma_e(\text{g})}{\gamma_e(\text{b})} \prod_{e \in \delta v(\text{-})} \frac{\gamma_e(\text{g})}{\gamma_e(\text{y})},$$

$$\lambda_e(\text{r}) = \frac{\gamma_e(\text{r})}{\gamma_e(\text{b})} \text{ for all } e \in \delta v, \tag{2.30}$$

recalling that all the other weights are fixed at one.

We now briefly review the belief propagation (BP) method in the context of the weighted color model. We shall apply BP only when the underlying bipartite factor graph is a finite tree — it is well known that BP is exact for this setting, though it is only a heuristic on more general graphs. The reader is referred to [MM09] for a detailed introduction to BP for a far broader class of models.

Let $T = (V, F, E)$ be a finite bipartite factor tree. For simplicity we assume that the leaves of T are all variables. Consider a weighted color model (2.29) on T. Provided this measure has positive mass, we can normalize it to be a probability distribution ν supported on valid colorings $\underline{\sigma} \equiv \underline{\sigma}_T$ of the tree T. The measure ν is an example of what is more generally termed a **Gibbs measure**. Belief propagation is a way of computing local marginals of a Gibbs measure ν on a finite tree T: for any subgraph $U \subseteq T$, write ν_U for the marginal of ν on U. The local marginals ν_U are efficiently computed in terms of the solution to a system of equations, known as the BP recursions or BP fixed-point equations, as we now describe.

Let $\mathscr{M}$ denote the space of probability measures over $\{\text{r}, \text{y}, \text{g}, \text{b}\}$. The BP recursions are equations defined in terms of the BP **messages** $\dot{q}, \hat{q} \in \mathscr{M}$ indexed by edges $e = (av)$:

$$\dot{q}_e \equiv \dot{q}_{va} \equiv \dot{q}_{v \to a} \equiv \text{variable-to-clause message from } v \text{ to } a$$
$$= \text{marginal law of } \sigma_{va} \text{ "in absence of } a\text{"},$$
$$\hat{q}_e \equiv \hat{q}_{av} \equiv \hat{q}_{a \to v} \equiv \text{clause-to-variable message from } a \text{ to } v$$
$$= \text{marginal law of } \sigma_{va} \text{ "in absence of } v\text{"}.$$

(The messages $\dot{q}, \hat{q}$ represent distributions over warnings $\dot{\text{w}}, \hat{\text{w}}$; see [MM09, Ch. 19].) It is well understood how to relate these messages by BP equations, which express the message outgoing from a variable (clause) across an edge as a function of the messages incoming to the variable (clause) across the other incident edges. These mappings are

parametrized by the relevant weights:

$$\dot{q}_{va}(\tau) = \text{BP}_{va}[\hat{q};\Lambda_v](\tau) \equiv \frac{1}{\dot{z}_{va}} \sum_{\underline{\sigma}_{\delta v}:\sigma_{av}=\tau} \varphi_v(\underline{\sigma}_{\delta v};\Lambda_v) \prod_{b\in\partial v\setminus a} \hat{q}_{bv}(\sigma_{bv})$$

$$\hat{q}_{av}(\tau) = \text{BP}_{av}[\dot{q};\Gamma_a](\tau) \equiv \frac{1}{\hat{z}_{av}} \sum_{\underline{\sigma}_{\delta a}:\sigma_{av}=\tau} \hat{\varphi}_a(\underline{\sigma}_{\delta a};\Gamma_a) \prod_{u\in\partial a\setminus v} \dot{q}_{ua}(\sigma_{ua}) \tag{2.31}$$

where $\dot{z}_{va}, \hat{z}_{av}$ are the normalizing constants making the output of BP a probability measure. We drop Λ,Γ from the notation to indicate the unweighted recursions. Although we generally do not write it explicitly, the normalizing constants $\dot{z}_{va}, \hat{z}_{av}$ also depend on the choice of weights.

For a leaf variable v incident to a single edge $e = (av)$, recall from above that Λ_v simply denotes the edge weight on e. In this case, the variable-to-clause BP equation (2.31) simplifies to

$$\dot{q}_{va}(\tau) = \text{BP}_{va}[\hat{q};\Lambda_v](\tau) = \frac{\lambda_e(\tau)}{\sum_{\tau'} \lambda_e(\tau')}, \tag{2.32}$$

i.e., the message from v into the graph is simply the weight on v. It follows, by recursing inwards from the leaves, that for any Gibbs measure ν on a finite tree, there is a unique solution $(\dot{q},\hat{q})$ of the BP equations. Local marginals of ν can be expressed in terms of this solution, for example, the marginal on the edges incident to a single vertex can be expressed as

$$(\text{variable } v) \quad \nu_{\delta v}(\underline{\sigma}_{\delta v}) = \nu_{\delta v}[\Lambda_v; \hat{q}_{\partial v\to v}](\underline{\sigma}_{\delta v}) \equiv \frac{1}{\dot{z}_v} \varphi_v(\underline{\sigma}_{\delta v};\Lambda_v) \prod_{a\in\partial v} \hat{q}_{av}(\sigma_{av}), \tag{2.33}$$

$$(\text{clause } a) \quad \nu_{\delta a}(\underline{\sigma}_{\delta a}) = \nu_{\delta a}[\Gamma_a; \dot{q}_{\partial a\to a}](\underline{\sigma}_{\delta a}) \equiv \frac{1}{\hat{z}_a} \hat{\varphi}_a(\underline{\sigma}_{\delta a};\Gamma_a) \prod_{v\in\partial a} \dot{q}_{va}(\sigma_{av}). \tag{2.34}$$

Taking the edge marginal from either of these gives

$$\nu_{av}(\sigma_{av}) = \nu_{av}[\dot{q}_{va}, \hat{q}_{av}](\sigma_{av}) \equiv \frac{\dot{q}_{va}(\sigma_{av})\hat{q}_{av}(\sigma_{va})}{\bar{z}_{av}}, \tag{2.35}$$

where the normalizing constant $\bar{z}_{av}$ satisfies the relations

$$\bar{z}_{av} = \frac{\dot{z}_v}{\dot{z}_{va}} = \frac{\hat{z}_a}{\hat{z}_{av}}. \tag{2.36}$$

The expressions (2.33), (2.34), and (2.35) will be used many times throughout the paper. In particular, for an edge $e = (av)$ we will often write "$\nu_e \cong \dot{q}_e\hat{q}_e$" to remind the reader of (2.35).

To conclude the section, we now explain how the tree recursions for the frozen model (§2.5) can be retrieved as a special case of the color model BP recursions. This is the only place where we will make use of the warning propagation model, as an intermediary between the frozen model and color model (cf. (2.21)). We shall keep the discussion here brief, and refer the reader to our previous works [DSS13, DSS16a] where we covered analogous correspondences in substantial detail.

In §2.5 we considered a finite tree T with variables at the leaves, and defined the "frozen model on T with rigid boundary" (Definition 2.13). On each edge (av) we defined a probability measure η_{va} on $\{+,-,\mathtt{f}\}$, as well as a probability measure $\hat{u}_{av}$ on $\{+,\mathtt{f}\}$. We now demonstrate that this corresponds to the color model on T with all vertices unweighted, except for the leaf variables where we put weights $\lambda_v(\mathtt{f}) = 0$. As before, we use $q = (\dot{q},\hat{q})$ to denote the messages in this weighted color model. We now write $h = (\dot{h},\hat{h})$ for the analogous BP messages in the warning propagation model — thus both $\dot{h}_{va}$ and $\hat{h}_{av}$ are probability measures over message pairs $\mathtt{w}_{av} \equiv (\dot{\mathtt{w}}_{va},\hat{\mathtt{w}}_{av})$. Given η, we can define h by setting $\dot{h}_{va}(\dot{\mathtt{w}}_{va},\hat{\mathtt{w}}_{av}) \cong \eta_{va}(\dot{\mathtt{w}}_{va})$ and $\hat{h}_{av}(\dot{\mathtt{w}}_{va},\hat{\mathtt{w}}_{av}) \cong \hat{u}_{av}(\hat{\mathtt{w}}_{av})$. Working out the proper normalization gives

$$\dot{h}_{va}(\dot{\mathtt{w}}_{va},\hat{\mathtt{w}}_{av}) \cong \frac{\eta_{va}(\dot{\mathtt{w}}_{va})}{2 - \eta_{va}(-)},$$

$$\hat{h}_{av}(\dot{\mathtt{w}}_{va},\hat{\mathtt{w}}_{av}) = \frac{\hat{u}_{av}(\hat{\mathtt{w}}_{av})}{3 - \hat{u}_{av}(+)} \tag{2.37}$$

Projecting down to the color model, we can define

$$\left(\dot{q}_{va}(\mathtt{r}), \dot{q}_{va}(\mathtt{y}), \dot{q}_{va}(\mathtt{g}), \dot{q}_{va}(\mathtt{b}) \right) = \left(\frac{\eta_{va}(\mathtt{+}) + \eta_{va}(\mathtt{f})}{2 - \eta_{va}(\mathtt{-})}, \frac{\eta_{va}(\mathtt{-})}{2 - \eta_{va}(\mathtt{-})}, \frac{\eta_{va}(\mathtt{f})}{2 - \eta_{va}(\mathtt{-})}, \frac{\eta_{va}(\mathtt{+})}{2 - \eta_{va}(\mathtt{-})} \right),$$

$$\left(\hat{q}_{av}(\mathtt{r}), \hat{q}_{av}(\mathtt{y}), \hat{q}_{av}(\mathtt{g}), \hat{q}_{av}(\mathtt{b}) \right) = \left(\frac{\hat{u}_{av}(\mathtt{+})}{3 - 2\hat{u}_{av}(\mathtt{+})}, \frac{\hat{u}_{av}(\mathtt{f})}{3 - 2\hat{u}_{av}(\mathtt{+})}, \frac{\hat{u}_{av}(\mathtt{f})}{3 - 2\hat{u}_{av}(\mathtt{+})}, \frac{\hat{u}_{av}(\mathtt{f})}{3 - 2\hat{u}_{av}(\mathtt{+})} \right). \tag{2.38}$$

In the above, $2 - \eta_{va}(\mathtt{-})$ and $3 - \hat{u}_{av}(\mathtt{+})$ are the normalizing constants. It is straightforward to verify that if $\boldsymbol{\eta}$ and $\hat{\boldsymbol{u}}$ are as defined in §2.5, then $\dot{q}$ and $\hat{q}$ as defined by (2.38) solve the BP equations (2.31) for the color model on T with weights $\lambda_v(\mathtt{f}) = 0$ at leaf variables v.[6] In particular, it follows from (2.32) that the message $\dot{q}_{va}$ from a leaf variable v to its neighboring clause a will be uniform over $\{\mathtt{r}, \mathtt{y}, \mathtt{b}\}$. This corresponds via (2.38) precisely to the boundary conditions $\eta_{ua(u)}(\mathtt{+}) = 1/2 = \eta_{ua(u)}(\mathtt{-})$, as specified in §2.5.

3. VARIABLE TYPES, PREPROCESSING, AND PROOF OUTLINE

In this section we formally describe the preprocessing algorithm and give a more detailed outline of the proof of the main result Theorem 1. This section is organized as follows:

– In §3.1 we define the notion of **simple types**, based on R-neighborhoods of variables in the original k-SAT graph. We also define **"canonical"** edge marginals based on r-neighborhoods of variables (for $r = R/10^4$), and define a **coherence** condition for edge marginals around a clause to be mutually compatible.
– In §3.2 we define a series of properties of simple types.
– In §3.3 we define a procedure which takes a k-SAT instance $\mathscr{G}'$, and produces a processed graph $\mathscr{G} = \mathrm{pr}\, \mathscr{G}'$ in which all variables satisfy certain desirable properties. The most important property of the processed graph is that it can be covered by regions ("enclosures") where a certain estimate (3.24) holds.
– In §3.4 we give the basic outline of the proof of the main result Theorem 1. In this subsection we will state the first and second moment results that are needed for the proof.
– In §3.5 we present preliminary results towards the first moment calculation.
– In §3.6 we present preliminary results towards the second moment calculation. This subsection also presents some of the conceptual ideas behind the second moment calculation.
– In §3.7 we give the technical details of the coherence condition from §3.1.

3.1. **Simple types and coherence.** As mentioned above, a key step in [CP16] is to condition on the degree profile of $\mathscr{G}$. In this work we condition on the empirical distribution of depth-R neighborhood types, which we regard as a "generalized degree profile" in the manner of [BC15]. We establish a satisfiability lower bound $\alpha_{\mathrm{lbd}}(R)$ for each fixed R, and show that $\alpha_{\mathrm{lbd}}(R) \to \alpha_\star$ in the limit $R \to \infty$. To be precise, the order of limits taken throughout this paper is the following: for each $k \geq k_0$ (where k_0 is a large absolute constant), take $n \to \infty$ with R fixed, then take $R \to \infty$. To this end, we fix r a (large) positive integer, and define R', R so that

$$\frac{R}{10^4} = \frac{R'}{10^2} = r. \tag{3.1}$$

Recall from Definition 2.3 that we work with the measure $\mathbb{P} \equiv \mathbb{P}^{n,\alpha}$ on bipartite factor graphs $\mathscr{G} = (V, F, E)$ (with edge signs) where $V = [n] \equiv \{1, \ldots, n\}$. We now further assign a **marking** of the variables, which is simply a uniformly random mapping

$$l \equiv l_R : V \to \left\{ 1, 2, \ldots, \left\lceil \exp(4^k R) \right\rceil \right\}. \tag{3.2}$$

From now on, the graph $\mathscr{G}$ is always understood to come equipped with the random marking l_R. (In the random graph $\mathscr{G} \sim \mathbb{P}$, a typical variable has an R-neighborhood of volume at most $(k^{O(1)} 2^k)^R$. The choice of $\lceil \exp(4^k R) \rceil$ for the range of the random markings ensures that in any given R-neighborhood, all variables receive distinct markings with chance $1 - o_R(1)$.)

---

[6]For the purposes of this paper, it suffices merely to note that (2.37) defines a valid BP solution for the warning propagation model, and projects to a BP solution for the color model. We omit the derivation of (2.37) since it is not central here, and similar correspondences were already explained in detail in [DSS13, DSS16a].

In the limit $n \to \infty$, the random graph $\mathcal{G} \sim \mathbb{P}^{n,\alpha}$ converges locally in distribution (in the sense of [BS01, AL07], and reviewed in Remark 4.2 below) to a Poisson Galton–Watson tree $\mathcal{T} = (V_{\mathcal{T}}, F_{\mathcal{T}}, E_{\mathcal{T}})$ which can be generated as follows: start with a single root variable v_{RT}, then generate offspring according to the rule that each variable independently generates $\text{Pois}(\alpha k)$ child clauses, and each clause generates $k-1$ child variables. Each edge is labelled with a literal $\mathtt{L}$ which takes values + or - with equal probability, independently over all the edges. Finally, for each variable $v \in V_{\mathcal{T}}$ assign an independent uniformly random mark

$$l(v) \equiv l_R(v) \in \left\{1, 2, \ldots, \left\lceil \exp(4^k R) \right\rceil \right\}.$$

We write $\mathbf{PGW}^\alpha$ for the resulting probability measure on rooted trees.[7]

To generalize the notion of degree distribution, we begin with a preliminary definition of neighborhood type, as follows. This definition will be augmented over the course of this section.

Definition 3.1 (simple type)**.** In a graph $\mathcal{G}$, the **simple type** t_e of a clause-variable edge $e \equiv (av) \equiv (va) \in E$ is the isomorphism class of $(B_R(v), e)$, the R-neighborhood around v rooted at edge e.[8] We write $j(t_e) \equiv j(v; a)$ to indicate the position of the edge in the clause. The **simple type** of a vertex $x \in V \cup F$ is the multi-set of simple types of all incident edges, $\{t_e : e \in \delta x\}$. Note that this has a different meaning depending on whether x is a clause or a variable:

(i) If $x \in F$ is a clause, then its simple type — which we hereafter denote L_x — is a multi-set with no repeated elements, since each edge $e \in \delta x$ has a distinct index $j(t_e) \in [k]$. Thus L_x is equivalently represented as the ordered k-tuple $(L_x(1), \ldots, L_x(k))$ where $L_x(j)$ is the type of the j-th edge in δx.
(ii) If $x \in V$ is a variable, then its simple type — which we hereafter denote T_x — may have repeated elements. It is equivalently represented as the isomorphism class of $B_R(v)$ regarded as a graph rooted at v.

We say that an edge e is **acyclic** if its simple type t_e is acyclic. We say that e is **proper** if it is acyclic, and moreover no two variables $u \neq w$ in t_e receive the same mark $l_R(u) = l_R(w)$. A vertex will be termed **acyclic** (resp. **proper**) if all its incident edges are acyclic (resp. proper). If $v \in V$ is a proper variable, then its simple type T_v has no repeated elements.

Remark 3.2. In the graph $\mathcal{G}$ sampled according to $\mathbb{P} \equiv \mathbb{P}^{n,\alpha}$, the fraction of cyclic variables will typically be around

$$\frac{2^{O(kR)}}{n} = o_n(1)$$

while the fraction of improper variables will typically be around

$$\frac{2^{O(kR)}}{\exp\{4^k R\}} = o_R(1).$$

During processing we will remove all improper variables from the graph, ensuring that the final graph will have girth greater than $2R$, since all variables remaining will be proper (hence acyclic). This further ensures that for any surviving variable v, its simple type in the initial graph is a multi-set with no repeated elements.

Definition 3.3 (directed trees)**.** If T is any bipartite factor tree and $e = (av)$ is any edge in T:

– We let T_{va} be the connected component of $T \setminus a$ that contains v. We regard T_{va} as being rooted at v, where v has parent edge e that points to the deleted clause a. We call T_{va} a **variable-to-clause tree**.
– Similarly we let T_{av} be the connected component of $T \setminus v$ that contains a. We regard T_{av} as being rooted at a, where a has parent edge e that points to the deleted clause v. We call T_{av} a **clause-to-variable tree**.

We will refer to both T_{va} and T_{av} as **directed trees**. Note that trees of this kind have already appeared previously, in the discussions of §2.5.

---

[7]The measure $\mathbf{PGW}^\alpha$ differs from the most standard definition of the Poisson Galton–Watson law only in minor details: the alternation between variables and clauses, and the presence of random edge signs and variable marks.

[8]The edge-rooted graphs (T_i, e_i), $i = 1, 2$, are isomorphic if there is a bijective graph homomorphism $\iota : T_1 \to T_2$ which maps e_1 to e_2, preserves all edge labels $\mathtt{L}_{av} \in \{\texttt{+}, \texttt{-}\}$ and indices $j(v; a) \in [k]$, and preserves all variable marks $l(v) \in \{1, \ldots, \exp(\lceil 4^k R \rceil)\}$.

Definition 3.4 (canonical messages and marginals). Recall the definition of $\mathbf{F}$ from the discussion following (2.24). For an acyclic edge $(av) \in E$, let $T = B_r(v)$, and let $_{\square}T = B_{r-1/2}(a)$ (equivalently, $_{\square}T$ is the union of $B_{r-1}(u)$ over $u \in \partial a$), so that $_{\square}T \subseteq T$ (and both are trees). Let

$$
\begin{aligned}
{}_{\star}\eta_{va} &\equiv \mathbf{F}_{T,va} = \mathbf{F}_r(T_{va}) \longleftrightarrow {}_{\star}\dot{q}_{va} \,, \\
{}_{\star}\hat{u}_{av} &\equiv \mathbf{F}_{T,av} = \mathbf{F}_{r-1/2}(T_{av}) \longleftrightarrow {}_{\star}\hat{q}_{av} \,, \\
{}_{\square}\eta_{va} &\equiv \mathbf{F}_{{}_{\square}T,va} = \mathbf{F}_{r-1}({}_{\square}T_{va}) = \mathbf{F}_{r-1}(T_{va}) \longleftrightarrow {}_{\square}\dot{q}_{va} \,, \\
{}_{\square}\hat{u}_{av} &\equiv \mathbf{F}_{{}_{\square}T,va} = \mathbf{F}_{r-1/2}({}_{\square}T_{av}) = \mathbf{F}_{r-1/2}(T_{av}) \longleftrightarrow {}_{\square}\hat{q}_{av} \,,
\end{aligned}
\tag{3.3}
$$

where $\longleftrightarrow$ indicates the correspondence (2.38). Thus ${}_{\star}\eta_{va}$ and ${}_{\square}\eta_{va}$ are probability measures over $\{+,-,\mathtt{f}\}$; and ${}_{\star}\hat{u}_{av}$ and ${}_{\square}\hat{u}_{av}$ are probability measures over $\{+,\mathtt{f}\}$. Meanwhile ${}_{\star}\dot{q}_{va}, {}_{\star}\hat{q}_{av}, {}_{\square}\dot{q}_{va}$, and ${}_{\square}\hat{q}_{av}$ are all probability measures over $\{\mathtt{r},\mathtt{y},\mathtt{g},\mathtt{b}\}$. Note that ${}_{\star}\hat{u}_{av} = {}_{\square}\hat{u}_{av}$, so ${}_{\star}\hat{q}_{av} = {}_{\square}\hat{q}_{av}$. Recalling (2.35), we define

$$
{}_{\star}\pi_{av}(\sigma) \equiv \frac{{}_{\star}\dot{q}_{va}(\sigma)\, {}_{\star}\hat{q}_{av}(\sigma)}{\sum_{\sigma'} {}_{\star}\dot{q}_{va}(\sigma')\, {}_{\star}\hat{q}_{av}(\sigma')}\,, \qquad
{}_{\square}\pi_{av}(\sigma) \equiv \frac{{}_{\square}\dot{q}_{va}(\sigma)\, {}_{\square}\hat{q}_{av}(\sigma)}{\sum_{\sigma'} {}_{\square}\dot{q}_{va}(\sigma')\, {}_{\square}\hat{q}_{av}(\sigma')}\,,
\tag{3.4}
$$

where in each case the sum in the denominator goes over $\sigma' \in \{\mathtt{r},\mathtt{y},\mathtt{g},\mathtt{b}\}$. For an acyclic edge $e = (av)$, we call ${}_{\star}\dot{q}_{va}$ and ${}_{\star}\hat{q}_{av}$ the **canonical messages**, and ${}_{\star}\pi_{av}$ the **canonical marginal**, all based on $T = B_r(v)$. We call ${}_{\square}\dot{q}_{va}$ and ${}_{\square}\hat{q}_{av}$ the **clause-based messages**, and ${}_{\square}\pi_{av}$ the **clause-based marginal**, all based on $_{\square}T = B_{r-1/2}(a) \subseteq T$.

Remark 3.5. Let $e = (av)$ be an acyclic edge, so $T = B_r(v)$ and $_{\square}T = B_{r-1/2}(a) \subseteq T$ are both trees. Note that the correspondence (2.38) implies the relations

$$
{}_{\star}\hat{q}_{av}(\mathtt{b}) = {}_{\star}\hat{q}_{av}(\mathtt{g}) = {}_{\star}\hat{q}_{av}(\mathtt{y}) = \frac{\hat{u}_{av}(\mathtt{f})}{3 - 2\hat{u}_{av}(+)}\,,
$$

$$
{}_{\star}\dot{q}_{va}(\mathtt{r}) = {}_{\star}\dot{q}_{va}(\mathtt{b}) + {}_{\star}\dot{q}_{va}(\mathtt{g}) = \frac{\eta_{va}(+) + \eta_{va}(\mathtt{f})}{2 - \eta_{va}(-)}\,,
$$

and similarly for ${}_{\square}\hat{q}$ and ${}_{\square}\dot{q}$. The canonical messages ${}_{\star}\dot{q}_{va}$ and ${}_{\star}\hat{q}_{av}$ (for $a \in \partial v$) are all based on T, which means they satisfy the **variable** BP relation ${}_{\star}\dot{q}_{va} = \mathrm{BP}_{va}[{}_{\star}\hat{q}]$. By contrast, they need **not** satisfy clause BP relations: even if all edges incident to clause a are acyclic, it is **not** necessarily the case that ${}_{\star}\hat{q}_{av} = \mathrm{BP}_{av}[{}_{\star}\dot{q}]$, because for each $u \in \partial a \setminus v$ the message ${}_{\star}\dot{q}_{ua}$ is based on a different neighborhood $B_r(u)$. On the other hand, the clause-based messages ${}_{\square}\dot{q}_{ua}$ and ${}_{\square}\hat{q}_{au}$ (for $u \in \partial a$) are all based on $_{\square}T$, which means they do satisfy the **clause** BP relation ${}_{\square}\hat{q}_{av} = \mathrm{BP}_{av}[{}_{\square}\dot{q}]$. A related observation is that if we define (cf. (2.33))

$$
\nu_{\delta v}(\underline{\sigma}_{\delta v}) = \frac{1}{\dot{z}_v} \varphi_v(\underline{\sigma}_{\delta v}) \prod_{a \in \partial v} {}_{\star}\hat{q}_{av}(\sigma_{av})
$$

where $\dot{z}_v$ denotes the normalization that makes $\nu_{\delta v}$ a probability measure, then $\nu_{\delta v}$ has edge marginals ${}_{\star}\pi_e$ for all $e \in \delta v$. On the other hand, if we define

$$
\nu_{\delta a}(\underline{\sigma}_{\delta a}) = \frac{1}{\hat{z}_a} \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{u \in \partial a} {}_{\square}\hat{q}_{ua}(\sigma_{au})\,,
$$

then $\nu_{\delta a}$ has has edge marginals $_{\square}\pi_e$ for all $e \in \delta a$. The measure $\nu_{\delta v}$ is consistent with the frozen model with rigid boundary conditions (Definition 2.13) on T, while the measure $\nu_{\delta a}$ is consistent with the frozen model with rigid boundary conditions on $_{\square}T$.

In the limit $r \to \infty$ we expect the difference between quantities based on $T = B_r(v)$ versus $_{\square}T = B_{r-1/2}(a)$ to go away. However, when working at a fixed finite radius, a major technical difficulty is to handle inconsistencies between the two. In this paper we work primarily with the quantities based on $T = B_r(v)$ (hence our choice of the term "canonical" to describe those quantities). As discussed above, the canonical messages do not, in general, satisfy clause BP relations. A closely related issue is that for $u \in \partial a$, each ${}_{\star}\pi_{au}$ is based on a different neighborhood $B_r(u)$. An important technical result of this section (stated and proved in §3.7) shows that we can **reweight** clauses such that the BP equations do hold exactly — provided the clauses are "not excessively inconsistent," as we now formalize. Following [CP16], it is useful to define **composite colors**

$$
\begin{aligned}
\mathtt{c} &\equiv \mathtt{cyan} \equiv \{\mathtt{g},\mathtt{b}\}\,, \\
\mathtt{p} &\equiv \mathtt{purple} \equiv \{\mathtt{r},\mathtt{b}\}\,.
\end{aligned}
\tag{3.5}
$$

The clause factor (2.19) does not distinguish between g and b, while the variable factor (2.17) does not distinguish between r and b; introducing the composite colors helps to simplify some parts of the analysis. The following gives our formal criterion for "consistency" within clauses:

Definition 3.6 (coherence). For a clause $a \in F$, suppose $\pi = (\pi_e)_{e \in \delta a}$ where each π_e is any probability measure over $\{r, y, g, b\}$. We say that π is **weakly coherent** if it satisfies the following:

$$\text{coher}_e(\pi) \equiv \pi_e(y) - \sum_{e' \in \delta a \setminus e} \pi_{e'}(r) \geq 0 \,, \tag{3.6}$$

for each $e \in \delta a$; and with $\pi_e(c) \equiv \pi_e(b) + \pi_e(g)$ we have

$$\text{coher}_a(\pi) \equiv \sum_{e \in \delta a} \pi_e(c) - 2 \left\{ 1 - \sum_{e \in \delta a} \pi_e(r) \right\} \geq 0 \,. \tag{3.7}$$

We say that π is **strictly coherent** if condition (3.7) holds with strict inequality, and condition (3.6) holds with strict inequality for each $e \in \delta a$ where $\pi_e(y) > 0$. We then say that the clause a is **weakly (strictly) coherent** if it is acyclic, and its canonical edge marginals $_\star\pi = (_\star\pi_e)_{e \in \delta a}$ are weakly (strictly) coherent.

A full analysis of the (weak and strict) coherence conditions is deferred to §3.7. For the purposes of the upcoming discussion, the most immediately relevant result from §3.7 is Lemma 3.49 which says that $_\square\pi$ is strictly coherent for all $r \geq 2$, and so $_\star\pi$ will also be strictly coherent if it is "close enough" to $_\square\pi$. The remainder of §3.7 is occupied with showing that strictly coherent measures can be realized by appropriate reweighting systems (in particular, see Corollaries 3.55 and 3.56).

3.2. **Classification of simple types.** In this subsection we make several classifications of (simple, acyclic) types in order to identify the vertices that must be removed during preprocessing. Recall (3.1). One aim will be to ensure that all clauses in the final processed graph are strictly coherent (Definition 3.6). For an acyclic edge $e = (av)$, recall the canonical messages and marginals of Definition 3.4 ((3.3) and (3.4)), all based on the neighborhood $B_r(v)$. Since all these quantities can be determined from the edge type $t = t_e$ (which encodes the structure of the larger neighborhood $B_R(v)$), we will freely interchange e and t in the subscripts, so for instance $_\star\pi_e \equiv _\star\pi_t$.

Definition 3.7 (stable). Suppose the edge $e = (av)$ is acyclic (Definition 3.1), so that all the messages and marginals of Definition 3.4 are well-defined. As we noted above — and will prove in Lemma 3.49 below — the measure $_\square\pi$ is strictly coherent, i.e., it satisfies all the conditions of Definition 3.6. We say that the edge $e = (av)$ is **marginal-stable** if all the following bounds hold:

$$\left| _\star\pi_e(y) - _\square\pi_e(y) \right| \leq \frac{\text{coher}_e(_\square\pi)}{5k} \,, \tag{3.8}$$

$$\left| _\star\pi_e(c) - _\square\pi_e(c) \right| \leq \frac{\text{coher}_a(_\square\pi)}{5k} \,, \tag{3.9}$$

$$\left| _\star\pi_e(r) - _\square\pi_e(r) \right| \leq \frac{1}{5k} \min \left\{ \text{coher}_a(_\square\pi), \min_{e' \in \delta a \setminus e} \text{coher}_{e'}(_\square\pi) \right\} \,, \tag{3.10}$$

$$_\star\pi_e(\sigma) \geq \frac{1}{2} _\square\pi_e(\sigma) \quad \text{for all } \sigma \in \{r, y, b, c\} \,. \tag{3.11}$$

(We impose the last condition (3.11) because it forces $\text{supp}\, _\star\pi_e = \text{supp}\, _\star\pi_e$, which is convenient for the analysis.) We say that e is **message-stable** if

$$1 - _\star\eta_{va}(-) \geq \frac{1}{2^r} \,, \qquad \max_{\sigma \in \{r, b, yg\}} \left| \frac{_\star\hat{q}_{av}(\sigma)}{\text{BP}_{av}[_\star\dot{q}](\sigma)} - 1 \right| \leq \frac{1}{k^r} \,. \tag{3.12}$$

(We take the convention $0/0 = 1$, so in the above it is permitted to have both $_\star\hat{q}_{av}(\sigma)$ and $\text{BP}_{av}[_\star\dot{q}](\sigma)$ equal to zero.) We say that the edge is **stable** if it is both marginal- and message-stable. Finally, we say that an acyclic variable v is **stable** if all its incident edges $e \in \delta v$ are stable.

Remark. Note that if all the variables in a graph are marginal-stable, then $_\star\pi$ will be strictly coherent on all the clauses in the graph. Indeed, substituting (3.8) and (3.10) into (3.6) gives

$$\text{coher}_e(_\star\pi) \geq \left(1 - \frac{1}{5}\right)\text{coher}_e(_\square\pi),$$

for all edges e in the graph. Similarly, substituting (3.9) and (3.10) into (3.7) gives

$$\text{coher}_a(_\star\pi) \geq \left(1 - \frac{3}{5}\right)\text{coher}_a(_\square\pi)$$

for all clauses a in the graph. Since $_\square\pi$ is strictly coherent (Lemma 3.49), it follows that $_\star\pi$ is also.

Definition 3.8 (nice). An acyclic variable v is **nice** if it has degrees

$$\left||\partial v(\pm)| - \frac{2^k k \ln 2}{2}\right| \leq 2^{2k/3}, \tag{3.13}$$

and its incoming and outgoing canonical messages satisfy the bounds

$$\max\left\{\left|_\star\dot{q}_{va}(\mathtt{y}) - \frac{1}{3}\right|, \left|2^k {_\star\dot{q}_{va}}(\mathtt{g}) - \frac{1}{3}\right|\right\} \leq \frac{1}{2^{k/10}}, \tag{3.14}$$

$$\max\left\{\left|_\star\hat{q}_{av}(\mathtt{y}) - \frac{1}{3}\right|, \left|2^{|\partial a|-1} {_\star\hat{q}_{av}}(\mathtt{r}) - \frac{1}{3}\right|\right\} \leq \frac{1}{2^{k/10}}, \tag{3.15}$$

for all $a \in \partial v$.[9] (In condition (3.15) we write $|\partial a|$, rather than simply k, because we will also use this definition in the processed graph where some clauses will have degree $k - 1$.) Since the canonical messages are functions of $B_r(v)$, niceness is also a property of $B_r(v)$.

Definition 3.9 (1-stable and 1-nice). We say an acyclic variable v is **1-stable** if is stable, and remains stable after the removal of any one subtree descended from a variable $u \in B_r(v) \setminus v$. The **1-nice** property is analogously defined. We use **0-nice** to mean simply **nice**. For $\mathtt{i} \in \{0, 1\}$ we let

$$D^{*,\mathtt{i}} \equiv \left\{v \in V : v \text{ is acyclic but not } \mathtt{i}\text{-nice}\right\}. \tag{3.16}$$

Note that $D^{*,0} \subseteq D^{*,1}$.

We will now identify defective regions via the following bootstrap percolation process. In a general bipartite factor graph $\mathcal{G} = (V, F, E)$, given some subset of variables $D_0 \subseteq V$, for $t \geq 1$ set $D_t \supseteq D_{t-1}$ to be the union of D_{t-1} together with all variables having at least two neighboring variables in $D_{t-1} \cap V$. The set

$$\text{BSP}(D_0; \mathcal{G}) \equiv \bigcup_{t \geq 0} D_t \tag{3.17}$$

will be termed the **bootstrap percolation** of D_0 in $\mathcal{G}$.

Definition 3.10 (defective). Let κ_* be a large absolute constant, to be determined later.[10] Recall (3.1) that $R' = 10^2 r$. Recall (3.16) that $D^{*,\mathtt{i}}$ denotes the set of variables that are acyclic but not $\mathtt{i}$-nice, for $\mathtt{i} \in \{0, 1\}$. Let $D^{\kappa_*,\mathtt{i}} \equiv B_{\kappa_*}(D^{*,\mathtt{i}})$ be the union of $B_{\kappa_*}(v)$ over all $v \in D^{*,\mathtt{i}}$. The set of $\mathtt{i}$-**defective** variables is defined as

$$\text{DEF}^\mathtt{i} \equiv \text{DEF}^\mathtt{i}(\mathcal{G}) \equiv \left\{v : v \in \text{BSP}\left(D^{\kappa_*,\mathtt{i}} \cap B_{R'/2}(v); B_{R'/2}(v)\right)\right\}.$$

We use **0-defective** and **defective** interchangeably; note $\text{DEF}^0(\mathcal{G}) \subseteq \text{DEF}^1(\mathcal{G})$. Whether a variable is $\mathtt{i}$-defective can be determined from its R'-neighborhood. We say that a clause a is $\mathtt{i}$-**defective** if all its incident variables are $\mathtt{i}$-defective. A $\mathtt{i}$-**defect** of $\mathcal{G}$ is a (maximal) connected component of $\mathtt{i}$-defective variables and clauses. Note that each $\mathtt{i}$-defect has at its boundary a buffer of nice variables of depth at least κ_*.

---

[9]Recall that we automatically have the identities $_\star\hat{q}_{av}(\mathtt{y}) = {_\star\hat{q}_{av}}(\mathtt{g}) = {_\star\hat{q}_{av}}(\mathtt{b})$ and $_\star\dot{q}_{va}(\mathtt{r}) = {_\star\dot{q}_{va}}(\mathtt{b}) + {_\star\dot{q}_{va}}(\mathtt{g})$. Thus the conditions (3.14) and (3.15) constrain $_\star\dot{q}_{va}(\sigma)$ and $_\star\hat{q}_{av}(\sigma)$ for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$.

[10]Ultimately we will require $\kappa_* \geq (240/\zeta)^4$ (see Propositions 8.1 and 8.11) where ζ is the absolute constant from Definition 7.2. Finally we will take $\zeta = \zeta_*/4$ where ζ_* is the absolute constant from an *a priori* estimate, Proposition 8.4.

For the remainder of the paper, we say **path** (in the bipartite factor graph G) to mean a finite sequence of vertices

$$P \equiv \left(v_1, a_2, v_2, \ldots, a_\ell, v_\ell \right)$$

with no repeated elements, such that each entry of the sequence is a neighbor (in G) of the previous entry. The path P contains ℓ variables, and has **length** $\ell - 1$. For acyclic variables $u, v \in V$ at distance $d(u, v) \leq R$, and $\mathsf{i} \in \{0, 1\}$, let $\mathfrak{B}^{\mathsf{i}}(u, v)$ count the variables on the (unique) shortest path from u to v (inclusive) that are i-defective. The following property is essential to our contraction argument.

Definition 3.11 (contained). Let δ_* be a small absolute constant, to be determined later.[11] For an acyclic variable v, for $\mathsf{i} \in \{0, 1\}$, and for any integer $1 \leq t \leq 2R'$, we define

$$\mathfrak{R}^{\mathsf{i}}(v, t) \equiv \sum_{u : t \leq d(u,v) < 2R'} \frac{\exp\{k(\delta_*)^{-1} \mathfrak{B}^{\mathsf{i}}(u, v)\}}{\exp\{(k \ln 2)(1 + \delta_*) d(u, v)\}} \, . \tag{3.18}$$

Note that $\mathfrak{R}^{\mathsf{i}}(v, 2R')$ is zero, since it is an empty sum. We then define the i-**containment radius** of v to be

$$\mathrm{rad}^{\mathsf{i}}(v) \equiv \min \left\{ t \geq 1 : \mathfrak{R}^{\mathsf{i}}(v, t) \leq \frac{1}{4} \right\} \leq 2R' \, , \tag{3.19}$$

so that $1 \leq \mathrm{rad}^{\mathsf{i}}(v) \leq 2R'$. For $\mathsf{i} \in \{0, 1\}$, we say v is i-**self-contained** if we have $\mathrm{rad}^{\mathsf{i}}(u) \leq d(u, v)$ for all u with $1 \leq d(u, v) \leq R'$. Note that the containment radius of any variable can be determined from its $3R'$-neighborhood. Whether a variable is i-self-contained can be determined from its $4R'$-neighborhood.

The central aim of preprocessing is to ensure that it is possible to carve up the graph into "enclosures": the formal definition is given below, but roughly speaking these will be regions of diameter at most R' such that every variable in a given enclosure has containment radius less than or equal to its minimal distance from the enclosure boundary (in particular, all the boundary variables must be self-contained). At the same time we will require another desirable property, which will be applied in the proof of Proposition 3.31 below:

Definition 3.12 (orderly). For $\mathsf{i} \in \{0, 1\}$, we say that an acyclic variable v is i-**orderly** if along any path emanating from v of length at most R', at most $(\delta_*)^3$ fraction of variables along the path are i-defective. In particular, a path leaving v of length less than $(\delta_*)^{-3}$ cannot contain any i-defective variable: that is to say, an i-**orderly** variable cannot lie within distance $(\delta_*)^{-3}$ of any defect. Since being i-defective is a property of a variable's R'-neighborhood, being i-orderly is a property of a variable's $2R'$-neighborhood.

The following definitions will be of use in carving up the graph:

Definition 3.13 (perfect; fair). For $\mathsf{i} \in \{0, 1\}$, we say that an acyclic variable v is i-**perfect** if it is both i-orderly and i-self-contained; this is a property of the variable's $4R'$-neighborhood. We say v is i-**fair** if

(i) it is i-stable;
(ii) its $5R'$-neighborhood contains no more than $\exp\{k^2(5R')\}$ variables; and
(iii) every length-R' path emanating from v contains at least one i-perfect variable.

Whether a variable is i-fair is a property of its $5R'$-neighborhood.

Lastly we wish to ensure that every type appears a linear number of times in the preprocessed graph. The following definition is towards this purpose.

Definition 3.14 (good; excellent). A rooted tree T of depth $10R'$ is i-**excellent** if, with $\cong$ denoting isomorphism of rooted graphs, and with $\mathbf{PGW} \equiv \mathbf{PGW}^\alpha$ as defined above (prior to Definition 3.1), we have

$$p_{\mathrm{ex},\mathsf{i}}(T) \equiv \mathbf{PGW} \left(\left. \begin{array}{c} B_{20R'}(u) \text{ contains any} \\ \text{variable which is not } \mathsf{i}\text{-fair} \end{array} \, \right| \, B_{10R'}(u) \cong T \right) \leq \frac{1}{\exp(k^3 R)} \, . \tag{3.20}$$

An acylic variable v is termed i-**excellent** if its $10R'$-neighborhood $B_{10R'}(v)$ satisfies (3.20). Lastly, we say that v is i-**good** if (i) it is i-fair, and (ii) every length-$40R'$ path emanating from v contains at least one i-excellent variable. Note that i-excellent implies i-good which in turn implies i-fair.

---

[11]Ultimately we will require $\delta_* \leq \min\{\zeta/30, (\ln 2)/(2(\kappa_*)^{1/2})\}$ (see Proposition 8.1) where ζ is the absolute constant from Definition 7.2. Finally we will take $\zeta = \zeta_*/4$ where ζ_* is the absolute constant from Proposition 8.4.

3.3. **Preprocessing algorithm.** We now formally describe our preprocessing procedure, which depends on the parameter R. Recall from (3.1) that $R = 10^2 R' = 10^4 r$. Recall that we write $B_\ell(v)$ to indicate the ℓ-neighborhood of variable v. In what follows, we will often write $B_\ell(v; \mathcal{G})$ to emphasize that the ℓ-neighborhood is defined with respect to the graph $\mathcal{G}$.

Definition 3.15 (removal process BSP′). In a graph $\mathcal{G} = (V, F, E)$, define the **activated set**

$$\text{Act}(\mathcal{G}) \equiv \left\{ \begin{array}{c} \text{variables } v \in V \text{ such that } B_{3R/10}(v; \mathcal{G}) \text{ contains} \\ \text{at least two clauses of degree } k - 1, \text{ or} \\ \text{at least one clause of degree} \le k - 2. \end{array} \right\}. \tag{3.21}$$

Given an initial subset of variables $A \subseteq V$, let $\mathcal{G}_{-1} \equiv \mathcal{G}$, $A_{-1} \equiv A$, and denote

$$\mathcal{G}_0 \equiv \text{pr}_0 \mathcal{G} \equiv \mathcal{G} \setminus \left\{ \bigcup_{v \in A} B_R(v; \mathcal{G}) \right\} = \mathcal{G} \setminus B_R(A; \mathcal{G}).$$

Then, for $t \ge 0$, we use (3.21) to define inductively $A_t \equiv \text{Act}(\mathcal{G}_t)$ and

$$\mathcal{G}_{t+1} \equiv \text{pr}_{t+1} \mathcal{G} \equiv \mathcal{G}_t \setminus \left\{ \bigcup_{v \in A_t} B_R(v; \mathcal{G}_t) \right\} = \mathcal{G}_t \setminus B_R(A_t; \mathcal{G}_t).$$

The process terminates at the first time t that $\text{Act}(\mathcal{G}_t) = \varnothing$, and we denote the terminal graph $\mathcal{G}_t \equiv \mathcal{G}_\infty$. We let $\text{BSP}'(A; \mathcal{G})$ denote the set of all variables removed by this procedure,

$$\text{BSP}'(A; \mathcal{G}) \equiv \bigcup_{t=-1}^{\infty} \left\{ \bigcup_{v \in A_t} B_R(v; \mathcal{G}_t) \right\} = \bigcup_{t=-1}^{\infty} B_R(A_t; \mathcal{G}_t).$$

With a minor abuse of notation we will also let $\text{BSP}'(A; \mathcal{G})$ denote the subgraph of $\mathcal{G}$ induced by these variables. We hereafter refer to $\text{BSP}'(A; \mathcal{G})$ as the **removal process in $\mathcal{G}$ with initial set** A. We then define:

Preprocessing algorithm on $\mathcal{G} = (V, F, E)$:
Let $A \subseteq V$ be the set of all variables in $\mathcal{G}$ that are
improper (Definition 3.1) or not 1-good (Definition 3.14).
Delete $\text{BSP}'(A; \mathcal{G})$ and output the processed graph $\text{pr}\,\mathcal{G} \equiv \mathcal{G}_\infty$.

Throughout this paper, we use the terms "preprocessing procedure" or "preprocessing algorithm" to refer to this mapping $\mathcal{G} \mapsto \text{pr}\,\mathcal{G}$.

The processed graph $\text{pr}\,\mathcal{G}$ is guaranteed to have certain desirable properties, as we now describe. First note that if a variable v survives in $\text{pr}\,\mathcal{G}$, then its processed neighborhood $B_{3R/10}(v; \text{pr}\,\mathcal{G})$ must be obtainable by deleting at most one subtree from its original neighborhood $B_{3R/10}(v; \mathcal{G})$. (Otherwise, v would belong to $\text{Act}(\text{pr}\,\mathcal{G})$, a contradiction since $\text{Act}(\text{pr}\,\mathcal{G})$ is by definition empty.) This means that if v was 1-nice in the original graph $\mathcal{G}$ and survives in the processed graph $\text{pr}\,\mathcal{G}$, then it must be nice (i.e., 0-nice) with respect to $\text{pr}\,\mathcal{G}$. Recalling (3.16), this directly implies $D^{*,0}(\text{pr}\,\mathcal{G}) \subseteq D^{*,1}(\mathcal{G})$. Recalling Definition 3.10, taking κ_*-neighborhoods gives

$$D^{\kappa_*,0}(\text{pr}\,\mathcal{G})) = B_{\kappa_*}\Big(D^{*,0}(\text{pr}\,\mathcal{G}); \text{pr}\,\mathcal{G}\Big) \subseteq B_{\kappa_*}\Big(D^{*,1}(\mathcal{G}); GG\Big) = D^{\kappa_*,1}(\mathcal{G})).$$

If $D \subseteq D'$ and $\mathcal{H} \subseteq \mathcal{H}'$ then $\text{BSP}(D; \mathcal{H}) \subseteq \text{BSP}(D'; \mathcal{H}')$, so it follows that

$$\text{DEF}^0(\text{pr}\,\mathcal{G}) = \left\{ v \in \text{pr}\,\mathcal{G} : v \in \text{BSP}\Big(D^{\kappa_*,0}(\text{pr}\,\mathcal{G}) \cap B_{R'/2}\big(v; \text{pr}\,\mathcal{G}\big); B_{R'/2}\big(v; \text{pr}\,\mathcal{G}\big)\Big) \right\} \subseteq \text{DEF}^1(\mathcal{G}). \tag{3.22}$$

This means that if a graph is not 1-defective in $\mathcal{G}$, then it is not defective in $\text{pr}\,\mathcal{G}$.

The inclusion (3.22) has the following implications. First, recalling Definition 3.12, it follows that if a variable v was 1-orderly in the original graph $\mathcal{G}$, and survives to the processed graph $\text{pr}\,\mathcal{G}$, then it is orderly with respect to $\text{pr}\,\mathcal{G}$. Further, recalling (3.18) from Definition 3.11, it holds for all t that

$$\mathfrak{R}^0\Big(v, t; \text{pr}\,\mathcal{G}\Big) \le \mathfrak{R}^1\Big(v, t; \mathcal{G}\Big).$$

Consequently, recalling (3.19), any variable v in $\mathrm{pr}\,\mathscr{G}$ must satisfy

$$\mathrm{rad}^0\left(v;\mathrm{pr}\,\mathscr{G}\right) \le \mathrm{rad}^1\left(v;\mathscr{G}\right). \tag{3.23}$$

In particular, this implies that if a variable $v \in \mathrm{pr}\,\mathscr{G}$ was 1-self-contained in $\mathscr{G}$, then it is self-contained in $\mathrm{pr}\,\mathscr{G}$. It follows that if $v \in \mathrm{pr}\,\mathscr{G}$ was 1-perfect in $\mathscr{G}$, then it is perfect in $\mathscr{G}$.

Since A includes all improper variables (Definition 3.1), hence all acyclic variables, the processed graph $\mathrm{pr}\,\mathscr{G}$ will have girth greater than $2R$. By the same reasoning as for the 1-nice property, it follows from Definition 3.9 that if $v \in \mathrm{pr}\,\mathscr{G}$ was 1-stable in $\mathscr{G}$, then it is stable in $\mathrm{pr}\,\mathscr{G}$. It then follows from Definition 3.13 that if $v \in \mathrm{pr}\,\mathscr{G}$ was 1-fair in $\mathscr{G}$, then it is fair in $\mathrm{pr}\,\mathscr{G}$. Since the processing removes all variables that are not 1-good (Definition 3.14), any $v \in \mathrm{pr}\,\mathscr{G}$ must have been 1-good in $\mathscr{G}$, hence also 1-fair. It follows that **all variables in $\mathrm{pr}\,\mathscr{G}$ are fair** (with respect to $\mathrm{pr}\,\mathscr{G}$). It then follows by Definition 3.13 part (iii) that any connected component in $\mathrm{pr}\,\mathscr{G}$ of non-perfect variables must have diameter at most R', so in particular it must be a tree. We can therefore carve up the graph into small regions separated by perfect variables:

Definition 3.16 (compound enclosure). A **compound enclosure** is a subgraph $U \subseteq \mathrm{pr}\,\mathscr{G}$ induced by a subset of variables $U^\circ \cup \partial U^\circ \subseteq \mathrm{pr}\,V$ where U° is a nonempty, maximal connected component of variables that are not perfect in $\mathrm{pr}\,\mathscr{G}$, and

$$\partial U^\circ \equiv \partial_\circ U \equiv \left\{ u \in \mathrm{pr}\,V : d(u, U^\circ) = 1 \right\}$$

is the external boundary of U° (which, by definition, must consist entirely of variables that are perfect in $\mathrm{pr}\,\mathscr{G}$). By the above observations, the compound enclosure must be a tree of diameter at most R'.

By construction, all compound enclosures are pairwise edge-disjoint. It is possible for two compound enclosures to share one boundary variable; they cannot have more than one variable in common because $\mathrm{pr}\,\mathscr{G}$ has girth at least $2R$. In any compound enclosure U, if $u \in \partial_\circ U$ and $v \in U^\circ$ then it follows from (3.23) together with the self-containment condition on u (Definition 3.11) that

$$\mathrm{rad}^0\left(v;\mathrm{pr}\,\mathscr{G}\right) \le \mathrm{rad}^1\left(v;\mathscr{G}\right) \le d\left(v,u;\mathscr{G}\right) \le d\left(v,u;\mathrm{pr}\,\mathscr{G}\right).$$

Minimizing the right-hand side over all $u \in \partial_\circ U$ gives

$$\mathrm{rad}^0\left(v;\mathrm{pr}\,\mathscr{G}\right) \le d\left(v,\partial_\circ U;\mathrm{pr}\,\mathscr{G}\right) \tag{3.24}$$

for all $v \in U^\circ$. The property (3.24) is essential to our analysis for compound enclosures.

Definition 3.17 (simple total type). For each edge $e = (av)$ in the processed graph $\mathrm{pr}\,\mathscr{G}$, the **simple total type** of the edge records its simple type (Definition 3.1) both before and after preprocessing: that is to say, it is the ordered pair of edge-rooted trees

$$\left(\left(B_R(v;\mathscr{G}), e \right), \left(B_R(v;\mathrm{pr}\,\mathscr{G}), e \right) \right),$$

modulo (edge-rooted) isomorphism. (As discussed in a footnote to Definition 3.1, the isomorphism must also preserve edge labels $\mathtt{L}_{av}$, indices $j(v,a)$, and marks $l(v)$.)

Definition 3.18 (compound type). For any edge e appearing inside a compound enclosure U of the processed graph $\mathrm{pr}\,\mathscr{G}$, the **compound type** of the edge records the graph structure of U with the position of e marked, as well as the simple total type of every edge e' in U. In particular, different edges appearing in the same compound enclosure U must have different compound types (since they take different positions in U), even if their simple total types match.

Definition 3.19 (total type). The **total type** $t \equiv t_e$ of an edge e in $\mathrm{pr}\,\mathscr{G}$ is defined to be its compound type if it belongs to a compound enclosure, and its simple total type otherwise. The **total type** of a variable or clause in $\mathrm{pr}\,\mathscr{G}$ is the multi-set of incident edge total types — since improper variables were removed during processing, the multi-sets are now simply sets. We hereafter use T to denote variable total types, and L to denote clause total types. For an edge $e = (av)$ of type t we write $j(t) \equiv j(v;a)$ for the position of the variable within the clause. We write $t \in L$ to indicate compatibility of types in the sense that $L(j(t)) = t$.

Definition 3.20 (processed neighborhood profile). Let $\mathscr{G}' = (V', F', E')$ be a k-SAT problem instance, and denote its processed version by $\mathscr{G} = (V, F, E) = \operatorname{pr}\mathscr{G}'$. Write $V = \{v_1, \ldots, v_n\}$ and $F = \{a_1, \ldots, a_m\}$. We then define the **processed neighborhood sequence of $\mathscr{G}'$** (equivalently, the **neighborhood sequence of $\mathscr{G}$**) as

$$\mathcal{D} \equiv \mathcal{D}_{\mathscr{G}} \equiv \begin{pmatrix} (T_{v_1}, \ldots, T_{v_n}) \\ (L_{a_1}, \ldots, L_{a_m}) \end{pmatrix}.$$

We then define $\mathscr{D}$ to be the same as $\mathcal{D}$ except that we forget the ordering of V and of F. Thus $\mathscr{D}$ contains only the information of the empirical counts

$$n_T \equiv \text{number of variables in } \operatorname{pr}V \text{ of total type } T;$$
$$m_L \equiv \text{number of clauses in } \operatorname{pr}F \text{ of total type } L;$$
$$n_t \equiv \text{number of edges in } \operatorname{pr}E \text{ of total type } t. \tag{3.25}$$

The empirical count of edge types can be determined as a marginal of either the clause or variable counts:

$$\sum_T n_T \mathbf{1}\{t \in T\} = n_t = \sum_L m_L \sum_j \mathbf{1}\{L(j) = t\}. \tag{3.26}$$

(Recall from Remark 3.2 that the simple type of any variable which survives preprocessing is a multi-set with no repeated elements — thus, in (3.26), $\mathbf{1}\{t \in T\}$ is the same as the number of occurrences of t in T.) We will sometimes abuse notation and use $\mathscr{D}$ to denote the normalized empirical measures

$$\dot{\mathscr{D}}(T) = \frac{n_T}{|V|}, \quad \hat{\mathscr{D}}(L) = \frac{m_L}{|F|}, \quad \bar{\mathscr{D}}(t) = \frac{n_t}{|E|}.$$

However, we emphasize that $\mathscr{D}$ encodes $|V| = n$ and $|F| = m$ in addition to the normalized empirical measures. We call $\mathscr{D} \equiv \mathscr{D}_{\mathscr{G}}$ the **processed neighborhood profile of $\mathscr{G}'$** (equivalently, the **neighborhood profile of $\mathscr{G}$**).

The next remark is most relevant to the precise statement of Proposition 3.24 below:

Remark 3.21. The processed graph $\mathscr{G} = \operatorname{pr}\mathscr{G}'$ is a subgraph of the original graph $\mathscr{G}'$, and therefore carries less information in general. For the rest of this paper, whenever we refer to the processed graph $\mathscr{G}$, we assume that each vertex and each edge of $\mathscr{G}$ **carries the information of its total type**. In particular, each variable in $\mathscr{G}$ does carry the information concerning its simple type (Definition 3.1), even if part of its original R-neighborhood is deleted during processing. However, $\mathscr{G}$ does **not** carry the information of the entire original graph $\mathscr{G}'$, so we still do allow for the possibility that $\operatorname{pr}\mathscr{G}' = \operatorname{pr}\mathscr{G}''$ for $\mathscr{G}' \neq \mathscr{G}''$. We let $\mathbf{CM}(\mathcal{D})$ denote the set of all graphs $\mathscr{G}$ consistent with a given neighborhood sequence $\mathcal{D}$ (where $\mathscr{G}$ is interpreted as we have just described). In particular,

$$|\mathbf{CM}(\mathcal{D})| = \prod_t (n_t)!. \tag{3.27}$$

The notation $\mathbf{CM}$ stands for **configuration model**. A uniformly random element of $\mathbf{CM}(\mathcal{D})$ can be sampled by a generalization of the standard configuration model for graphs with given degree sequence — this type of construction goes back to [Bol80], and we refer to [Wor99] for a survey. Generalized configuration models were analyzed in detail in [BC15], and the sampling procedure can be described as follows. Start with a collection of isolated vertices — n variables V together with m clauses F — labelled with total types according to $\mathcal{D}$. Each vertex is then equipped with the appropriate number of "half-edges," all labelled with edge total types. The total number of half-edges incident to each vertex corresponds to its degree in the processed graph $\mathscr{G}$. Let δV denote the variable-incident half-edges, and δF the clause-incident half-edges. Then take a **uniformly random matching** between δV and δF that **respects edge total types**. (A full edge consists of one variable-incident half-edge matched together with one clause-incident half-edge.) This procedure generates a uniformly random element of $\mathbf{CM}(\mathcal{D})$. This discussion is most relevant to the statement of Proposition 3.24 below.

The next three propositions, all proved in Section 5, give the key properties of the processed graph:

Proposition 3.22 (proved in §5.4: processing removes a small fraction of variables). *Let $\mathscr{G} \sim \mathbb{P} \equiv \mathbb{P}_{n,m}$ for m such that $|m - n\alpha| \leq n^{1/2} \ln n$. Recall that the preprocessing algorithm (Definition 3.15) removes $\operatorname{BSP}'(A; \mathscr{G})$. We have*

$$\mathbb{P}\left(\left|B_{10R}\left(\operatorname{BSP}'(A; \mathscr{G}); \mathscr{G}\right)\right| \geq \frac{n}{\exp(2^{ck}R)}\right) = o_n(1).$$

where c is a positive absolute constant (depending only on the absolute constants κ_*, δ_*). Moreover, the probability that any connected component of $\mathrm{BSP}'(A; \mathcal{G})$ contains a bicycle is also $o_n(1)$.[12]

Proposition 3.23 (proved in §5.6: each surviving total type occurs linearly often). *Let $\mathcal{G} \sim \mathbb{P} \equiv \mathbb{P}_{n,m}$ for $|m - n\alpha| \leq n^{1/2} \ln n$. Let $\mathrm{pr}\,\mathcal{G}$ be the processed graph given by Definition 3.15, with neighborhood profile $\mathcal{D}_{\mathrm{pr}\,\mathcal{G}}$. Then*

$$\mathbb{P}\left(\min\left\{ \hat{\mathscr{D}}_{\mathrm{pr}\,\mathcal{G}}(L) : L \text{ is feasible} \right\} \geq c_1 \;\middle|\; \mathrm{girth}(\mathcal{G}) > 8R \right) \geq 1 - o_n(1)$$

where c_1 is a positive constant depending on κ_*, δ_*, k, R, and the class of "feasible" L is given by Definition 5.23.

Proposition 3.24 (proved in §5.7: the processed graph is uniformly random given the neighborhood sequence). *Let $\mathcal{G} \sim \mathbb{P} \equiv \mathbb{P}_{n,m}$ for any m. Let $\mathrm{pr}\,\mathcal{G}$ be the processed graph given by Definition 3.15, with neighborhood sequence $\mathcal{D}_{\mathrm{pr}\,\mathcal{G}}$. For any $\mathcal{D}$ such that $\mathcal{D}_{\mathrm{pr}\,\mathcal{G}} = \mathcal{D}$ with positive probability under $\mathbb{P}$, we have*

$$\mathbb{P}\left(\mathrm{pr}\,\mathcal{G} = H \;\middle|\; \mathcal{D}_{\mathrm{pr}\,\mathcal{G}} = \mathcal{D} \right) = \frac{\mathbf{1}\{H \in \mathbf{CM}(\mathcal{D})\}}{\mathbf{CM}(\mathcal{D})}$$

for all $H \in \mathbf{CM}(\mathcal{D})$. Moreover, conditional on the neighborhood profile $\mathcal{D}_{\mathrm{pr}\,\mathcal{G}} = \mathscr{D}$, the law of the sequence $\mathcal{D}_{\mathrm{pr}\,\mathcal{G}}$ is uniformly random among all sequences $\mathcal{D}$ with empirical counts $\mathscr{D}$.

Write $\mathbb{P}_{\mathcal{D}}$ for the uniform measure over the set $\mathbf{CM}(\mathcal{D})$ from Remark 3.21, and write $\mathbb{E}_{\mathcal{D}}$ for expectation with respect to $\mathbb{P}_{\mathcal{D}}$. Then Proposition 3.24 tells us that

$$\mathbb{P}\left(\mathrm{pr}\,\mathcal{G} \in \cdot \;\middle|\; \mathcal{D}_{\mathrm{pr}\,\mathcal{G}} = \mathcal{D} \right) = \mathbb{P}_{\mathcal{D}}(\cdot).$$

This has the important consequence that the processed graph $\mathrm{pr}\,\mathcal{G}$, conditional on $\mathcal{D}$, can be sampled by the simple procedure described in Remark 3.21 — this is essential to our analysis of the k-SAT model on the processed graph.. Next, write $\mathcal{D} \sim \mathscr{D}$ if $\mathcal{D}$ has empirical counts given by $\mathscr{D}$. Then Proposition 3.24 also tells us that

$$\mathbb{P}\left(\mathrm{pr}\,\mathcal{G} \in \cdot \;\middle|\; \mathscr{D}_{\mathcal{G}} = \mathscr{D} \right) = \mathbb{P}_{\mathscr{D}}(\cdot) \equiv \sum_{\mathcal{D}:\mathcal{D}\sim\mathscr{D}} \frac{\mathbb{P}_{\mathcal{D}}(\cdot)}{|\{\mathcal{D}' : \mathcal{D}' \sim \mathscr{D}\}|}.$$

We will work sometimes with $\mathbb{P}_{\mathscr{D}}$ and sometimes with $\mathbb{P}_{\mathcal{D}}$, depending on convenience; they are equivalent modulo the ordering of the vertices.

3.4. **Proof outline for main theorem.** In this subsection we give a more detailed outline of the proof of the main result Theorem 1.

Definition 3.25 (empirical measures of colors conditional on types). Take any k-SAT instance $\mathcal{G}'$, and let $\mathcal{G} \equiv \mathrm{pr}\,\mathcal{G}'$ be its processed version (Definition 3.15), with total types as in Definition 3.19. Given a valid $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$-coloring $\underline{\sigma}$ of $\mathcal{G}$ (Definition 2.10), let π be the **empirical measure of colors conditioned on edge type**: for each $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$ and each edge total type t,

$$\pi_t(\sigma) \equiv \frac{|\{e \in E : t_e = t \text{ and } \sigma_e = \sigma\}|}{|\{e \in E : t_e = t\}|} = \frac{|\{e \in E : t_e = t \text{ and } \sigma_e = \sigma\}|}{n_t}, \tag{3.28}$$

with n_t as in (3.25). Further, let ω be the **empirical measure of colors conditioned on incident clause type**: for each $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$, each clause total type L, and each index $1 \leq j \leq k(L)$ (where $k(L) \in \{k-1, k\}$ is the degree in $\mathcal{G} = \mathrm{pr}\,\mathcal{G}'$ of a clause of type L), let

$$\omega_{L,j}(\sigma) \equiv \frac{|\{(av) \in E : L_a = L, j(v;a) = j, \sigma_{av} = \sigma\}|}{m_L}, \tag{3.29}$$

with m_L as in (3.25). Note that π can be obtained as a marginal of ω:

$$n_t \pi_t(\sigma) = \sum_{j \in [k]} \sum_L m_L \mathbf{1}\{L(j) = t\} \omega_{L,j}(\sigma) \tag{3.30}$$

for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$, since both sides count the total number of edges of type t with color σ in $\mathcal{G} = \mathrm{pr}\,\mathcal{G}'$.

The next two definitions are of essential importance, and are adapted from [CP16]:

---

[12]We use "bicycle" to refer to any connected graph $G' = (V', F', E')$ with $|V'| + |F'| - |E'| = -1$.

Definition 3.26 (judicious; adapted from [CP16])**.** Let $\mathscr{G}'$ be a k-sat instance, $\mathscr{G} \equiv \mathrm{pr}\,\mathscr{G}'$ the processed graph, and $\underline{\sigma}$ any valid coloring of $\mathscr{G}$ (Definition 2.10). We say that $\underline{\sigma}$ is **judicious** if holds for all L and all j that

$$\omega_{L,j} = {}_\star\pi_{L(j)}$$

— i.e., the empirical distribution of the edge color conditional on (L, j) (the clause type and edge index) **depends only on** $L(j)$ (the edge type itself, which carries less information), and moreover agrees (up to rounding) with the canonical edge marginal ${}_\star\pi$ of Definition 3.4. For this definition, ${}_\star\pi_{av}$ should be based on the r-neighborhood of v with respect to the processed graph $\mathscr{G}$.

Definition 3.27 (separable; adapted from [CP16])**.** Let $\mathscr{G}'$ be a k-sat instance, $\mathscr{G} \equiv (V, F, E) \equiv \mathrm{pr}\,\mathscr{G}'$ the processed graph, and $\underline{\sigma} \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^E$ any judicious coloring of $\mathscr{G}$ (Definition 3.26). Let $\underline{x} \equiv \underline{x}(\underline{\sigma}) \in \{\mathtt{-}, \mathtt{+}, \mathtt{f}\}^V$ denote the frozen configuration on $\mathscr{G}$ that corresponds to $\underline{\sigma}$ via (2.21). We say that $\underline{\sigma}$ is **separable** if there are not too many other judicious configurations $\underline{\sigma}'$ that are significantly correlated with $\underline{\sigma}$, that is, if

$$\left| \left\{ \text{judicious } \underline{\sigma}' : \frac{|v \in V : x(\underline{\sigma})_v = x(\underline{\sigma}')_v|}{|V|} \notin \left[\frac{1}{2}\left(1 - \frac{k^4}{2^{k/2}}\right), \frac{1}{2}\left(1 + \frac{k^4}{2^{k/2}}\right) \right] \equiv I_0 \right\} \right| \le \exp\left\{ (\ln n)^5 \right\}. \tag{3.31}$$

It is simpler and more convenient for our purposes that the correlation between $\underline{\sigma}$ and $\underline{\sigma}'$ is measured through the corresponding frozen configurations $\underline{x}(\sigma)$ and $\underline{x}(\sigma')$, rather than the colorings themselves.

Definition 3.28 (extendible; adapted from [CP16])**.** Let $\mathscr{G} = (V, F, E)$ be a (processed) k-sat instance, and $\underline{\sigma}$ any valid coloring of $\mathscr{G}$. Let $\underline{x} \equiv \underline{x}(\underline{\sigma})$ be the frozen configuration corresponding to $\underline{\sigma}$ via (2.21). We say that $\underline{\sigma}$ is **extendible** if there exists an "extension" of $\underline{x} \in \{\mathtt{-}, \mathtt{+}, \mathtt{f}\}^V$ to a satisfying assignment $\underline{\acute{x}} \in \{\mathtt{-}, \mathtt{+}\}^V$ of $\mathscr{G}$, meaning $\acute{x}_v = x_v$ for all $v \in V$ where $x_v \in \{\mathtt{-}, \mathtt{+}\}$, and $\underline{\acute{x}}$ satisfies Definition 2.2.

For the remainder of this section, $\mathscr{G}' \sim \mathbb{P} \equiv \mathbb{P}^{n,\alpha}$ is a random k-sat instance, and $\mathscr{G} \equiv \mathrm{pr}\,\mathscr{G}'$ is its processed version from Definition 3.15. We define the following quantities based on $\mathscr{G}$:

$$\begin{aligned} \boldsymbol{Z} &\equiv \boldsymbol{Z}(\mathscr{G}) \equiv \text{number of judicious colorings } \underline{\sigma} \text{ of } \mathscr{G}; \\ \boldsymbol{Z}_{\mathrm{sep}} &\equiv \boldsymbol{Z}_{\mathrm{sep}}(\mathscr{G}) \equiv \text{number of judicious separable colorings } \underline{\sigma} \text{ of } \mathscr{G}; \\ \boldsymbol{Z}_{\mathrm{ext}} &\equiv \boldsymbol{Z}_{\mathrm{ext}}(\mathscr{G}) \equiv \text{number of judicious extendible colorings } \underline{\sigma} \text{ of } \mathscr{G}. \end{aligned} \tag{3.32}$$

Clearly, $\boldsymbol{Z}_{\mathrm{ext}} \le \boldsymbol{Z}$ and $\boldsymbol{Z}_{\mathrm{sep}} \le \boldsymbol{Z}$. We emphasize that $\boldsymbol{Z}_{\mathrm{ext}}$ counts judicious extendible colorings of the processed graph $\mathscr{G}$: by Definition 3.28, such colorings extend to satisfying assignments of $\mathscr{G}$. We are ultimately interested in whether they extend to satisfying assignments of the original instance $\mathscr{G}'$; this discrepancy will be addressed below in the proof of Theorem 1 (at the end of this subsection).

Let $\mathscr{G}' \sim \mathbb{P}$, and let $\mathscr{D} = \mathscr{D}_{\mathrm{pr}\,\mathscr{G}'}$ be its processed neighborhood profile from Definition 3.20. We let $\mathscr{L}(\cdot) \equiv \mathbb{P}(\mathscr{D}_{\mathrm{pr}\,\mathscr{G}'} \in \cdot)$ denote the law of $\mathscr{D}$ itself, and let

$$\mathscr{L}_{\mathrm{girth}}(\cdot) \equiv \mathbb{P}\left(\mathscr{D}_{\mathrm{pr}\,\mathscr{G}'} \in \cdot \,\middle|\, \mathrm{girth}(\mathscr{G}') > 8R \right) \tag{3.33}$$

denote the law conditional on $\mathrm{girth}(\mathscr{G}') > 8R$ Throughout this paper, when we say that an event holds "with high probability over the random neighborhood profile $\mathscr{D}$," we mean that the $\mathscr{L}$-measure of the event is $1 - o_n(1)$. We add the caveat "conditional on $\mathrm{girth}(\mathscr{G}') > 8R$" to mean that the $\mathscr{L}_{\mathrm{girth}}$-measure of the event is $1 - o_n(1)$. We have the following propositions (assuming always that $k \ge k_0$ and α satisfies (1.7)):

Proposition 3.29 (proved in §4.6: first moment of judicious colorings matches 1-rsb formula)**.** *Let* Φ *be the* 1-rsb *free energy from Proposition 1.2. There exists* ϵ_R *which tends to zero as* $R \to \infty$, *such that*

$$\mathbb{E}_{\mathscr{D}} \boldsymbol{Z} \ge \exp\left\{ n\left[\Phi(\alpha) - \epsilon_R \right] \right\}$$

holds with high probability over the random R*-neighborhood profile* $\mathscr{D}$.

Proposition 3.30 (proved in §6.3: first moment of judicious colorings dominated by extendible colorings)**.** *For the random variables* $\boldsymbol{Z}_{\mathrm{ext}} \le \boldsymbol{Z}$ *as in* (3.32)*, we have*

$$\mathbb{E}_{\mathscr{D}}(\boldsymbol{Z}_{\mathrm{ext}}) = \left\{ 1 - o_n(1) \right\} \mathbb{E}_{\mathscr{D}} \boldsymbol{Z}$$

with high probability over the random R*-neighborhood profile* $\mathscr{D}$.

Proposition 3.31 (proved in §6.4: first moment of judicious colorings dominated by separable colorings)**.** *For the random variables* $\mathbf{Z}_{\mathrm{sep}} \leq \mathbf{Z}$ *as in* (3.32)*, we have*

$$\mathbb{E}_{\mathscr{D}}(\mathbf{Z}_{\mathrm{sep}}) = \left\{1 - o_n(1)\right\} \mathbb{E}_{\mathscr{D}} \mathbf{Z}$$

with high probability over the random R-neighborhood profile $\mathscr{D}$.

The most difficult result of this paper is a second moment estimate. To state it, let us decompose

$$\mathbf{Z}^2 \equiv \sum_z \mathbf{Z}^2[z] \tag{3.34}$$

where $\mathbf{Z}^2[z]$ denotes the contribution from pairs $(\underline{\sigma}^1, \underline{\sigma}^2)$ whose corresponding frozen configurations $\underline{x}(\underline{\sigma}^i)$ agree on exactly z fraction of the variables in $\mathscr{G}$. For any subset $I \subseteq [0,1]$, let $\mathbf{Z}^2[I]$ denote the sum of $\mathbf{Z}^2[z]$ over $z \in I$. The central part of the paper is concerned with the following estimate:

Proposition 3.32 (proved in §8.1: **main technical result**)**.** *Recall from* (3.31) *the definition of the interval $I_0 \subset [0,1]$. There is a constant $C \equiv C(k, R)$ such that*

$$\mathbb{E}_{\mathscr{D}}\left(\mathbf{Z}^2[I_0]\right) \leq C\left(\mathbb{E}_{\mathscr{D}} \mathbf{Z}\right)^2$$

with high probability over the random R-neighborhood profile $\mathscr{D}$, conditional on $\mathrm{girth}(\mathscr{G}') > 8R$.

In §3.5 and §3.6 we give further discussion on ideas of the proofs of Propositions 3.29 and 3.32. For now we turn to explaining how the above propositions imply the main result. We rely on the following well-known (and elementary) bound: if Y is any non-negative random variable with finite second moment, then for any $0 < \delta < 1$,

$$\mathbb{E}\left[Y; Y \geq \delta \,\mathbb{E}Y\right] = \mathbb{E}Y - \mathbb{E}\left[Y; Y < \delta \,\mathbb{E}Y\right] \geq (1 - \delta)\mathbb{E}Y.$$

On the other hand, by the Cauchy–Schwarz inequality,

$$\mathbb{E}\left[Y; Y \geq \delta \,\mathbb{E}Y\right] \leq \left\{\mathbb{E}(Y^2)\mathbb{P}\left(Y \geq \delta \,\mathbb{E}Y\right)\right\}^{1/2}.$$

Combining the bounds and rearranging gives

$$\mathbb{P}\left(Y \geq \delta \,\mathbb{E}Y\right) \geq (1 - \delta)^2 \frac{(\mathbb{E}Y)^2}{\mathbb{E}(Y^2)}. \tag{3.35}$$

Thus, any estimate of the form $\mathbb{E}(Y^2) \leq O((\mathbb{E}Y)^2)$ gives a lower bound $\mathbb{P}(Y \geq \delta \,\mathbb{E}Y) \geq \Omega(1)$.

Proof of Theorem 1. Propositions 1.1 and 1.2 together show that for random k-SAT with $k \geq k_0$, the 1-RSB free energy $\Phi(\alpha)$ is well-defined, with a unique root $\alpha_\star$ in the interval (1.7). It follows from Proposition 1.3 that $\alpha_\star$ upper bounds the satisfiability regime, so it remains to show the lower bound. To this end, let $\alpha < \alpha_\star$ (still within the regime (1.7)), so that $\Phi(\alpha) > 0$. We divide the rest of the argument into two parts:

Step 1. Lower bound on separable colorings. Analogously to (3.34), decompose

$$(\mathbf{Z}_{\mathrm{sep}})^2 = \sum_z (\mathbf{Z}_{\mathrm{sep}})^2[z].$$

Let $(\mathbf{Z}_{\mathrm{sep}})^2[I]$ denote the sum of $(\mathbf{Z}_{\mathrm{sep}})^2[z]$ over $z \in I \subseteq [0,1]$. By the separability condition (Definition 3.27),

$$(\mathbf{Z}_{\mathrm{sep}})^2\left[[0,1] \setminus I_0\right] \leq e^{o(n)} \mathbf{Z}_{\mathrm{sep}} \leq e^{o(n)} \mathbf{Z}$$

almost surely. Recall the definition of $\mathscr{L}_{\mathrm{girth}}$ from (3.33). Proposition 3.32 gives the bound

$$\mathscr{L}_{\mathrm{girth}}\left(\mathbb{E}_{\mathscr{D}}\left[(\mathbf{Z}_{\mathrm{sep}})^2[I_0]\right] \leq \mathbb{E}_{\mathscr{D}}\left[\mathbf{Z}^2[I_0]\right] \leq C\left(\mathbb{E}_{\mathscr{D}} \mathbf{Z}\right)^2\right) = 1 - o_n(1)$$

for $C = C(k, R)$. Combining the above bounds gives

$$\mathscr{L}_{\text{girth}}\left(\mathbb{E}_{\mathscr{D}}\left((\mathbf{Z}_{\text{sep}})^2\right) \le C\left(\mathbb{E}_{\mathscr{D}}\mathbf{Z}\right)^2 + e^{o(n)}\,\mathbb{E}_{\mathscr{D}}\mathbf{Z}\right) = 1 - o_n(1)\,. \tag{3.36}$$

On the other hand, it follows by Propositions 3.29 and 3.31 that

$$\mathscr{L}\left(\mathbb{E}_{\mathscr{D}}\mathbf{Z}_{\text{sep}} \ge \frac{\mathbb{E}_{\mathscr{D}}\mathbf{Z}}{2} \ge \frac{1}{2}\exp\left\{n\left[\Phi(\alpha) - \epsilon_R\right]\right\}\right) = 1 - o_n(1)\,. \tag{3.37}$$

Since we chose $\alpha < \alpha_\star$, we know that $\Phi(\alpha)$ is positive, and it follows for large enough R (depending on k, α) that $\mathbb{E}_{\mathscr{D}}\mathbf{Z}$ is exponentially large in n. It follows that (3.36) and (3.37) combine to yield

$$\mathscr{L}_{\text{girth}}\left(\mathbb{E}_{\mathscr{D}}\left((\mathbf{Z}_{\text{sep}})^2\right) \le 5C\left(\mathbb{E}_{\mathscr{D}}(\mathbf{Z}_{\text{sep}})\right)^2\right) = 1 - o_n(1)\,.$$

Applying (3.35), we see that there exists a positive constant $\delta = \delta(k, R)$ such that

$$\mathscr{L}_{\text{girth}}\left(\mathbb{P}_{\mathscr{D}}\left[\mathbf{Z}_{\text{sep}} \ge \delta\,\mathbb{E}_{\mathscr{D}}(\mathbf{Z}_{\text{sep}})\right] \ge \delta\right) = 1 - o_n(1)\,. \tag{3.38}$$

Step 2. Lower bound on extendible colorings. It follows from Proposition 3.30 and Markov's inequality that, with $\delta = \delta(k, R)$ as above, we have

$$\mathbb{P}_{\mathscr{D}}\left(\mathbf{Z} - \mathbf{Z}_{\text{ext}} \ge \delta^2\,\mathbb{E}_{\mathscr{D}}\mathbf{Z}\right) \le \frac{\mathbb{E}_{\mathscr{D}}(\mathbf{Z} - \mathbf{Z}_{\text{ext}})}{\delta^2\,\mathbb{E}_{\mathscr{D}}\mathbf{Z}} = o_n(1)$$

with high probability over the random neighborhood profile $\mathscr{D}$. Combining with (3.37) and (3.38) gives

$$\mathscr{L}_{\text{girth}}\left(\mathbb{P}_{\mathscr{D}}\left[\mathbf{Z}_{\text{ext}} \ge \mathbf{Z} - \delta^2\,\mathbb{E}_{\mathscr{D}}\mathbf{Z} \ge \mathbf{Z}_{\text{sep}} - \delta^2\,\mathbb{E}_{\mathscr{D}}\mathbf{Z} \ge \frac{\delta(1 - 2\delta)}{2}\exp\left\{n\left[\Phi(\alpha) - \epsilon_R\right]\right\} > 0\right] \ge \frac{\delta}{2}\right) = 1 - o_n(1)\,.$$

Recall the definition of $\mathscr{L}_{\text{girth}}$ from (3.33): since the graph $\mathscr{G}' \sim \mathbb{P}$ has $\text{girth}(\mathscr{G}') > 8R$ with asymptotically positive probability, we deduce from the above that

$$\mathscr{L}\left(\mathbb{P}_{\mathscr{D}}\left[\mathbf{Z}_{\text{ext}} > 0\right] \ge \frac{\delta}{2}\right) \ge \delta' = \delta'(k, R)\,. \tag{3.39}$$

As long as no connected component of $\mathscr{G}' \setminus \text{pr}\,\mathscr{G}'$ contains a bicycle, any satisfying assignment of $\mathscr{G} = \text{pr}\,\mathscr{G}'$ extends to a satisfying assignment of $\mathscr{G}'$. It follows by combining with Definition 3.28 that

$$\mathbf{1}\left\{\begin{array}{c}\text{no connected component of}\\ \mathscr{G}' \setminus \text{pr}\,\mathscr{G}' \text{ contains a bicycle}\end{array}\right\}\mathbf{1}\{\mathbf{Z}_{\text{ext}} > 0\} \le \mathbf{1}\{\mathscr{G}' \text{ is satisfiable}\}\,.$$

From Proposition 3.22 that, with high probability, no connected component of $\mathscr{G}' \setminus \text{pr}\,\mathscr{G}'$ contains a bicycle. Combining with (3.39) gives

$$\mathbb{P}^{n,\alpha}\left(\mathscr{G}' \text{ is satisfiable}\right) \ge \mathbb{E}\left(\mathbf{1}\left\{\begin{array}{c}\text{no connected component of}\\ \mathscr{G}' \setminus \text{pr}\,\mathscr{G}' \text{ contains a bicycle}\end{array}\right\} \cdot \mathbb{P}_{\mathscr{D}}\left[\mathbf{Z}_{\text{ext}} > 0\right]\right) \ge \delta''$$

for any $\alpha < \alpha_\star$ (where δ'' depends on k and R; and R depends on k and α). Combining with Friedgut's theorem [Fri99] gives $\alpha_{\text{sat}} \ge \alpha$. The result follows by taking $\alpha \uparrow \alpha_\star$ and $R \uparrow \infty$. $\qquad\square$

Remark 3.33. Note that Propositions 3.22–3.24 (stated in §3.3, proved in Section 5) are not directly referenced in the above proof of Theorem 1. Instead, they are used indirectly via the other propositions. In particular, for the first moment (Proposition 3.29), the uniformity result (Proposition 3.24) provides a simple combinatorial formula for $\mathbb{E}_{\mathscr{D}}\mathbf{Z}$. We then need the result that only $o_R(1)$ fraction of the variables are removed (Proposition 3.22) to relate the combinatorial formula to the 1-RSB free energy (1.11). The second moment (Proposition 3.32) similarly relies on Proposition 3.24 for the combinatorial formula, but additionally requires the result that each type occurs linearly many times (Proposition 3.23) to estimate the combinatorial formula up to a *constant* multiplicative error.

This concludes our overview of the proof of Theorem 1. The remainder of this section is organized as follows:

- In §3.5 we describe the basic ideas in the proof of of Proposition 3.29.
- In §3.6 we elaborate on some of the principles in the proof of Proposition 3.32.
- In §3.7 we prove our main claims regarding coherent clauses, thereby concluding the current section.

At the end of this section (page 57), an outline of the remainder of the paper is provided.

3.5. **First moment and the Bethe formula for colorings.** In this subsection we give some basic calculations for the proofs of Propositions 3.29 and 3.32. Again, let $\mathscr{G}'$ be a k-SAT instance, and $\mathscr{G} \equiv \mathrm{pr}\,\mathscr{G}'$ the processed graph. Recall from Definition 3.25 that if $\underline{\sigma}$ is any valid coloring of $\mathscr{G}$, then it defines two conditional empirical measures — π as in (3.28), and ω as in (3.29). We now further define:

Definition 3.34 (empirical measures of vertex-incident colorings). As before, let $\mathscr{G}'$ be a k-SAT instance, and denote its processed version $\mathscr{G} \equiv \mathrm{pr}\,\mathscr{G}'$. Given a valid coloring $\underline{\sigma}$ of $\mathscr{G}$, we define $\dot{v}$ to be the **empirical measure of variable colorings conditional on type**,

$$\dot{v}_T(\underline{\sigma}_{\delta v}) \equiv \frac{1}{n_T}\left|\left\{w \in V : T_w = T \text{ and } \underline{\sigma}_{\delta w} = \underline{\sigma}_{\delta v}\right\}\right|, \tag{3.40}$$

with n_T as in (3.25). We define $\hat{v}$ **empirical measure of clause colorings conditional on type**,

$$\hat{v}_L(\underline{\sigma}_{\delta a}) \equiv \frac{1}{m_L}\left|\left\{b \in F : L_b = L \text{ and } \underline{\sigma}_{\delta b} = \underline{\sigma}_{\delta a}\right\}\right|, \tag{3.41}$$

with m_L as in (3.25). As a shorthand, we call $v \equiv (\dot{v}, \hat{v})$ the **empirical measure (of vertex colorings)** associated to $\underline{\sigma}$. Recall from (3.30) that π can be obtained as a marginal of ω. Now note that π can also be obtained as a marginal of $\dot{v}$. Similarly, ω can be obtained as a marginal of $\hat{v}$. We say that v is **judicious** if its marginal ω is judicious in the sense of Definition 3.26. For Z as in (3.32) and any judicious empirical measure $v \equiv (\dot{v}, \hat{v})$, we let $Z[v]$ denote the contribution to Z from colorings with empirical measure v.

Remark 3.35. We sometimes abuse terminology slightly by conflating edges with edge types, and vertices with vertex types. For example, when we say that we fix a type-L clause a and consider the distribution of colorings $\underline{\sigma}_{\delta a}$, we are referring to the empirical measure of these colorings among all type-L clauses in the graph. We will use the notations $\hat{v}_L$ and $\hat{v}_a$ interchangeably for this measure.

Let $\mathscr{G}' \sim \mathbb{P}_{n,m}$ be the random k-SAT instance, and let $\mathscr{G} \equiv (V, F, E) \equiv \mathrm{pr}\,\mathscr{G}'$ be the processed graph, with neighborhood profile $\mathscr{D}$ as in Definition 3.20. For $v = (\dot{v}, \hat{v})$ as in Definition 3.34, define

$$\Phi_{\mathscr{D}}(v) = \frac{1}{n}\left\{|V|\,\mathbb{E}_{\dot{\mathscr{D}}}[\mathcal{H}(\dot{v}_T)] + |F|\,\mathbb{E}_{\hat{\mathscr{D}}}[\mathcal{H}(\hat{v}_L)] - |E|\,\mathbb{E}_{\bar{\mathscr{D}}}[\mathcal{H}(\pi_t)]\right\} \tag{3.42}$$

(where $\mathcal{H}$ denotes the usual entropy function, $\mathcal{H}(p) = -\sum_x p_x \ln p_x$). In the above expression, the expectations on the right-hand side refer to sampling total types according to $\mathscr{D}$: for instance,

$$|E|\,\mathbb{E}_{\bar{\mathscr{D}}}[\mathcal{H}(\pi_t)] \equiv |E| \sum_t \bar{\mathscr{D}}(t)\mathcal{H}(\pi_t) = -\sum_t n_t \sum_\sigma \pi_t(\sigma) \ln \pi_t(\sigma), \tag{3.43}$$

with n_t as in (3.25). For given R there are only finitely many total types possible, so v lies in a simplex of bounded dimension. Let $s_t = |\mathrm{supp}\,{}_\star\pi_t|$. Then let s_T denote the number of colorings $\underline{\sigma}_{\delta v}$ that can appear on a variable v of type T. Define likewise s_L for clause types L, and let

$$s \equiv s(\mathscr{D}) \equiv \sum_T (s_T - 1) + \sum_L (s_L - 1) - \sum_t (s_t - 1). \tag{3.44}$$

Next, recalling the statement of Proposition 3.23, we will say that "$\mathscr{D}$ **is bounded away from zero**" to mean that

$$\min\left\{\hat{\mathscr{D}}(L) : L \text{ is feasible}\right\} \geq c_1, \tag{3.45}$$

where "feasible" is specified by Definition 5.23. We then have the following:

Lemma 3.36. *Let $\mathscr{G}' \sim \mathbb{P}_{n,m}$ be the random k-SAT instance, and let $\mathscr{G} = \mathrm{pr}\,\mathscr{G}'$ be its processed version, with neighborhood profile $\mathscr{D}$ as in Definition 3.20. Let $\mathbb{P}_{\mathscr{D}}$ be the uniform measure over $\mathbf{CM}(\mathscr{D})$ (Remark 3.21), and $\mathbb{E}_{\mathscr{D}}$ the expectation with respect to $\mathbb{P}_{\mathscr{D}}$. For any $\mathscr{D}$ and any judicious ν, we have, with $\mathbf{\Phi}_{\mathscr{D}}(\nu)$ as in (3.42),*

$$\mathbb{E}_{\mathscr{D}} Z[\nu] = \frac{1}{n^{O(1)}} \exp\left\{ n\mathbf{\Phi}_{\mathscr{D}}(\nu) \right\}. \tag{3.46}$$

If $\mathscr{D}$ is bounded away from zero and ν lies in the interior of its simplex, then

$$\mathbb{E}_{\mathscr{D}} Z[\nu] \asymp_R \frac{1}{n^{s/2}} \exp\left\{ n\mathbf{\Phi}_{\mathscr{D}}(\nu) \right\} \tag{3.47}$$

for $s \equiv s(\mathscr{D})$ as in (3.44).

Proof. It follows from the description of the configuration model (Remark 3.21) that

$$\mathbb{E}_{\mathscr{D}} Z[\nu] = \underbrace{\left\{ \prod_T \binom{n_T}{n_T \dot{\nu}_T} \prod_L \binom{m_L}{m_L \hat{\nu}_L} \right\}}_{\substack{\text{number of colorings} \\ \text{prior to matching}}} \underbrace{\left\{ \prod_t \binom{n_t}{n_t \pi_t} \right\}^{-1}}_{\substack{\text{probability of matching} \\ \text{to respect colorings}}}. \tag{3.48}$$

The first estimate (3.47) then follows by Stirling's formula, ignoring polynomial corrections. The second estimate (3.47) follows by taking the polynomial corrections into account. $\qquad\square$

Returning to the form of $\mathbf{\Phi}_{\mathscr{D}}(\nu)$ in (3.42), note that if ω is fixed, then π is determined (from Definition 3.25), and we see that $\mathbf{\Phi}_{\mathscr{D}}(\nu)$ is a **strictly concave** function of ν for fixed ω. As a result, there is a unique maximizer: we let

$$\mathbf{\Psi}_{\mathscr{D}}(\omega) \equiv \mathbf{\Phi}_{\mathscr{D}}(\nu^{\mathrm{op}}[\omega]), \quad \nu^{\mathrm{op}}[\omega] \equiv \arg\max_\nu \left\{ \mathbf{\Phi}_{\mathscr{D}}(\nu) : \nu \text{ is consistent with } \omega \right\}. \tag{3.49}$$

In fact, as we discuss below, $\nu^{\mathrm{op}}[\omega]$ takes a rather explicit form (3.55), which eventually allows us to relate $\mathbf{\Psi}_{\mathscr{D}}(\omega)$ to the free energy $\Phi(\alpha)$ from (1.11). We defer this for the moment, and proceed with the calculation of $\mathbb{E}_{\mathscr{D}} Z$ in terms of $\mathbf{\Psi}_{\mathscr{D}}(\star\omega)$. Let $x_T \in \{0, 1, 2\}$ count the number of frozen spins ($+$ or $-$) that can appear on a variable v of type T. (Thus $x_T = 0$ indicates that variables of type T must always be free, while $x_T = 2$ indicates that variables of type T can take any spin in $\{-, +, \mathtt{f}\}$.) Then define

$$\wp \equiv \wp(\mathscr{D}) \equiv \sum_L \left\{ \sum_j (s_{L(j)} - 1) \right\} - \sum_T x_T \left(|\delta v| - 1 \right). \tag{3.50}$$

In the first term on the right-hand side of (3.50), the outer sum goes over all clause total types L, while the inner sum goes over $1 \le j \le k(L)$ where $k(L)$ is the degree in $\mathscr{G}$ of a clause of type L.

Corollary 3.37. *Let $\mathscr{G}' \sim \mathbb{P}_{n,m}$ be the random k-SAT instance, and let $\mathscr{G} = \mathrm{pr}\,\mathscr{G}'$ be its processed version, with neighborhood profile $\mathscr{D}$ as in Definition 3.20. If $\mathscr{D}$ is bounded away from zero in the sense of (3.45), then*

$$\mathbb{E}_{\mathscr{D}} Z \asymp_R \frac{\exp\{n\mathbf{\Psi}_{\mathscr{D}}(\star\omega)\}}{n^{\wp/2}} \tag{3.51}$$

for $\mathbf{\Psi}_{\mathscr{D}}$ as in (3.49), $\wp = \wp(\mathscr{D})$ as in (3.50), and $\star\omega$ defined by $(\star\omega)_{L,j} = \star\pi_{L(j)}$.

Proof. Again recall from (3.32) that Z counts judicious colorings (Definition 3.26) $\underline{\sigma}$ of $\mathscr{G}$. For any such $\underline{\sigma}$, the empirical measure ω (Definition 3.25) is completely fixed by the judicious condition: it must agree (up to rounding) with the measure $\star\omega$ defined by $(\star\omega)_{L,j} = \star\pi_{L(j)}$.[13] Let us abbreviate $\nu \sim \star\omega$ if ν is consistent with $\star\omega$. We will show in Lemma 3.52 (this lemma is deferred to §3.7) that

$$d_1(\mathscr{D}) \equiv \dim\left\{ \nu : \nu \sim \star\omega \right\} = s(\mathscr{D}) - \wp(\mathscr{D}), \tag{3.52}$$

---

[13]Recall from Definition 3.19 that $L(j)$ encodes the neighborhoods in both the initial graph $\mathscr{G}'$ and the final graph $\mathscr{G} = \mathrm{pr}\,\mathscr{G}'$. However, $\star\pi_{L(j)}$ should be given by Definition 3.4 applied to the r-neighborhood of the variable **in the final graph** $\mathscr{G}$.

for $s(\mathcal{D})$ as in (3.44) and $\wp(\mathcal{D})$ as in (3.50). It follows from Lemma 3.36 that

$$\mathbb{E}_{\mathcal{D}} Z = \sum_{v:v\sim_\star\omega} \mathbb{E}_{\mathcal{D}} Z[v] \asymp_R \sum_{v:v\sim_\star\omega} \frac{\exp\{n\boldsymbol{\Phi}_{\mathcal{D}}(v)\}}{n^{s/2}} \tag{3.53}$$

Now recall the general fact that for any fixed dimension d, in the limit $n \to \infty$ we have

$$\sum_{x\in(\mathbb{Z}/n)^d} \frac{1}{\exp(n\|x\|^2)} = \sum_{x\in(\mathbb{Z}/n^{1/2})^d} \frac{1}{\exp(\|x\|^2)} \asymp_d n^{d/2}. \tag{3.54}$$

Since we already noted above that $\boldsymbol{\Phi}_{\mathcal{D}}(v)$ is strictly concave in v for fixed ω, the claimed result follows by applying (3.54) to (3.53), with dimension d given by $s(\mathcal{D}) - \wp(\mathcal{D})$ as in (3.52). $\qquad\square$

Let us now return to the explicit optimization problem (3.49). Suppose ω is given, so that π is determined by the relation (3.30). The optimal $\dot{v}^{\mathrm{op}}[\omega]$ depends on ω only through π: it is given by finding, for each variable type T, the measure $\dot{v}_T$ (on valid colorings of such a variable) that maximizes entropy and is consistent with marginals π. As in Remark 3.35, let us abuse notation and write $\dot{v}_T \equiv \dot{v}_v$ where v is a variable of type T, and $\pi_t \equiv \pi_e$ where e is an edge of type t. Then

$$(\dot{v}^{\mathrm{op}}[\omega])_v = \arg\max_v \left\{ \mathcal{H}(\dot{v}_v) : \pi_e(\sigma) = \sum_{\underline{\sigma}_{\delta v}} \mathbf{1}\{\sigma_e = \sigma\}\dot{v}_v(\underline{\sigma}_{\delta v}) \text{ for all } e \in \delta v, \sigma \in \{\mathtt{r},\mathtt{y},\mathtt{g},\mathtt{b}\} \right\}.$$

The associated Lagrangian is given by

$$-\sum_{\underline{\sigma}_{\delta v}} \dot{v}_v(\underline{\sigma}_{\delta v}) \ln \dot{v}_v(\underline{\sigma}_{\delta v}) + \sum_{e\in\delta v} \sum_\sigma \lambda_e(\sigma)\left[\sum_{\underline{\sigma}_{\delta v}} \mathbf{1}\{\sigma_e = \sigma\}\dot{v}_v(\underline{\sigma}_{\delta v}) - \pi_e(\sigma) \right],$$

where the first sum goes over $\underline{\sigma}_{\delta v}$ for which $\varphi_v(\underline{\sigma}_{\delta v})$, as defined by (2.20), equals one. It follows that

$$(\dot{v}^{\mathrm{op}}[\omega])_v(\underline{\sigma}_{\delta v}) \cong \varphi_v(\underline{\sigma}_{\delta v}) \prod_{e\in\delta v} q_e(\sigma_e) \tag{3.55}$$

where $\cong$ indicates the overall normalization, and the q_e are probability measures on $\{\mathtt{r},\mathtt{y},\mathtt{g},\mathtt{b}\}$ such that $\dot{v}^{\mathrm{op}}[\omega]$ satisfies the constraint of having marginals consistent with π. Note the clear resemblance between (2.33) and (3.55). The above is for general ω; in the main special case of interest we have:

Lemma 3.38. *For any ω with marginals $_\star\pi$, $\dot{v}^{\mathrm{op}}[\omega]$ is given by (3.55) with $q_e = {}_\star\hat{q}_e$ for all e, that is to say, by*

$$_\star\dot{v}_v(\underline{\sigma}_{\delta v}) \equiv \frac{1}{\dot{z}_v} \varphi_v(\underline{\sigma}_{\delta v}) \prod_{e\in\delta v} {}_\star\hat{q}_e(\sigma_e) \tag{3.56}$$

where $\dot{z}_v$ is the normalizing constant.

Proof. It suffices to verify that

$$_\star\dot{v}_v(\sigma_{av} = \sigma) \cong {}_\star\hat{q}_{av}(\sigma_{av}) \sum_{\underline{\sigma}_{\delta v}} \mathbf{1}\{\sigma_{av} = \sigma\}\varphi_v(\underline{\sigma}_{\delta v}) \prod_{b\in\partial v\setminus a} {}_\star\hat{q}_{bv}(\sigma_{bv})$$

$$\cong {}_\star\hat{q}_{av}(\sigma_{av})\left\{ \left(\mathrm{BP}_{va}[{}_\star\hat{q}]\right)(\sigma_{av}) \right\} = {}_\star\hat{q}_{av}(\sigma_{av}){}_\star\dot{q}_{va}(\sigma_{av}) \cong {}_\star\pi_{av}(\sigma), \tag{3.57}$$

where BP is the belief propagation mapping for the color model (the unweighted version of (2.31)), and the last two identities follow from the fact that $_\star\pi_{av}$, $_\star\dot{q}_{va}$, and $_\star\hat{q}_{av}$ are all based on the **same** tree, namely, the r-neighborhood of v in $\mathscr{G}$ (see also Remark 3.5). This verifies that the marginals of $_\star\dot{v}$ are indeed consistent with $_\star\pi$. Since $_\star\dot{v}$ takes the form (3.55) given by the Lagrangian calculation, it follows that $_\star\dot{v} = \dot{v}^{\mathrm{op}}[\omega]$. $\qquad\square$

In the optimization (3.49), if ω is given, then the optimal $\hat{v}^{\mathrm{op}}[\omega]$ solves a similar problem of maximizing entropy subject to marginals ω. It is important to note, however, that the analogue of Lemma 3.38 does **not** hold for the clause measure: for $(_\star\omega)_{L,j} = {}_\star\pi_{L(j)}$, it is **not** necessarily the case that $\hat{v}^{\mathrm{op}}[_\star\omega]$ is given by (cf. (2.34))

$$_\star\hat{v}_a(\underline{\sigma}_{\delta a}) = \frac{1}{\hat{z}_a}\hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{v\in\partial a} {}_\star\dot{q}_{va}(\sigma_{av}). \tag{3.58}$$

This is simply because $_\star\hat{v}$ need not be consistent with $_\star\omega$: a similar calculation as (3.57) gives

$$_\star\hat{v}_a(\sigma_{av} = \sigma) \cong {}_\star\dot{q}_{va}(\sigma_{av})\left\{\left(\mathrm{BP}_{av}[_\star\dot{q}]\right)(\sigma_{av})\right\}.$$

This generally does not match $_\star\pi_{av}(\sigma_{av})$ since, as noted in Remark 3.5, $\mathrm{BP}_{av}[_\star\dot{q}]$ need not equal $_\star\hat{q}_{av}$. A significant part of the proof of Proposition 3.29 is concerned with this discrepancy between the explicit measure $_\star\hat{v}$ of (3.58) and the optimizer $\hat{v}^{\mathrm{op}}[_\star\omega]$. This calculation is deferred to §4.6.

Given the above discussion, it is now easy to *guess* what value $n^{-1}\ln\mathbb{E}_{\mathscr{D}}Z$ should concentrate around, in the limit $n \to \infty$ followed by $R \to \infty$. Since the fraction of variables removed by preprocessing is $o_R(1)$ (Proposition 3.22), the measure $\mathscr{D}$ should concentrate around the Galton–Watson measure that is the local weak limit of the original k-SAT graph $\mathscr{G}'$ (for the details see Definition 4.1). The discrepancy between $_\star\hat{v}$ and $\hat{v}^{\mathrm{op}}[_\star\omega]$ should also go away as $R \to \infty$. The canonical messages $_\star\dot{q}$ and $_\star\hat{q}$ should converge, in a distributional sense, to the limiting measures on the Galton–Watson tree. We formalize this as follows:

Definition 3.39 (random messages for the color model). Let μ be the measure given by Proposition 1.1, and let $\underline{\eta}$ be an array of i.i.d. samples from μ. Let $\underline{d} \equiv (d^+, d^-)$ be two independent samples from the $\mathrm{Pois}(\alpha k/2)$ distribution. Define $\Pi^\pm \equiv \Pi^\pm(\underline{d}, \underline{\eta})$ as in (1.9), and define (cf. (1.10))

$$\eta(+) = \frac{(1 - \Pi^+)\Pi^-}{\Pi^+ + \Pi^- - \Pi^+\Pi^-}, \quad \eta(-) = \frac{(1 - \Pi^-)\Pi^+}{\Pi^+ + \Pi^- - \Pi^+\Pi^-}, \quad \eta(\mathtt{f}) = \frac{\Pi^+\Pi^-}{\Pi^+ + \Pi^- - \Pi^+\Pi^-},$$

so that η is a (random) probability measure over $\{-, +, \mathtt{f}\}$. Now let η_j be i.i.d. copies of η, and define (cf. (2.25))

$$\hat{u} \equiv \left(\hat{u}(+), \hat{u}(\mathtt{f})\right) = \left(\prod_{j=1}^{k-1}\eta_j(-), 1 - \prod_{j=1}^{k-1}\eta_j(-)\right). \tag{3.59}$$

We substitute the random measures η and $\hat{u}$ into (2.38) to define random messages for the color model:

$$\dot{q}(\eta) = (\dot{q}(\mathtt{r}), \dot{q}(\mathtt{y}), \dot{q}(\mathtt{g}), \dot{q}(\mathtt{b})) = \frac{(\eta(+) + \eta(\mathtt{f}), \eta(-), \eta(\mathtt{f}), \eta(+))}{2 - \eta(-)}, \tag{3.60}$$

$$\hat{q}[(\eta_j)_{j\geq 1}] = (\hat{q}(\mathtt{r}), \hat{q}(\mathtt{y}) = \hat{q}(\mathtt{g}) = \hat{q}(\mathtt{b})) = \frac{(\hat{u}(+), \hat{u}(\mathtt{f}))}{3 - 2\hat{u}(+)}.$$

Let μ^{col} denote the law of $\dot{q}$, and let $\hat{\mu}^{\mathrm{col}}$ denote the law of $\hat{q}$.

Definition 3.40 (Bethe free energy of color model). Given a sequence of probability measures $\underline{\dot{q}} = (\dot{q}_1, \ldots, \dot{q}_k)$ on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$, we define a probability measure on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^k$ by (cf. (2.34) and (3.58))

$$\left(\hat{v}[\underline{\dot{q}}]\right)(\sigma_1, \ldots, \sigma_k) \equiv \frac{\hat{\varphi}_a(\sigma_1, \ldots, \sigma_k)}{\hat{z}[\underline{\dot{q}}]} \prod_{i=1}^{k}\dot{q}_i(\sigma_i),$$

with $\hat{\varphi}_a$ given by (2.19). Similarly, given integers $\underline{d} \equiv (d^+, d^-)$ with $d = d^+ + d^-$, and an array $\underline{\hat{q}} \equiv ((\hat{q}^+)_i, (\hat{q}^-)_i)_i$ of probability measures on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$, we define a probability measure on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^d$ by (cf. (2.34) and (3.58))

$$\left(\dot{v}[\underline{d}, \underline{\hat{q}}]\right)(\sigma_1, \ldots, \sigma_d) \equiv \frac{\varphi_v(\sigma_1, \ldots, \sigma_d)}{\dot{z}[\underline{d}, \underline{\hat{q}}]}\left(\prod_{i=1}^{d^+}(\hat{q}^+)_i(\sigma_i)\right)\left(\prod_{i=d^++1}^{d}(\hat{q}^-)_i(\sigma_i)\right),$$

with φ_v given by (2.20) for a variable in which the first d^+ incident edges have the + sign, and the remaining d^- incident edges have the – sign. Finally, given two measures $\dot{q}$ and $\hat{q}$ on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$, define (cf. (2.35) and (3.4))

$$\left(\bar{v}[\dot{q}, \hat{q}]\right)(\sigma) \equiv \frac{\dot{q}(\sigma)\hat{q}(\sigma)}{\bar{z}[\dot{q}, \hat{q}]}.$$

Abbreviate $\mathrm{po}_{\pm}$ for the law of $\underline{d}$, and recall μ^{col} and $\hat{\mu}^{\mathrm{col}}$ from Definition 3.39. Define

$$H^{\mathrm{col,v}}(\alpha) \equiv \int \mathcal{H}\Big(\dot{v}[\underline{d}, \hat{\underline{q}}]\Big)\, d\mathrm{po}_{\pm}(\underline{d})\, d(\hat{\mu}^{\mathrm{col}})^{\otimes}(\hat{\underline{q}})\,,$$

$$H^{\mathrm{col,cl}}(\alpha) \equiv \int \mathcal{H}\Big(\hat{v}[\dot{\underline{q}}]\Big)\, d(\mu^{\mathrm{col}})^{\otimes}(\dot{\underline{q}})\,,$$

$$H^{\mathrm{col,e}}(\alpha) \equiv \int \int \mathcal{H}\Big(\bar{v}[\dot{q}, \hat{q}]\Big)\, d\mu^{\mathrm{col}}(\dot{q})\, d\hat{\mu}^{\mathrm{col}}(\hat{q})\,.$$

(The dependence on α is through the law of $\underline{d}$ as well as the measures μ^{col} and $\hat{\mu}^{\mathrm{col}}$.) The **Bethe free energy of the color model** is given by (cf. (3.42))

$$\Phi^{\mathrm{col}}(\alpha) \equiv H^{\mathrm{col,v}}(\alpha) + \alpha\Big\{H^{\mathrm{col,cl}}(\alpha) - k\, H^{\mathrm{col,e}}(\alpha)\Big\}\,. \tag{3.61}$$

An equivalent (and more commonly seen) expression is given by defining

$$\Phi^{\mathrm{col,v}}(\alpha) \equiv \int \ln \dot{z}[\underline{d}, \hat{\underline{q}}]\, d\mathrm{po}_{\pm}(\underline{d})\, d(\hat{\mu}^{\mathrm{col}})^{\otimes}(\hat{\underline{q}})\,,$$

$$\Phi^{\mathrm{col,cl}}(\alpha) \equiv \int \ln \hat{z}[\dot{\underline{q}}]\, d(\mu^{\mathrm{col}})^{\otimes}(\dot{\underline{q}})\,,$$

$$\Phi^{\mathrm{col,e}}(\alpha) \equiv \int \int \ln \bar{z}[\dot{q}, \hat{q}]\, d\mu^{\mathrm{col}}(\dot{q})\, d\hat{\mu}^{\mathrm{col}}(\hat{q})\,. \tag{3.62}$$

It is straightforward to check that

$$\Phi^{\mathrm{col}}(\alpha) = \Phi^{\mathrm{col,v}}(\alpha) + \alpha\Big\{\Phi^{\mathrm{col,cl}}(\alpha) - k\, \Phi^{\mathrm{col,e}}(\alpha)\Big\}\,. \tag{3.63}$$

We will show in §4.6 (Lemma 4.31) that $\Phi^{\mathrm{col}}(\alpha)$ is exactly the same as the 1-RSB free energy $\Phi(\alpha)$ from (1.11).

In §4.6 we will complete the proof of Proposition 3.29 by showing that the quantity $\mathbf{\Psi}_{\mathscr{D}}(\omega)$ (as defined by (3.49), and appearing in the estimate of Corollary 3.37) is lower bounded by a quantity that tends to $\Phi^{\mathrm{col}}(\alpha) = \Phi(\alpha)$ in the limit $n \to \infty$ and $R \to \infty$.

3.6. **Second moment and constrained entropy maximization.** In this subsection we introduce some of the core principles of the proof of Proposition 3.32. As before, let $\mathscr{G}'$ denote the original k-SAT instance, and $\mathscr{G} \equiv \mathrm{pr}\,\mathscr{G}'$ its processed version. Recall (3.32) that $\mathbf{Z} \equiv \mathbf{Z}(\mathscr{G})$ counts all colorings of $\mathscr{G}$ that are judicious (Definition 3.26). In Proposition 3.32 we seek to calculate the expected value, under the measure $\mathbb{P}_{\mathscr{D}}$, of

$$\mathbf{Z}^2[I_0] \equiv \sum_{z \in I_0} \mathbf{Z}^2[z]\,, \tag{3.64}$$

for I_0 as defined by (3.31), and $\mathbf{Z}^2[z]$ as defined by (3.34).

Throughout this subsection, we let $\underline{\sigma}$ denote a pair $(\underline{\sigma}^1, \underline{\sigma}^2)$ where each $\underline{\sigma}^i$ is a judicious coloring of $\mathscr{G}$. Given any such $\underline{\sigma}$, let π and ω be defined analogously to (3.28) and (3.29) from Definition 3.25, except that now edge spins take values in $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$ rather than $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$. We hereafter refer to $\omega \equiv (\omega_{L,j})_{L,j}$ as the **pair empirical measure (on edges)**. Let $\mathbf{Z}^2(\omega)$ denote the contribution to $\mathbf{Z}^2$ from configurations $\underline{\sigma} \equiv (\underline{\sigma}^1, \underline{\sigma}^2)$ with empirical measure ω. Given $\underline{\sigma}$, we can also define the vertex pair empirical measure $\nu \equiv (\dot{\nu}, \hat{\nu})$ as in (3.40) and (3.41), except with pairs of colors instead of single colors on each edge. Let $\mathbf{Z}^2[\nu]$ denote the contribution to $\mathbf{Z}^2$ from pairs $\underline{\sigma}$ that are consistent with ν. In the pair coloring model, as in the single-copy model, the edge empirical measure ω can be determined as a function of the vertex empirical measure ν (in fact, ω can be determined from $\hat{\nu}$ alone).

Definition 3.41 (judicious pair empirical measures). As $\mathbf{Z}$ is defined to count only judicious configurations, in order for the contribution $\mathbf{Z}^2(\omega)$ to be non-zero, ω must satisfy two properties. First, the single-copy marginals of ω must agree with the measure $_\star\omega$ defined in the statement of Corollary 3.37: that is to say, for both $i = 1, 2$, we must have

$$_\star\omega_{L,j}(\sigma^i) = (\omega_{L,j})^j(\sigma^i) \equiv \sum_{\tau \in \{\mathtt{r},\mathtt{y},\mathtt{g},\mathtt{b}\}^2} \mathbf{1}\{\tau^i = \sigma^i\}\omega_{L,j}(\tau) \tag{3.65}$$

for all L, j. Secondly, ω must arise as the marginal of a valid vertex measure $\nu = (\dot{\nu}, \hat{\nu})$. We say that ω is **judicious** if it satisfies both these properties. (The second property is important mainly for computing the dimension of the space of feasible ω, as we will see in the proof of Lemma 3.53 below.)

With the above notations, we can refine the above decomposition (3.64) as

$$\mathbf{Z}^2[I_0] = \sum_{\omega \in I_0} \mathbf{Z}^2(\omega)\,,$$

where I_0 denotes the subset of measures ω that are judicious and consistent with $z \in I_0$. Throughout what follows, we will use the term "pair coloring model" for the two-copy version of (2.18), with factors

$$\varphi_{v,2}(\underline{\sigma}_{\delta v}) \equiv \prod_{j=1,2} \varphi_v((\underline{\sigma}^j)_{\delta v}))\,,$$

$$\varphi_{a,2}(\underline{\sigma}_{\delta a}) \equiv \prod_{j=1,2} \hat{\varphi}_a((\underline{\sigma}^j)_{\delta a})\,. \tag{3.66}$$

Let $\mathbf{\Phi}_{\mathscr{D},2}(\nu)$ and $\mathbf{\Psi}_{\mathscr{D},2}(\omega)$ be the analogues of (3.42) and (3.49) for the pair coloring model. We then have the following extension of Lemma 3.36:

Lemma 3.42. *Denote $_*\omega \equiv {}_\star\omega \otimes {}_\star\omega$ for $_\star\omega$ as in Corollary 3.37. Suppose that if we restrict to $\omega \in I_0$, the function $\mathbf{\Psi}_{\mathscr{D},2}$ is uniquely maximized at $_*\omega$, with negative-definite Hessian. Then*

$$\mathbb{E}_{\mathscr{D}}\left(\mathbf{Z}^2[I_0]\right) \asymp_R \left(\mathbb{E}_{\mathscr{D}}\mathbf{Z}\right)^2\,,$$

i.e., the conclusion of Proposition 3.32 holds.

Proof. The same calculation leading to (3.47) gives (in the interior of the simplex of feasible ν) a similar formula

$$\mathbb{E}_{\mathscr{D}}\mathbf{Z}^2[\nu] \asymp_R \frac{1}{n^{s_2}} \exp\left\{n\mathbf{\Phi}_{\mathscr{D},2}(\nu)\right\}\,,$$

where $s_2 = s_2(\mathscr{D})$ takes into account the polynomial corrections from the Stirling approximation, and is the analogue of (3.44) for the pair model:

$$s_2 = s_2(\mathscr{D}) = \sum_T \left((s_T)^2 - 1\right) + \sum_L \left((s_L)^2 - 1\right) - \sum_t \left((s_t)^2 - 1\right)\,. \tag{3.67}$$

Let us write $\nu \sim \omega$ if ν is consistent with ω. In Lemma 3.53 (deferred to §3.7), we will calculate

$$d_2(\mathscr{D}) \equiv \dim\left\{\nu : \nu \sim \omega\right\}\,. \tag{3.68}$$

for ω close to $_*\omega$. Then, similarly to (3.51) and (3.53), we have

$$\mathbb{E}_{\mathscr{D}}\mathbf{Z}^2(\omega) = \sum_{\nu:\nu\sim\omega} \mathbb{E}_{\mathscr{D}}\mathbf{Z}^2[\nu] \asymp_R \sum_{\nu:\nu\sim\omega} \frac{1}{n^{s_2/2}} \exp\left\{n\mathbf{\Phi}_{\mathscr{D},2}(\nu)\right\} \asymp_R \frac{n^{d_2/2}}{n^{s_2/2}} \exp\left\{\mathbf{\Psi}_{\mathscr{D},2}(\omega)\right\}\,.$$

Finally, we will show in Lemma 3.54 (also deferred to §3.7) that

$$j_2 \equiv j_2(\mathscr{D}) \equiv \dim\left\{\omega : \omega \text{ is judicious}\right\} = s_2(\mathscr{D}) - d_2(\mathscr{D}) - 2\wp(\mathscr{D}) \tag{3.69}$$

for $\wp$ as in (3.50). If $\mathbf{\Psi}_{\mathscr{D},2}$ satisfies the conditions of the lemma, then the gaussian summation estimate (3.54) gives

$$\mathbb{E}_{\mathscr{D}}\mathbf{Z}^2[I_0] \asymp_R \frac{1}{n^\wp} \exp\left\{\mathbf{\Psi}_{\mathscr{D},2}(_*\omega)\right\} \stackrel{\odot}{=} \left(\frac{1}{n^{\wp/2}} \exp\left\{\mathbf{\Psi}_{\mathscr{D}}(_\star\omega)\right\}\right)^2 \asymp_R \left(\mathbb{E}_{\mathscr{D}}\mathbf{Z}\right)^2\,.$$

(The step marked $\odot$ uses the identity $\mathbf{\Psi}_{\mathscr{D},2}(_*\omega) = 2\mathbf{\Psi}_{\mathscr{D}}(_\star\omega)$ which is easy to verify.) $\qquad\square$

Thus, in order to prove Proposition 3.32, it suffices to verify the condition of Lemma 3.42. In the remainder of this subsection, we show that this condition can be reduced to solving a family of constrained entropy maximization problems on finite trees, which we define next. We will separately consider two cases, one for compound enclosures and one for non-compound variables. The solution of the optimization problems occupies most of Sections 7–9.

Entropy maximization problem for compound enclosures. We discuss the case of compound enclosures first. Although these regions are more complicated in the sense that they contain defective variables, we have an important advantage in that the notion of compound type (Definition 3.18) encodes the structure of the entire enclosure. This allows us to reduce the analysis of each type of compound enclosure to an optimization problem concerning colorings of a fixed tree with fixed edge types (Proposition 3.44 below). By contrast, to obtain an analogous statement for non-compound variables, we will have to consider a more complicated optimization problem where some of the edge types on the tree can vary (Proposition 3.46 below).

Definition 3.43 (judicious measures on trees). Let U be a finite bipartite factor tree, with all vertices and edges labelled by (mutually compatible) **compound** total types, such that all the leaves of U are variables. Let $\partial_\circ U$ be a designated nonempty subset of leaf variables. We use δU to denote the edges incident to $\partial_\circ U$. Define the simplex of probability measures

$$\Delta(U) \equiv \left\{ \begin{array}{c} \text{probability measures } \nu \text{ over pair} \\ \text{colorings } \underline{\sigma} \equiv (\underline{\sigma}^1, \underline{\sigma}^2) \text{ of } U \end{array} \right\} .$$

We say that $\nu \in \Delta(U)$ is **judicious** if all of its edge marginals match the canonical marginals of Definition 3.4: that is, for all edges e in U and for both $j = 1, 2$, we have

$$\star \pi_e(\sigma^j) = (\nu_e)^j(\sigma^j) \equiv \nu((\sigma_e)^j = \sigma^j) \equiv \sum_\tau \mathbf{1}\{\tau^j = \sigma^j\}\nu(\tau_e = \tau)$$

for all $\sigma^j \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$. Note that, since all the edges in U are assumed to be of compound type, this is equivalent to saying that all the edge marginals ν_e satisfy the condition (3.65) from Definition 3.41. Let ω denote a tuple $(\omega_e)_e$ where ω_e is a judicious probability measure on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$ for each edge e in U. We then let

$$J(U; \omega_{\delta U}) \equiv \left\{ \begin{array}{c} \nu \in \Delta(U) : \nu \text{ is judicious,} \\ \text{and } \nu_e = \omega_e \text{ for all } e \in \delta U \end{array} \right\} ,$$

where ν_e denotes the marginal of ν on edge e. We also let

$$\Delta(U; \omega_U) \equiv \left\{ \nu \in \Delta(U) : \nu_e = \omega_e \text{ for all } e \in E_U \right\} ,$$

and note that $\Delta(U; \omega_U) \subseteq J(U; \omega_{\delta U}) \subseteq \Delta(U)$ for any judicious ω.

Proposition 3.44 (block optimization for compound regions). *In the processed graph* $\mathscr{G} = (V, F, E) = \mathrm{pr}\,\mathscr{G}'$, *let* $U \equiv (V_U, F_U, E_U)$ *be any compound enclosure. Recall from Definition 3.16 that* $U = U^\circ \cup \partial U^\circ$ *where* $\partial U^\circ \equiv \partial_\circ U$ *is the set of perfect variables in* U. *Let* δU *denote the edges in* U *that are incident to* $\partial_\circ U$, *and decompose*

$$\omega \equiv \begin{pmatrix} \omega_{\mathrm{in}} \\ \omega_{\delta U} \\ \omega_{\mathrm{out}} \end{pmatrix} \equiv \begin{pmatrix} (\omega_{L,j} : \text{some edge in } E_U \setminus \delta U \text{ has type } (L, j)) \\ (\omega_{L,j} : \text{some edge in } \delta U \text{ has type } (L, j)) \\ (\omega_{L,j} : \text{some edge in } E \setminus E_U \text{ has type } (L, j)) \end{pmatrix} .$$

(An edge $e = (av)$ *has type* (L, j) *if* $L_a = L$ *and* e *is the* j-th *edge in* δa.) *We also abbreviate* $\omega_U \equiv (\omega_{\mathrm{in}}, \omega_{\delta U})$ *and* $\omega_{\mathrm{bd}} \equiv (\omega_{\delta U}, \omega_{\mathrm{out}})$. *Assume the neighborhood profile* $\mathscr{D}$ *of* $\mathscr{G}$ *is bounded away from zero in the sense of* (3.45). *Then, using the notation of Definition 3.43, we have*

$$\max_{\acute{\omega}} \left\{ \mathbf{\Psi}_{\mathscr{D},2}(\acute{\omega}) - \mathbf{\Psi}_{\mathscr{D},2}(\omega) : \omega \text{ and } \acute{\omega} \text{ are judicious}, \acute{\omega}_{\mathrm{bd}} = \omega_{\mathrm{bd}} \right\} \tag{3.70}$$

$$= c_{\mathrm{in}} \left(\max_\nu \left\{ \mathcal{H}(\nu) : \nu \in J(U, \omega_{\delta U}) \right\} - \max_\nu \left\{ \mathcal{H}(\nu) : \nu \in \Delta(U; \omega_U) \right\} \right), \tag{3.71}$$

where c_{in} *depends only on* $\mathscr{D}$ *and on* U, *and is lower bounded by* c_1.

Proof. Recalling Remark 3.20, let $\mathcal{D}$ be any neighborhood sequence that is consistent with $\mathscr{D}$. Since for any ω we have $\mathbb{E}_{\mathscr{D}} \mathbf{Z}^2(\omega) = \mathbb{E}_{\mathcal{D}} \mathbf{Z}^2(\omega)$, rather than working under $\mathbb{P}_{\mathscr{D}}$ we can instead work under $\mathbb{P}_{\mathcal{D}}$, which can be sampled by the (generalized) configuration model described in Remark 3.21. Let $t(U)$ denote the collection of all edge types t appearing on edges inside U; recall that these are all compound types.

Let $\mathfrak{M}$ denote the random matching of δV to δF (which defines the graph $\mathcal{G}$). We let $\mathfrak{m}$ denote the restriction of $\mathfrak{M}$ to half-edges with types in $t(U)$; this defines a subgraph $\mathcal{U} \equiv (V_{\mathcal{U}}, F_{\mathcal{U}}, E_{\mathcal{U}}) \subseteq \mathcal{G}$ which consists of $n_U = nc_{\mathrm{in}}$ disjoint copies of U. Moreover, since $\mathcal{U}$ consists of n_U disjoint copies of U for any valid realization of $\mathfrak{m}$, we can condition on $\mathfrak{m}$ without changing the expected value of the partition function:

$$\frac{1}{n^{O(1)}} \exp\left\{ n \mathbf{\Psi}_{\mathscr{D},2}(\omega) \right\} = \mathbb{E}_{\mathscr{D}} \mathbf{Z}^2(\omega) = \mathbb{E}_{\mathcal{D}} \mathbf{Z}^2(\omega) = \mathbb{E}_{\mathcal{D}}\left(\mathbf{Z}^2(\omega) \,\middle|\, \mathfrak{m} \right).$$

Let $\partial_\circ \mathcal{U}$ denote the n_U copies of $\partial_\circ U$ inside $\mathcal{U}$, and let $\delta \mathcal{U}$ denote the edges in $\mathcal{U}$ incident to $\partial_\circ \mathcal{U}$. Any (pair) coloring $\underline{\sigma}$ on $\mathcal{G}$ can be decomposed as

$$\underline{\sigma} \equiv \begin{pmatrix} \underline{\sigma}_{\mathrm{in}} \\ \underline{\sigma}_{\delta \mathcal{U}} \\ \underline{\sigma}_{\mathrm{out}} \end{pmatrix} \equiv \begin{pmatrix} (\sigma_e : e \in E_{\mathcal{U}} \setminus \delta \mathcal{U}) \\ (\sigma_e : e \in \delta \mathcal{U}) \\ (\sigma_e : e \in E \setminus E_{\mathcal{U}}) \end{pmatrix}.$$

Note that this corresponds precisely to the above decomposition of ω: a coloring $\underline{\sigma}$ has empirical measure ω (which we abbreviate $\underline{\sigma} \sim \omega$) if and only if $\underline{\sigma}_{\mathrm{in}}, \underline{\sigma}_{\delta \mathcal{U}}$, and $\underline{\sigma}_{\mathrm{out}}$ have empirical measures $\omega_{\mathrm{in}}, \omega_{\delta U}$, and ω_{out} respectively. We also abbreviate $\underline{\sigma}_{\mathcal{U}} \equiv (\underline{\sigma}_{\mathrm{in}}, \underline{\sigma}_{\delta \mathcal{U}})$ and $\omega_U \equiv (\omega_{\mathrm{in}}, \omega_{\delta U})$. With this notation, we can decompose $\mathbf{Z}^2(\omega)$ as

$$\mathbf{Z}^2(\omega) = \sum_{\underline{\sigma}_{\mathcal{U}}} \mathbf{1}\{\underline{\sigma}_{\mathcal{U}} \sim \omega_U\} \varphi_{\mathcal{U}}(\underline{\sigma}_{\mathcal{U}}) \left\{ \sum_{\underline{\sigma}_{\mathrm{out}}} \mathbf{1}\{\underline{\sigma}_{\mathrm{out}} \sim \omega_{\mathrm{out}}\} \varphi_{\mathrm{out}}(\underline{\sigma}_{\delta \mathcal{U}}, \underline{\sigma}_{\mathrm{out}}) \right\}, \tag{3.72}$$

where, recalling the definition (3.66) of the factors for the pair coloring model, we define

$$\varphi_{\mathcal{U}}(\underline{\sigma}_{\mathcal{U}}) \equiv \prod_{v \in V_{\mathcal{U}} \setminus \partial_\circ U} \varphi_v(\underline{\sigma}_{\delta v}) \prod_{a \in F_{\mathcal{U}}} \hat{\varphi}_a(\underline{\sigma}_{\delta a}),$$

$$\varphi_{\mathrm{out}}(\underline{\sigma}_{\delta \mathcal{U}}, \underline{\sigma}_{\mathrm{out}}) \equiv \prod_{v \in (V \setminus V_U) \cup \partial_\circ U} \varphi_v(\underline{\sigma}_{\delta v}) \prod_{a \in F \setminus F_{\mathcal{U}}} \hat{\varphi}_a(\underline{\sigma}_{\delta a}).$$

Taking the expectation conditional on $\mathfrak{m}$ gives

$$\mathbb{E}_{\mathcal{D}}\left(\mathbf{Z}^2(\omega) \,\middle|\, \mathfrak{m} \right) = \sum_{\underline{\sigma}_{\mathcal{U}}} \mathbf{1}\{\underline{\sigma}_{\mathcal{U}} \sim \omega_U\} \varphi_{\mathcal{U}}(\underline{\sigma}_{\mathcal{U}}) \overbrace{\mathbb{E}_{\mathcal{D}}\left(\sum_{\underline{\sigma}_{\mathrm{out}}} \mathbf{1}\{\underline{\sigma}_{\mathrm{out}} \sim \omega_{\mathrm{out}}\} \varphi_{\mathrm{out}}(\underline{\sigma}_{\delta \mathcal{U}}, \underline{\sigma}_{\mathrm{out}}) \,\middle|\, \mathfrak{m} \right)}^{X \equiv X(\omega_{\delta U}, \omega_{\mathrm{out}})},$$

where we emphasize that the value of X is constant over all $\underline{\sigma}_{\delta \mathcal{U}} \sim \omega_{\delta U}$. Now recall that $\mathcal{U}$ consists of a disjoint union of copies of U, which we denote U_i for $1 \le i \le n_U$. For any $\underline{\sigma}_{\mathcal{U}}$ we can define its empirical measure over the copies of U, that is to say,

$$\nu(\underline{\sigma}_U) = \frac{1}{n_U} \sum_{i=1}^{n_U} \mathbf{1}\{\underline{\sigma}_{U_i} = \underline{\sigma}_U\}$$

We write $\underline{\sigma}_{\mathcal{U}} \sim \nu$ if $\underline{\sigma}_{\mathcal{U}}$ has empirical measure ν. Then, with $\Delta(U; \omega_U)$ as in Definition 3.43, we have

$$\sum_{\underline{\sigma}_{\mathcal{U}}} \mathbf{1}\{\underline{\sigma}_{\mathcal{U}} \sim \omega_U\} \varphi_{\mathcal{U}}(\underline{\sigma}_{\mathcal{U}}) = \sum_{\nu \in \Delta(U; \omega_U)} \sum_{\underline{\sigma}_{\mathcal{U}}} \mathbf{1}\{\underline{\sigma}_{\mathcal{U}} \sim \nu\} \varphi_{\mathcal{U}}(\underline{\sigma}_{\mathcal{U}}) = \sum_{\nu \in \Delta(U; \omega_U)} \binom{n_U}{n_U \nu}$$

$$= n^{O(1)} \exp\left\{ n_U \max_\nu \left\{ \mathcal{H}(\nu) : \nu \in \Delta(U; \omega_U) \right\} \right\},$$

where the last step is by Stirling's approximation. Altogether we conclude

$$\exp\left\{ n \mathbf{\Psi}_{\mathscr{D},2}(\omega) \right\} = n^{O(1)} X \exp\left\{ n_U \max_\nu \left\{ \mathcal{H}(\nu) : \nu \in \Delta(U; \omega_U) \right\} \right\} \tag{3.73}$$

On the other hand, summing over all possibilities of ω_{in} gives

$$\frac{1}{n^{O(1)}} \exp\left\{n \max\left\{\boldsymbol{\Psi}_{\mathscr{D},2}(\acute{\omega}) : \acute{\omega}_{\text{bd}} = \omega_{\text{bd}}\right\}\right\} = \sum_{\acute{\omega}} \mathbf{1}\{\acute{\omega}_{\text{bd}} = \omega_{\text{bd}}\} \mathbb{E}_{\mathscr{D}}\left(\mathbf{Z}^2(\acute{\omega}) \,\middle|\, \mathfrak{m}\right)$$

$$= X \sum_{\underline{\sigma}_{\mathscr{U}}} \mathbf{1}\{\underline{\sigma}_{\delta\mathscr{U}} \sim \omega_{\delta U}\}\varphi_{\mathscr{U}}(\underline{\sigma}_{\mathscr{U}}) = n^{O(1)} X \exp\left\{n_U \max_{v}\left\{\mathcal{H}(v) : v \in J(U; \omega_{\delta U})\right\}\right\}. \tag{3.74}$$

The claim follows by combining (3.73) with (3.74). $\qquad\square$

Entropy maximization problem for non-compound variables. We now give the analogues of Definition 3.43 and Proposition 3.44 for the case of non-compound variables. As noted above, we now have the added difficulty that for a variable of non-compound type T, the clause types L neighboring to the variable are not uniquely determined by T. This difficulty will be countervailed by the fact that non-compound variables are perfect, and as a result it will be sufficient to consider only the depth-one neighborhood of the variable.

Definition 3.45 (judicious measures on trees with augmented alphabet). Let v be a variable of total type T, which we assume is not of compound type. Let $U \equiv U_T$ be the depth-one neighborhood of v, in which each edge $e \in \delta v$ is labelled with its corresponding type t_e. However, we **forget the total type labellings on the clauses and other edges of U,** and each boundary edge $e \in \delta U$ is labelled only with its index $j(e) \in [k]$. An **augmented (pair) coloring** on U is a configuration $(\underline{\sigma}, \underline{L})$ which assigns to each edge $e \in U$ a spin (σ_e, L_e) where $\sigma_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$ and L_e is a clause total type. Recall (3.66) where we defined the variable and clause factors φ_v and $\hat{\varphi}_a$ for the pair coloring model. The factors for the augmented pair coloring model are

$$\varphi_v(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \equiv \varphi_v(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \mathbf{1}\{L_e \ni t_e\},$$

$$\hat{\varphi}_a(\underline{\sigma}_{\delta a}, \underline{L}_{\delta a}) \equiv \hat{\varphi}_a(\underline{\sigma}_{\delta a})\mathbf{1}\left\{\text{the } L_e \text{ are the same for all } e \in \delta a\right\}. \tag{3.75}$$

We say that $(\underline{\sigma}, \underline{L})$ is a valid augmented coloring on U as long as

$$\varphi_v(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \prod_{a \in \partial v} \hat{\varphi}_a(\underline{\sigma}_{\delta a}, \underline{L}_{\delta a}) = 1.$$

Analogously to the simplex $\boldsymbol{\Delta}(U)$ from Definition 3.43, we now let

$$\boldsymbol{\Delta}_{\text{aug}}(U) \equiv \left\{\begin{array}{c}\text{probability measures } v \text{ on valid} \\ \text{augmented pair colorings } (\underline{\sigma}, \underline{L}) \text{ of } U\end{array}\right\}.$$

For each edge e in U, let $p(e)$ denote the edge in δv that is closest to e. (If $e \in \delta v$ then $p(e) = e$.) We then say that a measure $v \in \boldsymbol{\Delta}_{\text{aug}}(U)$ is **fully judicious with respect to $\mathscr{D}$** if it holds for all edges e in U that

$$v_e(\sigma, L) = \overbrace{\frac{\hat{\mathscr{D}}(L)\mathbf{1}\{L_{j(t_{p(e)})} = t_{p(e)}\}}{\underbrace{\sum_{L'} \hat{\mathscr{D}}(L')\mathbf{1}\{(L')_{j(t_{p(e)})} = t_{p(e)}\}}_{\text{denote this } \pi_{\mathscr{D}}(L \,|\, t_{p(e)})}}}^{\text{denote this } (\text{aug}_{\mathscr{D}}(\tilde{\omega}))_e(\sigma, L)} \tilde{\omega}_{L,j(e)}(\sigma) \tag{3.76}$$

for some $\tilde{\omega}$ that is judicious (in the sense of Definition 3.41). Given any particular ω, we now write

$$\omega_{\delta U} \equiv \left(\begin{array}{c}\omega_{L,j} : \text{some edge } e \in \delta U \text{ has } j(e) = j, \text{ and can take clause} \\ \text{type } L_e = L \text{ in the augmented coloring model}\end{array}\right).$$

With this notation, we can define (compare with $J(J; \omega_{\delta U}($ from Definition 3.43)

$$J_{\mathscr{D}}(U; \omega_{\delta U}) \equiv \left\{\begin{array}{c}v \in \boldsymbol{\Delta}_{\text{aug}}(U) : v \text{ is fully judicious with respect to } \mathscr{D}, \\ \text{and } v_e = (\text{aug}_{\mathscr{D}}(\omega))_e \text{ for all } e \in \delta U\end{array}\right\},$$

where ν_e now denotes the marginal law under ν of $(\sigma_e, \boldsymbol{L}_e)$. We also let

$$\boldsymbol{\Delta}_{\mathscr{D}}(U; \omega_U) \equiv \left\{ \nu \in \boldsymbol{\Delta}_{\mathrm{aug}}(U) : \nu_e = (\mathrm{aug}_{\mathscr{D}}(\omega))_e \text{ for all } e \in U \right\},$$

and note that $\boldsymbol{\Delta}_{\mathscr{D}}(U; \omega_U) \subseteq \boldsymbol{J}_{\mathscr{D}}(U; \omega_{\delta U}) \subseteq \boldsymbol{\Delta}_{\mathrm{aug}}(U)$ for any judicious ω.

Proposition 3.46 (block optimization for non-compound variables). *Let $\boldsymbol{T}$ be any variable total type that appears in the processed graph $\mathscr{G} = (V, F, E) = \mathrm{pr}\,\mathscr{G}'$. Assume that $\boldsymbol{T}$ is not of compound type. Define the subgraph*

$$\mathscr{U} \equiv \bigcup_{v : T_v = \boldsymbol{T}} B_1(v; \mathscr{G}) \equiv (V_{\mathscr{U}}, F_{\mathscr{U}}, E_{\mathscr{U}}) \subseteq \mathscr{G},$$

and let $\delta\mathscr{U}$ denote the leaf edges of $\mathscr{U}$. Decompose

$$\omega \equiv \begin{pmatrix} \omega_{\mathrm{in}} \\ \omega_{\delta U} \\ \omega_{\mathrm{out}} \end{pmatrix} \equiv \begin{pmatrix} (\omega_{L,j} : \text{some edge in } E_{\mathscr{U}} \setminus \delta\mathscr{U} \text{ has type } (L, j)) \\ (\omega_{L,j} : \text{some edge in } \delta\mathscr{U} \text{ has type } (L, j)) \\ (\omega_{L,j} : \text{some edge in } E \setminus E_{\mathscr{U}} \text{ has type } (L, j)) \end{pmatrix}.$$

We also abbreviate $\omega_U \equiv (\omega_{\mathrm{in}}, \omega_{\delta U})$ and $\omega_{\mathrm{bd}} \equiv (\omega_{\delta U}, \omega_{\mathrm{out}})$. Assume the neighborhood profile $\mathscr{D}$ of $\mathscr{G}$ is bounded away from zero in the sense of (3.45). Then, using the notation of Definition 3.45, we have

$$\max_{\acute{\omega}} \left\{ \boldsymbol{\Psi}_{\mathscr{D},2}(\acute{\omega}) - \boldsymbol{\Psi}_{\mathscr{D},2}(\omega) : \omega \text{ and } \acute{\omega} \text{ are judicious}, \acute{\omega}_{\mathrm{bd}} = \omega_{\mathrm{bd}} \right\}$$

$$= c_{\mathrm{in}} \left(\max_{\nu} \left\{ \mathcal{H}(\nu) : \nu \in \boldsymbol{J}_{\mathscr{D}}(U, \omega_{\delta U}) \right\} - \max_{\nu} \left\{ \mathcal{H}(\nu) : \nu \in \boldsymbol{\Delta}_{\mathscr{D}}(U; \omega_U) \right\} \right),$$

where c_{in} depends only on $\mathscr{D}$ and on $\boldsymbol{T}$, and is lower bounded by c_1.

Proof. As in the proof of Proposition 3.44, we can fix a neighborhood sequence $\mathcal{D}$ that is consistent with $\mathscr{D}$, and work under the measure $\mathbb{P}_{\mathcal{D}}$. Let $t(\boldsymbol{T})$ denote the set of all edge types that can be incident to a variable of type $\boldsymbol{T}$. Let $L(\boldsymbol{T})$ denote the set of all clause types that can neighbor a variable of type $\boldsymbol{T}$:

$$L(\boldsymbol{T}) = \left\{ L : t \in L \text{ for some } t \in t(\boldsymbol{T}) \right\}.$$

Let $\mathfrak{M}$ denote the random matching of δV to δF (which defines the graph $\mathscr{G}$). We let $\mathfrak{m}_1$ denote the restriction of $\mathfrak{M}$ to the half-edges of types belonging to $t(\boldsymbol{T})$. We then let $\mathfrak{m}_2$ denote the edges in $\mathfrak{M} \setminus \mathfrak{m}_1$ that are incident to the edges of $\mathfrak{m}_1$. Finally, we let $\mathfrak{m}_3 = \mathfrak{M} \setminus (\mathfrak{m}_1 \cup \mathfrak{m}_2)$. The partial matching $(\mathfrak{m}_1, \mathfrak{m}_2)$ defines the subgraph $\mathscr{U} \subseteq \mathscr{G}$ consisting of the depth-one neighborhoods of all the variables of type $\boldsymbol{T}$. Analogously to (3.72), we have the decomposition

$$\boldsymbol{Z}^2(\omega) = \sum_{\underline{\sigma}_{\mathscr{U}}} \mathbf{1}\{\underline{\sigma}_{\mathscr{U}} \sim \omega_U\} \varphi_{\mathscr{U}}(\underline{\sigma}_{\mathscr{U}}) \left\{ \sum_{\underline{\sigma}_{\mathrm{out}}} \mathbf{1}\{\underline{\sigma}_{\mathrm{out}} \sim \omega_{\mathrm{out}}\} \varphi_{\mathrm{out}}(\underline{\sigma}_{\delta\mathscr{U}}, \underline{\sigma}_{\mathrm{out}}) \right\}.$$

Note that $\varphi_{\mathscr{U}}$ depends on $(\mathfrak{m}_1, \mathfrak{m}_2)$, while φ_{out} depends on $(\mathfrak{m}_2, \mathfrak{m}_3)$. However, we have

$$\mathbb{E}_{\mathcal{D}} \left(\sum_{\underline{\sigma}_{\mathrm{out}}} \mathbf{1}\{\underline{\sigma}_{\mathrm{out}} \sim \omega_{\mathrm{out}}\} \varphi_{\mathrm{out}}(\underline{\sigma}_{\delta\mathscr{U}}, \underline{\sigma}_{\mathrm{out}}) \,\middle|\, \mathfrak{m}_2 \right) = X(\omega_{\delta U}, \omega_{\mathrm{out}}) \equiv X$$

for any $\underline{\sigma}_{\delta\mathscr{U}} \sim \omega_{\delta U}$. Under $\mathbb{P}_{\mathcal{D}}$, the matching $\mathfrak{m}_1$ is independent of the pair $(\mathfrak{m}_2, \mathfrak{m}_3)$. The expected value of $\boldsymbol{Z}^2(\omega)$ under $\mathbb{P}_{\mathcal{D}}$ remains the same even after conditioning on $\mathfrak{m}_2$, so we have

$$\mathbb{E}_{\mathcal{D}} \boldsymbol{Z}^2(\omega) = \mathbb{E}_{\mathcal{D}}\left(\boldsymbol{Z}^2(\omega) \,\middle|\, \mathfrak{m}_2 \right) = X \cdot \mathbb{E}_{\mathcal{D}}\left(\sum_{\underline{\sigma}_{\mathscr{U}}} \mathbf{1}\{\underline{\sigma}_{\mathscr{U}} \sim \omega_U\} \varphi_{\mathscr{U}}(\underline{\sigma}_{\mathscr{U}}) \,\middle|\, \mathfrak{m}_2 \right)$$

$$= X \cdot \sum_{\mathfrak{m}_1} \mathbb{P}_{\mathcal{D}}(\mathfrak{m}_1) \underbrace{\sum_{\underline{\sigma}_{\mathscr{U}}} \mathbf{1}\{\underline{\sigma}_{\mathscr{U}} \sim \omega_U\} \varphi_{\mathscr{U}}(\underline{\sigma}_{\mathscr{U}})}_{\text{depends on } (\mathfrak{m}_1, \mathfrak{m}_2)}.$$

Recall that $(\mathfrak{m}_1, \mathfrak{m}_2)$ defines the subgraph $\mathscr{U}$ consisting of the depth-one neighborhoods of all the variables of type T; without loss we suppose those variables are labelled $\{v_1, \ldots, v_{n_T}\}$. On the other hand, let $\mathcal{U}$ denote the graph consisting of n_T disjoint copies $U_1, \ldots, U_{n_T}$ of the tree U from Definition 3.45. A tuple $(\mathfrak{m}_1, \mathfrak{m}_2, \underline{\sigma}_{\mathscr{U}})$ with $\underline{\sigma}_{\mathscr{U}} \sim \omega_U$ can be mapped bijectively to an augmented coloring $(\underline{\sigma}_{\mathcal{U}}, L_{\mathcal{U}})$ of $\mathcal{U}$ with edge empirical measures $\mathrm{aug}(\omega_U)$. (The bijection goes as follows: given $(\mathfrak{m}_1, \mathfrak{m}_2)$ there is a unique graph isomorphism $g : \mathscr{U} \to \mathcal{U}$ which maps v_i to the root of U_i for each i, respects edge types t_e for all $e \in \delta v_i$ for all i, and respects the edge indices $j(e)$ for all $e \in \delta \mathscr{U}$. The coloring $\underline{\sigma}_{\mathscr{U}}$ is mapped under g to a coloring on $\mathcal{U}$. Finally, for each edge $e = (au)$ in U, we set L_e to be the clause type of $g^{-1}(a)$.) It follows that

$$\sum_{\mathfrak{m}_1} \mathbb{P}_{\mathcal{D}}(\mathfrak{m}_1) \sum_{\underline{\sigma}_{\mathscr{U}}} \mathbf{1}\{\underline{\sigma}_{\mathscr{U}} \sim \omega_U\} \varphi_{\mathscr{U}}(\underline{\sigma}_{\mathscr{U}}) = \frac{1}{|\{\mathfrak{m}_1\}|} \sum_{(\underline{\sigma}_{\mathcal{U}}, L_{\mathcal{U}})} \mathbf{1}\{(\underline{\sigma}_{\mathcal{U}}, L_{\mathcal{U}}) \sim \mathrm{aug}(\omega_U)\} \varphi_{\mathcal{U}}(\underline{\sigma}_{\mathcal{U}}, L_{\mathcal{U}}),$$

where $\{\mathfrak{m}_1\}$ denotes the set of all matchings $\mathfrak{m}_1$ that are consistent with $\mathcal{D}$. It follows analogously to (3.73) that

$$\exp\left\{n \Psi_{\mathcal{D},2}(\omega)\right\} = \frac{n^{O(1)} X}{|\mathfrak{m}_1|} \exp\left\{n_T \max_v \left\{\mathcal{H}(\nu) : \nu \in \Delta_{\mathcal{D}}(U; \omega_U)\right\}\right\}.$$

On the other hand, summing over all possibilities of ω_{in} gives, analogously to (3.74),

$$\exp\left\{n \max\left\{\Psi_{\mathcal{D},2}(\acute{\omega}) : \acute{\omega}_{\mathrm{bd}} = \omega_{\mathrm{bd}}\right\}\right\} = \frac{n^{O(1)} X}{|\mathfrak{m}_1|} \exp\left\{n_T \max_v \left\{\mathcal{H}(\nu) : \nu \in J_{\mathcal{D}}(U; \omega_{\delta U})\right\}\right\}.$$

The claim follows. $\qquad \square$

Contraction estimates and coordinate descent. We now give some informal discussion of how Propositions 3.44 and 3.46 are used in the proof of the key second moment estimate Proposition 3.32. The details of the proof are rather complicated, and are laid out in Sections 7–9. However, some of the high-level ideas are rather simple, and we point them out here. As above, let U denote a compound enclosure, or the depth-one neighborhood of a perfect variable. For the purposes of this discussion, we will express the pair empirical measure $\omega = (\omega_{L,j})_{L,j}$ as a tuple $y = (y_1, \ldots, y_\ell)$ where each y_i denotes a subset of entries of ω corresponding to edges in the interior U° of U. Thus ℓ is the number of distinct choices of U, where "distinct" here means that the types are distinct. Write $f(y) \equiv \Psi_{\mathcal{D},2}(\omega)$, and then let $_*y \equiv (_*y_1, \ldots, _*y_\ell)$ denote the y that corresponds to $_*\omega$. As explained in Lemma 3.42, the conclusion of Proposition 3.32 follows if we can show that in a neighborhood of $_*y$ (corresponding to $\omega \in I_0$) the function f is uniquely maximized at $_*y$, with negative-definite Hessian.

With the above notation, we see that Propositions 3.44 and 3.46 explain how to optimize $f(y)$ in a single coordinate y_i, keeping the other coordinates $y_{-i} \equiv (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_\ell)$ fixed. Moreover, if y_{-i} is fixed, then the optimization over y_i is a entropy maximization problem (constrained to affine subspaces corresponding to $J(U; \omega_{\delta U})$ or $J_{\mathcal{D}}(U; \omega_{\delta U})$), which means the function is strictly concave in y_i if y_{-i} is fixed. Given y, let

$$\tilde{y}_i \equiv \tilde{y}_i(y_{-i}) = \arg\max_{x_i}\left\{f(x_i, y_{-i})\right\}, \qquad (3.77)$$

i.e., $\tilde{y}_i$ is the result of optimizing the i-th coordinate keeping the others fixed. A key estimate that we will prove is that if y is close enough to $_*y$, then this update brings $\tilde{y}_i$ closer to $_*y_i$:

$$\|\tilde{y}_i - _*y_i\| \le \frac{\|y_i - _*y_i\|}{2}. \qquad (3.78)$$

We have not specified the norm $\| \cdot \|$ for which the above holds — in fact we obtain contraction estimates for a "discrepancy" measure (7.4) which is not quite a norm, but is close enough to serve our purpose. Even ignoring this issue, the bound (3.78) is a simplification of the precise contraction results that we obtain, which are characterized by Propositions 7.4 and 8.1.

Assuming the simplified estimate (3.78), it is straightforward to argue that in a small neighborhood of $y = _*y$, the function $f(y)$ is uniquely maximized at $_*y$ — for any $y \ne _*y$, we can apply the update (3.77) in some coordinate

$y_i \neq {}_\star y_i$, and the value of f will increase because f is strictly convex in each individual coordinate y_i. Having shown this, we can proceed as follows: for $\tilde{y}_i$ as defined by (3.77), we have

$$\left(f({}_\star y) - f(y)\right)^{1/2} \geq \left(f(\tilde{y}_i, y_{-i}) - f(y)\right)^{1/2} \gtrsim \|\tilde{y}_i - y_i\| \geq \|y_i - {}_\star y_i\| - \|\tilde{y}_i - {}_\star y_i\| \overset{(3.78)}{\geq} \frac{\|y_i - {}_\star y_i\|}{2},$$

which shows that f has negative-definite Hessian at ${}_\star y$. (We remark again that this is a simplified sketch of the actual proof, which is more complicated because we do not have such a simple estimate as (3.78). The detailed proof of the negative-definite Hessian condition appears in §8.1.)

Based on the above discussion, we have the following proof strategy. First, show that if we restrict to $\omega \in I_0$, then the maximum of $\Psi_{\mathcal{D},2}(\omega)$ can only be attained in a small neighborhood of $\omega = {}_\star\omega$. We call this step the "a priori estimate," and it is deferred to Section 9. Then show that in the small neighborhood of $\omega = {}_\star\omega$ (i.e., $y = {}_\star y$), we have (some version of) the contraction estimate (3.78). The contraction estimate occupies the majority of Sections 7 and 8, and is the most essential technical ingredient of the proof. Combining these steps with the above argument will show that if we restrict to $\omega \in I_0$, then $\Psi_{\mathcal{D},2}(\omega) = f(y)$ is uniquely maximized at $\omega = {}_\star\omega$ with negative-definite Hessian. Thus the conditions of Lemma 3.42 are satisfied, and Proposition 3.32 follows.

We conclude this subsection with some discussion of the key contraction estimate (3.78). Recall that $\tilde{y}_i$ is defined by (3.77), or equivalently, by the constrained entropy maximization problems from Propositions 3.44 and 3.46. To be concrete, consider the compound case from Proposition 3.44. Roughly speaking, the approach will be to find a weight $\Lambda(\underline{\sigma})$ (where $\underline{\sigma}$ is a pair coloring of the enclosure U) which is "Lagrangian" in the sense that $\langle \ln \Lambda, \nu \rangle$ is constant over $\nu \in J(U; \omega_{\delta U})$, and which turns the constrained maximization problem into an unconstrained one:

$$\nu = \arg\max\left\{\mathcal{H}(\nu) : \nu \in J(U; \omega_{\delta U})\right\} = \arg\max\left\{\mathcal{H}(\nu) + \langle \ln \Lambda, \nu \rangle : \nu \in \Delta(U)\right\}.$$

By calculus, the solution is given simply by $\nu(\underline{\sigma}) \cong \Lambda(\underline{\sigma})$, the Λ-weighted Gibbs measure on U. Thus, if q denotes the BP messages for the Λ-weighted model, we can easily read off edge marginals of ν from the usual formula $\nu_e \cong \dot{q}_e \hat{q}_e$. The basic strategy of the proof of (3.78) is to show that **in the pair coloring model, the the BP recursion contracts towards the product message ${}_\star q \equiv {}_\bullet q \otimes {}_\bullet q$, provided we start close enough to ${}_\star q$.** As a consequence, if the boundary condition $\omega_{\delta U}$ is close to product, we will be able to construct weights Λ that are close to ${}_\star\Lambda \otimes {}_\star\Lambda$ (for ${}_\star\Lambda$ from Corollary 3.56). Then, for e in the interior of U, the messages $\dot{q}_e$ and $\hat{q}_e$ will be closer to product than the messages on δU, so the discrepancy between $\omega_e \cong {}_\star\dot{q}_e{}_\star\hat{q}_e$ and ${}_\star\omega_e$ will be closer than the discrepancies on δU. This gives the rough idea of the proof of (3.78), and we leave the details to Sections 7–8.

3.7. **Weights for coherent clauses.**
In this subsection we prove a series of claims concerning weakly and strictly coherent clauses (Definition 3.6). The key implications are that ${}_\square\pi$ is always strictly coherent, so ${}_\star\pi$ will also be strictly coherent if it is "close enough" to ${}_\square\pi$. When ${}_\star\pi$ is strictly coherent, we will show that clauses can be reweighted such that BP equations hold. First, however, we make a remark on the possibilities of $\text{supp}\,{}_\star\pi_e$:

Remark 3.47. The purpose of this remark is to emphasize that the canonical marginal ${}_\star\pi_e$ can in general be supported a strict subset of $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$. In fact, although $\text{supp}\,{}_\star\pi_e$ always contains the green spin, $(\text{supp}\,{}_\star\pi_e) \setminus \mathtt{g}$ can be any subset of $\{\mathtt{r}, \mathtt{y}, \mathtt{b}\}$. For example, suppose in a k-SAT instance $\mathcal{G} = (V, F, E)$ that a clause $a \in F$ has among its neighbors two leaf variables $u \neq v$. Then both u and v must always be free, so

$$\text{supp}\,{}_\star\pi_{au} = \text{supp}\,{}_\star\pi_{av} = \{\mathtt{g}\}.$$

The clause a can never be forcing, so for any other variable $w \in \partial a \setminus \{u, v\}$ we must have $\mathtt{r} \notin \text{supp}\,{}_\star\pi_{aw}$. However, such w can be forced by other clauses $b \in \partial w \setminus a$ — indeed, $\text{supp}\,{}_\star\pi_{aw}$ contains $\{\mathtt{b}\}$ if and only if w may be forced by some $b \in \partial w(\mathtt{+L}_{aw}) \setminus a$; and it contains $\{\mathtt{y}\}$ if and only if w may be forced by some $b \in \partial w(\mathtt{-L}_{aw})$. Therefore, in this scenario, $\text{supp}\,{}_\star\pi_{aw} \setminus \mathtt{g}$ can be any subset of $\{\mathtt{y}, \mathtt{b}\}$. Similarly, on edges $e \in E$ where forcing can occur (meaning $\mathtt{r} \in \text{supp}\,{}_\star\pi_e$), it is easy to construct examples to see that $\text{supp}\,{}_\star\pi_e \setminus \{\mathtt{g}, \mathtt{r}\}$ can be any subset of $\{\mathtt{y}, \mathtt{b}\}$.

We now turn to the main results of the subsection.

Lemma 3.48. *For a clause $a \in F$, a tuple $\pi = (\pi_e)_{e \in \delta a}$ is weakly coherent if and only if there exists a probability measure $\nu_{\delta a}$ over valid colorings of δa with edge marginals π: that is,*

$$\pi_e(\sigma) = \sum_{\underline{\sigma}_{\delta a}} \mathbf{1}\{\sigma_e = \sigma\} \nu_{\delta a}(\underline{\sigma}_{\delta a})$$

for all $e \in \delta a$ and all $\sigma \in \{r, y, g, b\}$.

Proof. Recall that a valid coloring of δa is a tuple $\underline{\sigma} \equiv \underline{\sigma}_{\delta a} \in \{r, y, g, b\}^{\delta a}$ for which $\hat{\varphi}_a(\underline{\sigma}) = 1$, as defined by (2.19) with the grouping $c = \{g, b\}$. If $|\delta a| = k$, then a valid coloring $\underline{\sigma} \in \{r, y, g, b\}^k$ falls in one of two cases: (i) it has exactly one r entry with the remaining entries y; or (ii) all entries are in $\{c, y\}$ and at least two entries are c. If ν is a probability measure on valid colorings $\underline{\sigma} \in \{r, y, g, b\}^k$, let (R_i, Y_i, G_i, B_i) be the associated marginal probabilities of r, y, g, b on each coordinate $i \in \{1, \ldots, k\}$, so for instance

$$R_i \equiv \nu(\sigma_i = r) = \sum_{\underline{\sigma} \in \{r,y,g,b\}^k} \mathbf{1}\{\sigma_i = r\} \nu(\underline{\sigma}).$$

Thus, for each $i \in [k] \equiv \{1, \ldots, k\}$, the quantities R_i, Y_i, G_i, B_i are nonnegative and sum to one. This lemma is purely a claim about the feasible polytope of edge marginals coming from probability measures over valid colorings: the assertion is that for any $k \geq 3$, the following are equivalent:

(a) The tuple $(R_i, Y_i, G_i, B_i)_{1 \leq i \leq k}$ satisfies the following constraints: for each $i \in [k]$, the quantities R_i, Y_i, G_i, B_i are nonnegative and sum to one. Moreover, with $R \equiv R_1 + \ldots + R_k$, we have $Y_i \geq R - R_i$ for each i. Lastly, writing $C_i \equiv G_i + B_i$, and $C \equiv C_1 + \ldots + C_k$, we have $C \geq 2(1 - R)$.

(b) The marginals $(R_i, Y_i, G_i, B_i)_{1 \leq i \leq k}$ can be realized by a probability measure on valid colorings $\underline{\sigma} \in \{r, y, g, b\}^k$.

Let us first verify the straightforward direction, that (b) implies (a). If ν is a probability measure on valid colorings $\underline{\sigma} \in \{r, y, g, b\}^k$, then for each coordinate i it is clear that the marginal probabilities R_i, Y_i, G_i, B_i are nonnegative and sum to one. Next, since the color r can only occur together with $k - 1$ entries y, we have $R_i = \nu(y^{i-1} r y^{k-i})$ for each $i \in [k]$. This implies

$$Y_i \geq \sum_{j \in [k] \backslash i} \nu(y^{j-1} r y^{k-j}) = \sum_{j \in [k] \backslash i} R_j = R - R_i.$$

On the other hand, since any valid $\underline{\sigma}$ with no r entry must be in $\{y, c\}^k$ with at least two c entries, we have

$$\sum_{i=1}^{k} C_i = \sum_{\underline{\sigma} \in \{y,c\}^k} \nu(\underline{\sigma}) \sum_{i=1}^{k} \mathbf{1}\{\sigma_i = c\} \geq 2 \sum_{\underline{\sigma} \in \{y,c\}^k} \nu(\underline{\sigma}) = 2(1 - R). \qquad (3.79)$$

This proves that (b) implies (a).

In the converse direction, given $(R_i, Y_i, G_i, B_i)_{1 \leq i \leq k}$ satisfying the conditions of (a), we now describe one particular construction of a measure ν on valid colorings that realizes these marginals.[14] First, it is clear that we must set $\nu(y^{i-1} r y^{k-i}) = R_i$ for each $i \in [k]$. This step assigns R of the probability mass of ν, where the conditions in (a) ensure that $0 \leq R \leq R_i + Y_i \leq 1$. If $R = 1$ then we are done, so assume otherwise, meaning there is a positive amount $1 - R$ of mass left over that remains to be assigned. On each i let y_i denote the marginal weight of y left over after the first step:

$$y_i = Y_i - \sum_{j \in [k] \backslash i} R_j = Y_i + R_i - R,$$

which is nonnegative by (a). The total mass left over on each edge i is $C_i + y_i = 1 - R$. Let $\mathcal{C}_1, \ldots, \mathcal{C}_k$ be consecutive intervals of length C_i,

$$\mathcal{C}_i = \left[\sum_{j=1}^{i-1} C_j, \sum_{j=1}^{i} C_j \right).$$

---

[14]The total number of valid colorings $\underline{\sigma} \in \{r, y, g, b\}^k$ is $3^k - k - 1$, so ν ranges over a $(3^k - k - 2)$-dimensional simplex. On the other hand, since $R_i + Y_i + G_i + B_i = 1$ for all i, the tuple $(R_i, Y_i, G_i, B_i)_{1 \leq i \leq k}$ is restricted to an affine space of dimension $3k$. Thus, simply by comparing dimensions, we would expect that any generic feasible tuple $(R_i, Y_i, G_i, B_i)_{1 \leq i \leq k}$ can be realized by an uncountable (and convex) family of measures ν.

These intervals give a partition of $[0, C)$, which in turn is a subset of $[0, k(1 - R))$. Let $J \equiv [0, 1 - R)$, and let p be the mapping from $\mathbb{R}$ to J which sends each real number to its representative modulo $(1 - R)\mathbb{Z}$ in J. Let $J_{i,c}$ be the image of C_i under p (i.e., the consecutive intervals get "wrapped around"), and note that the restriction of p to C_i is one-to-one since $|C_i| = C_i \leq 1 - R$. Let $J_{i,g}$ be any subset of $J_{i,c}$ with Lebesgue measure G_i, and let $J_{i,b} \equiv J_{i,c} \setminus J_{i,g}$. Let $J_{i,y} \equiv J \setminus J_{i,c}$. For $t \in J$ let $\sigma_i(t)$ be the unique element $\sigma \in \{y, g, b\}$ such that $t \in J_{i,\sigma}$. Let $\underline{\sigma}(t) \equiv (\sigma_i(t))_{i \in [k]}$. For $\underline{\sigma} \in \{y, g, b\}^k$, let

$$\nu(\underline{\sigma}) \equiv \mathrm{Leb}\left(\left\{t \in J : \underline{\sigma}(t) = \underline{\sigma}\right\}\right) = \mathrm{Leb}\left(\bigcap_{i=1}^{k} J_{i,\sigma_i}\right),$$

where Leb denotes Lebesgue measure. This completes the definition of ν. It is immediate from the construction that ν has marginals $(R_i, Y_i, G_i, B_i)_{1 \leq i \leq k}$, and all configurations in its support are of form $y^{i-1}ry^{k-i}$, or lie in $\{y, g, b\}^k$. It remains to check that all colorings of the latter case are valid, which is to say that $\underline{\sigma}(t)$ has at least two c entries for every $t \in J$. By definition, $\sigma_i(t) = c$ if and only if $t \in J_{i,c} = p(C_i)$, which occurs if and only if $t + \ell(1 - R) \in C_i$ for some integer ℓ. It follows that the number of c entries in $\underline{\sigma}(t)$ is

$$\sum_{i=1}^{k} \sum_{\ell \in \mathbb{Z}} \mathbf{1}\left\{t + \ell(1 - R) \in C_i\right\} = \sum_{\ell \in \mathbb{Z}} \mathbf{1}\left\{t + \ell(1 - R) \in [0, C)\right\} \geq 2,$$

where the last inequality holds for all $t \in J \equiv [0, 1 - R)$ using the final condition $C \geq 2(1 - R)$ from (a). This shows that ν is supported on valid colorings, thereby concluding our proof that (a) implies (b). $\square$

Lemma 3.49. *For an acyclic clause $a \in F$, the clause-based marginals $_\square\pi = (_\square\pi_e)_{e \in \delta a}$ are strictly coherent for all $r \geq 2$.*

Proof. Recall from Remark 3.5 that the probability measure

$$\nu_{\delta a}(\underline{\sigma}_{\delta a}) = \frac{1}{\hat{z}_a} \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{e \in \delta a} {}_\square \dot{q}(\sigma_e)$$

has marginals $_\square\pi$. It therefore follows immediately from Lemma 3.48 that $_\square\pi$ is weakly coherent. Recall from (3.4) that $_\square\pi_e(\sigma)$ is proportional $_\square\dot{q}_e(\sigma)_\square\hat{q}_e(\sigma)$. It follows from the correspondence (2.38) that $_\square\dot{q}_e(g)$ and $_\square\hat{q}_e(g)$ are positive for all $r \geq 2$, so $_\square\pi_e(g)$ is positive on any edge e. Now suppose $_\square\pi_e(y)$ is positive: this implies that $_\square\dot{q}_e(y)$ and $_\square\hat{q}_e(y)$ must both be positive, therefore

$$_\square\pi_e(y) - \sum_{e' \in \delta a \setminus e} {}_\square\pi_{e'}(r) \geq \nu_{\delta a}\left(\sigma_e = y, \text{ and } \sigma_{e'} = g \text{ for all } e' \in \delta a \setminus e\right) = \frac{1}{\hat{z}_a} {}_\square\dot{q}_e(y) \prod_{e' \in \delta a \setminus e} {}_\square\dot{q}_{e'}(g) > 0,$$

which shows that $_\square\pi$ satisfies (3.6) with strict inequality whenever $_\square\pi_e(y)$. Next, recalling the proof of Lemma 3.48, we see that (3.79) holds with equality if and only if $\nu_{\delta a}$ gives zero mass to configurations $\underline{\sigma}_{\delta a}$ with more than two c entries. This does not happen in the current situation, since

$$\nu_{\delta a}(\sigma_e = g \text{ for all } e \in \delta a) = \frac{1}{\hat{z}_a} \prod_{e \in \delta a} {}_\square\dot{q}_e(g) > 0.$$

This proves that $_\square\pi$ always satisfies (3.7) with strict inequality. The claim follows. $\square$

Lemma 3.50. *For a clause $a \in F$, if the tuple $\pi \equiv (\pi_e)_{e \in \delta a}$ is strictly coherent, then there exists a measure $\bar{\nu}_{\delta a}$ with edge marginals π_e for all $e \in \delta a$, and $\mathrm{supp}\, \bar{\nu}_{\delta a} = \mathrm{COLS}_{\delta a}$ where*

$$\mathrm{COLS}_{\delta a} \equiv \left\{\text{valid colorings } \underline{\sigma}_{\delta a} \,\middle|\, \sigma_e \in \mathrm{supp}\, \pi_e \text{ for all } e \in \delta a\right\}.$$

Proof. It follows from Lemma 3.48 that there is a probability measure $\nu_{\delta a}$ with edge marginals $\pi \equiv (\pi_e)_{e \in \delta a}$, which of course implies $\mathrm{supp}\, \nu_{\delta a} \subseteq \mathrm{COLS}_{\delta a}$. If $\mathrm{COLS}_{\delta a}$ is a singleton then we must have $\mathrm{supp}\, \nu_{\delta a} \subseteq \mathrm{COLS}_{\delta a}$, in which case the assertion follows immediately by taking $\bar{\nu}_{\delta a} = \nu_{\delta a}$. We thus assume for the remainder of the proof that $\mathrm{COLS}_{\delta a}$ has at least two elements.

Now let $\mathrm{MARG}_{\delta a}$ denote the space of *all* edge marginals $\pi \equiv (\pi_e)_{e \in \delta a}$ which can arise from probability measures over $\mathrm{COLS}_{\delta a}$. It follows from Lemma 3.48 that $\mathrm{MARG}_{\delta a}$ is characterized by the conditions (3.6) and (3.7), together with the constraints imposed by the supports of the π_e. To describe this more explicitly, let

$$R(\pi) \equiv \sum_{e \in \delta a} \pi_e(\mathbf{r}), \quad C(\pi) \equiv \sum_{e \in \delta a} \pi_e(\mathbf{c}).$$

We then divide the scenarios into three cases, according to the number of edges $e \in \delta a$ with $\pi_e(\mathbf{r}) > 0$:

(i) If $|\{e \in \delta a : \pi_e(\mathbf{r}) > 0\}| \geq 2$, then we must have $\pi_e(\mathbf{y}) > 0$ for all $e \in \delta a$, and

$$\mathrm{MARG}_{\delta a} \equiv \left\{ \pi \,\middle|\, \begin{array}{l} \pi_e(\mathbf{y}) \geq R(\pi) - \pi_e(\mathbf{r}) \text{ for all } e \in \delta a\,, \\ C(\pi) \geq 2[1 - R(\pi)]\,, \\ \mathrm{supp}\,\pi_e \subseteq \mathrm{supp}\,\pi_e \text{ for all } e \in \delta a \end{array} \right\}.$$

(ii) If there is a unique edge $e' \in \delta a$ with $\pi_{e'}(\mathbf{r}) > 0$, then

$$\mathrm{MARG}_{\delta a} \equiv \left\{ \pi \,\middle|\, \begin{array}{l} \pi_e(\mathbf{y}) \geq R(\pi) = \pi_{e'}(\mathbf{r}) \text{ for all } e \in \delta a \setminus \{e'\}\,, \\ C(\pi) \geq 2[1 - R(\pi)] = 2[1 - \pi_{e'}(\mathbf{r})]\,, \\ \mathrm{supp}\,\pi_e \subseteq \mathrm{supp}\,\pi_e \text{ for all } e \in \delta a \end{array} \right\}.$$

(iii) If $|\{e \in \delta a : \pi_e(\mathbf{r}) > 0\}| = 0$, then

$$\mathrm{MARG}_{\delta a} \equiv \left\{ \pi \,\middle|\, \begin{array}{l} C(\pi) \geq 2[1 - R(\pi)] = 2\,, \\ \mathrm{supp}\,\pi_e \subseteq \mathrm{supp}\,\pi_e \text{ for all } e \in \delta a \end{array} \right\}.$$

In all cases, the definition of strict coherence ensures that $\pi \in \mathrm{ri}\,\mathrm{MARG}_{\delta a}$, the relative interior of $\mathrm{MARG}_{\delta a}$. Now let $u_{\delta a}$ be the uniform measure on $\mathrm{COLS}_{\delta a}$, and let π^{unif} be the marginals resulting from $u_{\delta a}$. Since $\pi^{\mathrm{unif}} \in \mathrm{MARG}_{\delta a}$ and $\pi \in \mathrm{ri}\,\mathrm{MARG}_{\delta a}$, it must hold for sufficiently small ϵ that

$$\pi^\epsilon \equiv \frac{\pi - \epsilon \pi^{\mathrm{unif}}}{1 - \epsilon} \in \mathrm{MARG}_{\delta a}\,.$$

By Lemma 3.48 there is a probability measure ν^ϵ on $\mathrm{COLS}_{\delta a}$ with edge marginals π^ϵ. Then $\bar{\nu}_{\delta a} \equiv (1 - \epsilon)\nu^\epsilon + \epsilon u_{\delta a}$ is fully supported on $\mathrm{COLS}_{\delta a}$, with edge marginals

$$(1 - \epsilon)\pi^\epsilon + \epsilon \pi^{\mathrm{unif}} = (1 - \epsilon)\frac{\pi - \epsilon \pi^{\mathrm{unif}}}{1 - \epsilon} + \epsilon \pi^{\mathrm{unif}} = \pi\,.$$

This concludes the proof. $\qquad\square$

Corollary 3.51. *For a clause $a \in F$, if the tuple $\pi \equiv (\pi_e)_{e \in \delta a}$ is strictly coherent and further satisfies $\pi_e(\mathbf{g}) > 0$ for all $e \in \delta a$, then there exists a set of edge weights $w_e : \{\mathbf{r}, \mathbf{y}, \mathbf{g}, \mathbf{b}\} \to [0, \infty)$ such that the w-weighted measure on valid colorings of δa has edge marginals consistent with π. Moreover, if we fix $w_e(\mathbf{g}) \equiv 1$ for all $e \in \delta a$ and require that $\mathrm{supp}\,w_e \subseteq \mathrm{supp}\,\pi_e$, then w is unique.*

Proof. We will obtain the weights under the assumption that $w_e(\mathbf{g}) \equiv 1$, and $w_e(\sigma) = 0$ if $\sigma \notin \mathrm{supp}_\star \pi_e$; this will imply the result. We can thus restrict our attention to $\mathrm{PR}_{\delta a}$, the space of all probability measures over $\mathrm{COLS}_{\delta a}$. In the case that $\mathrm{COLS}_{\delta a}$ consists of a single element, that element must be the all-$\mathbf{g}$ coloring. Then $\mathrm{PR}_{\delta a}$ also consists of a single element, which is the probability measure fully supported on the all-$\mathbf{g}$ coloring, and setting $w_e(\sigma) \equiv \mathbf{1}\{\sigma = \mathbf{g}\}$ gives the unique weights satisfying the stated conditions.

We therefore assume for the remainder of the proof that $\mathrm{COLS}_{\delta a}$ has at least two elements. Recall that $\mathrm{MARG}_{\delta a}$ denotes the space of all edge marginals $\pi \equiv (\pi_e)_{e \in \delta a}$ which can arise from probability measures over $\mathrm{COLS}_{\delta a}$. We now claim that the space $\mathrm{MARG}_{\delta a}$ has dimension

$$\dim \mathrm{MARG}_{\delta a} = D_a = \sum_{e \in \delta a} \left\{ |\mathrm{supp}_\star \pi_e| - 1 \right\}.$$

Indeed, it is clear that $\dim \mathrm{MARG}_{\delta a} \leq D_a$. Equality follows from the strict coherence assumption: let $\acute{\pi}$ be any tuple of measures $(\acute{\pi}_e)_{e \in \delta a}$, subject *only* to the constraints that $\acute{\pi}_e$ and π_e have the same support for all $e \in \delta a$, and moreover that $\|\pi - \acute{\pi}\|_\infty \leq \delta$. The set of such perturbations $\acute{\pi}$ has dimension D_a. Since π is strictly coherent, we have that $\acute{\pi}$ is also coherent for small enough δ. It then follows from Lemma 3.48 that there is a probability measure

$\acute{\nu}$ on $\text{COLS}_{\delta a}$ with marginal $\acute{\pi}$, which means that $\acute{\pi} \in \text{MARG}_{\delta a}$. Since $\acute{\pi}$ goes over a D_a-dimensional space, this proves $\dim \text{MARG}_{\delta a} = D_a$.

Let A be the linear mapping that takes $\nu \in \text{PR}_{\delta a}$ to its edge marginals $\pi_e(\sigma)$, for $e \in \delta a$ and $\sigma \in \text{supp}\, \pi_e \setminus \{\mathbf{g}\}$. We can regard A as a matrix with row indices $\{(e, \sigma') : e \in \delta a, \sigma' \in \text{supp}\, \pi_e \setminus \{\mathbf{g}\}\}$, and column indices $\text{COLS}_{\delta a}$. The entry of A with row index (e, σ') and column index $\underline{\sigma}_{\delta a}$ is the indicator that $\sigma_e = \sigma'$. The image of $\text{PR}_{\delta a}$ under A is in one-to-one correspondence with $\text{MARG}_{\delta a}$, whose dimension D_a exactly equals the number of rows of A. This shows that A is of full rank. Moreover, let $\text{MSR}_{\delta a}$ be the space of all nonnegative *measures* (not necessarily normalized to unit mass) over $\text{COLS}_{\delta a}$. The image of $\text{MSR}_{\delta a}$ under the matrix

$$A' \equiv \begin{pmatrix} A \\ \mathbf{1}^{\text{t}} \end{pmatrix} \tag{3.80}$$

is the space of all nonnegative scalar multiples of elements of $\text{MARG}_{\delta a}$, and has dimension $D_a + 1$. It follows that the matrix A' is also of full rank.

Now consider the constrained entropy maximization problem

$$\text{OPT} \equiv \max \left\{ - \sum_{\underline{\sigma}_{\delta a} \in \text{COLS}_{\delta a}} \nu(\underline{\sigma}_{\delta a}) \ln \nu(\underline{\sigma}_{\delta a}) \,\middle|\, \nu \in [0, \infty)^{\text{COLS}_{\delta a}} \text{ with } A\nu = \pi \text{ and } \langle \mathbf{1}, \nu \rangle = 1 \right\}, \tag{3.81}$$

as well as its Lagrange dual: for $(\mathbf{s}, \bar{s})$ where $\mathbf{s} \equiv (s_{e, \sigma'} : e \in \delta a, \sigma' \in \text{supp}\, \pi_e \setminus \{\mathbf{g}\}) \in \mathbb{R}^{D_a}$ and $\bar{s} \in \mathbb{R}$,

$$g(\mathbf{s}, \bar{s}) = \max \left\{ - \sum_{\underline{\sigma}_{\delta a} \in \text{COLS}_{\delta a}} \nu(\underline{\sigma}_{\delta a}) \ln \nu(\underline{\sigma}_{\delta a}) + \langle \mathbf{s}, A\nu - \pi \rangle + \bar{s}(\langle \mathbf{1}, \nu \rangle - 1) \,\middle|\, \nu \in [0, \infty)^{\text{COLS}_{\delta a}} \right\}. \tag{3.82}$$

By the strict coherence condition together with Lemma 3.50, there exists $\bar{\nu}_{\delta a} \in (0, \infty)^{\text{COLS}_{\delta a}}$ satisfying $A\bar{\nu}_{\delta a} = \pi$. This implies that the constraints of the (primal) optimization problem (3.81) are feasible, so OPT is well-defined. Since the domain $(0, \infty)^{\text{COLS}_{\delta a}}$ is an open set, it follows by a classical result (see e.g. [Roc97, Thm. 28.2]) that this problem enjoys strong Lagrange duality. This means that there exists $(\mathbf{S}, \bar{S})$ (*a priori*, not necessarily unique) such that

$$\text{OPT} = g(\mathbf{S}, \bar{S}) = \min_{\mathbf{s}, \bar{s}} \left\{ g(\mathbf{s}, \bar{s}) : \mathbf{s} \in \mathbb{R}^{D_a}, \bar{s} \in \mathbb{R} \right\}, \tag{3.83}$$

that is to say, the Lagrange dual problem (the right-hand side of (3.83)) achieves the same value as the constrained (primal) problem (3.81). Now, for any fixed $(\mathbf{s}, \bar{s})$, the value (3.82) of $g(\mathbf{s}, \bar{s})$ is given by optimizing a strictly concave objective over all $\nu \in [0, \infty)^{\text{COLS}_{\delta a}}$, so it is attained by a unique maximizer $\nu[\mathbf{s}, \bar{s}]$. Likewise, in the primal problem (3.81), since the entropy is strictly concave, the value of OPT must be uniquely attained by some measure ν^{op}. In the definition 3.82 of $g(\mathbf{s}, \bar{s})$, by substituting ν^{op} into the objective, we see that $g(\mathbf{s}, \bar{s}) \geq \text{OPT}$. If equality holds, then $\nu[\mathbf{s}, \bar{s}]$ must equal ν^{op}. This shows that $\nu[\mathbf{S}, \bar{S}]$ solves (3.81).

We next argue that the pair $(\mathbf{S}, \bar{S})$ is unique. To this end, note that for any $(\mathbf{s}, \bar{s})$, the maximizer $\nu[\mathbf{s}, \bar{s}]$ cannot occur on the boundary of $\nu \in [0, \infty)^{\text{COLS}_{\delta a}}$, because

$$\lim_{\nu(\underline{\sigma}_{\delta a}) \downarrow 0} \frac{-\nu(\underline{\sigma}_{\delta a}) \ln \nu(\underline{\sigma}_{\delta a})}{\nu(\underline{\sigma}_{\delta a})} = \infty.$$

Therefore $\nu[\mathbf{s}, \bar{s}]$ is the unique stationary point of the objective in $(0, \infty)^{\text{COLS}_{\delta a}}$, and we can solve for it by differentiation with respect to ν. It gives, for all $\underline{\sigma}_{\delta a} \in \text{COLS}_{\delta a}$,

$$\left(\nu[\mathbf{s}, \bar{s}] \right)(\underline{\sigma}_{\delta a}) = \exp \left\{ (A^{\text{t}} \mathbf{s})_{\underline{\sigma}_{\delta a}} + \bar{s} - 1 \right\} = \frac{\exp(\bar{s})}{e} \prod_{e \in \delta a} \exp(s_{e, \sigma_e}),$$

where for the final expression to make sense we define $s_{e, \mathbf{g}} = 0$. The corresponding objective value is

$$g(\mathbf{s}, \bar{s}) = \sum_{\underline{\sigma}_{\delta a} \in \text{COLS}_{\delta a}} \exp \left\{ (A^{\text{t}} \mathbf{s})_{\underline{\sigma}_{\delta a}} + \bar{s} - 1 \right\} - \langle \mathbf{s}, \pi \rangle - \bar{s} = \sum_{\underline{\sigma}_{\delta a} \in \text{COLS}_{\delta a}} \exp \left\{ (A')^{\text{t}} \begin{pmatrix} \mathbf{s} \\ \bar{s} \end{pmatrix} - 1 \right\} - \langle \mathbf{s}, \pi \rangle - \bar{s},$$

which is clearly strictly convex as a function of $(A')^{\text{t}}(\mathbf{s}, \bar{s})$. Since $(A')^{\text{t}}$ is a $|\{\text{COLS}_{\delta a}\}| \times (D_a + 1)$ matrix of rank $D_a + 1$, it defines an injective mapping, so we conclude that g is also strictly convex as a function of the Lagrange variables $(\mathbf{s}, \bar{s})$. This shows that the Lagrange dual problem (3.83) has a unique minimizer $(\mathbf{S}, \bar{S})$.

In summary, we have shown that $v[\mathbf{S}, \bar{S}] = v^{\mathrm{op}}$, the unique solution of the primal problem (3.81). Moreover, for any $(\mathbf{s}, \bar{s}) \neq (\mathbf{S}, \bar{S})$ the measure $v[\mathbf{s}, \bar{s}]$ does not satisfy the constraints of (3.81): supposing that it did, the fact that $(\mathbf{S}, \bar{S})$ is the unique minimizer of g would give

$$g(\mathbf{S}, \bar{S}) < g(\mathbf{s}, \bar{s}) = - \sum_{\underline{\sigma}_{\delta a} \in \mathrm{COLS}_{\delta a}} \left\{ \left(v[\mathbf{s}, \bar{s}] \right) (\underline{\sigma}_{\delta a}) \right\} \ln \left\{ \left(v[\mathbf{s}, \bar{s}] \right) (\underline{\sigma}_{\delta a}) \right\},$$

which gives a contradiction since $v[\mathbf{S}, \bar{S}] = v^{\mathrm{op}}$. The claimed result follows by setting $w_e(\sigma) = \exp(s_{e,\sigma})$ for all $e \in \delta a$, $\sigma \in \mathrm{supp}\,\pi_e$. □

As a byproduct of the considerations in the above proof, we can also compute the dimensions claimed in the proofs of Corollary 3.37 and Lemma 3.42:

Lemma 3.52 (proof of (3.52)). *In the setting of Corollary 3.37 we have $d_1(\mathscr{D}) = s(\mathscr{D}) - \wp(\mathscr{D})$.*

Proof. Recall from (3.52) that $d_1(\mathscr{D})$ counts the dimension of the space $\{v : v \sim {}_{\star}\omega\}$. Recall from Definition 3.34 that $v \equiv (\dot{v}, \hat{v})$ is the empirical measure of vertex colorings. Since ${}_{\star}\omega$ is fixed, we can treat each vertex type separately: if $d_1(T)$ is the dimension of the space of variable empirical measures $\dot{v}_T$ that are consistent with ${}_{\star}\omega$, and $d_1(L)$ is the dimension of the space of clause empirical measures $\hat{v}_L$ that are consistent with ${}_{\star}\omega$, then

$$d_1(\mathscr{D}) = \sum_{L} d_1(L) + \sum_{T} d_1(T). \tag{3.84}$$

For any given clause type L, let A' be the matrix from (3.80), which we showed to be of full rank in the proof of Corollary 3.51. The space of $\hat{v}_L$ that are consistent with ${}_{\star}\omega$ is an affine shift of the kernel of A' (intersected with the simplex of probability measures). It follows that

$$d_1(L) = \dim \ker A' = s_L - \mathrm{rk}\,A' = s_L - 1 - \sum_{j} (s_{L(j)} - 1). $$

To make the analogous calculation for a variable type T, we argue in a few steps:

a. First we need the analogue of the (strict) coherence condition for a collection of empirical measures around a variable v (rather than around a clause, as in Definition 3.6). We shall say that $(\pi_e)_{e \in \delta v}$ is **weakly coherent** if there exists a probability measure μ on $\{\text{-}, \text{+}, \text{f}\}$ such that

$$\left(\pi_e(\mathrm{p}), \pi_e(\mathrm{y}), \pi_e(\mathrm{g}) \right) = \left(\mu(\text{+L}_e), \mu(\text{-L}_e), \mu(\text{f}) \right)$$

for all $e \in \delta v$, and for both $\mathrm{L} \in \{\text{-}, \text{+}\}$ we have

$$\sum_{e \in \delta v(\mathrm{L})} \pi_e(\mathrm{r}) \geq \mu(\mathrm{L}). \tag{3.85}$$

We say that $(\pi_e)_{e \in \delta v}$ is **strictly coherent** if $\mu(\text{f})$ is strictly positive and (3.85) holds with strict inequality for both $\mathrm{L} \in \{\text{-}, \text{+}\}$. By a very similar argument as in Lemma 3.49, the canonical measures $({}_{\star}\pi_e)_{e \in \delta v}$ are strictly coherent.

b. By a similar argument as in Lemma 3.50, if $(\pi_e)_{e \in \delta v}$ is weakly coherent then there exists a measure $\dot{v}$ on variable colorings $\underline{\sigma}_{\delta v}$ with edge marginals π_e for all $e \in \delta v$. Indeed, clearly, $\dot{v}$ should give weight $\mu(\text{f})$ to the all-green coloring. For $\mathrm{L} \in \{\text{-}, \text{+}\}$, the measure $\dot{v}$ must give weight $\mu(\mathrm{L})$ to the colorings that are purple on all of $\delta v(\text{+L})$ and yellow on all of $\delta v(\text{-L})$. The only non-trivial step is to separate purple into red and blue, and this can be done by repeating the "wrapping around consecutive intervals" construction from Lemma 3.48.

c. Similarly as in Corollary 3.51, we define a matrix (the variable analogue of (3.80))

$$B' = \begin{pmatrix} B \\ \mathbf{1}^{\mathrm{t}} \end{pmatrix} \tag{3.86}$$

which encodes the edge marginal constraints. The columns of the matrix are indexed by the valid colorings $\underline{\sigma}_{\delta v}$ around the variable. For each frozen spin $\mathrm{L} \in \{\text{-}, \text{+}\}$ that the variable can take, the matrix B also has a row which is the indicator on the colorings that are purple on $\delta v(\text{+L})$. For each edge $e \in \delta v$ where both colors $\{\mathrm{r}, \mathrm{b}\}$ can appear, the matrix B has a row with entries $\mathbf{1}\{\sigma_e = \mathrm{r}\}$. Since $({}_{\star}\pi_e)_{e \in \delta v}$ is strictly coherent, we can repeat the argument of Corollary 3.51 to show that B' is a full rank matrix.

For each edge type $t \in T$, let $r_t = \mathbf{1}\{\{\mathtt{r}, \mathtt{b}\} \subseteq \mathrm{supp}\,_\star\pi_t\}$. It follows that

$$d_1(T) = \dim \ker B' = s_T - \mathrm{rk}\, B' = s_T - 1 - x_T - \sum_{t \in T} r_t \,,$$

where we recall from the discussion around (3.50) that $x_T \in \{0, 1, 2\}$ counts the number of frozen spins that can appear on a variable of type T. Substituting into (3.84) gives an expression for $d_1(\mathscr{D})$. Combining with our earlier calculation (3.44) of $s(\mathscr{D})$ gives

$$s(\mathscr{D}) - d_1(\mathscr{D}) = s(\mathscr{D}) \equiv \sum_L \sum_j (s_{L(j)} - 1) - \sum_T \left(-x_T + \sum_{t \in T} (s_t - 1 - r_t) \right)$$

$$= \sum_L \sum_j (s_{L(j)} - 1) - \sum_T x_T \Big(|\{t : t \in T\}| - 1 \Big),$$

where the last equality uses that for all $t \in T$ we have $s_t = |\,\mathrm{supp}\,_\star\pi_t| = 1 + x_T + r_t$. This matches the expression for $\wp(\mathscr{D})$ from (3.50), and concludes the proof. $\qquad\square$

Lemma 3.53 (calculation of (3.68)). *In the setting of Lemma 3.42 we have*

$$d_2(\mathscr{D}) = \sum_L \left((s_L)^2 - 1 - \sum_j \left[(s_{L(j)})^2 - 1 \right] \right) + \sum_T \left((s_T)^2 - 1 - \left[(x_T)^2 + 2x_T + 2x_T r_T + 3r_T \right] \right)$$

where r_T denotes the sum of $r_t = \mathbf{1}\{\{\mathtt{r}, \mathtt{b}\} \subseteq \mathrm{supp}\,_\star\pi_t\}$ over all edge types $t \in T$.

Proof. Recall that $d_2(\mathscr{D})$ counts the dimension of $\{v : v \sim \omega\}$ where ω is the pair empirical measure. Since ω is fixed, similarly to (3.84) we can decompose

$$d_2(\mathscr{D}) = \sum_L d_2(L) + \sum_T d_2(T). \tag{3.87}$$

Given a clause type L, let A' be the matrix from (3.80). Let $A[j]$ denote the submatrix of rows in A' concerning the j-th edge incident to the clause, so that A' consists of the submatrices $A[j]$ (for $1 \le j \le k(L)$) together with the all-ones row. Let A'' be the matrix whose rows are given by the pairwise tensor products $u \otimes w$ of rows u, w from A', **except** if u and w come from distinct blocks $A[j] \ne A[k]$. Since A' is of full rank (as was shown in the proof of Corollary 3.51), so is A''. It follows that

$$d_2(L) = \dim \ker A'' = (s_L)^2 - \mathrm{rk}\, A'' = (s_L)^2 - 1 - \sum_j \left[(s_{L(j)})^2 - 1 \right].$$

Similarly given a variable type T, let B' be the matrix from (3.86). Let $B^{\mathtt{r}}$ denote the submatrix of its first r_T rows, which concern the edges around the variable that can take on both colors $\{\mathtt{r}, \mathtt{b}\}$. Let B'' be the matrix whose rows are given by the pairwise tensor products $u \otimes w$ of rows u, w from B', **except** if u and w are distinct rows from $B^{\mathtt{r}}$. Since B' is of full rank (as was shown in the proof of Lemma 3.52), so is B''. It follows that

$$d_2(T) = \dim \ker B'' = (s_T)^2 - \mathrm{rk}\, B'' = (s_T)^2 - \left[\Big(1 + x_T + r_T \Big)^2 - r_T(r_T - 1) \right].$$

Substituting into (3.87) proves the claim. $\qquad\square$

Lemma 3.54 (proof of (3.69)). *In the setting of Lemma 3.42 we have $j_2(\mathscr{D}) = s_2(\mathscr{D}) - d_2(\mathscr{D}) - 2\wp(\mathscr{D})$.*

Proof. Recall that $j_2(\mathscr{D})$ counts the dimension of the space of judicious pair empirical measures ω around the product measure $_\star\omega$. Given a pair empirical measure ω (which we assume throughout this proof to be close to $_\star\omega$), let π_t be the average of the entries $\omega_{L,j}$ such that $L(j) = t$:

$$\pi_t = \sum_L \pi_{\mathscr{D}}(L \,|\, t)\omega_{L,j} \tag{3.88}$$

where $\pi_{\mathscr{D}}(L \,|\, t)$ is as in (3.76). If ω is judicious, then π_t must also be judicious. In order for ω to be feasible, for each T there must exist a measure μ_T on $\{\text{-}, \text{+}, \mathtt{f}\}^2$ which is consistent with π_t for every $t \in T$. The measure μ_T must also

be judicious, in the sense that its single-copy marginals are consistent with $_\star\pi_t$ for $t \in T$. Recall that $x_T \in \{0, 1, 2\}$ counts the number of frozen spins that can appear at a variable of type T. Let

$$C' = \begin{pmatrix} C \\ \mathbf{1}^t \end{pmatrix} \in \mathbb{R}^{(1+x_T) \times (1+x_T)},$$

where for each frozen spin that can appear at T, the matrix C contains a row which is the indicator of that spin. Then the matrix

$$C'' = \begin{pmatrix} C \otimes \mathbf{1}^t \\ \mathbf{1}^t \otimes C \\ \mathbf{1}^t \otimes \mathbf{1}^t \end{pmatrix} \in \mathbb{R}^{(2x_T+1) \times (1+x_T)^2}$$

encodes the judicious constraints on μ_T in the pair model. It follows that

$$d_2(\mu_T) \equiv \dim \left\{ \mu_T : \mu_T \text{ is judicious} \right\} = \dim \ker C'' = (1 + x_T)^2 - \left(2x_T + 1 \right) = (x_T)^2.$$

Now note that for each $t \in T$, the choice of μ_T already fixes π_t as a measure on $\{\mathrm{p}, \mathrm{y}, \mathrm{g}\}^2$. It remains to separate purple into red and blue in such a way that the resulting π_t has single-copy marginals $_\star\pi_t$:

$$d_2(\pi_t \mid \mu_T) \equiv \dim \left\{ \pi_t : \pi_t \text{ is judicious and consistent with } \mu_T \right\}$$

$$= \dim \left\{ \pi_t : \pi_t \text{ is judicious} \right\} - \dim \left\{ \mu_T : \mu_T \text{ is judicious} \right\} = (s_t - 1)^2 - (x_T)^2.$$

In order for ω to be consistent with π_t, it must satisfy the linear constraints (3.88). Note that some of these are redundant with the judicious constraints on the individual measures $\omega_{L,j}$. Altogether

$$d_2((\omega_{L,j})_{L(j)=t} \mid \pi_t) \equiv \dim \left\{ (\omega_{L,j})_{L(j)=t} : \omega \text{ is judicious and consistent with } \pi_t \right\}$$

$$= \sum_{L,j:L(j)=t} (s_{L(j)} - 1)^2 - (s_t - 1)^2$$

Combining these calculations gives

$$j_2(\mathscr{D}) = \sum_T \left(d_2(\mu_T) + \sum_{t \in T} \left[d_2(\pi_t \mid \mu_T) + d_2((\omega_{L,j})_{L(j)=t} \mid \pi_t) \right] \right)$$

$$= -\sum_T (x_T)^2 \left(|\{t : t \in T\}| - 1 \right) + \sum_{L,j} (s_{L(j)} - 1)^2.$$

Now recall that $s_2(\mathscr{D})$ is given in (3.67), and $d_2(\mathscr{D})$ was calculated in Lemma 3.53. Thus, if we abbreviate $\wp_2(\mathscr{D}) \equiv s_2(\mathscr{D}) - d_2(\mathscr{D}) - j_2(\mathscr{D})$, then we obtain

$$\wp_2(\mathscr{D}) = \sum_T \left((x_T)^2 |\{t : t \in T\}| + 2x_T + (2x_T + 3)r_T - \sum_{t \in T} \left[(s_t)^2 - 1 \right] \right) + 2 \sum_{L,j} (s_{L(j)} - 1).$$

For $t \in T$, we can expand $s_t = 1 + x_t + r_t$, so

$$\sum_{t \in T} \left((s_t)^2 - 1 \right) = \sum_{t \in T} \left((x_T)^2 + 2x_T + 2x_T r_t + 3r_t \right) = \left((x_T)^2 + 2x_T \right) |\{t : t \in T\}| + (2x_T + 3)r_T,$$

where we recall that r_T is the sum of r_t over all $t \in T$. Substituting into the previous calculation gives

$$\wp_2(\mathscr{D}) = -2 \sum_T x_T (|\{t : t \in T\}| - 1) + 2 \sum_{L,j} (s_{L(j)} - 1),$$

which is exactly twice the value of $\wp(\mathscr{D})$ from (3.50). This proves the claim. $\qquad\square$

We conclude this section with the main applications of the result of Corollary 3.51:

Corollary 3.55. *Let U' be any finite bipartite tree, and suppose $U \subseteq U'$ is a tree containing at least one clause, whose leaves are all variables. For each edge $e = (av)$ in U', let $_\star\pi_e$ be the canonical marginal on e based on $B_r(v)$, the r-neighborhood of v relative to U'. If all clauses in U are strictly coherent, then there is a set of edge weights $\gamma_e : \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\} \to [0, \infty)$ (for e in U, and with the convention $\gamma_e(\mathtt{g}) \equiv 1$), such that in the Gibbs measure on valid colorings of U where the edges are weighted by γ and the boundary variables are weighted by $_\star\dot{q}$, all edge marginals agree with the canonical ones $_\star\pi$. (The Gibbs measure is explicitly given by* (3.90) *below.)*

Proof. Let $_\star\dot{q}$ and $_\star\hat{q}$ be the canonical messages (based on r-neighborhoods) on the edges of U, as in Definition 3.4. For each variable v of U, let $\mathrm{COLS}_{\delta v}$ be the set of valid colorings $\underline{\sigma}_{\delta v}$ such that $\sigma_e \in \mathrm{supp}\,_\star\pi_e$ for all $e \in \delta v$. Now recall the discussion around (2.33): the probability measure

$$\nu_{\delta v}(\underline{\sigma}_{\delta v}) = \frac{1}{\dot{z}_v} \varphi_v(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} {}_\star\hat{q}_e(\sigma_e) \tag{3.89}$$

is supported on valid colorings of δv, and has edge marginals on δv that are consistent with $_\star\pi$. Moreover, it is clear from the above expression that if $e \in \delta v$ and $_\star\pi_e(\sigma) > 0$, then we must have $_\star\hat{q}_e(\sigma) > 0$. Consequently, for any $\underline{\sigma}_{\delta v} \in \mathrm{COLS}_{\delta v}$, we have $\varphi_v(\underline{\sigma}_{\delta v}) = 1$ and $_\star\hat{q}_e(\sigma_e) > 0$ for all $e \in \delta a$, which means that (3.89) is positive. That is to say, the measure $\nu_{\delta v}$ defined by (3.89) has support exactly equal to $\mathrm{COLS}_{\delta v}$.

We next make a simple observation: for any edge e of U and any vertex x incident to e, we have $\sigma \in \mathrm{supp}\,_\star\pi_e$ if and only if there exists $\underline{\sigma}_{\delta x} \in \mathrm{COLS}_{\delta x}$ which has value σ on edge e. The "if" direction is obvious. For the "only if" direction, recall that we have constructed $\nu_{\delta x}$ with marginals $_\star\pi$ (this follows from Lemma 3.48 if x is a clause, and from the above discussion if x is a variable). Since $_\star\pi_e$ is the marginal of $\nu_{\delta x}$, if $\sigma \in \mathrm{supp}\,_\star\pi_e$ then the measure $\nu_{\delta x}$ must give positive weight to some $\underline{\sigma}_{\delta x}$ which has value σ on edge e. Then $\underline{\sigma}_{\delta x} \in \mathrm{supp}\,\nu_{\delta x} \subseteq \mathrm{COLS}_{\delta x}$, as desired.

Now define COLS_U to be the set of valid colorings $\underline{\sigma}_U$ of U such that $\sigma_e \in \mathrm{supp}\,_\star\pi_e$ for all edges e of U. This is simply because U is a tree, so we can construct an element of COLS_U as follows: start from any vertex x of U, and choose $\underline{\sigma}_{\delta x} \in \mathrm{COLS}_{\delta x}$. For each vertex $y \in \partial x$, by the preceding observation, the set $\mathrm{COLS}_{\delta y}$ must contain an element $\underline{\sigma}_{\delta y}$ which agrees with $\underline{\sigma}_{\delta y}$ on the edge $e = (xy)$. Choose this $\underline{\sigma}_{\delta y}$, and proceed in the same way to color the edges incident to the neighbors of y, and so on until all of U is colored. This shows $\mathrm{COLS}_U \neq \emptyset$.

Now recall from (3.4) that $_\star\pi_e(\sigma)$ is proportional to $_\star\dot{q}_e(\sigma)_\star\hat{q}_e(\sigma)$. In particular, if $\sigma \in \mathrm{supp}\,_\star\pi_e$, then $_\star\dot{q}_e(\sigma)$ must be positive, in which case we can define

$$\gamma_e(\sigma) \equiv \frac{w_e(\sigma)}{_\star\dot{q}_e(\sigma)} {}_\star\dot{q}_e(\mathtt{g}) .$$

for w_e as given by Corollary 3.51. Recall that $\mathtt{g} \in \mathrm{supp}\,_\star\pi_e$ and we took the convention $w_e(\mathtt{g}) \equiv 1$, so now $\gamma_e(\mathtt{g}) \equiv 1$ also. For $\sigma \notin \mathrm{supp}\,_\star\pi_e$ we simply define $\gamma_e(\sigma) = 0$. We write $U = (V, F, E)$. Write $\mathcal{L}U$ for all the leaves of U, which were assumed to be variables. For each $u \in \mathcal{L}U$ let $a(u)$ denote the unique clause in U that neighbors u. The Gibbs measure described in the statement of this lemma can be expressed as

$$\nu(\underline{\sigma}_U) \cong \prod_{v \in V \setminus \mathcal{L}U} \varphi_v(\underline{\sigma}_{\delta v}) \prod_{a \in F} \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{e \in E} \gamma_e(\sigma_e) \prod_{u \in \mathcal{L}U} {}_\star\dot{q}_{ua(u)}(\sigma_{ua(u)}) , \tag{3.90}$$

where $\cong$ denotes the normalization constant, which is positive since $\mathrm{supp}\,\gamma_e = \mathrm{supp}\,_\star\pi_e$, $\mathrm{supp}\,_\star\dot{q}_e \supseteq \mathrm{supp}\,_\star\pi_e$, and $\mathrm{COLS}_U \neq \emptyset$. We claim that this measure ν has marginals $_\star\pi$ on the edges of U. To this end, note that the canonical messages $_\star\dot{q}$, $_\star\hat{q}$ solve the belief propagation equations for ν_U:

 (i) If $u \in \mathcal{L}U$, then we regard u as being weighted only by $_\star\dot{q}_{ua(u)}$, so then the BP message from u to $a(u)$ will be $_\star\dot{q}_{ua(u)}$ simply by the standard conventions of BP at the leaves of trees.
 (ii) If v is an internal variable of U, then we regard v as being unweighted, and the equation $_\star\dot{q}_{va}(\sigma) = \mathrm{BP}_{va}[_\star\hat{q}]$ is satisfied for all $a \in \partial v$ by definition of $_\star\dot{q}$ and $_\star\hat{q}$ (cf. Remark 3.5).
(iii) Finally, if a is a clause of U, then we regard a as being weighted by the product of γ_e over $e \in \delta a$. Consider

$$\nu_{\delta a}(\underline{\sigma}_{\delta a}) \cong \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{e \in \delta a} \left\{ \gamma_e(\sigma_e)_\star\dot{q}_e(\sigma_e) \right\} \cong \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{e \in \delta a} w_e(\sigma_e) . \tag{3.91}$$

(We are not yet claiming that $\nu_{\delta a}$ is the marginal of ν.) By the result of Corollary 3.51, the measure $\nu_{\delta a}$ has marginals $\star\pi$, so for all $e = (av) \in \delta a$ we have

$$\star\pi_e(\sigma) \cong \sum_{\underline{\sigma}_{\delta a}} \mathbf{1}\{\sigma_e = \sigma\}\hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{e' \in \delta a} \left\{\gamma_{e'}(\sigma_{e'})\star\dot{q}_{e'}(\sigma_{e'})\right\} \cong \star\dot{q}_e(\sigma) \cdot \mathrm{BP}_{av}[\star\dot{q}; \gamma](\sigma).$$

On the other hand we know $\star\pi_e(\sigma) \cong \star\dot{q}_e(\sigma)\star\hat{q}_e(\sigma)$, so it must be that $\mathrm{BP}_{av}[\star\dot{q}; \gamma](\sigma) = \star\hat{q}_{av}(\sigma)$.

This proves our claim that $(\star\dot{q}, \star\hat{q})$ solves the belief propagation equations for ν. It follows from this that the marginal of ν on each δa is indeed correctly described by (3.91), which in turn implies that the edge marginals are consistent with $\star\pi$.   □

In §6.1 we will show that if the clauses are nice (Definition 3.8), then we can estimate the weights of Corollary 3.55, which will be needed in the second half of the paper. For the precise statement see Corollary 6.2.

Corollary 3.56. *In the setting of Corollary 3.55, there is a system $\star\Lambda_U$ of positive variable weights*

$$\Lambda_v = \begin{pmatrix} \lambda_v \equiv (\lambda_v(+) = 1, \lambda_v(-), \lambda_v(\mathtt{f})) \\ \lambda_e \equiv \lambda_e(\mathtt{r}) \end{pmatrix}$$

such that the $\star\Lambda_U$-weighted Gibbs measure on valid colorings of U has all edge marginals agreeing with the canonical edge marginals $\star\pi$. (However, since the weights were shifted from clauses to variables, the associated BP messages will no longer be $\star\dot{q}, \star\hat{q}$.)

Proof. The Gibbs measure (3.90) is equivalent to the Gibbs measure in which all variables are unweighted; all internal edges e of U are weighted by $\tilde{\gamma}_e = \gamma_e$; and all leaf edges e of U are weighted by

$$\tilde{\gamma}_e(\sigma) = \frac{\gamma_e(\sigma)\star\dot{q}_e(\sigma)}{\sum_\tau \gamma_e(\tau)\star\dot{q}_e(\tau)}.$$

To define $\star\Lambda_U$, we redistribute the weights as follows. For each internal variable v of U, we let $x_v \in \{-, +, \mathtt{f}\}$ denote the frozen spin corresponding corresponding to $\underline{\sigma}_{\delta v}$, and put the weight

$$\star\Lambda_v(\underline{\sigma}_{\delta v}) \equiv \star\lambda_v(x_v) \prod_{e \in \delta v} \star\lambda_e(\sigma_e),$$

where $\star\lambda_v$ and $\star\lambda_e$ are defined by (2.30). For each leaf variable $v \in \mathcal{L}U$, the set δv consists of a single edge e, and in this case we will simply take the weight

$$\star\lambda_v(\sigma_e) \equiv \star\lambda_e(\sigma_e) \equiv \tilde{\gamma}_e(\sigma_e).$$

For the redistributed weights, the BP messages on all edges e of U are given by

$$\bullet\dot{q}_e(\sigma_e) \cong \star\dot{q}_e(\sigma_e)\gamma_e(\sigma_e), \quad \bullet\hat{q}_e(\sigma_e) \cong \frac{\star\hat{q}_e(\sigma_e)}{\gamma_e(\sigma_e)}. \tag{3.92}$$

Thus the messages for the redistributed weights $\star\Lambda_U$ also satisfy the familiar identity $\star\pi_e(\sigma_e) \cong \bullet\dot{q}_e(\sigma_e)\bullet\hat{q}_e(\sigma_e)$. Moreover, if e is a leaf edge then $\star\lambda_e = \tilde{\gamma}_e = \bullet\dot{q}_e$.   □

Outline of remaining sections. The remainder of this paper is organized as follows:

- In Section 4 we analyze the distributional 1-RSB recursion (1.10) to prove Propositions 1.1 and 1.3.
- In Section 5 we analyze the random k-SAT graph and the preprocessing algorithm to prove Propositions 3.22–3.24.
- In Section 6 we prove Propositions 3.30 and 3.31.
- Sections 7–9 are devoted to the proof of Proposition 3.32.
- In Section 10 we prove Proposition 1.2.

4. One-step RSB threshold

In this section we analyze the distributional 1-RSB recursion (1.10) to prove Propositions 1.1 and 1.3. The section is organized as follows:

– In §4.1 we formally define the random Galton–Watson tree that arises as a local weak limit of the k-SAT graph. We prove a basic estimate, Lemma 4.7, on the volume of neighborhoods in the random tree. We also formalize the natural coupling between the 1-RSB recursion and the random tree.
– In §4.2 we prove preliminary concentration bounds on the effect of the distributional recursion.
– In §4.3 we use the bounds from §4.2 to show that the root of the Galton–Watson tree is very likely to be **nice** in a strong sense (Proposition 4.15). This will be used in §5.1 to bound the occurrence of **defective** regions.
– In §4.4 we use the coupling of the distributional recursion with the tree to show Proposition 4.17, which says that the root of the Galton–Watson tree is very likely to be **stable**. From this we will deduce the result of Proposition 1.1. Proposition 4.17 will furthermore be used in the proof of Proposition 3.29, and also in §5.4 to assure that the initial set A of preprocessing (Definition 3.15) is small.
– In §4.5 we prove Proposition 1.3, the 1-RSB upper bound on the satisfiability threshold. This is deduced from the interpolation bounds of [FL03, PT04].
– Lastly, in §4.6 we prove Proposition 3.29, showing that the first moment of judicious colorings asymptotically is lower bounded (in the exponent) by the 1-RSB free energy $\Phi(\alpha)$ of (1.11).

It is assumed throughout the section, even when not explicitly stated, that $k \geq k_0$ and α satisfies (1.7).

4.1. **Random trees and the distributional recursion.** We already mentioned in §3.1 that the random k-SAT graph converges "locally in distribution" (also termed "locally in law," or "locally weakly") to the Galton–Watson measure $\mathbf{PGW} \equiv \mathbf{PGW}^\alpha$. We begin this section by formally restating the definition of $\mathbf{PGW}$, and then reviewing some basic notions of local weak convergence which will be used later.

Definition 4.1 (Galton–Watson measure, $\mathbf{PGW}$). We define the **bipartite Poisson Galton–Watson tree** to be the random bipartite factor tree $\mathscr{T}$ which is generated as follows: starting from a root variable v_{RT}, each variable independently generates $\text{Pois}(\alpha k)$ child clauses, and each clause generates $k-1$ child variables. Each edge is labelled with a literal L which takes values + or - with equal probability, independently over all the edges. We write $\mathbf{PGW}$ for the law of this tree.

Remark 4.2. The measure $\mathbf{PGW}$ is the "local weak limit" of the random k-SAT graph in the following formal sense: if $\mathscr{G}_n = (V_n, F_n, E_n)$ denotes an instance of random k-SAT at clause density α on n variables, and I_n is a uniformly random element of $V_n \equiv [n]$, then it holds for any finite r that

$$\lim_{n \to \infty} \mathbb{P}\left(B_r(I_n; \mathscr{G}_n) \cong T\right) = \mathbf{PGW}^\alpha\left(B_r(v_{\text{RT}}; \mathscr{T}) \cong T\right)$$

for every T. In the above, $B_r(I_n; \mathscr{G}_n)$ is the r-neighborhood of I_n in $\mathscr{G}_n$, viewed as a graph rooted at I_n. Likewise $B_r(v_{\text{RT}}; \mathscr{T})$ is the r-neighborhood of v_{RT} in $\mathscr{T}$, viewed as a graph rooted at v_{RT}. Finally T is any graph with a root o, and $\cong$ denotes isomorphism of rooted graphs. The measure $\mathbf{PGW}$ is **unimodular** (cf. [LPP95] and [AL07, Defn. 2.1]):

$$\int \left[\sum_u f(\mathscr{T}, v_{\text{RT}}, u)\right] d\mathbf{PGW}(\mathscr{T}) = \int \left[\sum_u f(\mathscr{T}, u, v_{\text{RT}})\right] d\mathbf{PGW}(\mathscr{T}) \tag{4.1}$$

where the sum goes over all variables u in $\mathscr{T}$, and f is any nonnegative Borel function on the space $\mathscr{G}_{\star\star}$ of bi-rooted graphs.[15] To see why (4.1) should hold, note that changing the order of summation gives

$$\mathbb{E} \sum_{u \in \partial I_n} f(\mathscr{G}_n, I_n, u) = \frac{1}{|V_n|} \sum_{v \in V_n} \left[\sum_{u \in N(v)} f(\mathscr{G}_n, v, u)\right] = \mathbb{E} \sum_{u \in \partial I_n} f(\mathscr{G}_n, u, I_n). \tag{4.2}$$

In the limit $n \to \infty$, the left-hand side of (4.2) converges to the left-hand side of (4.1), while the right-hand side of (4.2) converges to the right-hand side of (4.1).

---

[15]A bi-rooted graph is a graph rooted at an ordered pair of vertices. For a detailed account and careful discussion of topological considerations (in particular, what it means to be a Borel function on the space $\mathscr{G}_{\star\star}$), we refer to [AL07].

We emphasize that under the measure $\mathbf{PGW}$, the root v_{RT} plays a special role: in the local weak limit interpretation, v_{RT} represents a uniformly random variable, while other vertices in the tree represent random *neighbors*. Thus the root degree $|\partial v_{\mathrm{RT}}|$ is distributed as $\mathrm{Pois}(\alpha k)$, while for variables $u \neq v_{\mathrm{RT}}$, the degree $|\partial u|$ is distributed as a **size-biased** $\mathrm{Pois}(\alpha k)$, with probability mass function

$$\overline{\mathrm{pois}}_{\alpha k}(j) = \frac{j\,\mathrm{pois}_{\alpha k}(j)}{\alpha k}\,.$$

The most obvious distinguishing feature of v_{RT} is that it has degree zero with positive probability, while any other vertex of the random tree must have positive degree since it connects to its parent. In the above description of the $\mathbf{PGW}$ law, we used the fact that

$$\overline{\mathrm{pois}}_{\alpha k}(j) = \frac{e^{-\alpha k}(\alpha k)^j \cdot j}{j! \cdot \alpha k} = \frac{e^{-\alpha k}(\alpha k)^{j-1}}{(j-1)!} = \mathrm{pois}_{\alpha k}(j-1)\,,$$

meaning that a size-biased $\mathrm{Pois}(\alpha k)$ random variable is equidistributed as a $[1 + \mathrm{Pois}(\alpha k)]$ random variable — this explains why variables $u \neq v_{\mathrm{RT}}$ generate a $\mathrm{Pois}(\alpha k)$ number of children. To understand how (canonical) messages behave in the random k-SAT graph, we study how they behave in the limiting random tree $\mathscr{T} \sim \mathbf{PGW}$. For this purpose, we make the following definitions:

Definition 4.3 (variable-to-clause measure, PGW). Let $\mathscr{T} \sim \mathbf{PGW}$ with root variable v_{RT}. Let $\mathscr{T}_{v_{\mathrm{RT}} a_{\mathrm{RT}}}$ be the tree $\mathscr{T}$ together with one additional edge $e_{\mathrm{RT}} \equiv (v_{\mathrm{RT}} a_{\mathrm{RT}})$ incident to the root, equipped with a random sign $\mathsf{L}_{e_{\mathrm{RT}}}$. We think of e_{RT} as the parent edge of v_{RT}, pointing to a deleted clause a_{RT}. Then $\mathscr{T}_{v_{\mathrm{RT}} a_{\mathrm{RT}}}$ is a random variable-to-clause tree, and we denote its law PGW.

Definition 4.4 (clause-to-variable measure, $\widehat{\mathrm{PGW}}$). Let $\mathscr{T}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}$ be the random tree formed as follows: start with a root clause a_{RT}, and attach $k-1$ independent samples of PGW (Definition 4.3) as subtrees to a_{RT}. Then attach to a_{RT} one additional incident edge e_{RT}, which we think of as the parent edge of a_{RT}, pointing to a deleted variable v_{RT}. The result is a random clause-to-variable tree $\mathscr{T}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}$, whose law we denote as $\widehat{\mathrm{PGW}}$.

The definitions of PGW and $\widehat{\mathrm{PGW}}$ can be viewed in this way: let $\mathscr{T} \sim \mathbf{PGW}$, $a \in \partial v_{\mathrm{RT}}$, and $u \in \partial a \setminus v_{\mathrm{RT}}$. Recall from Definition 2.13 and the subsequent discussion that the message from u to a is defined in terms of $\mathscr{T}_{ua}$ — the component of $\mathscr{T} \setminus a$ containing u, including the edge (ua) but not including a itself. It follows from the above discussion that if we also remove (ua) from $\mathscr{T}_{ua}$, then the resulting tree (rooted at u) is equidistributed as the original tree $\mathscr{T} \sim \mathbf{PGW}$. It follows that the law of $\mathscr{T}_{ua}$ is precisely the measure PGW. Similarly, if $\mathscr{T} \sim \mathbf{PGW}$ and $a \in \partial v_{\mathrm{RT}}$, then the law of $\mathscr{T}_{av_{\mathrm{RT}}}$ is given by $\widehat{\mathrm{PGW}}$. In summary, in order to understand messages in the random k-SAT graph, we will study messages from v_{RT} to a_{RT} under the law PGW, and messages from a_{RT} to v_{RT} under the law $\widehat{\mathrm{PGW}}$. In fact, for technical reasons (which will become apparent in the proofs of this section and the next one) we will study messages on slightly more general trees, defined as follows:

Definition 4.5 (Galton–Watson based on fixed tree, $\mathbf{PGW}(T)$). Let T be any fixed tree rooted at a variable v_{RT}, such that all clauses in T have at most $k-1$ children. At every variable v of T (including $v = v_{\mathrm{RT}}$), attach an independent subtree $\mathscr{T}_v \sim \mathbf{PGW}$ (by identifying v with the root of $\mathscr{T}_v$). At every clause a of T, if a has $k-1-j$ children, attach j independent samples of PGW as subtrees to a. Let $\mathbf{PGW}(T)$ be the law of the resulting tree.

Similarly, if $T_{v_{\mathrm{RT}} a_{\mathrm{RT}}}$ is a fixed variable-to-clause tree where every clause has at most $k-1$ children, we can perform the same procedure as above to obtain a random variable-to-clause tree, whose law we denote $\mathrm{PGW}(T_{v_{\mathrm{RT}} a_{\mathrm{RT}}})$. Likewise, if $T_{a_{\mathrm{RT}} v_{\mathrm{RT}}}$ is any fixed clause-to-variable tree where every clause has at most $k-1$ children, we use $\widehat{\mathrm{PGW}}(T_{a_{\mathrm{RT}} v_{\mathrm{RT}}})$ to denote the law of the random clause-to-variable tree based on $T_{a_{\mathrm{RT}} v_{\mathrm{RT}}}$.

Definition 4.6 (Galton–Watson with random deletions, $\mathbf{PGW}_\epsilon$). Let $\mathbf{PGW}_\epsilon$ be the law of the random tree generated as follows: starting from a root variable v_{RT}, each variable independently generates $\mathrm{Pois}(\alpha k)$ clauses, and each clause independently generates either $k-1$ child variables (with probability $1-\epsilon$) or $k-2$ child variables (with probability ϵ). When $\epsilon = 0$, the measure $\mathbf{PGW}_\epsilon$ coincides with the measure $\mathbf{PGW}$ of Definition 4.1. The measure $\mathbf{PGW}_\epsilon$ is also unimodular, because it can also be obtained via a local weak limit: let $\hat{\epsilon}$ and $\hat{\alpha}$ be defined by

$$\epsilon = \frac{(k-1)\hat{\epsilon}}{k - \hat{\epsilon}}\,, \quad \hat{\alpha} = \frac{\alpha}{1 - \hat{\epsilon}/k}$$

Start with n isolated variables and $n\hat{\alpha}$ isolated clauses. For each of the first $1 - \hat{\epsilon}$ fraction of the clauses, put k edges to randomly chosen variables. For the last $\hat{\epsilon}$ fraction of the clauses, put $k - 1$ edges to randomly chosen variables. The total number of edges in the resulting random graph is

$$n\hat{\alpha}\left((1 - \hat{\epsilon})k + \hat{\epsilon}(k - 1)\right) = n\alpha k.$$

In the limit $n \to \infty$, this random graph converges locally weakly to the $\mathbf{PGW}_\epsilon$ measure, justifying our claims that $\mathbf{PGW}_\epsilon$ is unimodular. Define likewise PGW_ϵ and $\widehat{\mathrm{PGW}}_\epsilon$ to be the obvious generalizations of the measures PGW and $\widehat{\mathrm{PGW}}$ from Definitions 4.3 and 4.4. We hereafter write cl_ϵ for the probability measure which puts weight $1 - \epsilon$ on k, and weight ϵ on $k - 1$.

The next lemma gives a simple bound on the growth of balls under the $\mathbf{PGW}$ measure:

Lemma 4.7. *Let $\mathscr{T} \sim \mathbf{PGW}$, and consider the ℓ-neighborhood of the root, $B_\ell(v_{\mathrm{RT}}) \equiv B_\ell(v_{\mathrm{RT}}; \mathscr{T})$. We have*

$$e_\ell \equiv \int \exp\left\{\frac{|B_\ell(v_{\mathrm{RT}}; \mathscr{T})|}{(\alpha k^2)^\ell}\right\} d\mathbf{PGW}(\mathscr{T}) \le e \tag{4.3}$$

for all $\ell \ge 0$.

Proof. Let S_ℓ be the number of variables in $\mathscr{T}$ at distance exactly ℓ from the root v_{RT}, with $S_0 = 1$. Then

$$e_\ell = \mathbb{E}\left[\exp\left\{\frac{1}{(\alpha k^2)^\ell}\sum_{j=0}^{\ell} S_j\right\}\right].$$

Let $\mathscr{F}_\ell$ denote the σ-field generated by $B_{\ell-1}(v_{\mathrm{RT}})$: then

$$\mathbb{E}\left[\exp\left\{\frac{S_\ell}{(\alpha k^2)^\ell}\right\}\ \middle|\ \mathscr{F}_{\ell-1}\right] = \exp\left\{S_{\ell-1}\left[\exp\left(\frac{k-1}{(\alpha k^2)^\ell}\right) - 1\right]\alpha k\right\} \le \exp\left\{\frac{S_{\ell-1}[1 - \Theta(1/k)]}{(\alpha k^2)^{\ell-1}}\right\}.$$

It follows by iterated expectations that

$$e_\ell \le \mathbb{E}\left[\exp\left\{\frac{1}{(\alpha k^2)^\ell}\sum_{j=0}^{\ell-2} S_j + \frac{S_{\ell-1}}{(\alpha k^2)^\ell} + \frac{S_{\ell-1}[1 - \Theta(1/k)]}{(\alpha k^2)^{\ell-1}}\right\}\right] \le e_{\ell-1} \le e_0 = e.$$

This proves the claim. $\square$

We now review the distributional recursion introduced in §1.6, and which we saw in §2.5 is related to the frozen model on finite trees. (Recall, in particular, the similarity between (1.10) and (2.27).) Let d^+, d^- be independent samples from the $\mathrm{Pois}(\alpha k/2)$ distribution, and write $\underline{d} \equiv (d^+, d^-)$. We denote their probability mass function by

$$\mathrm{po}_\pm(\underline{d}) \equiv \frac{e^{-\alpha k}(\alpha k/2)^{d^+ + d^-}}{(d^+)!(d^-)!}.$$

Note that if $\mathscr{T}_{v_{\mathrm{RT}}a_{\mathrm{RT}}} \sim \mathrm{PGW}$ as discussed above, then (with the notation of (2.22)) the pair

$$\left(\left|\delta v(+a_{\mathrm{RT}})\right|, \left|\delta v(-a_{\mathrm{RT}})\right|\right)$$

has law exactly $\mathrm{po}_\pm$. We next set some notations for laws on messages:

– **Variable-to-clause messages.** As before, we use $\boldsymbol{\eta}$ to denote a probability measure over $\{+, -, \mathtt{f}\}$, interpreted as a message from some variable v to one of its neighboring clauses a. We will often summarize $\boldsymbol{\eta}$ by the scalar value $\eta = \boldsymbol{\eta}(-) \in [0, 1)$, which is interpreted as the chance (according to the message) for v not to satisfy a. We write $\mathcal{P}$ for the space of probability measures over $\boldsymbol{\eta}$, and use $\boldsymbol{\mu}$ to denote elements of $\mathcal{P}$. We write $\mathscr{P}$ for the space of probability measures over $\eta \in [0, 1)$, and use μ to denote elements of $\mathscr{P}$. The mapping from $\boldsymbol{\eta}$ to $\eta = \boldsymbol{\eta}(-)$ naturally induces a mapping from $\boldsymbol{\mu} \in \mathcal{P}$ to $\mu \in \mathscr{P}$.

– **Clause-to-variable messages.** Likewise, we use $\hat{\boldsymbol{u}}$ to denote a probability measure over $\{+, \mathtt{f}\}$, with the interpretation of a message from some clause a to one of its neighboring variables v. We will often summarize $\hat{\boldsymbol{u}}$ by the scalar value $\hat{u} = \hat{\boldsymbol{u}}(+) \in [0, 1)$, which is interpreted as the chance (according to the message) that a is forcing to v. Of course, since $\hat{\boldsymbol{u}}$ is a probability measure over only two elements, the correspondence between $\hat{\boldsymbol{u}}$ and $\hat{u}$ is a bijection. We write $\hat{\mathcal{P}}$ for the space of probability measures over $\hat{\boldsymbol{u}}$, and use $\hat{\boldsymbol{\mu}}$ to denote elements of $\hat{\mathcal{P}}$. The (one-to-one) mapping from $\hat{\boldsymbol{u}}$ to $\hat{u} = \hat{\boldsymbol{u}}(+)$ naturally induces a one-to-one mapping from $\hat{\boldsymbol{\mu}} \in \hat{\mathcal{P}}$ to $\hat{\mu} \in \mathscr{P}$.

It may be useful to keep in mind that, ultimately, the measures μ and $\hat{\boldsymbol{\mu}}$ will capture the influence of the random local geometry in the graph. To construct μ and $\hat{\boldsymbol{\mu}}$ appropriately, we next define the mappings that will capture the distributional effect of the frozen model recursions (2.25) and (2.26).

Definition 4.8 (distributional effect of clause recursion (2.25)). Given $\mu \in \mathscr{P}$, let $\underline{\eta}' \equiv (\eta_j)_{j \geq 1}$ denote a sequence of i.i.d. samples from μ. Independently, let $K \sim \mathrm{cl}_\epsilon$ (as in Definition 4.6). Let $\hat{\boldsymbol{u}} \equiv \hat{\boldsymbol{u}}(\underline{\eta}')$ be the probability measure on $\{+, \mathtt{f}\}$ defined by

$$\left(\hat{\boldsymbol{u}}(+), \hat{\boldsymbol{u}}(\mathtt{f}) \right) = \left(\prod_{j=1}^{K-1} \eta_j, 1 - \prod_{j=1}^{K-1} \eta_j \right).$$

Let $\hat{\mathcal{R}}_\epsilon \mu$ denote the law of $\hat{\boldsymbol{u}}$. Thus $\hat{\mathcal{R}}_\epsilon$ defines a mapping from $\mathscr{P}$ to $\hat{\mathcal{P}}$, which captures how randomness is passed through the clause update (2.25) of the frozen model recursions. We write $\hat{\mathcal{R}} \equiv \hat{\mathcal{R}}_0$ for the $\epsilon = 0$ case.

Definition 4.9 (distributional effect of variable recursion (2.26)). Given $\hat{\boldsymbol{\mu}} \in \hat{\mathcal{P}}$, let $\underline{\hat{\boldsymbol{u}}} \equiv (\hat{\boldsymbol{u}}_i^+, \hat{\boldsymbol{u}}_i^-)_{i \geq 1}$ be an array of i.i.d. samples from $\hat{\boldsymbol{\mu}}$, and write $\hat{u}_i^\pm \equiv \hat{\boldsymbol{u}}_i^\pm(+) \in [0, 1)$. Let $\underline{d} \sim \mathrm{po}_\pm$, and define

$$\Pi^+ \equiv \Pi^+(\underline{d}, \underline{\hat{\boldsymbol{u}}}) \equiv \prod_{i=1}^{d^+} \left(1 - \hat{u}_i^+ \right), \quad \Pi^- \equiv \Pi^-(\underline{d}, \underline{\hat{\boldsymbol{u}}}) \equiv \prod_{i=1}^{d^+} \left(1 - \hat{u}_i^- \right),$$

and note that $\Pi^\pm \in (0, 1]$. Let $\eta \equiv \eta(\underline{\hat{\boldsymbol{u}}})$ be the probability measure defined by

$$\left(\eta(+), \eta(-), \eta(\mathtt{f}) \right) = \left(\frac{\Pi^-(1 - \Pi^+)}{\Pi^+ + \Pi^- - \Pi^+\Pi^-}, \frac{\Pi^+(1 - \Pi^-)}{\Pi^+ + \Pi^- - \Pi^+\Pi^-}, \frac{\Pi^+\Pi^-}{\Pi^+ + \Pi^- - \Pi^+\Pi^-} \right). \tag{4.4}$$

Let $\dot{\mathcal{R}}\hat{\boldsymbol{u}}$ denote the law of this η. Thus $\dot{\mathcal{R}}$ defines a mapping from $\hat{\mathcal{P}}$ to $\mathscr{P}$, which captures how randomness is passed through the variable BP recursion.

Definition 4.10 (full distributional recursion). Let $\mathcal{R}_\epsilon$ denote the composition $\dot{\mathcal{R}} \circ \hat{\mathcal{R}}_\epsilon$: this gives a mapping from $\mathscr{P}$ to $\mathscr{P}$, which captures how randomness is passed through one full update (clause updates follows by variable update) of variable-to-clause messages. We can explicitly describe $\mathcal{R}_\epsilon$ as follows: given $\mu \in \mathscr{P}$, let $\underline{\eta}$ be an array of i.i.d. samples from μ, as in (1.8). Independently, let $\underline{d} \sim \mathrm{po}_\pm$, and let $\underline{K} \equiv (K_i)_{i \geq 1}$ be a sequence of i.i.d. samples from cl_ϵ (as in Definition 4.6). From this, define the random variables

$$\Pi^\pm \equiv \Pi^\pm(\underline{d}, \underline{\eta}) \equiv \prod_{i=1}^{d^\pm} \left(1 - \prod_{j=1}^{K_i - 1} \eta_{ij}^\pm \right) = \prod_{i=1}^{d^\pm} \left(1 - \hat{u}_i^\pm \right).$$

Substitute these $\Pi^\pm$ into (4.4) to define $\eta = \eta(\underline{d}, \underline{K}, \underline{\eta})$. Then the law of $\eta(\underline{d}, \underline{K}, \underline{\eta})$ is $\mathcal{R}_\epsilon \mu \in \mathscr{P}$. As in (1.10) we let

$$R(\underline{d}, \underline{K}, \underline{\eta}) \equiv \left[\eta(\underline{d}, \underline{K}, \underline{\eta}) \right](-).$$

As we noted in §1.6, since all the η's must lie in $[0, 1)$, we must have $\Pi^\pm \in (0, 1]$ and therefore $R(\underline{d}, \underline{K}, \underline{\eta}) \in [0, 1)$. We finally define $\boldsymbol{R}_\epsilon \mu \in \mathscr{P}$ to be the law of $R(\underline{d}, \underline{K}, \underline{\eta})$, so $\boldsymbol{R}_\epsilon$ gives a mapping from $\mathscr{P}$ to itself.[16] We denote $\mathcal{R} \equiv \mathcal{R}_\epsilon$ and $\boldsymbol{R} \equiv \boldsymbol{R}_\epsilon$ for the $\epsilon = 0$ case.

---

[16]The formal characterization of $\boldsymbol{R}_\epsilon$ is that for any measurable $B \subseteq [0, 1)$,

$$(\boldsymbol{R}_\epsilon \mu)(\eta \in B) = \sum_{\underline{d}} \mathrm{po}_\pm(\underline{d}) \int \left[\int \mathbf{1}\{R(\underline{d}, \underline{K}, \underline{\eta}) \in B\} \, d\mu^\otimes(\underline{\eta}) \right] d(\mathrm{cl}_\epsilon)^\otimes(\underline{K}),$$

where we abbreviate $\mu^\otimes$ for the law of the array $\underline{\eta}$ of i.i.d. samples from μ, and similarly we abbreviate $(\mathrm{cl}_\epsilon)^\otimes$ for the law of the sequence $\underline{K}$ of i.i.d. samples from cl_ϵ.

Definition 4.11 (coupled sequences of messages). Recalling the statement of Proposition 1.1, let $\mu^0 = \delta_{1/2}$. For $\ell \geq 1$ let $\mu^{\ell,\epsilon} \equiv (R_\epsilon)^\ell \mu^0$ and $\mu^{\ell,\epsilon} \equiv \mathcal{R}_\epsilon \mu^{\ell-1,\epsilon}$. Moreover, for all $\ell \geq 0$ let $\hat{\mu}^{\ell+1/2,\epsilon} \equiv \hat{\mathcal{R}}_\epsilon \mu^{\ell,\epsilon}$. The natural mapping from $\mathscr{P}$ to $\mathscr{P}$ (as discussed above) takes $\mu^{\ell,\epsilon}$ to $\mu^{\ell,\epsilon}$. Recalling the discussion below (2.24), if we take $\mathscr{T}_{v_{\text{RT}} a_{\text{RT}}} \sim \text{PGW}_\epsilon$ and define the random sequence

$$(\eta^\ell)_{\ell \geq 0} \equiv \left(\mathbf{F}_\ell(\mathscr{T}_{v_{\text{RT}} a_{\text{RT}}}) \right)_{\ell \geq 0}, \tag{4.5}$$

then the marginal law of each η^ℓ is precisely $\mu^{\ell,\epsilon}$. The marginal law of $\eta^\ell \equiv \eta^\ell(\text{-}) \in [0, 1)$ is $\mu^{\ell,\epsilon}$. Likewise, if we take $\hat{\mathscr{T}}_{a_{\text{RT}} v_{\text{RT}}} \sim \widehat{\text{PGW}}_\epsilon$ and define

$$(\hat{u}^{\ell+1/2})_{\ell \geq 0} \equiv \left(\mathbf{F}_{\ell+1/2}(\hat{\mathscr{T}}_{a_{\text{RT}} v_{\text{RT}}}) \right)_{\ell \geq 0},$$

then the marginal law of each $\hat{u}^{\ell+1/2}$ is precisely $\hat{u}^{\ell+1/2,\epsilon}$. We drop ϵ from the notation when $\epsilon = 0$.

4.2. **Concentration bounds for the distributional recursion.** In this subsection we prove two lemmas on the distributional recursion described above: Lemma 4.13 studies the distributional effect of a single clause update, and Lemma 4.12 gives concentration bounds on the variable-to-clause messages. The lemmas will be applied below in §4.3 to show that under the **PGW** measure, the root is 1-nice with very good probability.

Lemma 4.12. *Let μ be any probability measure over $\eta \in [0, 1)$ that satisfies the bounds*

(I) $\mu(\eta \geq 1/2 + s) \leq \mu(\ln[\eta/(1 - \eta)] \geq 4s) \leq \exp(-9s2^{k/4})$ *for all $s \geq 2^{-k/4}$;*

(II) $\mu(\eta \leq 1/2 - s) \leq \exp(-9s2^{k/4})$ *for all $2^{-k/4} \leq s \leq 1/2$.*

Then the measure $\mu^{\ell,\epsilon} \equiv (R_\epsilon)^\ell \mu$ satisfies the same bounds (I) and (II) for all $\ell \geq 0$. In addition, for all $\ell \geq 1$, the measure $\mu^{\ell,\epsilon} = \mathcal{R}_\epsilon \mu^{\ell-1,\epsilon}$ satisfies the bound

$$\mu^{\ell,\epsilon}\left(\left| 2^{k+1} \eta(\mathfrak{f}) - 1 \right| \geq \frac{1}{2^{k/4}} \right) \leq \frac{1}{\exp(2^{k/4})}. \tag{4.6}$$

The estimates of this lemma hold for all $0 \leq \epsilon \leq 1$.

Remark. In Lemma 4.12, we remark that the right tail bound (I) tends to zero in the limit $s \to \infty$, which is consistent with $\{\eta = 1\}$ having zero measure under μ. On the other hand, the left tail bound (II) does not go below $\exp\{-2^{k/4}\}$, which means that $\mu(\eta = 0)$ can be positive. Although the precise bound (II) is likely suboptimal, we point out that, under the conditions of Lemma 4.12, $\mu^\ell(\eta = 0)$ must indeed be positive for all $\ell \geq 1$. This is simply because it holds with positive probability that d^- is zero, which implies that the empty product Π^- is one, which in turn implies that $R(\underline{d}, \underline{\eta}) = 0$. The probability for d^- to be zero is

$$\mathbb{P}\left(\text{Pois}\left(\frac{\alpha k}{2} \right) = 0 \right) = \frac{1}{\exp(\alpha k/2)} \geq \frac{1}{\exp(k2^k)},$$

so $\mu^\ell(\eta = 0) \geq \exp(-k2^k)$ for all $\ell \geq 1$.

Lemma 4.13 (bounds on clause recursion). *Let μ be any probability measure over $\eta \in [0, 1)$, and let $\hat{\mu} = \hat{\mathcal{R}}_\epsilon \mu$ for $\hat{\mathcal{R}}_\epsilon$ as given by Definition 4.8. Sample $\hat{u} \sim \hat{\mu}$, and note that $\hat{u}(+) \in [0, 1)$ almost surely. Let*

$$X \equiv \ln \frac{1}{1 - \hat{u}(+)} \geq \hat{u}(+) \geq 0.$$

Write $\mathbb{P}$ and $\mathbb{E}$ for probability and expectation over the law of X. If μ satisfies the conditions (I) and (II) of Lemma 4.12, then X satisfies the following estimates:

(a) $\mathbb{P}(X \leq (2^{-\xi}/2)^{k-1}) \leq \exp\{-\Omega(\xi 2^{k/4})\}$ *for all $k/2^{k/4} \leq \xi \leq 1$.*

(b) *For all $x \geq 1$, $\mathbb{P}(X \geq x) \leq \exp\{-\Omega(kx2^{k/4})\}$;*

(c) $\mathbb{P}(X \geq (2^{1/10}/2)^{k-1}) \leq \exp\{-\Omega(k2^{k/4})\})$;

(d) $\mathbb{E}X = [1 + O(k/2^{k/8})]/2^{k-1}$;

(e) $\mathbb{E}[X \mathbf{1}\{X \geq (2^{1/10}/2)^{k-1}\}] \leq \exp\{-\Omega(k2^{k/4})\}$.

The estimates of this lemma hold uniformly over all $0 \leq \epsilon \leq 1$.

Proof. We can sample X as follows: let $(\eta_j)_{j\geq 1}$ be a sequence of i.i.d. samples from μ, let $K \sim \mathrm{cl}_\epsilon$, and let

$$X \equiv -\ln\left(1 - \prod_{j=1}^{K-1} \eta_j\right) \geq \prod_{j=1}^{K-1} \eta_j \equiv \hat{u} \geq \prod_{j=1}^{k-1} \eta_j \geq 0.$$

If $0 \leq \xi \leq 1$ then $2^{-\xi}/2 \leq 1/2 - \xi/4$. Consequently, for all $4/2^{k/4} \leq \xi \leq 1$, we find

$$\mathbb{P}\left(\hat{u} \leq \left(\frac{2^{-\xi}}{2}\right)^{k-1}\right) \leq \mathbb{P}\left(\min_{j \in [k-1]} \eta_j \leq \frac{2^{-\xi}}{2}\right) \leq k\mu\left(\eta \leq \frac{2^{-\xi}}{2}\right) \leq k\mu\left(\eta \leq \frac{1}{2} - \frac{\xi}{4}\right) \leq \frac{k}{\exp\{\Omega(\xi 2^{k/4})\}},$$

where the last step is by the assumed lower bound (II). Since $X \geq \hat{u}$, we obtain, for all $4/2^{k/4} \leq \xi \leq 1$, that

$$\mathbb{P}\left(X \leq \left(\frac{2^{-\xi}}{2}\right)^{k-1}\right) \leq \mathbb{P}\left(\hat{u} \leq \left(\frac{2^{-\xi}}{2}\right)^{k-1}\right) \leq \frac{k}{\exp\{\Omega(\xi 2^{k/4})\}}. \tag{4.7}$$

This implies (a). We next consider the right tail of X. We have $x + \ln(1 - e^{-x}) \geq \Omega(x)$ uniformly over $x \geq 1$, so

$$\mathbb{P}(X \geq x) = \mathbb{P}\left(\hat{u} = \prod_{j=1}^{K-1} \eta_j \geq 1 - e^{-x}\right) \leq \mathbb{P}\left(\max_{1 \leq j \leq k-2} \eta_j \geq 1 - e^{-x}\right)$$

$$= \mu\left(\ln\frac{\eta}{1-\eta} \geq x + \ln(1 - e^{-x})\right)^{k-2} \leq \frac{1}{\exp\{\Omega(kx2^{k/4})\}}$$

for all $x \geq 1$, where the last step follows by using the assumed upper bound (I). This proves (b); and we obtain

$$\mathbb{E}\left(X\mathbf{1}\{X \geq 1\}\right) \leq \frac{1}{\exp\{\Omega(k2^{k/4})\}} \tag{4.8}$$

by integrating (b) over $x \geq 1$. Next, for $0 < \xi < 1$, we have

$$\mathbb{P}\left(\hat{u} \geq \left(\frac{2^\xi}{2}\right)^{k-1}\right) \leq \mathbb{P}\left(\eta_j \geq \frac{2^{\xi/2}}{2} \text{ for at least } \frac{(k-2)\xi}{2} \text{ indices } j \in [k-2]\right) \leq k^{k\xi}\mu\left(\eta \geq \frac{2^{\xi/2}}{2}\right)^{(k-2)\xi/2},$$

where the factor $k^{k\xi}$ upper bounds the choice of indices j. We then note that

$$s(\xi) \equiv \ln\frac{2^\xi/2}{1 - 2^\xi/2} \geq (2\ln 2)\xi$$

by calculus. Substituting into the previous bound gives, for all $2^{-k/4} \leq \xi < 1$,

$$\mathbb{P}\left(\hat{u} \geq \left(\frac{2^\xi}{2}\right)^{k-1}\right) \leq k^{k\xi}\mu\left(\ln\frac{\eta}{1-\eta} \geq (2\ln 2)\xi\right)^{(k-2)\xi/2} \leq \frac{k^{k\xi}}{\exp\{\Omega(k\xi^2 2^{k/4})\}}, \tag{4.9}$$

having again used the assumed upper bound (I). We next note that if $X \geq (2^{2\xi}/2)^{k-1}$, then

$$\hat{u} = 1 - e^{-X} \geq 1 - \exp\left\{-\left(\frac{2^{2\xi}}{2}\right)^{k-1}\right\} = \left(\frac{2^{2\xi}}{2}\right)^{k-1} \exp\left\{-O\left(\frac{2^{2\xi}}{2}\right)^{k-1}\right\} \geq \left(\frac{2^\xi}{2}\right)^{k-1},$$

where the last inequality holds if we restrict to $\xi \leq 1/4$. Combining with (4.9) gives

$$\mathbb{P}\left(X \geq \left(\frac{2^{2\xi}}{2}\right)^{k-1}\right) \leq \frac{k^{k\xi}}{\exp\{\Omega(k\xi^2 2^{k/4})\}} \leq \frac{1}{\exp\{\Omega(k\xi^2 2^{k/4})\}} \tag{4.10}$$

for all $k/2^{k/4} \leq \xi \leq 1/4$, of which a special case is the claimed bound (c). Combining (4.7), (4.8), and (4.10) gives (d). Finally, we can combine (c) with (4.8) to bound

$$0 \leq \mathbb{E}\left[X\mathbf{1}\left\{X \geq \left(\frac{2^{1/10}}{2}\right)^{k-1}\right\}\right] \leq \mathbb{P}\left(X \geq \left(\frac{2^{1/10}}{2}\right)^{k-1}\right) + \mathbb{E}\left(X\mathbf{1}\{X \geq 1\}\right) \leq \frac{1}{\exp\{\Omega(k2^{k/4})\}},$$

which gives (e). □

Proof of Lemma 4.12. First we recall that if D is a Poisson random variable with mean λ, then a standard Chernoff bound gives, with $f(u) \equiv (1+u)\ln(1+u) - u$,

$$\mathbb{P}(|D - \lambda| \geq y) \leq 2\exp\left\{ -\lambda\min\left\{ f\left(\frac{y}{\lambda}\right), f\left(-\frac{y}{\lambda}\right)\right\}\right\}. \tag{4.11}$$

Suppose inductively that the bounds (I) and (II) hold for μ^ℓ. Sample $\underline{d} = (d^+, d^-)$ from $\mathrm{po}_\pm$ and $\underline{\eta}$ from $(\mu^\ell)^{\otimes\infty}$, so that $\eta^{\ell+1} = R(\underline{d}, \underline{\eta})$ has law $\mu^{\ell+1}$. We then expand

$$\ln\frac{\eta^{\ell+1}}{1-\eta^{\ell+1}} = \ln\frac{\Pi^+(1-\Pi^-)}{\Pi^-} = \ln(1-\Pi^-) + \ln\frac{1}{\Pi^-} - \ln\frac{1}{\Pi^+}. \tag{4.12}$$

Conditional on $\underline{d}$ we can decompose $\Sigma^\pm \equiv \ln(1/\Pi^\pm)$ as a sum of i.i.d. terms:

$$\Sigma^\pm \equiv \ln\frac{1}{\Pi^\pm} = \sum_{i=1}^{d^\pm}\left\{ -\ln\left(1 - \prod_{j=1}^{k-1}\eta_{ij}^\pm\right)\right\} \equiv \sum_{i=1}^{d^\pm} X_i^\pm \geq 0, \tag{4.13}$$

where the $X_i^\pm$ are equidistributed as the random variable X of Lemma 4.13. The remainder of the proof is divided into a few steps.

Step 1. Concentration bounds for $\Sigma^\pm$. Consider Σ^+, and define the truncated random variables

$$Y_i \equiv X_i^+ \mathbf{1}\left\{ X_i^+ \leq \left(\frac{2^{1/10}}{2}\right)^{k-1}\right\}, \quad Y \equiv X\mathbf{1}\left\{ X \leq \left(\frac{2^{1/10}}{2}\right)^{k-1}\right\}.$$

Next define the events

$$\mathrm{BIG} \equiv \left\{ \max_{i\leq\alpha k} X_i^+ \geq \left(\frac{2^{1/10}}{2}\right)^{k-1}\right\}, \quad \mathrm{DEG} \equiv \left\{ \left|d^+ - \frac{\alpha k}{2}\right| \geq k^2 2^{5k/8}\right\}.$$

On the complement of the event $\mathrm{DEG}\cup\mathrm{BIG}$, it follows using Lemma 4.13(e) that

$$\left|\sum_{i=1}^{d^+} X_i^+ - \frac{\alpha k}{2}\mathbb{E}X\right| = \left|\sum_{i=1}^{d^+} Y_i - \frac{\alpha k}{2}\mathbb{E}X\right| \leq \left|\sum_{i=1}^{\alpha k/2} Y_i - \frac{\alpha k}{2}\mathbb{E}Y\right| + \frac{k^2 2^{5k/8}}{2^{9k/10}} + \alpha k\mathbb{E}(X - Y)$$

$$\leq \left|\sum_{i=1}^{\alpha k/2} Y_i - \frac{\alpha k}{2}\mathbb{E}Y\right| + \frac{k^2 2^{5k/8}}{2^{9k/10}} + \frac{\alpha k}{\exp\{\Omega(k2^{k/4})\}} \leq \left|\sum_{i=1}^{\alpha k/2} Y_i - \frac{\alpha k}{2}\mathbb{E}Y\right| + \frac{1}{k^5 2^{k/4}}.$$

Consequently, for $t \geq 1/(k2^{k/4})$, we obtain by a union bound that

$$\mathbb{P}\left(\left|\sum_{i=1}^{d^+} X_i^+ - \frac{\alpha k}{2}\mathbb{E}X\right| \geq t\right) \leq \mathbb{P}(\mathrm{BIG}\cup\mathrm{DEG}) + \mathbb{P}\left(\left|\sum_{i=1}^{\alpha k/2} Y_i - \frac{\alpha k}{2}\mathbb{E}Y\right| \geq \frac{t}{2}\right).$$

By Lemma 4.13(c) and a trivial union bound over $1 \leq i \leq \alpha k$, we have

$$\mathbb{P}(\mathrm{BIG}) \leq \frac{\alpha k}{\exp\{\Omega(k2^{k/4})\}} \leq \frac{1}{\exp\{\Omega(k2^{k/4})\}}.$$

In the Poisson Chernoff bound (4.11), for small u we have $f(u) \asymp u^2$, so $\mathbb{P}(\mathrm{DEG}) \leq \exp\{-\Omega(k^3 2^{k/4})\}$. Finally, by the Azuma–Hoeffding inequality, it holds for all $t \geq 0$ that

$$\mathbb{P}\left(\left|\sum_{i=1}^{\alpha k/2}(Y_i - \mathbb{E}Y_i)\right| \geq t\right) \leq \frac{O(1)}{\exp\{\Omega(t^2 2^{4k/5}/k)\}}. \tag{4.14}$$

Recall from (4.13) that the sum of X_i^+ over $1 \leq i \leq d^+$ is exactly Σ^+. Combining the above bounds gives

$$\mathbb{P}\left(\left|\Sigma^+ - \frac{\alpha k}{2}\mathbb{E}X\right| \geq t\right) \leq \frac{1}{\exp\{\Omega(k2^{k/4})\}} + \frac{1}{\exp\{\Omega(t^2 2^{4k/5}/k)\}}. \tag{4.15}$$

The same bound holds for Σ^-, which is equidistributed as Σ^+.

Step 2. Proof of right tail bound in moderate deviations regime. We now rewrite the decomposition (4.12) as

$$\ln \frac{\eta^{\ell+1}}{1-\eta^{\ell+1}} = \ln\left(1 - \frac{1}{\exp\Sigma^-}\right) + \Sigma^- - \Sigma^+ \leq |\Sigma^- - \Sigma^+| \leq \left|\Sigma^+ - \frac{\alpha k}{2}\mathbb{E}X\right| + \left|\Sigma^- - \frac{\alpha k}{2}\mathbb{E}X\right|. \tag{4.16}$$

We can then apply the preceding bound (4.15) to obtain that for all $k/2^{11k/40} \leq s \leq 1$,

$$\mu^{\ell+1}\left(\ln\frac{\eta^{\ell+1}}{1-\eta^{\ell+1}} \geq s\right) \leq \mathbb{P}\left(|\Sigma^- - \Sigma^+| \geq s\right) \leq 2\mathbb{P}\left(\left|\Sigma^- - \frac{\alpha k}{2}\mathbb{E}X\right| \geq \frac{s}{2}\right) \leq \frac{1}{\exp\{\Omega(sk2^{k/4})\}}.$$

This implies the desired upper bound (I) for $2^{-k/4} \leq s \leq 1$.

Step 3. Proof of left tail bound. We will show $\mu(\eta \leq 1/2 - s) \leq \exp(-\Omega(sk2^{k/4}))$ for all $2^{-k/4} \leq s \leq 1/2$; the bound (II) then follows. For $0 \leq s \leq 1/2$, note that $\eta \leq 1/2 - s$ if and only if

$$\ln\frac{\eta}{1-\eta} \leq \ln\frac{1-2s}{1+2s} \leq -4s.$$

Recall the decomposition (4.16) for $\ln[\eta^{\ell+1}/(1-\eta^{\ell+1})]$. It follows that

$$\mu^{\ell+1}\left(\eta^{\ell+1} \leq \frac{1}{2} - s\right) \leq \mathbb{P}\left(\ln\left(1 - \frac{1}{\exp\Sigma^-}\right) \leq -2s\right) + \mathbb{P}\left(\Sigma^- - \Sigma^+ \leq -2s\right). \tag{4.17}$$

On the right-hand side of (4.17), we bound the first term by noting that

$$\mathbb{P}\left(\ln\left(1 - \frac{1}{\exp\Sigma^-}\right) \leq -2s\right) = \mathbb{P}\left(\Sigma^- \leq \ln\frac{1}{1-\exp(-2s)} \leq \ln\frac{1}{s}\right),$$

again for all $0 \leq s \leq 1/2$. Recall Lemma 4.13(d), and let $\bar{s}$ be the solution to the equation

$$\ln\frac{1}{\bar{s}} = \frac{\alpha k}{2}\mathbb{E}X = \frac{2^k k \ln 2}{2}\frac{1+O(k/2^{k/8})}{2^{k-1}} \geq \frac{k\ln 2}{2},$$

so $\bar{s} \leq 2^{-k/2}$. Then for $2^{-k/4} \leq s \leq 1/2$ we have $s' \equiv s - \bar{s} \geq s/2$, It follows that the first term on the right-hand side of (4.17) is upper bounded by

$$\mathbb{P}\left(\Sigma^- \leq \ln\frac{1}{s}\right) = \mathbb{P}\left(\Sigma^- - \frac{\alpha k}{2}\mathbb{E}X \leq \ln\frac{1}{\bar{s}+s/2} - \ln\frac{1}{\bar{s}} \leq -s\right) \leq \frac{1}{\exp\{\Omega(sk2^{k/4})\}}, \tag{4.18}$$

where the last step uses the earlier estimate on Σ^-. The same estimate implies that the second term on the right-hand side of (4.17) is also at most $\exp\{-\Omega(sk2^{k/4})\}$; and this concludes the verification of the lower bound (II).

Step 4. Proof of right tail bound in large deviations regime. We now verify (I) for $s \geq 1$. Define the event

$$\text{DEG}^s \equiv \left\{\left|d^+ - \frac{\alpha k}{2}\right| \geq s2^{3k/4}\right\}.$$

Now note that in the Poisson Chernoff bound (4.11), the function $f(u) = (1+u)\ln(1+u) - u \geq 0$ is strictly convex with respect to $u \in (-1, \infty)$. Consequently, if $u \geq 2u_0 > 0$, then $f(u) \geq f(u_0) + f'(u_0)(u - u_0) \geq f'(u_0)u/2$. Likewise, if $u \leq 2u_0 < 0$, then $f(u) \geq f'(u_0)u/2$. We note also that $f'(u) = \ln(1+u)$, so if u_0 is small then $f'(u_0) \asymp u_0$. We can therefore conclude that for all $s \geq 1$, we have

$$\mathbb{P}(\text{DEG}^s) \leq \frac{1}{\exp\{\Omega(s2^{k/2}/k)\}}.$$

Recall from above that Y_i denotes the truncated version of X_i^+. On the event DEG^s we have

$$\left|\sum_{i=1}^{d^+} Y_i - \frac{\alpha k}{2}\mathbb{E}X\right| \leq \left|\sum_{i=1}^{\alpha k/2}(Y_i - \mathbb{E}Y_i)\right| + \left|\sum_{i=1}^{d^+} Y_i - \sum_{i=1}^{\alpha k/2} Y_i\right| + \frac{\alpha k}{2}\mathbb{E}(X - Y)$$

$$\leq \left|\sum_{i=1}^{\alpha k/2}(Y_i - \mathbb{E}Y_i)\right| + s2^{k/4}\left(\frac{2^{1/10}}{2}\right)^{k-1} + \frac{1}{\exp\{\Omega(k2^{k/4})\}} \leq \left|\sum_{i=1}^{\alpha k/2}(Y_i - \mathbb{E}Y_i)\right| + \frac{s}{2^{k/2}},$$

where we used Lemma 4.13(e) to bound $\mathbb{E}(X - Y)$, and the last bound holds for all $s \geq 1$. Then, by combining the bound on $\mathbb{P}(\text{DEG}^s)$ with the Azuma–Hoeffding bound (4.14), we find

$$\mathbb{P}\left(\left|\sum_{i=1}^{d^+} Y_i - \frac{\alpha k}{2}\mathbb{E}X\right| \geq \frac{s}{k}\right) \leq \frac{1}{\exp\{\Omega(s2^{k/2}/k)\}} + \frac{1}{\exp\{\Omega(s^2 2^{4k/5}/k^3)\}} \leq \frac{1}{\exp\{\Omega(s2^{k/2}/k)\}}. \tag{4.19}$$

Next we account for the difference between Y_i and X_i^+. Recall Lemma 4.13 parts (b) and (c). Together they imply, with $\preccurlyeq$ denoting stochastic domination, that

$$\sum_{i=1}^{d^+}(X_i^+ - Y_i) \preccurlyeq \sum_{i=1}^{d^+} I_i(1 + Z_i) \tag{4.20}$$

where I_i are i.i.d. $\text{Ber}(e^{-\theta})$ indicators, and the Z_i are i.i.d. $\theta^{-1}\text{Exp}$ random variables (where Exp denotes a standard exponential random variable), with $\theta = ck2^{k/4}$ for some absolute constant $c > 0$. By Poisson thinning,

$$A \equiv \sum_{i=1}^{d^+} I_i \sim \text{Pois}\left(\frac{\alpha k}{2e^\theta} \equiv \lambda'\right).$$

Note that λ' is very small, and it follows from (4.11) that for all $s \geq 1$,

$$\mathbb{P}\left(A \geq \frac{s}{k^{1/2}}\right) \leq \exp\left\{-\Omega\left(\frac{s}{k^{1/2}}\ln\frac{s/k^{1/2}}{\lambda'}\right)\right\} \leq \frac{1}{\exp\{\Omega(s\theta/k^{1/2})\}} \leq \frac{1}{\Omega(sk^{1/2}2^{k/4})}.$$

Conditioned on A, the other term of (4.20) is distributed as a gamma random variable with shape parameter A,

$$B \equiv \sum_{i=1}^{d^+} I_i Z_i \sim \frac{\text{Gamma}(A)}{\theta}.$$

Let E be a standard exponential random variable, independent of A. The moment-generating function of B is

$$m(t) = \mathbb{E}(e^{tB}) = \mathbb{E}\left[\left\{\mathbb{E}\exp\left(\frac{tE}{\theta}\right)\right\}^A\right] = \mathbb{E}\left[\left\{\frac{1}{1 - t/\theta}\right\}^A\right] = \exp\left\{\frac{\lambda' t}{\theta - t}\right\},$$

for $t < \theta$. For $x > \mathbb{E}B = \lambda'/\theta$, optimizing over t gives

$$\mathbb{P}(B \geq x) \leq \exp\left\{-\theta x\left[1 - \left(\frac{\lambda'}{\theta x}\right)^{1/2}\right]^2\right\},$$

which implies $\mathbb{P}(B \geq s/k^{1/2}) \leq \exp\{-\Omega(sk^{1/2}2^{k/4})\}$ for all $s \geq 1$. Combining these estimates on A and B with our earlier bound (4.19) gives altogether

$$\mathbb{P}\left(\left|\Sigma^+ - \frac{\alpha k}{2}\mathbb{E}X\right| \geq \frac{s}{2}\right) \leq \frac{1}{\exp\{\Omega(sk^{1/2}2^{k/4})\}}.$$

The same bound holds for Σ^-. Substituting into (4.16), we obtain the desired bound (I) for all $s \geq 1$.

Step 5. Conclusion. Having verified (I) and (II), it remains to show (4.6). Recall that we can express

$$\eta_{\ell+1}(\mathfrak{f}) = \frac{\Pi^+\Pi^-}{\Pi^+ + \Pi^- - \Pi^+\Pi^-} = \frac{1}{\exp(\Sigma^+) + \exp(\Sigma^-) - 1}.$$

It follows from (4.15) that Σ^+ is concentrated near $(\alpha k/2)\mathbb{E}X$, which by Lemma 4.13(d) is close to $k\ln 2$:

$$\mathbb{P}\left(\left|\Sigma^+ - k\ln 2\right| \geq \frac{k}{2^{11k/40}}\right) \leq \frac{1}{\exp\{\Omega(k2^{k/4})\}}, \tag{4.21}$$

and likewise for Σ^-. It follows that

$$\mu^{\ell+1}\left(\left|2^{k+1}\eta_{\ell+1}(\mathfrak{f}) - 1\right| \geq \frac{1}{2^{k/4}}\right) \leq 2\mathbb{P}\left(\left|\Sigma^+ - k\ln 2\right| \geq \frac{k}{2^{11k/40}}\right) \leq \frac{1}{\exp(2^{k/4})},$$

which concludes the proof of (4.6). $\qquad\square$

4.3. **Niceness in the Galton–Watson tree.** The main goal of this subsection is to prove that under the **PGW** measure, the root variable fails to be 1-nice with very small probability. For technical reasons (which will emerge in the proof of Lemma 5.2), we will prove a version of this statement which is slightly stronger in two ways. First, we generalize from **PGW** to **PGW**(T) (Definition 4.5) where T is very sparse, with maximum degree $O(1)$ (not growing with k) − since typical degrees in **PGW** diverge with k, it is intuitively plausible that planting the sparse subtree T cannot have a large effect. Second, we replace 1-nice with a more restrictive property which we now define:

Definition 4.14 (robustness). If U is any rooted tree and x is any vertex in U, we let $U(x)$ denote the subtree of U that lies below x. Given any tree $\mathscr{T}$ rooted at a variable v_{RT}, we say that $\mathcal{Q}$ is a $\mathfrak{c}$-**modification** of $\mathscr{T}$ if $\mathcal{Q}$ can be obtained from $\mathscr{T}$ by

– deleting at most one subtree $\mathscr{T}(u)$ for $u \in N(v_{\mathrm{RT}})$, and
– changing at most $\mathfrak{c}$ subtrees $\mathscr{T}(u)$ for $u \in \partial_2 v_{\mathrm{RT}}$.

The new subtrees $\mathcal{Q}(u)$, for $u \in \partial_2 v_{\mathrm{RT}}$, can be arbitrary. We then say that an acyclic variable v is $\mathfrak{c}$-**robust** if every $\mathfrak{c}$-modification of $B_r(v)$ is nice. Note that being $\mathfrak{c}$-robust is stronger than being 1-nice.

Proposition 4.15. *Let $\mathfrak{c}$ be an absolute constant. Let $\mathbb{P} = \mathbf{PGW}(T)$ where T is a fixed bipartite factor tree, rooted at a variable, with maximum vertex degree at most $\mathfrak{c}$. For $\mathscr{T} \sim \mathbb{P}$,*

$$\mathbb{P}\Big(\mathscr{T} \text{ is not } \mathfrak{c}\text{-robust}\Big) \leq \frac{1}{\exp(\Omega(k2^{k/4}))} \, .$$

(We allow $T = \varnothing$, in which case the statement is for $\mathbb{P} = \mathbf{PGW}(\varnothing) = \mathbf{PGW}$.)

In fact, this result is a relatively straightforward consequence of the technical lemmas of the previous §4.2. We first consider the effect of changing a small number of subtrees. If $\mathscr{T}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$ and $\mathcal{Q}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$ are both variable-to-clause trees, we will say that they are $\mathfrak{c}$-**perturbations** of one another if $\mathcal{Q}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$ can be obtained from $\mathscr{T}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$ by only changing subtrees $\mathscr{T}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}(u)$ for at most $\mathfrak{c}$ variables $u \in N(v_{\mathrm{RT}})$. The new subtrees $\mathcal{Q}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}(u)$ are allowed to be arbitrary.[17] The next lemma says, essentially, that $\mathfrak{c}$-perturbations have very little effect on the outgoing variable-to-clause message. The formal statement is as follows:

Lemma 4.16. *Let $\mathbb{P} = \mathrm{PGW}(T)$, where $T \equiv T_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$ is any fixed variable-to-clause tree of maximum vertex degree at most $\mathfrak{c}$. We allow $T = \varnothing$, in which case $\mathbb{P} = \mathrm{PGW}(\varnothing) = \mathrm{PGW}$. If $\mathscr{T} \equiv \mathscr{T}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$ is sampled from $\mathbb{P}$, then*

$$\mathbb{P}\left(\max\left\{\sum_{x \in \{+,-,\mathbf{f}\}}\left|2^{k\mathbf{1}\{x=\mathbf{f}\}}\eta'(x) - \frac{1}{2}\right| \, \Bigg| \, \begin{matrix} \eta' = \mathbf{F}_\ell(\mathcal{Q}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}), \text{ for } \mathcal{Q}_{v_{\mathrm{RT}}a_{\mathrm{RT}}} \\ \text{a } \mathfrak{c}\text{-perturbation of } \mathscr{T}_{v_{\mathrm{RT}}a_{\mathrm{RT}}} \end{matrix} \right\} \geq \frac{1}{2^{k/4}}\right) \leq \frac{1}{\exp(\Omega(k2^{k/4}))}$$

for all $\ell \geq 0$.

Proof. Throughout this proof we will abbreviate $T \equiv T_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$, $\mathscr{T} \equiv \mathscr{T}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$, and $\mathcal{Q} \equiv \mathcal{Q}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$. From the definitions (see, in particular, the discussion below (2.24)), the measure $\eta \equiv \mathbf{F}_\ell(\mathscr{T})$ only depends on the tree up to depth ℓ below v_{RT}. For $0 \leq \ell \leq 1$, any $\mathfrak{c}$-perturbation has $\eta' = \eta$, so there is nothing more to prove. We therefore assume $\ell \geq 2$ for the rest of the proof.

We next describe a procedure to generate a sample of η. In the fixed tree T, we partition the first layer of clauses (at depth $1/2$ below v_{RT}) according to the signs on the edges from these clauses to v_{RT}:

$$(\partial^+ T, \partial^- T) \equiv \Big(T \cap \partial v_{\mathrm{RT}}(+a_{\mathrm{RT}}), T \cap \partial v_{\mathrm{RT}}(-a_{\mathrm{RT}})\Big).$$

Let $d^{\pm}(T) \equiv |\partial^{\pm} T|$; these are both upper bounded by $\mathfrak{c}$. (If $T = \varnothing$ we define $d^{\pm}(T) \equiv 0$.) For $1 \leq i \leq d^+(T)$, the i-th clause in $\partial^+ T$ has $\mathfrak{c}^+(i)$ child variables in T. Let $T(+, i, j)$ be the subtree of T descended from the j-th child variable of the i-th clause in $\partial^+(T)$. Define similarly $\mathfrak{c}^-(i)$ and $T(-, i, j)$. Now define η by the following steps:

(i) Sample an array of independent random trees $\mathscr{T}_{ij}^{\pm} \sim \mathrm{PGW}(T(\pm, i, j))$, and define the corresponding messages

$$\eta_{ij}^{\pm}(T) \equiv \Big[\mathbf{F}_{\ell-1}(\mathscr{T}_{ij}^{\pm})\Big](-) \, ,$$

for $1 \leq i \leq d^{\pm}(T)$ and $1 \leq j \leq \mathfrak{c}^{\pm}(i)$.

---

[17]Note that the $\mathfrak{c}$-perturbation defined here is different from the $\mathfrak{c}$-modification of Definition 4.14.

(ii) Let $\underline{\eta}$ be an array (as in (1.8)) of i.i.d. samples from $\mu^{\ell-1} = R^{\ell-1}\mu^0$, where $\mu^0 = \delta_{1/2}$ as before. Let

$$u_i^+(T) \equiv \prod_{j=1}^{c^+(i)} \eta_{ij}^+(T) \prod_{j=c^+(i)+1}^{k-1} \eta_{ij}^+, \quad u_i^+ \equiv \prod_{j=1}^{k-1} \eta_{ij}^+.$$

Define similarly $u_i^-(T)$ and u_i^-.

(iii) Let $\underline{d} \equiv (d^+, d^-)$ be an independent sample from $\mathrm{po}_\pm$, and define

$$\Pi^+(T) \equiv \prod_{i=1}^{d^+(T)} \left(1 - u_i^+(T)\right), \quad \Pi^+ \equiv \prod_{i=d^+(T)+1}^{d^+(T)+d^+} \left(1 - u_i^+\right), \quad \Pi^{*+} \equiv \Pi^+(T)\Pi^+. \tag{4.22}$$

Define similarly $\Pi^-(T)$, Π^-, and $\Pi^{*-} = \Pi^-(T)\Pi^-$.

(iv) Substitute the $\Pi^{*\pm}$ into (4.4) to define $\underline{\eta}$:

$$\left(\eta(+), \eta(-), \eta(\mathtt{f})\right) = \left(\frac{\Pi^{*-}(1 - \Pi^{*+})}{\Pi^{*+} + \Pi^{*-} - \Pi^{*+}\Pi^{*-}}, \frac{\Pi^{*+}(1 - \Pi^{*-})}{\Pi^{*+} + \Pi^{*-} - \Pi^{*+}\Pi^{*-}}, \frac{\Pi^{*+}\Pi^{*-}}{\Pi^{*+} + \Pi^{*-} - \Pi^{*+}\Pi^{*-}}\right).$$

The $\underline{\eta}$ that results from this construction can be regarded as a sample of $\mathbf{F}_\ell(\mathscr{T})$ for $\mathscr{T} \sim \mathrm{PGW}(T)$, although we did not explicitly generate all of $\mathscr{T}$. Moreover, if $\underline{\eta}' = \mathbf{F}_\ell(\mathcal{Q})$ where $\mathcal{Q}$ is any $\mathfrak{c}$-perturbation of $\mathscr{T}$, then $\underline{\eta}'$ can also be obtained from the above procedure by modifying at most $\mathfrak{c}$ of the messages from depth one.

For convenience we will let $\underline{\theta}$ denote the messages that are actually used in the definition of $\underline{\eta}$, so

$$\theta_{ij}^+ \equiv \begin{cases} \eta_{ij}^+(T) & \text{if } 1 \le i \le d^+(T) \text{ and } 1 \le j \le c^+(i); \\ \eta_{ij}^+ & \text{otherwise,} \end{cases}$$

and similarly θ_{ij}^-. To bound the different between $\underline{\eta}$ and $\underline{\eta}'$, we first assume a *fixed* set of affected indices: without loss of generality, let $\underline{\vartheta}$ denote a new set of messages such that

$$\left\{(i,j) : \vartheta_{ij}^\pm \ne \theta_{ij}^\pm\right\} \subseteq \left\{(i,j) : 1 \le i \le 2\mathfrak{c}, 1 \le j \le 2\mathfrak{c}\right\}.$$

Returning to (4.22), let $\Sigma^\pm(T) \equiv -\ln\Pi^\pm(T)$, $\Sigma^\pm \equiv -\ln\Pi^\pm$, and $\Sigma^{*\pm} \equiv -\ln\Pi^{*\pm}$. Let $\Xi^\pm(T)$, $\Xi^\pm$, and $\Xi^{*\pm}$ be defined analogously as the Σ quantities, but with $\underline{\vartheta}$ in place of $\underline{\theta}$. It is then straightforward to check that

$$|\Xi^{*+} - \Sigma^+| \le \sum_{i=1}^{d^+(T)} -\ln\left(1 - \prod_{j=1}^{k-1} \vartheta_{ij}^+\right) + \sum_{i=d^+(T)+1}^{2\mathfrak{c}} \left|\ln\left(1 - \prod_{j=1}^{k-1} \vartheta_{ij}^+\right) - \ln\left(1 - \prod_{j=1}^{k-1} \theta_{ij}^+\right)\right|$$

$$\le \sum_{i=1}^{2\mathfrak{c}} \underbrace{\left[-\ln\left(1 - \prod_{j=2\mathfrak{c}+1}^{k-1} \theta_{ij}^+\right)\right]}_{\Delta^+(i)} \equiv \sum_{i=1}^{2\mathfrak{c}} \Delta^+(i).$$

By the same argument as for the estimate (4.10) from the proof of Lemma 4.13, we have

$$\mathbb{P}\left(|\Xi^{*+} - \Sigma^+| \ge \left(\frac{2^{1/20}}{2}\right)^k\right) \le 2\mathfrak{c} \cdot \mathbb{P}\left(|\Delta^+(i)| \ge \left(\frac{2^{1/21}}{2}\right)^k\right) \le \frac{1}{\exp\{\Omega(k2^{k/4})\}},$$

and the same bound applies for $|\Xi^{*-} - \Sigma^-|$. Now recall the bound (4.21) from the proof of Lemma 4.12, which says that the $\Sigma^\pm$ are well concentrated around $k\ln 2$. In particular, it holds with probability at least $1 - \exp(-\Omega(k2^{k/4}))$ that $|\Sigma^\pm - k\ln 2| \le k/2^{11k/40}$ and $|\Xi^{*\pm} - \Sigma^\pm| \le (2^{1/20}/2)^k$, in which case

$$\eta'(\pm) = \frac{\exp(\Xi^{*\pm}) - 1}{\exp(\Xi^{*\pm}) + \exp(\Xi^{*\mp}) - 1} = \frac{1 + O(k/2^{11k/40})}{2},$$

$$\eta'(\mathtt{f}) = \frac{1}{\exp(\Xi^{*\pm}) + \exp(\Xi^{*\mp}) - 1} = \frac{1 + O(k/2^{11k/40})}{2^{k+1}}.$$

Thus, given a *fixed* choice of at most $\mathfrak{c}$ perturbed indices, it holds with probability at least $1 - \exp(-\Omega(k2^{k/4}))$ that any resulting message $\underline{\eta}'$ satisfies the above estimates, regardless of how those $\mathfrak{c}$ indices are perturbed. To conclude,

we note that by (4.11) the event $\text{DEG}^+ \equiv \{\{d^+, d^-\} \geq 4^k/k\}$ has probability upper bounded by $\exp(-\Omega(4^k))$. On the complement of DEG^+, the number of distinct choices for the $\mathfrak{c}$ perturbed indices is $\exp(O(k))$, so we can take a union bound over all choices to obtain the result. □

Proof of Proposition 4.15. Let $\mathscr{T}$ be a sample of $\mathbf{PGW}(T)$, rooted at variable $v \equiv v_{\text{RT}}$. Write $\partial v(+)$ and $\partial v(-)$ for the clauses neighboring the root v in $\mathscr{T}$, and let

$$\text{DEG}' \equiv \max\left\{\left||\partial v(+)| - \frac{2^k k \ln 2}{2}\right|, \left||\partial v(-)| - \frac{2^k k \ln 2}{2}\right|\right\} \geq k 2^{5k/8}$$

It follows from (4.11) that $\mathbb{P}(\text{DEG}') \leq \exp(-\Omega(k 2^{k/4}))$.

Now, recalling Definition 4.14, we want to show that (with very good probability) any $\mathfrak{c}$-modification $\mathcal{Q}$ of $\mathscr{T}$ is nice in the sense of Definition 3.8. On the complement of the event DEG', it is clear that every $\mathfrak{c}$-modification $\mathcal{Q}$ of $\mathscr{T}$ will satisfy the degree condition (3.13). It remains to determine whether the canonical messages $_\star\dot{q}(\mathcal{Q})$ and $_\star\hat{q}(\mathcal{Q})$ satisfy the bounds (3.14) and (3.15).

To this end, let us fix $\mathcal{Q}$ momentarily, and abbreviate $q \equiv {}_\star\dot{q}(\mathcal{Q})$ and $\hat{q} \equiv {}_\star\hat{q}(\mathcal{Q})$. Note that q and $\hat{q}$ are based the r-neighborhood $\mathcal{Q}_r$ of the root v in $\mathcal{Q}$; see Definition 3.4. For any edge (au) in $\mathcal{Q}$, let us abbreviate

$$\eta_{ua} \equiv \eta(\mathcal{Q})_{ua} \equiv \mathbf{F}_{\mathcal{Q}_r, ua} ,$$
$$\hat{u}_{au} \equiv \hat{u}(\mathcal{Q})_{au} \equiv \mathbf{F}_{\mathcal{Q}_r, au} .$$

Recall from (2.38) the correspondence between $(\dot{q}, \hat{q})$ and $(\eta, \hat{u})$ for edges (av) incident to the root v:

$$\left(q_{va}(\mathtt{r}), q_{va}(\mathtt{y}), q_{va}(\mathtt{g}), q_{va}(\mathtt{b})\right) = \left(\frac{\eta_{va}(+) + \eta_{va}(\mathtt{f})}{2 - \eta_{va}(-)}, \frac{\eta_{va}(-)}{2 - \eta_{va}(-)}, \frac{\eta_{va}(\mathtt{f})}{2 - \eta_{va}(-)}, \frac{\eta_{va}(+)}{2 - \eta_{va}(-)}\right), \tag{4.23}$$

$$\left(\hat{q}_{av}(\mathtt{r}), \hat{q}_{av}(\mathtt{y}), \hat{q}_{av}(\mathtt{g}), \hat{q}_{av}(\mathtt{b})\right) = \left(\frac{\hat{u}_{av}(+)}{3 - 2\hat{u}_{av}(+)}, \frac{\hat{u}_{av}(\mathtt{f})}{3 - 2\hat{u}_{av}(+)}, \frac{\hat{u}_{av}(\mathtt{f})}{3 - 2\hat{u}_{av}(+)}, \frac{\hat{u}_{av}(\mathtt{f})}{3 - 2\hat{u}_{av}(+)}\right). \tag{4.24}$$

Recall also from §2.5 that $\hat{u}_{av}$ can be recursively computed as (cf. (2.25))

$$\left(\hat{u}_{av}(+), \hat{u}_{av}(\mathtt{f})\right) = \left(\prod_{u \in (\mathcal{Q} \cap \partial a) \setminus v} \eta_{ua}(-), 1 - \prod_{u \in (\mathcal{Q} \cap \partial a) \setminus v} \eta_{ua}(-)\right). \tag{4.25}$$

In view of these relations, for $\mathcal{Q}$ to be nice, it suffices to have

$$\max\left\{\left|\eta(\mathcal{Q})_{ua}(-) - \frac{1}{2}\right| \,\middle|\, \begin{array}{c} a \in \mathcal{Q} \cap \partial v, \\ u \in \partial a \setminus v \end{array}\right\} \leq \frac{1}{2^{k/5}} , \tag{4.26}$$

$$\max\left\{\sum_{x \in \{+,-,\mathtt{f}\}} \left|2^{k\mathbf{1}\{x=\mathtt{f}\}}\eta(\mathcal{Q})_{va}(x) - \frac{1}{2}\right| \,\middle|\, a \in \mathcal{Q} \cap \partial v\right\} \leq \frac{1}{2^{k/5}} . \tag{4.27}$$

Indeed, substituting (4.26) into (4.25) and (4.24) shows that $\hat{q} = {}_\star\hat{q}(\mathcal{Q})$ satisfies condition (3.15); while substituting (4.27) into (4.23) shows that $q = {}_\star\dot{q}(\mathcal{Q})$ satisfies condition (3.14).

It remains to bound, on the event DEG', the probability for (4.26) and (4.27) to hold for every $\mathfrak{c}$-modification $\mathcal{Q}$ of the original random tree $\mathscr{T} \sim \mathbf{PGW}(T)$. For any such $\mathcal{Q}$, for all $a \in \mathcal{Q} \cap \partial v$ and all $u \in \partial a \setminus v$, the subtree $\mathcal{Q}_{ua}$ is a $\mathfrak{c}$-perturbation of $\mathscr{T}_{ua}$, so Lemma 4.16 applies. For all $a \in \mathcal{Q} \cap \partial v$, the subtree $\mathcal{Q}_{va}$ is a $\mathfrak{c}$-perturbation of $\mathscr{T}_{va}$, so Lemma 4.16 again applies. It follows by a simple union bound that for $\mathscr{T} \sim \mathbf{PGW}(T)$, with probability lower bounded by $1 - \exp(-\Omega(k 2^{k/4}))$, every $\mathfrak{c}$-modification $\mathcal{Q}$ of $\mathscr{T}$ satisfies the degree condition (3.13) and the message conditions (4.26) and (4.27), and hence is nice. This concludes the proof. □

4.4. Stability in the Galton–Watson tree.

The main goal of this section is to prove the following:

Proposition 4.17. *For R exceeding a large absolute constant, we have*

$$\mathbf{PGW}\left(v_{\text{RT}} \text{ not 1-stable}\right) \leq \frac{1}{\exp(2^{k/20}R)} ,$$

where the 1-stable property is given by Definition 3.9.

The next few results (Corollaries 4.18–4.20, and Lemma 4.21) give some preliminary estimates, related to the concentration bounds from §4.2. They will be used in the proof of Lemma 4.27 below.

Corollary 4.18. *Let μ be any probability measure over $\eta \in [0, 1)$ that satisfies condition* (I) *from Lemma* 4.12. *If* $p = 2^{k\delta}$ *for any* $1/100 \le \delta \le 1/9$, *then we have the p-th moment bound*

$$\int \eta^p \, d\mu(\eta) \le \frac{1}{2^p} \left[1 + \frac{k}{2^{k(1/4-2\delta)}} \right].$$

Proof. Write $\delta' \equiv 1/4 - \delta$. We can decompose

$$E(p) \equiv \int \eta^p \, d\mu(\eta) \le \frac{1}{2^p} \left(1 + \frac{\ln k}{2^{k\delta'}} \right)^p + \left(\frac{2}{3} \right)^p \mu \left(\eta \ge \frac{1}{2} \left(1 + \frac{\ln k}{2^{k\delta'}} \right) \right) + \mu \left(\eta \ge \frac{2}{3} \right).$$

Note that $\delta' > \delta$, so $(p \ln k)/2^{k\delta'}$ is small. Combining with (I) gives

$$E(p) \le \frac{1}{2^p} \left\{ 1 + \frac{O(p \ln k)}{2^{k\delta'}} \right\} + \frac{(2/3)^p}{\exp(\Omega(2^{k/4}2^{-k\delta'} \ln k))} + \frac{1}{\exp(\Omega(2^{k/4}))}.$$

Substituting $p = 2^{k\delta}$ and $\delta' = 1/4 - \delta$ into the above gives

$$E(p) \le \frac{1}{2^p} \left[1 + \frac{O(\ln k)}{2^{k(1/4-2\delta)}} + \frac{\exp\{2^{k\delta} \ln(4/3)\}}{\exp(\Omega(2^{k\delta} \ln k))} + \frac{\exp\{2^{k\delta} \ln 2\}}{\exp(\Omega(2^{k/4}))} \right] \le \frac{1}{2^p} \left[1 + \frac{O(\ln k)}{2^{k(1/4-2\delta)}} \right],$$

which concludes the proof. $\qquad \square$

Corollary 4.19. *Let μ be any probability measure over $\eta \in [0, 1)$ that satisfies conditions* (I) *and* (II) *from Lemma* 4.12. *If* $p = 2^{k\delta}$ *for* $1/100 \le \delta \le 1/9$, *then*

$$\int \left| \frac{\eta - 1/2}{\eta + 1/2} \right|^p d\mu(\eta) \le \frac{k^{2p}}{2^{kp(1/4-\delta)}}.$$

Proof. It follows from (I) and (II) that for any $\delta < \delta' < 1/4$ we have

$$\int \left| \frac{\eta - 1/2}{\eta + 1/2} \right|^p d\mu(\eta) \le \frac{O(1)}{2^{kp\delta'}} + \frac{1}{\exp(2^{k/4}2^{-k\delta'})}.$$

The claim follows by setting $\delta' = 1/4 - \delta - (\ln k)/(k \ln 2)$. $\qquad \square$

Corollary 4.20. *Let μ be any probability measure over $\eta \in [0, 1)$ that satisfies condition* (II) *from Lemma* 4.12. *Write* $\mu^{\otimes}$ *for the law of a sequence* $\underline{\eta}' \equiv (\eta_j)_{j \ge 1}$ *of i.i.d. samples from μ. If* $p = 2^{k\delta}$ *for any* $1/100 \le \delta \le 1/9$, *then*

$$\int \left(1 - \prod_{j=1}^{k-1} \eta_j \right)^p d\mu^{\otimes}(\underline{\eta}') \le \exp \left\{ -\frac{p}{2^{k-1}} \left[1 - \frac{k^3}{2^{k/4}} \right] \right\}.$$

Proof. Write $\mathbb{P}$ and $\mathbb{E}$ for probability and expectation over the law of

$$X \equiv -\ln \left(1 - \prod_{j=1}^{k-1} \eta_j \right).$$

The quantity of interest is then $\mathbb{E}(e^{-pX})$. Recall that Lemma 4.13(a) gives

$$\mathbb{P} \left(X \le \left(\frac{2^{-\xi}}{2} \right)^{k-1} \right) \le \frac{1}{\exp\{\Omega(\xi 2^{k/4})\}}$$

for all $k/2^{k/4} \le \xi \le 1$. It follows that

$$\mathbb{E}(e^{-pX}) \le \exp \left\{ -\frac{p}{2^{k-1}} \left[1 - \frac{k^2 \ln k}{2^{k/4}} \right] \right\} + \mathbb{P} \left(X \le \frac{1}{2^{k-1}} \left[1 - \frac{k^2 \ln k}{2^{k/4}} \right] \right)$$

$$\le \exp \left\{ -\frac{p}{2^{k-1}} \left[1 - \frac{k^2 \ln k}{2^{k/4}} \right] \right\} + \frac{1}{\exp\{\Omega(k \ln k)\}} \le \exp \left\{ -\frac{p}{2^{k-1}} \left[1 - \frac{O(k^2 \ln k)}{2^{k/4}} \right] \right\},$$

and this implies the claim. $\qquad \square$

Lemma 4.21. *Let D be a* $\mathrm{Pois}(\alpha k/2)$ *random variable, and let $\mathbb{E}$ denote expectation over the law of D. If $p = 2^{k\delta}$ for $1/100 \leq \delta \leq 1/9$, then*

$$\mathbb{E}\left\{D^p \exp\left\{-\frac{Dp}{2^{k-1}}\left[1 - \frac{k^3}{2^{k/4}}\right]\right\}\right\} \leq \left(\frac{3k}{8}\right)^p.$$

Proof. The quantity of interest can be written as $\mathbb{E}[D^p/e^{\gamma D}]$ for

$$\gamma \equiv \frac{p}{2^{k-1}}\left[1 - \frac{k^3}{2^{k/4}}\right].$$

We can check by differentiation that $d^p/e^{\gamma d}$ is decreasing with respect to d for all

$$d > \frac{p}{\gamma} = 2^{k-1}\Big/\left[1 - \frac{k^3}{2^{k/4}}\right].$$

For all α in the regime (1.7), the mean $\mathbb{E}D = \alpha k/2$ is much larger than p/γ. As in (1.7) let $\alpha_{\mathrm{ubd}} \equiv 2^k \ln 2$. Then

$$\mathbb{E}\left[\frac{D^p}{\exp(\gamma D)}; D \geq \frac{k\alpha_{\mathrm{ubd}}}{2}\right] \leq \left(\frac{k\alpha_{\mathrm{ubd}}}{2}\right)^p \exp\left\{-\frac{\gamma k\alpha_{\mathrm{ubd}}}{2}\right\}$$

$$= \left(\frac{k2^k \ln 2}{2}\exp\left\{-k\ln 2\left[1 - \frac{k^3}{2^{k/4}}\right]\right\}\right)^p \leq O(1)\left(\frac{k\ln 2}{2}\right)^p.$$

On the other hand, we can use the Poisson moment-generating function to bound

$$\mathbb{E}\left[\frac{D^p}{\exp(\gamma D)}; D \leq \frac{k\alpha_{\mathrm{ubd}}}{2}\right] \leq \left(\frac{k\alpha_{\mathrm{ubd}}}{2}\right)^p \mathbb{E}\left[\frac{1}{\exp(\gamma D)}\right]$$

$$= \left(\frac{k\alpha_{\mathrm{ubd}}}{2}\right)^p \exp\left\{-\frac{\alpha k}{2}\left(1 - \frac{1}{e^\gamma}\right)\right\} \leq O(1)\left(\frac{k\ln 2}{2}\right)^p.$$

Combining the bounds gives the claim. $\qquad\square$

Lemma 4.22. *Let D be a* $\mathrm{Pois}(\alpha k/2)$ *random variable, and let F_i be events such that whenever $D \geq \alpha k/4$, we have $\mathbb{P}(F_i \mid D) \leq \exp(-2^{k/4})$ for all $1 \leq i \leq D$. If $p = 2^{k\delta}$ for $1/100 \leq \delta \leq 1/9$, then*

$$\mathbb{E}\left[D^{p-1}\sum_{i=1}^{D}\mathbf{1}_{F_i}\right] \leq \frac{1}{\exp(2^{k/5})}.$$

Proof. It follows from the Poisson Chernoff bound (4.11) that

$$\mathbb{E}\left[D^{p-1}\sum_{i=1}^{D}\mathbf{1}_{F_i}; D \leq \frac{\alpha k}{4}\right] \leq \left(\frac{\alpha k}{4}\right)^p \mathbb{P}\left(D \leq \frac{\alpha k}{4}\right) \leq \left(\frac{\alpha k}{4}\right)^p \frac{1}{\exp(\Omega(k2^k))} \leq \frac{1}{\exp(\Omega(k2^k))}.$$

We can use integration by parts and (4.11) to bound

$$\mathbb{E}\left(D^p; D \geq 3^k\right) \leq 3^{kp}\mathbb{P}(D \geq 3^k) + \int_{3^k}^{\infty} pt^{p-1}\mathbb{P}(D \geq t)\,dt \leq \frac{3^{kp}}{\exp(\Omega(k3^k))} + p\int_{3^k}^{\infty}\frac{t^{p-1}}{\exp(\Omega(kt))}\,dt$$

$$\leq \frac{3^{kp}}{\exp(\Omega(k3^k))} + \frac{p}{\exp(\Omega(k3^k))} \leq \frac{1}{\exp(\Omega(k3^k))},$$

so we have $\mathbb{E}(D^p) \leq O(3^{kp})$. It follows from the assumption on the F_i that

$$\mathbb{E}\left[D^{p-1}\sum_{i=1}^{D}\mathbf{1}_{F_i}; D \geq \frac{\alpha k}{4}\right] \leq \frac{\mathbb{E}(D^p)}{\exp(2^{k/4})} \leq \frac{1}{\exp(\Omega(2^{k/4}))},$$

which implies the claim. $\qquad\square$

Lemma 4.23. *Let Z, Z_i be i.i.d. random variables with symmetric distribution (meaning that Z is equidistributed as $-Z$). Then, for any positive integer d and positive even integer p,*

$$\mathbb{E}\left(\sum_{i=1}^{d}Z_i\right)^p \leq O(1)\left(\frac{dp}{e}\right)^{p/2}\mathbb{E}(|Z|^p).$$

Proof. Writing $\underline{i} \equiv (i_1, \ldots, i_p)$ for elements of $[d]^p$, we expand

$$\mathbb{E}\left(\sum_{i=1}^{d} Z_i\right)^p = \sum_{\underline{i} \in [d]^p} \mathbb{E}[Z_{i_1} \cdots Z_{i_p}].$$

On the right-hand side, $\mathbb{E}[Z_{i_1} \cdots Z_{i_p}]$ is zero unless every index appears an even number of times. The number of choices for $\underline{i}$ for which this holds is upper bounded by $d^{p/2}(p-1)!!$, where $(p-1)!!$ is the number of matchings on p elements. By Stirling's formula,

$$d^{p/2}(p-1)!! = \frac{d^{p/2}p!}{2^{p/2}(p/2)!} \leq O(1)\left(\frac{dp}{e}\right)^{p/2}.$$

By Jensen's inequality, for any $\underline{i}$ we have $\mathbb{E}[Z_{i_1} \cdots Z_{i_p}] \leq \mathbb{E}(|Z|^p)$. Combining these bounds gives the claim. □

The next three lemmas record some simple (deterministic) bounds, based on elementary calculus manipulations, which will also be used in the proof of Lemma 4.27 below. For $0 \leq x \leq 1$ and $y \in \mathbb{R}$, define the function

$$F(x, y) \equiv \frac{(1-x)e^y}{1 + (1-x)e^y}. \tag{4.28}$$

and note that F takes values in $[0, 1)$. Write $\nabla F \equiv (F_x, F_y) \equiv (\partial F/\partial x, \partial F/\partial y)$.

Lemma 4.24. *Let F be as defined by* (4.28). *If $0 \leq x_i \leq 1$ and $y_i \in \mathbb{R}$ with $x_i e^{y_i} \leq 1$, then*

$$\left|F(x_1, y_1) - F(x_2, y_2)\right| \leq \left|x_1 - x_2\right| + \left|y_1 - y_2\right|, \tag{4.29}$$

$$\left|F(x_1, y_1) - F(x_2, y_2) - \left\langle \nabla F(x_1, y_1), \binom{x_1 - x_2}{y_1 - y_2} \right\rangle\right| \leq c\left\{(x_1 - x_2)^2 + (y_1 - y_2)^2\right\}, \tag{4.30}$$

where c is an absolute constant.

Proof. Write F_{xx}, F_{xy}, F_{yy} for the second-order partial derivatives of F. It is straightforward to verify that for $0 \leq x \leq 1$ and all $y \in \mathbb{R}$ we have $|F_y| \leq 1$ and $|F_{yy}| \leq 1$. Further, under the additional restriction that $xe^y \leq 1$, we have $|F_x| \leq 1$, $|F_{xx}| \leq 2$, and $|F_{xy}| \leq 1$. The bounds (4.29) and (4.30) directly follow. □

Lemma 4.25. *Let F be as defined by* (4.28). *For any $0 \leq x_i \leq 1$ and $y_i \in \mathbb{R}$,*

$$\left|\frac{F(x_1, y_1) - F(x_2, y_2)}{F(x_1, y_1) + F(x_2, y_2)}\right| \leq \frac{|x_1 - x_2|}{2 - x_1 - x_2} + \frac{|y_1 - y_2|}{2}. \tag{4.31}$$

(Unlike Lemma 4.24, this does not require $x_i e^{y_i} \leq 1$.)

Proof. Note that the function $G(z) \equiv z/(1+z)$ satisfies

$$\frac{|G(w) - G(z)|}{G(w) + G(z)} = \frac{|w - z|}{w + z + 2wz} \leq \frac{|w - z|}{w + z} \tag{4.32}$$

for any $w, z \geq 0$. It follows from (4.32) that for any $0 \leq x_i \leq 1$ and any fixed $y \in \mathbb{R}$,

$$\frac{|F(x_2, y) - F(x_1, y)|}{F(x_2, y) + F(x_1, y)} = \frac{|G((1-x_2)e^y) - G((1-x_1)e^y)|}{G((1-x_2)e^y) + G((1-x_1)e^y)} \leq \frac{|x_2 - x_1|}{2 - x_1 - x_2}. \tag{4.33}$$

A further consequence of (4.32) is that for any $w, z \in \mathbb{R}$,

$$\frac{|G(e^w) - G(e^z)|}{G(e^w) + G(e^z)} \leq \frac{|e^w - e^z|}{e^w + e^z} = \text{th}\left(\frac{|w - z|}{2}\right) \leq \frac{|w - z|}{2}.$$

It follows from this that for any fixed $0 \leq x \leq 1$ and any $y_i \in \mathbb{R}$,

$$\frac{|F(x, y_1) - F(x, y_2)|}{F(x, y_1) + F(x, y_2)} = \frac{|G((1-x)e^{y_2}) - G((1-x)e^{y_1})|}{G((1-x)e^{y_2}) + G((1-x)e^{y_1})} \leq \frac{|y_1 - y_2|}{2}. \tag{4.34}$$

Combining (4.33) and (4.34) gives the claim. □

Lemma 4.26. *For any positive numbers a_j and b_j,*

$$\left| \frac{\prod_{j=1}^{k} a_j - \prod_{j=1}^{k} b_j}{\prod_{j=1}^{k} a_j + \prod_{j=1}^{k} b_j} \right| \leq \sum_{\varnothing \subsetneq J \subseteq [k]} 2^{|J|} \prod_{j \in J} \left| \frac{a_j - b_j}{a_j + b_j} \right| . \tag{4.35}$$

Proof. Partition the indices $[k] \equiv \{1, \dots, k\}$ into $A \equiv \{j : a_j \geq b_j\}$ and $B \equiv [k] \setminus A$. Write $\delta \equiv a - b$. Then

$$N \equiv \prod_{j=1}^{k} a_j - \prod_{j=1}^{k} b_j = \prod_{j \in A}(b_j + \delta_j) \prod_{j \in B} a_j - \prod_{j \in B}(a_j - \delta_j) \prod_{j \in A} b_j$$

$$= \sum_{\varnothing \subsetneq J \subseteq A} \prod_{j \in J} \delta_j \prod_{j \in A \setminus J} b_j \prod_{j \in B} a_j - \sum_{\varnothing \subsetneq J \subseteq B} \prod_{j \in J}(-\delta_j) \prod_{j \in B \setminus J} a_j \prod_{j \in A} b_j .$$

It follows using the definition of A and B that

$$|N| \leq \sum_{\substack{\varnothing \subsetneq J \subseteq [k], \\ J \subseteq A \text{ or } J \subseteq B}} \prod_{j \in J} |\delta_j| \prod_{j \in [k] \setminus J} \min\{a_j, b_j\} ,$$

On the other hand, for any $J \subseteq A$, we have

$$D \equiv \prod_{j=1}^{k} a_j + \prod_{j=1}^{k} b_j \geq \prod_{j=1}^{k} a_j \geq \prod_{j \in J} \max\{a_j, b_j\} \prod_{j \in [k] \setminus J} \min\{a_j, b_j\} \geq \prod_{j \in J} \frac{a_j + b_j}{2} \prod_{j \in [k] \setminus J} \min\{a_j, b_j\} ,$$

and the same bound holds for any $J \subseteq B$. Combining the last two bounds gives

$$\frac{|N|}{D} \leq \sum_{\substack{\varnothing \subsetneq J \subseteq [k], \\ J \subseteq A \text{ or } J \subseteq B}} \prod_{j \in J} \frac{2|\delta_j|}{a_j + b_j} ,$$

which implies the claim. $\qquad\square$

Now recall from Definition 4.6 the measure PGW_ϵ. Let $\mathscr{T} \equiv \mathscr{T}_{v_{\text{RT}} a_{\text{RT}}} \sim \text{PGW}_\epsilon$, and (as in (4.5)) let

$$(\eta^\ell)_{\ell \geq 0} \equiv \left(\mathbf{F}_\ell(\mathscr{T}_{v_{\text{RT}} a_{\text{RT}}}) \right)_{\ell \geq 0} . \tag{4.36}$$

The next lemma gives the main technical estimate which will be used to prove stability in the Galton–Watson tree.

Lemma 4.27. *Write $\mathbb{P}_\epsilon$ for the law of the random sequence defined by (4.36), and write $\mathbb{E}_\epsilon$ for expectation with respect to $\mathbb{P}_\epsilon$. Let $\eta^\ell \equiv \eta^\ell(\text{-})$. For $p = 2\lceil 2^{k/10} \rceil$ we have the bound*

$$\mathbb{E}_\epsilon \left[\left| \frac{\eta^{\ell+1} - \eta^\ell}{\eta^{\ell+1} + \eta^\ell} \right|^p \right] \leq \frac{1}{2^{kp/7} 2^{2kp\ell/5}} \leq \frac{1}{2^{kp(\ell+1)/7}}$$

for all $\ell \geq 0$. This holds for any $0 \leq \epsilon \leq 1$.

Proof. We will prove the bound by induction, starting from the base case $\ell = 0$: by definition, we have $\eta^0 = 1/2$ (with probability one), and so Corollary 4.19 gives

$$\mathbb{E}_\epsilon \left[\left| \frac{\eta^1 - \eta^0}{\eta^1 + \eta^0} \right|^p \right] = \mathbb{E}_\epsilon \left[\left| \frac{\eta^1 - 1/2}{\eta^1 + 1/2} \right|^p \right] \leq \frac{1}{2^{kp/7}} .$$

Now suppose inductively that for some $\ell \geq 1$ we have

$$\mathbb{E}_\epsilon \left[\left| \frac{\eta^\ell - \eta^{\ell-1}}{2} \right|^p \right] \leq \mathbb{E}_\epsilon \left[\left| \frac{\eta^\ell - \eta^{\ell-1}}{\eta^\ell + \eta^{\ell-1}} \right|^p \right] \leq I(\ell)^p \equiv \frac{1}{2^{kp/7} 2^{2kp\ell/5}} . \tag{4.37}$$

Let $A \equiv (H, h)$ denote a random variable with the same law as the pair $(\eta^\ell, \eta^{\ell-1})$. Let

$$\underline{A} \equiv \left(A_j \equiv \binom{H_j}{h_j}, A_{ij}^+ \equiv \binom{H_{ij}^+}{h_{ij}^+}, A_{ij}^- \equiv \binom{H_{ij}^-}{h_{ij}^-} \right)_{i,j \geq 1}$$

be an array of i.i.d. copies of A. Let K denote a random variable which takes value k with chance $1 - \epsilon$, and takes value $k - 1$ with chance ϵ. Let $\underline{K} \equiv (K_{ij}^+, K_{ij}^-)_{i,j \geq 1}$ be an array of i.i.d. copies of K, and let

$$S_i^\pm \equiv 1 - \prod_{j=1}^{K_i^\pm - 1} H_{ij}^\pm, \quad s_i^\pm \equiv 1 - \prod_{j=1}^{K_i^\pm - 1} h_{ij}^\pm.$$

Note that by construction the H and h random variables lie in $[0, 1)$ almost surely, so the S and s random variables lie in $(0, 1]$ almost surely. Let $\underline{d} \equiv (d^+, d^-) \sim \mathrm{po}_\pm$, and use this to define the random variables

$$\Pi^\pm \equiv \prod_{i=1}^{d^\pm} S_i^\pm, \quad \pi^\pm \equiv \prod_{i=1}^{d^\pm} s_i^\pm, \quad \Sigma \equiv \ln \frac{\Pi^+}{\Pi^-}, \quad \sigma \equiv \ln \frac{\pi^+}{\pi^-}. \tag{4.38}$$

For the remainder of the proof we will abbreviate $S_i \equiv S_i^-, s_i \equiv s_i^-, \Pi \equiv \Pi^-, \pi \equiv \pi^-$. With this notation, the pair $(\eta^{\ell+1}, \eta^\ell)$ is equidistributed as

$$\left(\frac{(1 - \Pi)\exp(\Sigma)}{1 + (1 - \Pi)\exp(\Sigma)}, \frac{(1 - \pi)\exp(\sigma)}{1 + (1 - \pi)\exp(\sigma)} \right) = \left(F(\Pi, \Sigma), F(\pi, \sigma) \right),$$

for F as defined by (4.28). To prove the result, it suffices to bound $\mathbb{E}(J^p)$ for

$$J^p \equiv \left| \frac{F(\Pi, \Sigma) - F(\pi, \sigma)}{F(\Pi, \Sigma) + F(\pi, \sigma)} \right|^p \leq \left(\frac{2|\Pi - \pi|}{2 - \Pi - \pi} \right)^p + |\Sigma - \sigma|^p,$$

where the last bound follows by Lemma 4.25. Since each of Π, π is a product over d^- terms, we can decompose their difference as a telescoping product:

$$\Pi - \pi = \prod_{i=1}^{d^-} S_i - \prod_{i=1}^{d^-} s_i = \sum_{i=1}^{d^-} \Pi[i](S_i - s_i), \quad \Pi[i] \equiv \prod_{j=1}^{i-1} S_j \prod_{j=i+1}^{d^-} s_j.$$

It follows from Hölder's inequality that for any $\mathbf{u} \in \mathbb{R}^d$, $|(\mathbf{1}, \mathbf{u})|^p \leq d^{p-1}(\|\mathbf{u}\|_p)^p$. This implies

$$|\Pi - \pi|^p \leq (d^-)^{p-1} \sum_{i=1}^{d^-} \Pi[i]^p |S_i - s_i|^p. \tag{4.39}$$

Note that $0 \leq \Pi, \pi \leq 1$, and $\Pi\pi \leq \Pi[i]$ for any i. Let F_i be the event that $\Pi[i] \geq 1/4$. If F_i does not occur, then we must have either $\Pi \leq 1/2$ or $\Pi \leq 1/2$, therefore $2 - \Pi - \pi \geq 1/2$. It follows that

$$J^p \leq \underbrace{\left[4^p (d^-)^{p-1} \sum_{i=1}^{d^-} \Pi[i]^p |S_i - s_i|^p \right]}_{J_1(p)} + \underbrace{\left[2^p (d^-)^{p-1} \sum_{i=1}^{d^-} \mathbf{1}_{F_i} \frac{\Pi[i]^p |S_i - s_i|^p}{(2 - \Pi - \pi)^p} \right]}_{J_2(p)} + |\Sigma - \sigma|^p. \tag{4.40}$$

Similarly to (4.39), we can expand $S - s \equiv S_i - s_i$ as a telescoping sum, then use Hölder's inequality to bound

$$|S - s|^p = \left| \prod_{j=1}^{K-1} H_j - \prod_{j=1}^{K-1} h_j \right|^p \leq k^{p-1} \sum_{j=1}^{K-1} U[j]^p |H_j - h_j|^p, \quad U[j] \equiv \prod_{t=1}^{j-1} H_t \prod_{t=j+1}^{K-1} h_t. \tag{4.41}$$

It follows by considerations of conditional independence that

$$\mathbb{E}J_1(p) \leq (4k)^p \mathbb{E}\left[(d^-)^p (\max\{\mathbb{E}(S^p), \mathbb{E}(s^p)\})^{d^- - 1} \right] \left(\max\{\mathbb{E}(H^p), \mathbb{E}(h^p)\} \right)^{k-3} \mathbb{E}(|H - h|^p).$$

Both $\mathbb{E}(S^p)$ and $\mathbb{E}(s^p)$ satisfy the bound from Corollary 4.20; and combining with Lemma 4.21 gives a bound on the first expectation on the right-hand side above. Both $\mathbb{E}(H^p)$ and $\mathbb{E}(h^p)$ satisfy the bound from Corollary 4.18. The last factor $\mathbb{E}(|H - h|^p)$ is bounded by the inductive hypothesis (4.37). Altogether it gives

$$\mathbb{E}J_1(p) \leq O(1)(4k)^p \left(\frac{3k}{8} \right)^p \frac{\mathbb{E}(|H - h|^p)}{2^{p(k-3)}} \leq \left(\frac{k^3}{2^k} \right)^p \frac{1}{2^{kp/7} 2^{2kp\ell/5}} = \left(\frac{k^3 I(\ell)}{2^k} \right)^p. \tag{4.42}$$

Turning to $J_2(p)$, we note that (very crudely) we have $0 \leq \Pi[i] \leq 1$ and $2 - \Pi - \pi \geq 2 - S_i - s_i$ for any $1 \leq i \leq d^-$. Therefore, again with considerations of conditional independence, we have

$$\mathbb{E}J_2(p) \leq 2^p \mathbb{E}\left[(d^-)^{p-1} \sum_{i=1}^{d^-} \mathbf{1}_{F_i}\right] \mathbb{E}\left[\left|\frac{S-s}{2-S-s}\right|^p\right]. \tag{4.43}$$

Recall that F_i is the event that $\Pi[i] \geq 1/4$, or equivalently that

$$\Sigma[i] \equiv \ln \frac{1}{\Pi[i]} = \sum_{j=1}^{i-1} \ln \frac{1}{S_i} + \sum_{j=i+1}^{d^-} \ln \frac{1}{s_i} \leq \ln 4.$$

By essentially the same argument as for (4.18), as long as $d^- \geq \alpha k/4$ we have $\mathbb{P}(F_i \mid d^-) \leq \exp(-\Omega(k 2^{k/4}))$. We can then apply Lemma 4.22 to bound the first expectation in (4.43):

$$\mathbb{E}\left[(d^-)^{p-1} \sum_{i=1}^{d^-} \mathbf{1}_{F_i}\right] \leq \frac{1}{\exp(2^{k/5})}. \tag{4.44}$$

For the second expectation in (4.43), Lemma 4.26 gives

$$\mathbb{E}\left[\left|\frac{S-s}{2-S-s}\right|^p\right] = \mathbb{E}\left[\left|\frac{\prod_{j=1}^{K-1} H_j - \prod_{j=1}^{K-1} h_j}{\prod_{j=1}^{K-1} H_j + \prod_{j=1}^{K-1} h_j}\right|^p\right] \leq \mathbb{E}\left[\left(\sum_{\varnothing \subsetneq J \subseteq [k-1]} 2^{|J|} \prod_{j \in J} \left|\frac{H_j - h_j}{H_j + h_j}\right|\right)^p\right].$$

Let $\underline{J} \equiv (J_1, \ldots, J_p)$ denote any p-tuple of nonempty subsets of $[k-1]$. We abbreviate $|\underline{J}| \equiv |J_1| + \ldots + |J_p| \geq p$. For any $j \in [k-1]$, let $n_j(\underline{J})$ denote the number of occurrences of j in the sets $J_1, \ldots, J_p$:

$$n_j(\underline{J}) = \sum_{t=1}^{p} \mathbf{1}\{j \in J_t\}, \quad \sum_{j=1}^{k-1} n_j(\underline{J}) = \sum_{t=1}^{p} |J_t| = |\underline{J}| \geq p.$$

With this notation, we can bound

$$\mathbb{E}\left[\left|\frac{S-s}{2-S-s}\right|^p\right] \leq \sum_{\underline{J}} 2^{|\underline{J}|} \prod_{j=1}^{k-1} \mathbb{E}\left[\left|\frac{H_j - h_j}{H_j + h_j}\right|^{n_j(\underline{J})}\right] \leq \sum_{\underline{J}} 2^{|\underline{J}|} \mathbb{E}\left[\left|\frac{H-h}{H+h}\right|^p\right]^{|\underline{J}|/p},$$

where the last step is by Jensen's inequality. The number of choices of $\underline{J} \equiv (J_1, \ldots, J_p)$ with $|\underline{J}| = b$ is upper bounded by k^b. Combining with the inductive hypothesis (4.37) gives

$$\mathbb{E}\left[\left|\frac{S-s}{2-S-s}\right|^p\right] \leq \sum_{b \geq p}\left(2k\mathbb{E}\left[\left|\frac{H-h}{H+h}\right|^p\right]^{1/p}\right)^b \leq \sum_{b \geq p}\left(\frac{2k}{2^{k/7} 2^{2k\ell/5}}\right)^b \leq \frac{k^{2p}}{2^{kp/7} 2^{2kp\ell/5}}.$$

Substituting this and (4.44) into (4.43) gives

$$\mathbb{E}J_2(p) \leq \frac{k^{2p}}{\exp(2^{k/5})} \frac{1}{2^{kp/7} 2^{2kp\ell/5}} \leq \left(\frac{k^2 I(\ell)}{\exp(2^{k/10})}\right)^p. \tag{4.45}$$

Finally we bound $\mathbb{E}(|\Sigma - \sigma|^p)$, where we recall from (4.38) that

$$\Sigma = \ln \frac{\Pi^+}{\Pi^-} = \sum_{i=1}^{d^+} \ln S_i^+ - \sum_{i=1}^{d^-} \ln S_i^-,$$

and similarly $\sigma = \ln(\pi^+/\pi^-)$. Let $D \sim \text{Pois}(\alpha k)$, and let $(\mathfrak{s}_i)_{i \geq 1}$ be a sequence of i.i.d. symmetric random signs. Then, recalling that p is an even integer, we have

$$\mathbb{E}(|\Sigma - \sigma|^p) = \mathbb{E}\left[\left(\sum_{i=1}^{D} \mathfrak{s}_i(\ln S_i - \ln s_i)\right)^p\right] \leq O(1)\left(\frac{dp}{e}\right)^{p/2} \mathbb{E}\left[\left|\ln \frac{S}{s}\right|^p\right], \tag{4.46}$$

where the last step uses Lemma 4.23. We then use a telescoping sum to bound

$$\left| \ln \frac{S}{s} \right| = \left| \ln \frac{1 - \prod_{j=1}^{K-1} H_j}{1 - \prod_{j=1}^{K-1} h_j} \right| \leq \sum_{j=1}^{K-1} \left| \ln \frac{1 - U[j]H_j}{1 - U[j]h_j} \right| \leq \sum_{j=1}^{K-1} \frac{U[j]}{1 - U[j]} \cdot |H_j - h_j|, \tag{4.47}$$

where the last step uses that the function $f_U(h) = \ln(1 - Uh)$ has $|(f_U)'(h)| \leq U/(1-U)$ for all $U, h \in [0,1]$. In the last expression above, we can further replace $U[j]$ (defined by (4.41)) with

$$U_\circ[j] \equiv \prod_{t=1}^{j-1} H_t \prod_{t=j+1}^{k-2} h_t \geq U[j].$$

It follows by Hölder's inequality, combined with the Cauchy-Schwarz inequality, that

$$\mathbb{E}\left[\left| \ln \frac{S}{s} \right|^p \right] \leq k^p \left(\max\{\mathbb{E}(H^{2p}), \mathbb{E}(h^{2p})\} \right)^{(k-3)/2} \max_{j \leq k-2} \mathbb{E}\left[\left(\frac{1}{1 - U_\circ[j]} \right)^{2p} \right]^{1/2} \mathbb{E}(|H - h|^p) \tag{4.48}$$

For $x \geq 3$, we have $\ln x + \ln(1 - 1/x) \geq (\ln x)/2$. It follows that

$$\mathbb{P}\left(\frac{1}{1 - U_\circ[j]} \geq x \right) = \mathbb{P}\left(U_\circ[j] \geq 1 - \frac{1}{x} \right) \leq \max\left\{ \mathbb{P}\left(H \geq 1 - \frac{1}{x} \right), \mathbb{P}\left(h \geq 1 - \frac{1}{x} \right) \right\}^{k-3}$$

$$\leq \max\left\{ \mathbb{P}\left(\ln \frac{H}{1 - H} \geq \frac{\ln x}{2} \right), \mathbb{P}\left(\ln \frac{h}{1 - h} \geq \frac{\ln x}{2} \right) \right\}^{k-3} \leq \frac{1}{\exp(k(\ln x)2^{k/4})},$$

where the last bound is by (I) from Lemma 4.12. Integrating this bound gives

$$\mathbb{E}\left[\left(\frac{1}{1 - U_\circ[j]} \right)^{2p} \right] \leq (3/2) \cdot 3^{2p} + \int_3^\infty \frac{pt^{p-1}}{t^{k2^{k/4}}} \, dt \leq 2 \cdot 3^{2p}.$$

The other factors in (4.48) are controlled by Corollary 4.18 and (4.37). Substituting into (4.46) gives

$$\mathbb{E}|\Sigma - \sigma|^p \leq O(1) \left(\frac{dp}{e} \right)^{p/2} \frac{k^{2p}}{2^{kp}} \frac{1}{2^{kp/7} 2^{2kp\ell/5}} \leq \left(\frac{k^2 I(\ell)}{2^{9k/20}} \right)^p. \tag{4.49}$$

Substituting (4.42), (4.45), and (4.49) into (4.40) gives

$$\mathbb{E}\left[\left| \frac{\eta^{\ell+1} - \eta^\ell}{\eta^{\ell+1} + \eta^\ell} \right|^p \right] = \mathbb{E}(J^p) \leq O(1) \left(\frac{k^2 I(\ell)}{2^{9k/20}} \right)^p,$$

which verifies the induction (4.37) and proves the result.   □

We will apply Lemma 4.27 below to obtain our final bound, Proposition 4.17, on 1-stability in the **PGW** tree. Before doing so, however, we note that Proposition 1.1 is essentially an immediate consequence of Lemma 4.27:

Proof of Proposition 1.1. Take the random sequence $(\eta^\ell)_{\ell \geq 0}$ as in (4.5) or (4.36), with $\epsilon = 0$. Then, as noted in the discussion around (4.5), the marginal law of each η^ℓ is precisely $\mu^\ell \equiv R^\ell \mu^0$, the same as the μ^ℓ appearing in the statement of this proposition. It follows from Lemma 4.27 that for $p = 2\lceil 2^{k/10} \rceil$ we have

$$\sum_{\ell \geq 0} \mathbb{E}\left[\left| \eta^{\ell+1} - \eta^\ell \right|^p \right]^{1/p} \leq 2 \sum_{\ell \geq 0} \mathbb{E}\left[\left| \frac{\eta^{\ell+1} - \eta^\ell}{\eta^{\ell+1} + \eta^\ell} \right|^p \right]^{1/p} < \infty.$$

Thus $(\eta^\ell)_{\ell \geq 0}$ is a Cauchy sequence in L^p, hence it must converge in L^p to a limiting random variable η as $\ell \to \infty$. It follows that μ^ℓ converges weakly to a limiting probability measure μ as $\ell \to \infty$. Since $\mu^\ell = R\mu^{\ell-1}$ and the mapping R is continuous with respect to the weak topology on the space of distributions, we conclude $\mu = R\mu$.   □

The remainder of this subsection is devoted to the proof of Proposition 4.17. The following lemma records the easy observation that the canonical messages can only become "more free" when the neighborhood is enlarged:

Lemma 4.28. *Let $\mathscr{T}_{va}$ be any variable-to-clause tree in the sense of Definition 3.3. Similarly as in (4.5) or (4.36), let*

$$(\eta^\ell)_{\ell \geq 0} \equiv \left(\mathbf{F}_\ell(\mathscr{T}_{va})\right)_{\ell \geq 0}.$$

It holds for any $\mathscr{T}_{va}$ and any $\ell \geq 0$ that if $\eta^\ell(+) = 0$ then $\eta^{\ell+1}(+) = 0$ also; likewise if $\eta^\ell(-) = 0$ then $\eta^{\ell+1}(-) = 0$ also. Moreover $\eta^\ell(\mathtt{f}) \in (0, 1]$ for all $\ell \geq 1$. As a consequence we always have

$$\{\mathtt{g}\} \subseteq \mathrm{supp}\, {}_\star\pi \subseteq \mathrm{supp}\, {}_\square\pi\,, \tag{4.50}$$

for all $r \geq 2$, for ${}_\star\pi$ and ${}_\square\pi$ as in Definition 3.4.

Proof. Thanks to the symmetry between + and -, it suffices to prove the first assertion for the quantities $\eta^\ell \equiv \eta^\ell(-)$. That is to say, we shall argue that if $\eta^\ell = 0$ then $\eta^{\ell+1} = 0$ also, for all $\ell \geq 0$. Note that $\eta^0 = 1/2 \neq 0$ by assumption, so the statement holds trivially for $\ell = 0$. Suppose inductively that it holds up to $\ell - 1$; we then compare η^ℓ with $\eta^{\ell+1}$. Similarly as in Definitions 4.8–4.10 and the proof of Lemma 4.27, we can express

$$\eta^{\ell+1} = \frac{\Pi^+(1 - \Pi^-)}{\Pi^+ + \Pi^- - \Pi^+\Pi^-}\,, \quad \Pi^\pm \equiv \prod_{b \in \partial v(\pm a)} \left(1 - \prod_{u \in \partial b \setminus v} \eta^\ell(\mathscr{T}_{ub})\right)$$

and similarly

$$\eta^\ell = \frac{\pi^+(1 - \pi^-)}{\pi^+ + \pi^- - \pi^+\pi^-}\,, \quad \pi^\pm \equiv \prod_{b \in \partial v(\pm a)} \left(1 - \prod_{u \in \partial b \setminus v} \eta^{\ell-1}(\mathscr{T}_{ub}).\right)$$

Now suppose $\eta^\ell = 0$. Since $\pi^\pm \in (0, 1]$, it must be that $\pi^- = 1$. This can occur in one of two possible ways:

- The first possibility is that $\partial v(-a)$. In this case $\Pi^- = 1$ also, and so $\eta^{\ell+1} = 0$.
- The only other possibility is that $\eta^{\ell-1}(\mathscr{T}_{ub}) = 0$ for all $u \in \partial b \setminus v$, for all $b \in \partial v(-a)$. In this case, it follows from the inductive hypothesis that $\eta^\ell(\mathscr{T}_{ub}) = 0$ for all $u \in \partial b \setminus v$, for all $b \in \partial v(-a)$. Therefore $\Pi^- = 1$ also, and so again we conclude $\eta^{\ell+1} = 0$.

This verifies the induction and thus proves the first assertion. It is easy to see from the form of the recursion (4.4) that $\eta^\ell(\mathtt{f}) \in (0, 1]$ for all $\ell \geq 1$, since $\Pi^\pm \in (0, 1]$. For the next assertion (4.50), recall from the discussion around (3.3) and (3.4) that ${}_\star\pi$ is proportional to the product of ${}_\star\hat{q}_{av}$ and ${}_\star\dot{q}_{va}$, while ${}_\square\pi$ is proportional to the product of ${}_\star\hat{q}_{av}$ and ${}_\square\dot{q}_{va}$. The only difference is that ${}_\star\dot{q}_{va}$ corresponds to η^r while ${}_\square\dot{q}_{va}$ corresponds to η^{r-1} (via (2.38)). It then follows from the prior assertions of this lemma that we have

$$\{\mathtt{f}\} \subseteq \mathrm{supp}\, \eta^r \subseteq \mathrm{supp}\, \eta^{r-1}$$

for all $r \geq 2$, and (4.50) follows straightforwardly by substituting into (2.38). $\square$

Lemma 4.29. *For the measure $\mathbf{PGW}_\epsilon$ of Definition 4.6, we have*

$$\mathbf{PGW}_\epsilon\left(v_{\mathrm{RT}} \text{ not stable}\right) \leq \frac{1}{\exp(2^{k/12}R)}\,,$$

where "stable" means 0-stable in the sense of Definition 3.9. This bound holds for all $0 \leq \epsilon \leq 1$.

Proof. For any tree $\mathscr{T}$ and any edge (au) of $\mathscr{T}$, we can consider the variable-to-clause tree $\mathscr{T}_{ua} \subseteq \mathscr{T}$, and define (cf. (4.5) and (4.36)) the measures $\eta^\ell_{ua} \equiv \mathbf{F}_\ell(\mathscr{T}_{ua})$. We define two functions on variable-to-clause trees,

$$f^+(\mathscr{T}_{ua}) \equiv \mathbf{1}\left\{\max\left\{\eta^r_{ua}(+), \eta^{r-1}_{ua}(+)\right\} \geq 1 - \frac{1}{k^r}\right\},$$

$$g^+(\mathscr{T}_{ua}) \equiv \mathbf{1}\left\{\left|\frac{\eta^r_{ua}(+) - \eta^{r-1}_{ua}(+)}{\eta^r_{ua}(+) + \eta^{r-1}_{ua}(+)}\right| \geq \frac{1}{2^{kr/8}}\right\},$$

and likewise $f^-(\mathscr{T}_{ua})$ and $g^-(\mathscr{T}_{ua})$. If u and w are neighboring variables on $\mathscr{T}$, we let $a(uw) \equiv a(wu)$ denote the unique clause that they share. For the rest of the proof, we let $\mathscr{T}$ be a sample from $\mathbf{PGW}_\epsilon$, rooted at $v \equiv v_{\mathrm{RT}}$. Since $\mathbf{PGW}_\epsilon$ is unimodular, we can apply (4.1) to evaluate

$$\int \left[\sum_{u \in N(v)} f^-(\mathscr{T}_{va(vu)})\right] d\mathbf{PGW}_\epsilon(\mathscr{T}) = \int \left[\sum_{u \in N(v)} f^-(\mathscr{T}_{ua(uv)})\right] d\mathbf{PGW}_\epsilon(\mathscr{T}). \tag{4.51}$$

Note that, given $B_1(v)$, the $f^-(\mathscr{T}_{va(vu)})$ are not conditionally independent over $u \in N(v)$, but the $f^-(\mathscr{T}_{ua(uv)})$ are. The above is therefore upper bounded by

$$k^2\alpha \int f^-(\mathscr{T}_{va})\, d\mathrm{PGW}_\epsilon(\mathscr{T}_{va}) \le \frac{k^2\alpha}{\exp(\Omega(r2^{k/4}))},\tag{4.52}$$

where the last bound is by the upper tail bound (I) from Lemma 4.12. By symmetry, the bound (4.52) also holds with f^+ in place of f^-. With g^- in place of f^-, the identity (4.51) also holds, and conditional independence gives the upper bound

$$k^2\alpha \int g^-(\mathscr{T}_{va})\, d\mathrm{PGW}_\epsilon(\mathscr{T}_{va}) \le k^2\alpha\, \mathbb{P}_\epsilon\left(\left|\frac{\eta^r - \eta^{r-1}}{\eta^r + \eta^{r-1}}\right| \ge \frac{1}{2^{kr/8}}\right),$$

where $\mathbb{P}_\epsilon$ is the measure from Lemma 4.27. From the bound of Lemma 4.27, the last expression is

$$\le k^2\alpha(2^{kr/8})^p \mathbb{E}_\epsilon\left[\left(\left|\frac{\eta^r - \eta^{r-1}}{\eta^r + \eta^{r-1}}\right|^p\right)\right] \le \frac{k^2\alpha 2^{kpr/8}}{2^{kpr/7}} \le \frac{4^k}{2^{kpr/56}} \le \frac{1}{\exp(kr2^{k/11})}.\tag{4.53}$$

Let $A = f^+ + f^- + g^+ + g^-$, and define the ($B_r(v;\mathscr{T})$-measurable) event

$$E \equiv \bigcap_{a\in\partial v}\left\{A(\mathscr{T}_{ua}) = 0 \text{ for all } u \in \partial a\right\}.$$

We then find by Markov's inequality together with (4.51), (4.52), and (4.53) that

$$\mathbf{PGW}_\epsilon(E^c) \le \int\left[\sum_{u\in\partial v}\left\{\frac{A(\mathscr{T}_{va(vu)})}{k-2} + A(\mathscr{T}_{ua(uv)})\right\}\right] d\mathbf{PGW}_\epsilon(\mathscr{T}) \le \frac{1}{\exp(r2^{k/11})}.$$

We will prove that the root v is stable on event E.

Recall from Definition 3.9 that v is stable if all its incident edges $e \in \delta v$ are both message-stable and marginal-stable (Definition 3.7). It is clear that E implies that every edge incident to v satisfies the message-stability condition (3.12). (The first part of (3.12) holds because E implies $f^\pm(\mathscr{T}_{ua}) = 0$ for all $a \in \partial v$ and $u \in \partial a$. The second part of (3.12) holds because E implies $g^\pm(\mathscr{T}_{ua}) = 0$ for all $a \in \partial v$ and $u \in \partial a$.) Therefore it remains only to check the marginal-stability conditions (3.8)–(3.11).

For the rest of the proof, for all $a \in \partial v$ and $u \in \partial a$, we will denote $H_{ua} \equiv \eta^r_{ua}(-)$, $h_{ua} \equiv \eta^{r-1}_{ua}(-)$, $P_{ua} \equiv \eta^r_{ua}(+)$, and $p_{ua} \equiv \eta^{r-1}_{ua}(+)$. For $a \in \partial v$ we also let

$$\hat{u}_{av} \equiv \prod_{u\in\partial a\backslash v} h_{ua}.$$

With this notation, and using the correspondence (2.38), we have for all $a \in \partial v$ that

$$\left(_\star\pi_{av}(\mathtt{r}),\, _\star\pi_{av}(\mathtt{y}),\, _\star\pi_{av}(\mathtt{b}),\, _\star\pi_{av}(\mathtt{c})\right) = \left(\frac{(1-H_{va})\hat{u}_{av}}{1-\hat{u}_{av}H_{va}},\, \frac{H_{va}(1-\hat{u}_{av})}{1-\hat{u}_{av}H_{va}},\, \frac{P_{va}(1-\hat{u}_{av})}{1-\hat{u}_{av}H_{va}},\, \frac{(1-H_{va})(1-\hat{u}_{av})}{1-\hat{u}_{av}H_{va}}\right).$$

We can obtain $_\square\pi_{av}$ from the same expressions, only h_{va} and p_{va} in place of H_{va} and P_{va} (keeping $\hat{u}_{av}$ the same). By Lemma 4.28, if $h_{va} = 0$ then $H_{va} = 0$, and if $p_{va} = 0$ then $P_{va} = 0$. On the event E, if $(H, h, \hat{u}) = (H_{va}, h_{va}, \hat{u}_{av})$ for any $a \in \partial v$, then either $h = H = 0$, or $h > 0$ and

$$\left|\frac{H}{h} - 1\right| = \frac{|H-h|}{h} \le 2\left|\frac{H-h}{H+h}\right|\Big/\left(1 - \left|\frac{H-h}{H+h}\right|\right) \le \frac{O(1)}{2^{kr/8}} \le \frac{1}{2^{kr/9}k^{10}},$$

$$\frac{|(1-\hat{u}H) - (1-\hat{u}h)|}{1-\hat{u}h} \le \frac{|H-h|}{1-h} \le k^r|H-h| \le 2k^r\left|\frac{H-h}{H+h}\right| \le \frac{2k^r}{2^{kr/8}} \le \frac{1}{2^{kr/9}k^{10}}.$$

The analogous bounds hold with $(P, p) = (P_{va}, p_{va})$ in place of (H, h). It follows that for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{b}\}$, we have either $_\square\pi_{av}(\sigma) = {}_\star\pi_{av}(\sigma) = 0$, or $_\square\pi_{av}(\sigma) > 0$ and

$$\left|\frac{_\star\pi_{av}(\sigma) - {}_\square\pi_{av}(\sigma)}{_\square\pi_{av}(\sigma)}\right| \le \frac{1}{2^{kr/9}}\tag{4.54}$$

For $\sigma = \mathtt{c}$, we know that $_\square\pi_{av}(\mathtt{c}) \ge {}_\square\pi_{av}(\mathtt{g})$ is always positive, and the estimate (4.54) holds on all of E. This implies the last stability condition (3.11). It remains to verify the other three conditions (3.8), (3.9), and (3.10). This essentially

amounts to a quantitative version of the argument of Lemma 3.49. Using (2.34) and the correspondence (2.38), we have

$$\mathrm{coher}_{au}({}_{\square}\pi) = \frac{h_{ua}}{\hat{z}_a}\left[1 - \prod_{w\in\partial a\setminus u} h_{wa} - \sum_{w\in\partial a\setminus u}(1-h_{wa})\prod_{z\in\partial a\setminus\{u,w\}} h_{za}\right],$$

where $\hat{z}_a$ is the clause normalization, given explicitly by

$$\hat{z}_a = 1 - \prod_{w\in\partial a} h_{wa} \le 1.$$

If w,z are any two distinct variables in $\partial a\setminus u$, on the event E we have the crude lower bound

$$\mathrm{coher}_{au}({}_{\square}\pi) \ge \frac{h_{ua}}{\hat{z}_a}(1-h_{wa})(1-h_{za}) \ge \frac{h_{ua}}{\hat{z}_a k^{2r}}.$$

Similarly, if u,w,z are any three distinct variables in ∂a, then on the event E we have

$$\mathrm{coher}_a({}_{\square}\pi) \ge \frac{1}{\hat{z}_a}(1-\eta_{ua})(1-\eta_{wa})(1-\eta_{za}) \ge \frac{1}{\hat{z}_a k^{3r}} \ge \frac{1}{k^{3r}}.$$

On the other hand, we have

$${}_{\square}\pi_{au}(\mathtt{y}) = \frac{h_{ua}}{\hat{z}_a}\left[1 - \prod_{w\in\partial a\setminus u} h_{wa}\right] \le \frac{h_{ua}}{\hat{z}_a} \le k^{2r}\mathrm{coher}_{au}({}_{\square}\pi). \tag{4.55}$$

Similarly, for all $w\in\partial a\setminus u$ we have

$${}_{\square}\pi_{aw}(\mathtt{r}) = \frac{(1-h_{wa})}{\hat{z}_a}\prod_{z\in\partial a\setminus w} h_{za} \le \frac{h_{ua}}{\hat{z}_a} \le k^{2r}\mathrm{coher}_{au}({}_{\square}\pi). \tag{4.56}$$

For all $u\in\partial a$, we have trivially

$$\left\{{}_{\square}\pi_{au}(\mathtt{c}) + {}_{\square}\pi_{au}(\mathtt{r})\right\} \le 1 \le k^{3r}\mathrm{coher}_a({}_{\square}\pi). \tag{4.57}$$

The bounds (3.8), (3.9), and (3.10) follow by combining (4.54) with (4.55), (4.56), (4.57). □

We now apply Lemma 4.29 to conclude the proof of the main result of this subsection:

Proof of Proposition 4.17. We can sample from $\mathbf{PGW}_\epsilon$ in the following way: first sample $\mathscr{T} \sim \mathbf{PGW}$. Declare each clause of $\mathscr{T}$ to be "open" with probability ϵ, independently over all clauses. For each open clause a, choose one of its child variables $u\in\partial a$ uniformly at random, and declare u to be "open" as well. Let V' be the (random) set of all open variables in $B_R(v_{\mathrm{RT}};\mathscr{T})$, and let $\mathscr{T}'$ be the connected component of $\mathscr{T}\setminus V'$ that contains v_{RT} — as a shorthand we will write $\mathscr{T}' \equiv \mathscr{T}'(V')$. Then $\mathscr{T}'$ is a sample from $\mathbf{PGW}_\epsilon$. Let $\gamma_\epsilon(\cdot\,|\,\mathscr{T})$ denote the law of V' given $\mathscr{T}$. Let Ω' denote the subspace of $\mathscr{T}$ for which $|B_R(v_{\mathrm{RT}};\mathscr{T})| \le M'$, where M' will be chosen below. Then

$$P(\epsilon) \equiv \mathbf{PGW}_\epsilon\Big(v_{\mathrm{RT}} \text{ is not stable}\Big) \ge \int_{\Omega'}\int \mathbf{1}\Big\{v_{\mathrm{RT}} \text{ not stable in } \mathscr{T}'(V')\Big\}\,d\gamma_\epsilon(V'\,|\,\mathscr{T})\,d\mathbf{PGW}(\mathscr{T})$$

$$\ge \int_{\Omega'}\sum_{u\in B_R(v_{\mathrm{RT}};\mathscr{T})\setminus v_{\mathrm{RT}}} \mathbf{1}\Big\{v_{\mathrm{RT}} \text{ not stable in } \mathscr{T}'(\{u\})\Big\}\int \mathbf{1}\Big\{V' = \{u\}\Big\}\,d\gamma_\epsilon(V'\,|\,\mathscr{T})\,d\mathbf{PGW}(\mathscr{T})$$

$$\ge \frac{\epsilon(1-\epsilon)^{M'}}{M'}\int_{\Omega'}\sum_{u\in B_R(v_{\mathrm{RT}};\mathscr{T})\setminus v_{\mathrm{RT}}} \mathbf{1}\Big\{v_{\mathrm{RT}} \text{ not stable in } \mathscr{T}'(\{u\})\Big\}\,d\mathbf{PGW}(\mathscr{T}),$$

where, for any $\mathscr{T}\in\Omega'$ and any $u\in B_R(v_{\mathrm{RT}};\mathscr{T})\setminus\{v_{\mathrm{RT}}\}$, the quantity $\epsilon(1-\epsilon)^{M'}/M'$ a crude lower bound on the chance that $V' = \{u\}$. If we rearrange the above and consider the contribution outside Ω', we obtain

$$P' \equiv \int\sum_{u\in B_R(v_{\mathrm{RT}};\mathscr{T})\setminus v_{\mathrm{RT}}} \mathbf{1}\Big\{v_{\mathrm{RT}} \text{ not stable in } \mathscr{T}'(\{u\})\Big\}\,d\mathbf{PGW}(\mathscr{T})$$

$$\le \frac{P(\epsilon)M'}{\epsilon(1-\epsilon)^{M'}} + \mathbb{E}\left[\Big|B_R(v_{\mathrm{RT}};\mathscr{T})\Big|;\Big|B_R(v_{\mathrm{RT}};\mathscr{T})\Big| \ge M'\right] \le \frac{1}{\exp(\Omega(2^{k/12}R))},$$

where the last bound follows by taking $M' = 1/\epsilon = \exp(k^2 R)$, and applying Lemma 4.7 with Lemma 4.29. Combining the above with the $\epsilon = 0$ case of Lemma 4.29 gives

$$\mathbf{PGW}\Big(v_{\mathrm{RT}} \text{ not 1-stable}\Big) \leq \mathbf{PGW}\Big(v_{\mathrm{RT}} \text{ not stable}\Big) + P' \leq \frac{O(1)}{\exp(\Omega(2^{k/12}R))} \,,$$

and the claimed bound follows. □

4.5. Threshold upper bound.

In this subsection we complete the proof of Proposition 1.3. Recall that in §4.4 we proved Proposition 1.1, saying that the sequence of measures μ^ℓ converges weakly to a limit $\mu \equiv \mu^{\infty,\alpha}$ which satisfies the distributional fixed point equation $R\mu = \mu$. Given this result, the 1-RSB free energy can be written as

$$\Phi(\alpha) \equiv \sum_{\underline{d}} \mathrm{po}_\pm(\underline{d}) \int \ln \frac{\Pi^+(\underline{d}, \underline{\eta}) + \Pi^-(\underline{d}, \underline{\eta}) - \Pi^+(\underline{d}, \underline{\eta})\Pi^-(\underline{d}, \underline{\eta})}{(1 - \prod_{j=1}^k \eta_j)^{(k-1)\alpha}} \, d\mu^\otimes\Big((\eta_j)_{j\geq 1}, \underline{\eta}\Big). \qquad (4.58)$$

(This is the same as (1.11).) We defer to Section 10 the proof of Proposition 1.2, which guarantees that Φ is decreasing in α so that the conjectured threshold $\alpha_\star$ is well-defined. In the current subsection, we give the proof of Proposition 1.3 assuming that Proposition 1.2 holds. The proof is an application of interpolation bounds [FL03, PT04] on the free energy of positive-temperature dilute spin glasses. We believe this argument was generally known, especially among the physics community; we include the proof of Proposition 1.3 only for the sake of completeness.

The **positive-temperature k-SAT model** can be formally defined as follows. First let $(\mathsf{L}_{aj})_{a,j\geq 0}$ be an array of i.i.d. symmetric random signs, $\mathsf{L}_{aj} \in \{\pm\}$ with $\mathbb{P}(\mathsf{L}_{aj} = +) = \mathbb{P}(\mathsf{L}_{aj} = -) = 1/2$. We then use these to define a vector $\underline{\theta} \equiv (\theta_a)_{a\geq 0}$ of i.i.d. random functions

$$\theta_a(x_1, \ldots, x_k) \equiv \mathbf{1}\Big\{\mathsf{L}_{aj}x_j = - \text{ for all } 1 \leq j \leq k\Big\}.$$

Let M be a Poisson random variable with mean $n\alpha$. Define the **random k-SAT Hamiltonian** $H_n : \{\pm\}^n \to [0, M]$,

$$H_n(\underline{x}) = \sum_{a=1}^M \theta_a(\underline{x}_{\partial a}),$$

where ∂a are chosen independently and uniformly at random from $[n]^k$. The **random k-SAT free energy** at inverse temperature β is then given by

$$F_n(\beta) = \frac{1}{n}\mathbb{E}_n \ln \sum_{\underline{x}\in\{\pm\}^n} \exp\{-\beta H_n(\underline{x})\}$$

with $\mathbb{E}_n$ denoting expectation over the random Hamiltonian H_n. The next bound is from [PT04, Theorem 3], edited only slightly to fit our notation. (In the remainder of this subsection, we will no longer refer to the random scalar η from §4.1, but will instead use η to denote a certain random measure which is closely related.)

Theorem 4.30 ([PT04, Theorem 3][18]). *Let $\mathcal{M}_1$ denote the space of probability measures on $\mathbb{R}$, and $\mathcal{M}_2$ the space of probability measures on $\mathcal{M}_1$. For $\zeta \in \mathcal{M}_2$, let $\underline{\eta} \equiv (\eta_{a,j})_{a,j\geq 0}$ be an array of i.i.d. samples from ζ. Conditioned on $\underline{\eta}$, let $\underline{\rho} \equiv (\rho_{a,j})_{a,j\geq 0}$ where each $\rho_{a,j}$ is a conditionally independent sample from $\eta_{a,j}$. For $x \in \{\pm\}$, define*

$$\boldsymbol{u}_a(x) = \sum_{\underline{x}\in\{\pm\}^k} \mathbf{1}\{x_k = x\} \exp\{-\beta\theta_a(\underline{x})\} \prod_{j=1}^{k-1} \frac{\exp\{\rho_{a,j}x_j\}}{2\,\mathrm{ch}\,\rho_{a,j}} \,,$$

with ch the hyperbolic cosine. Define also

$$\boldsymbol{u}_a = \sum_{\underline{x}\in\{\pm\}^k} \exp\{-\beta\theta_a(\underline{x})\} \prod_{j=1}^k \frac{\exp\{\rho_{a,j}x_j\}}{2\,\mathrm{ch}\,\rho_{a,j}} \,.$$

Note $\boldsymbol{u}_a(x)$ and $\boldsymbol{u}_a$ are random variables. Then, for any $0 < m < 1$ and for any $\zeta \in \mathcal{M}_2$,

$$F_n(\beta) \leq \Phi_{\text{1-RSB}}(\beta, \zeta, m) \equiv \frac{1}{m}\mathbb{E}\ln\mathbb{E}'\Bigg[\Big(\sum_{x\in\{\pm\}}\prod_{a=1}^d \boldsymbol{u}_a(x)\Big)^m\Bigg] - \frac{(k-1)\alpha}{m}\mathbb{E}\ln\mathbb{E}'[(\boldsymbol{u}_0)^m]$$

---

[18]Although the theorem in [PT04] is stated for even k, the same proof applies equally to odd k, as noted for example in [Tal11, Ch. 6].

where $\mathbb{E}'$ is expectation over ρ conditioned on $d, \underline{\theta}, \underline{\eta}$; and $\mathbb{E}$ is the overall expectation over $d, \underline{\theta}, \underline{\eta}, \rho$.

The threshold upper bound is a straightforward consequence:

Proof of Proposition 1.3. We will deduce the bound from Theorem 4.30 by taking a particular choice of ζ, m which is suggested by the survey propagation heuristic. Let $\mu \in \mathscr{P}$ be the fixed point given by Proposition 1.1. Let $\boldsymbol{\mu} = \mathcal{R}\mu$ for $\mathcal{R} \equiv \mathcal{R}_0$ as given by Definition 4.10. Now fix $\beta > 0$, and let

$$\left(\rho_+, \rho_-, \rho_{\mathtt{f}}\right) \equiv \left(\beta, -\beta, 0\right). \tag{4.59}$$

Let η be defined as the law of ρ_y where $y \in \{+, -, \mathtt{f}\}$ is distributed according to η, and η is distributed according to μ: formally, for any Borel set $B \subseteq \mathbb{R}$ (where it suffices to consider $B \subseteq \{\beta, -\beta, 0\}$),

$$\eta(B) = \int \eta(y) \mathbf{1}\left\{\rho_y \in B\right\} d\mu(\eta).$$

Thus $\eta \in \mathscr{M}_1$ is an η-measurable random measure supported on $\{\rho_+, \rho_-, \rho_{\mathtt{f}}\} = \{\beta, -\beta, 0\} \subseteq \mathbb{R}$. Let ζ denote the law of η, so $\zeta \in \mathscr{M}_2$ and the randomness in ζ is the randomness of η. For $x \in \{-, +\}$ and $\underline{y}_a \in \{+, -, \mathtt{f}\}^{k-1}$, let

$$u_a(x|\underline{y}_a) \equiv \sum_{\underline{x} \in \{\pm\}^k} \frac{\mathbf{1}\{x_k = x\}}{\exp(\beta \theta_a(\underline{x}))} \prod_{j=1}^{k-1} r(x_j|y_{a,j}), \quad r(x|y) \equiv \frac{\exp\{\rho_y x\}}{2 \operatorname{ch} \rho_y}.$$

Note it follows from the definition (4.59) that

$$r(+|+) = r(-|-) = \frac{e^\beta}{2 \operatorname{ch} \beta} = 1 - r(-|+) = 1 - r(+|-).$$

It follows by combining with the definition of θ_a that

$$u_a\left(-\mathtt{L}_{a,k} \,\middle|\, (-\mathtt{L}_{a,j})_{1 \le j \le k-1}\right) = \frac{1}{e^\beta} \prod_{j=1}^{k-1} r(-\mathtt{L}_{a,j}|-\mathtt{L}_{a,j}) + \left\{1 - \prod_{j=1}^{k-1} r(-\mathtt{L}_{a,j}|-\mathtt{L}_{a,j})\right\}$$

$$= \frac{1}{e^\beta}\left(\frac{e^\beta}{2\operatorname{ch}\beta}\right)^{k-1} + \left\{1 - \left(\frac{e^\beta}{2\operatorname{ch}\beta}\right)^{k-1}\right\} \le \frac{1}{e^{\beta/2}},$$

where the last bound holds assuming k is large enough and $\beta \ge k$. Thus, with $\mathbb{P}'$ denoting the law of ρ conditioned on $d, \underline{\theta}, \underline{\eta}$, for $x \in \{\pm\}$ we have

$$\mathbb{P}'\left(\underbrace{\prod_{a=1}^d u_a(x) > \frac{1}{e^{\beta/2}}}_{\text{event } F_x}\right) \le \Pi^x \equiv \prod_{a : \mathtt{L}_{a,k} = x} \left(1 - \prod_{j=1}^{k-1} \eta_{a,j}(-\mathtt{L}_{a,j})\right).$$

Since F_x can be decided by looking only at the clauses $\partial v(x)$, we have $\mathbb{P}'(F_+ \cup F_-) \le 1 - (1 - \Pi^+)(1 - \Pi^-)$, so

$$\mathbb{E}'\left[\left(\sum_{x \in \{\pm\}} \prod_{a=1}^d u_a(x)\right)^m\right] \le 2^m \mathbb{P}'\left(F_+ \cup F_-\right) + \left(\frac{2}{e^{\beta/2}}\right)^m \le 2^m \left(\Pi^+ + \Pi^- - \Pi^+ \Pi^- + \frac{1}{e^{m\beta/2}}\right).$$

Similar considerations give

$$\mathbb{E}'\left[(u_a)^m\right] \ge \frac{1}{2^m} \mathbb{P}'\left(u_a \ge \frac{1}{2}\right) \ge \frac{1}{2^m}\left(1 - \prod_{j=1}^k \eta_{a,j}(-\mathtt{L}_{a,j})\right).$$

Combining the above bounds gives

$$m\Phi_{\text{1-RSB}}(\beta, \zeta, m) \le m4^k + \mathbb{E} \ln \frac{\Pi^+ + \Pi^- - \Pi^+ \Pi^- + e^{-m\beta/2}}{[1 - \prod_{j=1}^{k-1} \eta_j]^{(k-1)\alpha}}.$$

Taking $m = 1/\beta^{1/2}$, the right-hand side converges to the function $\Phi(\alpha)$ from (4.58) in the limit $\beta \to \infty$. Now assume $\alpha > \alpha_\star$, so that $\Phi(\alpha) < 0$ (of course, this uses Proposition 1.2). Then, for sufficiently large β, we will have

$$F_n(\beta) \le \Phi_{\text{1-RSB}}\left(\beta, \zeta, \frac{1}{\beta^{1/2}}\right) \le 4^k + \frac{\beta^{1/2}\Phi(\alpha)}{2} \le \frac{\beta^{1/2}\Phi(\alpha)}{4} < 0. \tag{4.60}$$

To conclude, recall that the log-partition function $\ln Z_n(\beta)$ is well concentrated about its expected value $nF_n(\beta)$ (take the Doob martingale of $\ln Z_n(\beta)$ with respect to the k-SAT clause-revealing filtration, and apply the Azuma–Hoeffding inequality). Thus (4.60) implies that with high probability $\ln Z_n(\beta)$ is negative, i.e., the SAT instance is unsatisfiable.   □

4.6. **First moment of judicious colorings.** In this subsection we complete the proof of Proposition 3.29 — recall the statement of the proposition is that

$$\mathbb{E}_{\mathscr{D}}Z \ge \exp\left\{n\Big[\Phi(\alpha) - o_R(1)\Big]\right\}$$

with high probability over $\mathscr{D}$, where $\Phi(\alpha)$ is the 1-RSB free energy (1.11). We pick up from the discussion of §3.5 where several preliminary calculations were done. In Corollary 3.37 we saw that

$$\mathbb{E}_{\mathscr{D}}Z = \exp\left\{n\Psi_{\mathscr{D}}(\star\omega) - o(n)\right\}$$

where $\Psi_{\mathscr{D}}(\star\omega)$ is defined by the constrained optimization problem (3.49). Below Corollary 3.37, we discussed that the natural guess for the limiting value of $\Psi_{\mathscr{D}}(\star\omega)$ (as $n \to \infty$ and $R \to \infty$) is $\Phi^{\text{col}}(\alpha)$, the Bethe free energy of the coloring model (Definition 3.40). We first verify that this coincides precisely with the 1-RSB free energy $\Phi(\alpha)$ of the original model:

Lemma 4.31. *For α in the regime (1.7), the 1-RSB free energy of the k-SAT model ($\Phi(\alpha)$ from (1.11)) coincides with the replica symmetric (Bethe) free energy of the coloring model ($\Phi^{\text{col}}(\alpha)$ from Definition 3.40).*

Proof. We will show that the form of $\Phi^{\text{col}}(\alpha)$ given by (3.62) and (3.63) is equivalent to $\Phi(\alpha)$. As in Definition 3.39, let η_j be i.i.d. copies of η, and let $\hat{u}$ as in (3.59). Let $\dot{q}_j \equiv \dot{q}(\eta_j)$ via the correspondence (3.60). Then the $\bar{z}$ in (3.62) can be re-written as

$$\bar{z} = \sum_\sigma \dot{q}_k(\sigma)\hat{q}(\sigma) = \dot{q}_k(\mathtt{r})\hat{q}(\mathtt{r}) + [1 - \dot{q}_k(\mathtt{r})]\hat{q}(\mathtt{y}) = \frac{1}{[3 - 2\hat{u}(+)][2 - \eta_k(-)]}\left(1 - \prod_{j=1}^{k}\eta_j(-)\right).$$

Similarly, the $\hat{z}$ in (3.62) can be re-written as

$$\hat{z} = \sum_{\underline{\sigma}} \hat{\varphi}(\underline{\sigma})\prod_{j=1}^{k}\dot{q}_j(\sigma_j) = \prod_{j=1}^{k}[1 - \dot{q}_j(\mathtt{r})] - \prod_{j=1}^{k}\dot{q}_j(\mathtt{y}) + \sum_{j=1}^{k}[\dot{q}_j(\mathtt{r}) - \dot{q}_j(\mathtt{c})]\prod_{l \ne j}\dot{q}_l(\mathtt{y})$$

$$= \prod_{j=1}^{k}[1 - \dot{q}_j(\mathtt{r})] - \prod_{j=1}^{k}\dot{q}_j(\mathtt{y}) = \left(1 - \prod_{j=1}^{k}\eta_j(-)\right)\Bigg/\left(\prod_{j=1}^{k}[2 - \eta_j(-)]\right).$$

For the $\dot{z}$ in (3.62), given $\underline{d} = (d^+, d^-)$ let us abbreviate

$$\underline{\sigma} = \left(\left(\sigma_i^+\right)_{1 \le i \le d^+}, \left(\sigma_i^-\right)_{1 \le i \le d^-}\right),$$

and let $\varphi(\underline{\sigma}; \underline{d})$ be the indicator that $\underline{\sigma}$ gives a valid coloring of a variable where the first d^+ incident edges are of $+$ sign while the remaining d^- incident edges are of $-$ sign. Then we can rewrite

$$\dot{z} = \sum_{\underline{\sigma}} \varphi(\underline{\sigma}; \underline{d}) \prod_{i=1}^{d^+} \hat{q}_i^+(\sigma_i^+) \prod_{i=1}^{d^-} \hat{q}_i^-(\sigma_i^-) = \prod_{i=1}^{d^+} \hat{q}_i^+(\mathrm{g}) \prod_{i=1}^{d^-} \hat{q}_i^-(\mathrm{g})$$

$$+ \left\{ \prod_{i=1}^{d^+} \hat{q}_i^+(\mathrm{p}) - \prod_{i=1}^{d^+} \hat{q}_i^+(\mathrm{b}) \right\} \prod_{i=1}^{d^-} \hat{q}_i^-(\mathrm{y}) + \left\{ \prod_{i=1}^{d^+} \hat{q}_i^+(\mathrm{p}) - \prod_{i=1}^{d^+} \hat{q}_i^+(\mathrm{b}) \right\} \prod_{i=1}^{d^-} \hat{q}_i^-(\mathrm{y})$$

$$= \left(\Pi^+ + \Pi^- - \Pi^+ \Pi^- \right) \Big/ \left\{ \prod_{i=1}^{d^+} [3 - 2\hat{u}_i^+(+)] \prod_{i=1}^{d^-} [3 - 2\hat{u}_i^-(+)] \right\}, \quad \hat{u}_i^{\pm}(+) \equiv \prod_{j=1}^{k-1} \eta_{ij}^{\pm}(\text{-}).$$

Let us write simply $\mathbb{E}$ for expectation with respect to the law of the η_i. Substituting the above into (3.62) gives

$$\Phi^{\text{col,v}}(\alpha) = \mathbb{E}\left[\ln\left(\Pi^+ + \Pi^- - \Pi^+ \Pi^- \right) \right] - k\alpha\, \mathbb{E}\left[\ln\left(3 - 2\hat{u}(+) \right) \right],$$

$$\alpha\, \Phi^{\text{col,cl}}(\alpha) = \alpha\, \mathbb{E}\left[\ln\left(1 - \prod_{j=1}^{k} \eta_j(\text{-}) \right) \right] - k\alpha\, \mathbb{E}\left[\ln\left(2 - \eta_1(\text{-}) \right) \right],$$

$$-k\alpha\, \Phi^{\text{col,e}}(\alpha) \equiv -k\alpha\, \mathbb{E}\left[\ln\left(1 - \prod_{j=1}^{k} \eta_j(\text{-}) \right) \right] + k\alpha\, \mathbb{E}\left[\ln\left(3 - 2\hat{u}(+) \right) \right] + k\alpha\, \mathbb{E}\left[\ln\left(2 - \eta_1(\text{-}) \right) \right].$$

Substituting these into (3.63) we see that $\Phi^{\text{col}}(\alpha)$ agrees with $\Phi(\alpha)$ from (1.11). $\qquad\square$

To conclude the proof of Proposition 3.29, in light of Corollary 3.37 it remains to relate $\Psi_{\mathscr{D}}(\star\omega)$ to $\Phi^{\text{col}}(\alpha)$. The value $\Psi_{\mathscr{D}}(\star\omega)$ is defined by the constrained optimization (3.49). As we already noted in the discussion following Corollary 3.37, the optimal $\hat{v}^{\text{op}}[\star\omega]$ takes a simple form, given explicitly by the $\star\hat{v}$ in Lemma 3.38. However, we also noted that the analogous $\star\hat{v}$ (see (3.58)) is **not** the same as $\hat{v}^{\text{op}}[\star\omega]$, simply because $\star\hat{v}$ is not consistent with the given marginals $\star\omega$. However, we do have the following:

Lemma 4.32. *Let ω be as specified by Definition 3.26. If L is a stable clause type (Definition 3.7), then there exists a measure $\hat{v}_L$ which is consistent with ω and satisfies, for all $r \geq (\ln k)^2$,*

$$\mathcal{H}(\hat{v}_L) \geq \mathcal{H}(\star\hat{v}_L) - \frac{1}{k^{r/2}}$$

Proof. Throughout this proof we fix a clause a of type L, and suppress these from the notation when possible. The idea is to start from the explicit measure $\hat{v}_\star \equiv \star\hat{v}_L$ of (3.58) — which has incorrect edge marginals as commented above — and make small iterative adjustments to achieve the required edge marginals $\star\pi$. Indeed, for all $e \in \delta a$, by the message-stability condition (3.12) on e, we have

$$\frac{\hat{v}_\star(\sigma_e = \sigma)}{\star\pi_e(\sigma)} = \left(\frac{\tilde{q}_e(\sigma)}{\sum_\tau \star\dot{q}_e(\tau)\tilde{q}_e(\tau)} \right) \Big/ \left(\frac{\star\hat{q}_e(\sigma)}{\sum_\tau \star\dot{q}_e(\tau)\star\hat{q}_e(\tau)} \right) = 1 + \frac{O(1)}{k^r} \tag{4.61}$$

for all $\sigma \in \{\mathrm{r}, \mathrm{y}, \mathrm{g}, \mathrm{b}\}$. This shows that the edge marginals of $\hat{v}_\star$ are only slightly off from the canonical ones $\star\pi$.

Step 1. Correct marginal proportions of red spins. We first reweight $\hat{v}_\star$ to produce a measure $\hat{v}_0$ which has the correct marginal proportions of red spins. To this end, let CY denote the event that $\underline{\sigma}_{\delta a}$ has no red spins. It follows by using (2.19) and (3.58) that for any $e' \neq e''$ in δa,

$$\hat{v}_\star(\mathrm{CY}) = 1 - \sum_{e \in \delta a} \hat{v}_\star(\sigma_e = \mathrm{r}) \geq \hat{v}_\star\left(\sigma_{e'} = \mathrm{c} = \sigma_{e'} \right) = \frac{\star\dot{q}_{e'}(\mathrm{c})\star\dot{q}_{e''}(\mathrm{c})}{\hat{z}_\star} \prod_{f \in \delta a \setminus \{e', e''\}} \star\dot{q}_f(\mathrm{c}, \mathrm{y}),$$

where $\hat{z}_\star$ is the normalizing constant in (3.58). It is given by

$$\hat{z}_\star = \prod_{e \in \delta a} {}_\star \dot{q}_e(\mathtt{c},\mathtt{y}) - \prod_{e \in \delta a} {}_\star \dot{q}_e(\mathtt{y}) + \overbrace{\left[\sum_{e \in \delta a} \Big({}_\star \dot{q}_e(\mathtt{r}) - {}_\star \dot{q}_e(\mathtt{c}) \Big) \prod_{e' \in \delta a \setminus e} {}_\star \dot{q}_{e'}(\mathtt{y}) \right]}^{\text{zero}} \le \prod_{e \in \delta a} {}_\star \dot{q}_e(\mathtt{c},\mathtt{y}), \tag{4.62}$$

where the term in square brackets vanishes by the correspondence (2.38). It also follows from (2.38), in combination with the message-stability condition (3.12), that

$$\frac{{}_\star \dot{q}_e(\mathtt{c})}{{}_\star \dot{q}_e(\mathtt{c},\mathtt{y})} = 1 - {}_\star \eta_e(-) \ge \frac{1}{2^r} \,.$$

Combining these bounds gives, assuming $r \ge (\ln k)^2$,

$$\hat{v}_\star(\mathtt{CY}) \ge \frac{{}_\star \dot{q}_{e'}(\mathtt{c})}{{}_\star \dot{q}_{e'}(\mathtt{c},\mathtt{y})} \cdot \frac{{}_\star \dot{q}_{e''}(\mathtt{c})}{{}_\star \dot{q}_{e''}(\mathtt{c},\mathtt{y})} \ge \frac{1}{4^r} \,. \tag{4.63}$$

On the other hand, for any measure $\hat{v}$ that is consistent with ${}_\star \pi_e$ for all $e \in \delta a$, we must have

$$\frac{\hat{v}(\mathtt{CY})}{\hat{v}_\star(\mathtt{CY})} - 1 = \frac{1}{\hat{v}_\star(\mathtt{CY})} \left(\sum_{e \in \delta a} \hat{v}_\star(\sigma_e = \mathtt{r}) - \sum_{e \in \delta a} {}_\star \pi(\mathtt{r}) \right) = O\!\left(\frac{4^r}{k^r} \right), \tag{4.64}$$

where the last inequality follows by combining (4.61) with (4.63). It follows by using (4.61) and (4.64) that

$$\gamma_e \equiv \left(\frac{{}_\star \pi_e(\mathtt{r})}{\hat{v}(\mathtt{CY})} \right) \Big/ \left(\frac{\hat{v}_\star(\sigma_e = \mathtt{r})}{\hat{v}_\star(\mathtt{CY})} \right) = 1 + O\!\left(\frac{4^r}{k^r} \right), \tag{4.65}$$

for all $e \in \delta a$, where $\hat{v}$ again stands for any measure that is consistent with ${}_\star \pi$. Then define

$$\hat{v}_0(\underline{\sigma}_{\delta a}) \equiv \frac{1}{\hat{z}_\circ} \hat{v}_\star(\underline{\sigma}_{\delta a}) \prod_{e \in \partial a} (\gamma_e)^{\mathbf{1}\{\sigma_e = \mathtt{r}\}} \,,$$

where $\hat{z}_\circ$ is the normalization constant that makes $\hat{v}_0$ a probability measure. We then calculate

$$\hat{z}_\circ = \hat{v}_\star(\mathtt{CY}) + \sum_{e \in \delta a} \hat{v}_\star(\sigma_e = \mathtt{r}) \gamma_e = \frac{\hat{v}_\star(\mathtt{CY})}{\hat{v}(\mathtt{CY})}$$

(using the definition of γ with some simple algebra). It follows that

$$\hat{v}_0(\sigma_e = \mathtt{r}) = \frac{\hat{v}_\star(\sigma_e = \mathtt{r}) \gamma_e}{\hat{z}_\circ} = {}_\star \pi_e(\mathtt{r}) \,,$$

for all $e \in \delta a$, so $\hat{v}_0$ has the correct marginal proportions of red spins, as desired.

Step 2. Estimates for measure conditional on no red spins. Note that the above reweighting does not change the measure conditional on $\mathtt{CY}$, that is to say, $\hat{\mu}_0(\cdot) \equiv \hat{v}_0(\cdot \,|\, \mathtt{CY})$ is the same as $\hat{v}_\star(\cdot \,|\, \mathtt{CY})$. It remains to correct the edge marginals conditional on $\mathtt{CY}$. If $\hat{v}$ is any measure that is consistent with ${}_\star \pi$, we must have for $\sigma \in \{\mathtt{c},\mathtt{y}\}$ that

$$\hat{\mu}\Big(\sigma_e = \sigma \Big) \equiv \hat{v}\Big(\sigma_e = \sigma \,\Big|\, \mathtt{CY} \Big) = \frac{1}{\hat{v}(\mathtt{CY})} \left({}_\star \pi_e(\sigma) - \mathbf{1}\{\sigma = \mathtt{y}\} \sum_{e' \in \delta a \setminus e} {}_\star \pi_{e'}(\mathtt{r}) \right) \le \frac{{}_\star \pi_e(\sigma)}{\hat{v}(\mathtt{CY})}$$

It follows using (4.61), (4.63), and (4.64) that, for $\sigma \in \{\mathtt{c},\mathtt{y}\}$,

$$\hat{\mu}_0\Big(\sigma_e = \sigma \Big) - \hat{\mu}\Big(\sigma_e = \sigma \Big) = \hat{v}_\star\Big(\sigma_e = \sigma \,\Big|\, \mathtt{CY} \Big) - \hat{v}\Big(\sigma_e = \sigma \,\Big|\, \mathtt{CY} \Big) = O\!\left(\frac{4^r}{k^r} \cdot \frac{{}_\star \pi_e(\sigma)}{\hat{v}_\star(\mathtt{CY})} \right) = O\!\left(\frac{16^r}{k^r} \right). \tag{4.66}$$

Now, with respect to a given coloring $\underline{\sigma}_{\delta a}$, let us say that an edge $e \in \delta a$ is $\mathtt{w} \equiv$ white if at least two of the other edges in the clause are cyan: that is, if

$$\left| \Big\{ e' \in \delta a \setminus e : \sigma_{e'} \in \{\mathtt{g},\mathtt{b}\} \Big\} \right| \ge 2 \,.$$

Again recalling (2.19), the interpretation of e being white is that σ_e can be freely flipped between c and y without invalidating the coloring around a. For each $e \in \delta a$, we can choose any $e' \neq e''$ in $\delta a \setminus e$ and lower bound

$$W_e(\sigma) \equiv \frac{\mu_0(\sigma_e = \sigma; e \text{ is white})}{{}_\star\pi_e(\sigma)/\hat{v}_\star(\text{CY})} \geq \frac{{}_\star\dot{q}_e(\sigma){}_\star\dot{q}_{e'}(\text{c}){}_\star\dot{q}_{e''}(\text{c})}{{}_\star\pi_e(\sigma)\hat{z}_\star} \prod_{f \in \delta a \setminus \{e,e',e''\}} {}_\star\dot{q}_f(\text{c},\text{y}) .$$

By a similar calculation as in (4.62), we have

$$ {}_\star\pi_e(\sigma)\hat{z}_\star \leq {}_\star\dot{q}_e(\sigma) \prod_{f \in \delta a \setminus e} {}_\star\dot{q}_f(\text{c},\text{y}) .$$

It follows that, for $\sigma \in \{\text{c},\text{y}\}$ and any choice of three distinct edges $e, e', e'' \in \delta a$, we have

$$W_e(\sigma) \geq \frac{{}_\star\dot{q}_{e'}(\text{c})}{{}_\star\dot{q}_{e'}(\text{c},\text{y})} \cdot \frac{{}_\star\dot{q}_{e''}(\text{c})}{{}_\star\dot{q}_{e''}(\text{c},\text{y})} \geq \frac{1}{4^r} . \tag{4.67}$$

Recalling the definition of $W_e(\sigma)$ above, (4.67) says that $\mu_0(\sigma_e = \sigma; e \text{ is white})$ is large relative to (4.66).

Step 3. Correct marginals given no red spins. We can now define a sequence of measures $\hat{\mu}_0, \ldots, \hat{\mu}_{|\delta a|}$ where $\hat{\mu}_0$ is as defined above, and $\hat{\mu}_j$ will be consistent with ${}_\star\pi$ on the first j edges in δa. Given $\hat{\mu}_0$, we look at the first edge and define $\hat{\mu}_1$ to satisfying the following constraints:

(a) For any $\underline{\sigma}_{-1}$ that contains only one c spin, $\hat{\mu}_1(\text{c}, \underline{\sigma}_{-1}) = \hat{\mu}_0(\text{c}, \underline{\sigma}_{-1})$;
(b) For any $\underline{\sigma}_{-1}$ that contains at least two c spins, $\hat{\mu}_1(\text{c}, \underline{\sigma}_{-1}) + \hat{\mu}_1(\text{y}, \underline{\sigma}_{-1}) = \hat{\mu}_0(\text{c}, \underline{\sigma}_{-1}) + \hat{\mu}_0(\text{y}, \underline{\sigma}_{-1})$;
(c) The appropriate amount of mass is shifted between $\{\sigma_1 = \text{c}\}$ and $\{\sigma_1 = \text{y}\}$:

$$\hat{\mu}_1\left(\sigma_{e_1} = \sigma; e_1 \text{ is white}\right) = \hat{\mu}_0\left(\sigma_{e_1} = \sigma; e_1 \text{ is white}\right) + \left\{\hat{\mu}\left(\sigma_e = \sigma\right) - \hat{\mu}_0\left(\sigma_e = \sigma\right)\right\}.$$

Constraints (a) and (b) ensure that the marginals on the other edges are not affected. Finally, constraint (c) ensures that $\hat{\mu}_1$ will have marginals consistent with ${}_\star\pi$ on the first edge in δa, and it is feasible thanks to the estimates (4.66) and (4.67) from above. We can repeat the procedure with the remaining edges in δa to arrive at $\hat{\mu}_{|\delta a|} \equiv \hat{\mu}$. From the construction and from (4.66) we have the very crude bound

$$\left|\hat{\mu}(\underline{\sigma}_{\delta a}) - \hat{\mu}_0(\underline{\sigma}_{\delta a})\right| \leq \sum_{e \in \delta a} \left|\hat{\mu}_0\left(\sigma_e = \sigma\right) - \hat{\mu}\left(\sigma_e = \sigma\right)\right| \leq O\left(\frac{16^r k}{k^r}\right), \tag{4.68}$$

for all $\underline{\sigma}_{\delta a} \in \text{CY}$. We then finally define a measure on all valid colorings of δa by

$$\hat{v}(\underline{\sigma}_{\delta a}) \equiv \hat{v}_0\left(\underline{\sigma}_{\delta a}; \underline{\sigma}_{\delta a} \notin \text{CY}\right) + \hat{v}_0(\text{CY})\hat{\mu}(\underline{\sigma}_{\delta a}) .$$

By construction, $\hat{v}$ is consistent with the marginals ${}_\star\pi$, and satisfies

$$\left|\hat{v}(\underline{\sigma}_{\delta a}) - \hat{v}_\star(\underline{\sigma}_{\delta a})\right| \leq O\left(\frac{16^r k}{k^r}\right)$$

for all $\underline{\sigma}_{\delta a}$, by combining (4.68) with our earlier estimate (4.65) on the weights γ_e. This gives

$$\mathcal{H}(\hat{v}) \geq \mathcal{H}(\hat{v}_\star) - \frac{1}{k^{r/2}} ,$$

concluding the proof of the lemma. $\qquad\square$

Proof of Proposition 3.29. We have from Corollary 3.37 that

$$\mathbb{E}_{\mathscr{D}}Z = \exp\left\{n\Psi_{\mathscr{D}}({}_\star\omega) + o(n)\right\},$$

where $\Psi_{\mathscr{D}}$ is defined by (3.49). Let $v \equiv ({}_\star\dot{v}, \hat{v})$ for ${}_\star\dot{v}$ as in Lemma 3.38, and $\hat{v}$ as given by Lemma 4.32. Then, by construction, v is consistent with marginals ${}_\star\omega$, so $Z \geq Z[v]$. It follows by Lemma 3.36 that

$$\mathbb{E}_{\mathscr{D}}Z \geq \mathbb{E}_{\mathscr{D}}Z[v] = \exp\left\{n\Phi_{\mathscr{D}}(v) + o(n)\right\},$$

for $\boldsymbol{\Phi}_{\mathcal{D}}(v)$ as defined by (3.42). Now let $_\star v \equiv (_\star \dot{v}, _\star \hat{v})$ for $_\star \hat{v}$ as defined by (3.58). Then Lemma 4.32 gives

$$\boldsymbol{\Phi}_{\mathcal{D}}(v) \geq \boldsymbol{\Phi}_{\mathcal{D}}(_\star v) - \frac{1}{k^{r/3}} \, .$$

In the limit $n \to \infty$ followed by $R \to \infty$, the R-neighborhood empirical measure $\mathcal{D}$ concentrates around the limiting Galton–Watson measure, using the fact that an $o_R(1)$ fraction of variables are removed by preprocessing (Proposition 3.22). It then follows by the stability results (Lemma 4.27) that $_\star v$ converges in the distributional sense to the measures on the limiting random tree. It follows that

$$\lim_{R\to\infty} \left[\lim_{n\to\infty} \boldsymbol{\Phi}_{\mathcal{D}}(_\star v) \right] = \Phi^{\mathrm{col}}(\alpha)$$

in probability, and the result follows since $\Phi^{\mathrm{col}}(\alpha) = \Phi(\alpha)$ by Lemma 4.31. □

5. Analysis of preprocessing

In this section we prove Propositions 3.22–3.24. The section is organized as follows:

- In §5.1 we bound the occurrence of **defective** variables (Definition 3.10) in the **PGW** random tree (Proposition 5.1). We already have Proposition 4.15 from Section 4 which bounds the probability that a variable fails to be **nice**. The main task of §5.1 is to control the bootstrap percolation of non-nice variables that creates defective variables.

- In §5.2 we continue to analyze local properties of variables in the **PGW** random tree. To this end, recall that the **self-contained** (Definition 3.11) and **orderly** (Definition 3.12) properties both refer to being near relatively few defects (although each property captures a different notion of "relatively few"). In §5.2 we apply the estimates of §5.1 to bound the occurrence of variables that fail to be self-contained or orderly (Propositions 5.6 and 5.7). An immediate consequence of these two propositions is that most variables are **perfect**, i.e., both self-contained and orderly. Now recall from Definition 3.13 that a variable is **fair** if it is stable, it satisfies a loose bound on the volume of its local neighborhood, and it does not lie near many imperfect variables — thus we can deduce from the combination of Propositions 4.17, 5.6 and 5.7 that most variables are fair (Corollary 5.8). In fact we can further deduce (Corollary 5.9) that most variables are **excellent** — from Definition 3.14, excellent roughly means fair in a strong sense. Since **good** is weaker than excellent, it immediately implies that most variables are good.

- In §5.3 and §5.4 we prove Proposition 3.22, which controls the set of variables $\mathrm{BSP}'(A; \mathcal{G})$ that is removed by the preprocessing algorithm (Definition 3.15). Recall A is the set of variables in $\mathcal{G}$ that are improper (Definition 3.1) or not 1-good (Definition 3.14) — it is easy to bound the occurrence of improper variables, and the bound on variables that are not 1-good comes directly from Corollary 5.9 in §5.2. Thus we already know from §5.2 that variables are unlikely to be in A, and the task of §5.3 and §5.4 is to deduce a bound on $\mathrm{BSP}'(A; \mathcal{G})$. To this end, in §5.3 we prove a deterministic result (Proposition 5.14), saying that any large connected component of $\mathrm{BSP}'(A; \mathcal{G})$ must have certain characteristics. In §5.4 we show that those characteristics rarely occur in the random graph, thereby concluding the proof of Proposition 3.22.

- In §5.5 and §5.6 we prove Proposition 3.23, which says that, with high probability over the random graph $\mathcal{G}$, each surviving total type occurs linearly many times in the processed graph $\mathrm{pr}\,\mathcal{G}$. To this end, in §5.5 we prove a deterministic result (Corollary 5.27) which says that if the local neighborhood of a clause a in $\mathcal{G}$ satisfies certain properties, then a will have a particular total type L in the processed graph $\mathrm{pr}\,\mathcal{G}$. Since the conditions to have any given total type L are essentially local, in §5.6 we are able to argue that the total number of clauses satisfying the local conditions for type L concentrates around its mean, which grows linearly in n.

- Lastly, in §5.7 we prove Proposition 3.24 which says that the processed graph $\mathrm{pr}\,\mathcal{G}$ is uniform given its processed neighborhood profile $\mathcal{D}$, in the sense discussed in Remark 3.21 .

5.1. Bootstrap percolation of defects.
Recall from Definition 4.1 that $\mathcal{T}$ denotes a **PGW** tree, rooted at a variable v_{RT}. For $\mathfrak{i} \in \{0, 1\}$ we can use Definitions 3.9 and 3.10 to define the following subsets of $\mathcal{T}$:

$$D^{*,\mathfrak{i}} \equiv D^{*,\mathfrak{i}}(\mathcal{T}) \equiv \{\text{variables in } \mathcal{T} \text{ that are not } \mathfrak{i}\text{-nice}\}$$
$$\subseteq D^{\kappa_*,\mathfrak{i}} \equiv D^{\kappa_*,\mathfrak{i}}(\mathcal{T}) \equiv \{\text{variables in } \mathcal{T} \text{ within distance } \kappa_* \text{ of } D^{*,\mathfrak{i}}\}$$
$$\subseteq \mathrm{DEF}^{\mathfrak{i}} \equiv \mathrm{DEF}^{\mathfrak{i}}(\mathcal{T}) \equiv \{\mathfrak{i}\text{-defective variables in } \mathcal{T}\}.$$

The main goal of this subsection is to bound the occurrence of the i-defective set $\mathsf{DEF}^{\mathsf{i}}$ in $\mathscr{T}$. More precisely, we will bound the intersection of $\mathsf{DEF}^{\mathsf{i}}$ with **sparse subtrees** of $\mathscr{T}$. Let

$$\Lambda_{\mathfrak{c},\ell} \equiv \left\{ \begin{array}{c} \text{bipartite factor trees } T \equiv (V_T, F_T, E_T), \text{ rooted at a variable } v_{\mathrm{RT}}, \\ \text{with maximum degree at most } \mathfrak{c}, \text{ and number of variables } |V_T| = \ell \end{array} \right\}. \tag{5.1}$$

For any variable-rooted tree $\mathscr{T}$ we let

$$\Lambda_{\mathfrak{c},\ell}(\mathscr{T}) \equiv \left\{ \begin{array}{c} \text{bipartite factor trees } T \equiv (V_T, F_T, E_T) \text{ with } v_{\mathrm{RT}} \in T \subseteq \mathscr{T}, \\ \text{maximum degree at most } \mathfrak{c}, \text{ and number of variables } |V_T| = \ell \end{array} \right\}. \tag{5.2}$$

Without loss of generality we always assume $\mathfrak{c} \geq 2$. When considering $T \subseteq \mathscr{T}$, for any vertex $x \in \mathscr{T}$ we will always use ∂x to denote the immediate neighbors of x in $\mathscr{T}$, and $N(x) \equiv \partial_1 x$ the set of vertices at unit distance from x. We then denote

$$\partial_T x \equiv T \cap \partial x,$$
$$N_T(x) \equiv T \cap N(x).$$

If $T \in \Lambda_{\mathfrak{c},\ell}(\mathscr{T})$, then by definition we have $|\partial_T x| \leq \mathfrak{c}$ for all $x \in V_T \cup F_T$, so $N_T(x) \leq \mathfrak{c}(\mathfrak{c}-1) \leq \mathfrak{c}^2$. The main result of this subsection is the following:

Proposition 5.1. *Let $\mathsf{i} \in \{0, 1\}$, and suppose $\mathfrak{c} \geq 2$ is upper bounded by an absolute constant. Then, as long as k exceeds an absolute constant k_0 (depending only on the bound on $\mathfrak{c}$), we have the bound*

$$\mathbf{PGW}\left(|\mathsf{DEF}^{\mathsf{i}} \cap V_T| \geq \ell\epsilon \text{ for some } T \in \Lambda_{\mathfrak{c},\ell} \right) \leq \frac{\exp(k^2\ell)}{\exp(2^{k/4}\ell\epsilon/k^2)}$$

for all $\epsilon \geq 0$ and all integer $\ell \geq 0$.

Lemma 5.2. *Let $\mathsf{i} \in \{0, 1\}$, and suppose $\mathfrak{c} \geq 2$ is upper bounded by an absolute constant. Then, as long as k exceeds an absolute constant k_0 (depending only on the bound on $\mathfrak{c}$), we have the bound*

$$\mathbf{PGW}\left(|D^{\kappa_*,\mathsf{i}} \cap V_T| \geq \ell\epsilon \text{ for some } T \in \Lambda_{\mathfrak{c},\ell} \right) \leq \frac{\exp(k^2\ell)}{\exp(2^{k/4}\ell\epsilon)}$$

for all $\epsilon > 0$ and all integer-valued $\ell \geq 0$.

Proof. We divide the proof into three steps below. In the first step, we show that if some sparse subtree $T \subseteq \mathscr{T}$ has a nontrivial intersection with $D^{\kappa_*,\mathsf{i}}$, then there must exist $T \subseteq T' \subseteq \mathscr{T}$ such that T' is also sparse, and has a nontrivial intersection with $D^{*,\mathsf{i}}$ — this step is deterministic. In the second step, we explain a general identity concerning the expected number of subtrees of $\mathscr{T} \sim \mathbf{PGW}$ satisfying a given property. In the third step we bound this quantity to show that T' is unlikely to occur. If $\ell\epsilon = 0$ the bound is vacuous, so we assume without loss that $\ell\epsilon \geq 1$.

Step 1. Construction of T'. Take any $T \in \Lambda_{\mathfrak{c},\ell}(\mathscr{T})$ such that $A = D^{\kappa_*,\mathsf{i}} \cap V_T$ has size $|A| \geq \ell\epsilon$. For each variable $u \in A$, the number of variables in the $(4\kappa_* - 1)$-neighborhood of u in T is at most

$$|B_{4\kappa_*-1}(u) \cap V_T| = \sum_{\ell=0}^{4\kappa_*-1} \mathfrak{c}^{2\ell} = \frac{\mathfrak{c}^{8\kappa_*}}{\mathfrak{c}^2 - 1} \leq \mathfrak{c}^{8\kappa_*}.$$

Therefore we can extract $A' \subseteq A$ of size $|A'| = \lceil \ell\epsilon/\mathfrak{c}^{8\kappa_*} \rceil$ such that variables in A' lie at pairwise distance at least $4\kappa_*$. By definition, since $A' \subseteq A \subseteq D^{\kappa_*,\mathsf{i}}$, every variable in A' lies within distance κ_* of $D^{*,\mathsf{i}}$. For each $u \in A'$, let $g(u)$ be any of the variables in $D^{*,\mathsf{i}}$ which lies closest to u, and let $\gamma(u)$ be the path between u and $g(u)$. The paths $\gamma(u)$, $u \in A'$, are of length at most κ_*, and are mutually disjoint. Let T' be the union of T with the paths $\gamma(u)$. Then we have $T' \in \Lambda_{\ell',\mathfrak{c}+1}(\mathscr{T})$ for

$$\ell \leq \ell' \leq \ell + \left\lceil \frac{\ell\epsilon}{\mathfrak{c}^{8\kappa_*}} \right\rceil (\kappa_* - 1) \leq \ell\kappa_*.$$

Let $B' \equiv g(A') \subseteq D^{*,\mathsf{i}}$, and note that $g : A' \to B'$ is a one-to-one mapping. We have thus shown that if there exists $T \in \Lambda_{\mathfrak{c},\ell}(\mathscr{T})$ with at least $\ell\epsilon$ variables in $D^{\kappa_*,\mathsf{i}}$, then there must exist $T' \in \Lambda_{\mathfrak{c}+1,\ell'}(\mathscr{T})$, for $\ell \leq \ell' \leq \ell\kappa_*$, and a subset $B' \subseteq D^{*,\mathsf{i}} \cap V_{T'}$ such that $B' = \lceil \ell\epsilon/\mathfrak{c}^{8\kappa_*} \rceil$, and variables in B' lie at pairwise distance at least $2\kappa_*$.

Step 2. Expected number of embedded subtrees. Now fix any tree $T' \in \Lambda_{c+1,\ell'}$ for $\ell \leq \ell' \leq \ell\kappa_*$, and a subset B' of variables in T' such that $|B'| = \lceil \ell\epsilon/c^{8\kappa_*} \rceil$, and variables in B' lie at pairwise distance at least $2\kappa_*$. For this fixed pair (T', B') we will bound the probability

$$P(T', B') \equiv \mathbf{PGW}\left(\text{there exists an embedding } \zeta : T' \hookrightarrow \mathscr{T} \text{ such that } \zeta(B') \subseteq D^{*,\mathfrak{i}}(\mathscr{T}) \right)$$

$$\leq \int \left| \left\{ \text{embeddings } \zeta : T' \hookrightarrow \mathscr{T} \text{ such that } \zeta(B') \subseteq D^{*,\mathfrak{i}}(\mathscr{T}) \right\} \right| d\mathbf{PGW}(\mathscr{T}) . \tag{5.3}$$

(Later we will enumerate over (T', B').) To calculate (5.3), it is useful to take the following (equivalent) view of the random tree $\mathscr{T} \sim \mathbf{PGW}$: to each variable v we attach $D(v)$ which is an independently sampled Poisson point process of unit rate on the interval $[0, \alpha k]$. The atoms of $D(v)$ correspond to the child clauses of v, and we write them in ascending order as

$$D(v) = \left(a_i(v) : 1 \leq i \leq |D(v)| \right), \quad 0 \leq a_1(v) \leq \ldots \leq a_{|D(v)|}(v) \leq \alpha k .$$

Meanwhile, to each clause a we attach $D(a) = \{1, \ldots, k-1\}$ to indicate the $k-1$ child variables of a. In this view, an embedding $\zeta : T' \hookrightarrow \mathscr{T}$ can be specified by giving $D_{T'}(v) \subseteq D(v)$ and $D_{T'}(a) \subseteq D(a)$ for the variables and clauses of T'. Moreover, if D is a Poisson point process of unit rate on $[0, \alpha k]$, and $D_\circ$ is a finite subset of $[0, \alpha k]$, then (by standard properties of Poisson point processes) the law of D conditioned to contain $D_\circ$ is the same as the law of $D \cup D_\circ$. The volume (under Lebesgue measure) of all possible $D_{T'}(v)$ is given by

$$\left| \left\{ \left(a_i : 1 \leq i \leq |D_{T'}(v)| \right) : 0 \leq a_1 \leq \ldots \leq a_{|D_{T'}(v)|} \leq \alpha k \right\} \right| = \frac{(\alpha k)^{|D_{T'}(v)|}}{|D_{T'}(v)|!} .$$

The volume (under counting measure) of all possible $D_{T'}(a)$ is given by

$$\left| \left(v_j : 1 \leq j \leq D_{T'}(a) \right) : v_j \text{ integer-valued}, 1 \leq v_1 \leq \ldots \leq v_{|D_{T'}(a)|} \leq k \right| = \binom{k}{|D_{T'}(a)|} .$$

It follows that the expected volume of all embeddings $\zeta : T' \to \mathscr{T}$ is

$$\mathrm{emb}(T') \equiv \int \left| \left\{ \text{embeddings } \zeta : T' \hookrightarrow \mathscr{T} \right\} \right| d\mathbf{PGW}(\mathscr{T})$$

$$= \prod_{v \in V_{T'}} \frac{(\alpha k)^{|D_{T'}(v)|}}{|D_{T'}(v)|!} \prod_{a \in F_{T'}} \binom{k}{|D_{T'}(a)|} \leq (\alpha k)^{|F_{T'}|} k^{|V_{T'}|-1} \leq (\alpha k^2)^{\ell' c} . \tag{5.4}$$

The law of $\mathscr{T}$ conditioned on any such embedding is equivalent to its law under the measure $\mathbf{PGW}(T')$ described in Definition 4.5. The expectation (5.3) is then equal to

$$\mathrm{emb}(T') \cdot \left[\mathbf{PGW}(T') \right]\left(B' \subseteq D^{*,\mathfrak{i}} \right) \tag{5.5}$$

where $\mathrm{emb}(T') \leq (\alpha k^2)^{\ell' c} \leq (\alpha k^2)^{\ell\kappa_* c}$.

Step 3. Probability bound for T'. We now bound the last term in (5.5), recalling that $T' \in \Lambda_{c+1,\ell'}$ and $B' \subseteq V_{T'}$ with $|B'| = \lceil \ell\epsilon/c^{8\kappa_*} \rceil$. The main difficulty is that the neighborhoods $B_R(u) \subseteq \mathscr{T}$, for $u \in B'$, need not be disjoint, since we only ensure that variables in B' have pairwise distance at least $2\kappa_*$ for κ_* an absolute constant. This can be addressed by a simple modification: for $u \in B'$, consider $B_R(u)$ as a tree rooted as u. In this u-rooted tree, take each $w \in T' \cap \partial_2 u$ and delete the subtree descended from it (without deleting w itself). Denote the result $B_R(u)^- \subseteq B_R(u)$. The modified neighborhoods $B_R(u)^-$, for $u \in B'$, are mutually disjoint. Define the depth-two subtree $T_u \equiv T' \cap B_2(u)$, and note that $B_R(u)^-$ has the same law as the random tree $\mathscr{T}_R(u)^-$ defined as follows: first sample $\mathscr{T}' \sim \mathbf{PGW}(T_u)$ (rooted at u), then form $\mathscr{T}''$ by deleting from $\mathscr{T}'$ the subtrees of depth-two variables in T_u, then form $\mathscr{T}_R(u)^-$ by taking the R-neighborhood of the root of $\mathscr{T}''$. If the root of $\mathscr{T}'$ is $(c+1)^4$-robust, then $\mathscr{T}_R(u)^-$ must be 1-nice. It follows that

$$\left[\mathbf{PGW}(T') \right]\left(B' \subseteq D^{*,\mathfrak{i}} \right) \leq \prod_{u \in B'} \left[\mathbf{PGW}(T_u) \right]\left(\text{root of } \mathscr{T}' \text{ is not } (c+1)^4\text{-robust} \right) \leq \frac{1}{\exp(\Omega(|B'|k2^{k/4}))} \tag{5.6}$$

— in the intermediate expression we use $\mathscr{T}'$ to denote a random variable with law $\mathbf{PGW}(T_u)$, and the last bound is by Proposition 4.15. We can substitute this last bound into (5.5), which we recall is equal to (5.3). Enumerating over all pairs (T', B') (such that $T' \in \Lambda_{c+1,\ell}$ for $\ell \leq \ell' \leq \ell\kappa_*$, and $B' \subseteq V_{T'}$ with $|B'| = \lceil \ell\epsilon/c^{8\kappa_*} \rceil$) gives

$$\sum_{T',B'} \int \left| \left\{ \text{embeddings } \zeta : T' \hookrightarrow \mathscr{T} \text{ such that } \zeta(B') \subseteq D^{*,i}(\mathscr{T}) \right\} \right| d\mathbf{PGW}(\mathscr{T})$$

$$\leq \sum_{\ell \leq \ell' \leq \ell\kappa_*} \frac{(c+2)^{\ell'(c+2)} 2^{\ell'} (\alpha k^2)^{\ell'c}}{\exp(\Omega(\lceil \ell\epsilon/c^{8\kappa_*} \rceil \cdot k2^{k/4}))} \leq \frac{\exp(\ell\kappa_* ck)}{\exp(2^{k/4}\ell\epsilon)} ,$$

for k exceeding an absolute constant. The claimed bound follows. $\qquad\square$

Lemma 5.3. *Let $\mathscr{T}$ be any bipartite factor tree in which all clauses have degree at most k. Then*

$$|A| \geq \frac{|\mathrm{BSP}(A; \mathscr{T})|}{2k}$$

for any finite subset A of variables in $\mathscr{T}$.

Proof. Let $A_0 \equiv A$, and for $t \geq 1$ we let A_t be the union of A_{t-1} with all variables $u \in \mathscr{T}$ with $|N(u) \cap A_{t-1}| \geq 2$. Thus $A = A_0 \subseteq A_1 \subseteq \ldots$, and (as in (3.17)) we define

$$\mathrm{BSP}(A; \mathscr{T}) = \bigcup_{t \geq 0} A_t .$$

For all $t \geq 0$, let $G_t \equiv (A_t, F_t, E_t)$ be the subgraph of $\mathscr{T}$ induced by A_t — that is, F_t is the set of all clauses in $\mathscr{T}$ having at least two incident variables in A_t, and E_t is the set of all edges in $\mathscr{T}$ between A_t and F_t. As long as t is finite, G_t is a finite subgraph of $\mathscr{T}$. If we denote the (maximal) connected components of G_t by $G_{t,i} \equiv (A_{t,i}, F_{t,i}, E_{t,i})$ for $1 \leq i \leq i(t)$, and let $A_i \equiv A \cap G_{t,i}$, then we must have $\mathrm{BSP}(A_i; G_{t,i}) = A_{t,i}$. If we can show for all i that $|A_{t,i}| \leq 2k|A_i|$, then summing over i gives

$$|A_t| = \sum_{i=1}^{i(t)} |A_{t,i}| \leq 2k \sum_{i=1}^{i(t)} |A_i| = 2k|A| .$$

If the bound holds for all t, then the conclusion of the lemma follows. Since each $G_{t,i}$ is a finite tree, we conclude that it suffices to show the following special case of the lemma: for any finite bipartite factor tree $\mathscr{T} = (V, F, E)$ in which all clauses have width at most k, if $A \subseteq V$ such that $\mathrm{BSP}(A; \mathscr{T}) = V$, then we must have $|V| \leq 2k|A|$.

To show the last assertion, for all $v \in V$ let $\tau(v) \equiv \min\{j \geq 0 : v \in A_j\}$. For each edge $e = (av) \in E$, let

$$m(av) \equiv \min\left\{ 2, \left| \left\{ u \in \partial a : \tau(u) < \tau(v) \right\} \right| \right\} \in \{0, 1, 2\} .$$

Since $\mathrm{BSP}(A; \mathscr{T}) = V$, the definition of BSP implies that each variable $v \in V$ must have

$$\sum_{a \in \partial v} m(av) \geq 2 .$$

On the other hand, for each clause $a \in F$, if we reorder the variables in ∂a as $(v_1, \ldots, v_{|\partial a|})$ such that $\tau(v_i)$ is nondecreasing in i, then we must have $m(av_1) = 0$ and $m(av_2) \in \{0, 1\}$, so

$$\sum_{v \in \partial a} m(av) \leq 0 + 1 + (|\partial a| - 2)2 = 2|\partial a| - 3 .$$

Combining these inequalities gives

$$2|V \setminus A| \leq \sum_{v \in V} \sum_{a \in \partial v} m(av) = \sum_{a \in F} \sum_{v \in \partial a} m(av) \leq \sum_{a \in F} \left(2|\partial a| - 3 \right) = 2|E| - 3|F| .$$

Since $\mathscr{T} = (V, F, E)$ is by assumption a finite tree where all clauses have degree at most k, we must have

$$|V| - 1 = |E| - |F| = \sum_{a \in F} \left(|\partial a| - 1 \right) \leq (k-1)|F| .$$

Substituting into the previous gives

$$2|V \setminus A| \leq 2\Big(|V| - |1|\Big) - |F| \leq \Big(2 - \frac{1}{k-1}\Big)\Big(|V| - 1\Big).$$

Rearranging gives

$$|A| \geq 1 + \frac{|V| - 1}{2(k-1)} \geq \frac{|V|}{2k},$$

which proves the claim. The lemma follows as discussed above. □

Lemma 5.4. *Let $\mathscr{T}$ be any bipartite factor tree in which all clauses have degree at most k. For any $\mathfrak{c} \geq 2$ and $\ell \geq 1$, and any $T \in \Lambda_{\mathfrak{c},\ell}(\mathscr{T})$, there is a tree $T' \in \Lambda_{\mathfrak{c}',\ell'}(\mathscr{T})$, with $\mathfrak{c}' = \mathfrak{c} + 2$ and $\ell \leq \ell' \leq \ell(\mathfrak{c}')^{R'}$, such that*

$$|D^{\kappa_*,\mathfrak{i}} \cap V_{T'}| \geq \frac{1}{2k}\Big(|\mathsf{DEF}^{\mathfrak{i}} \cap V_T| + |V_{T'} \setminus V_T|\Big)$$

for $\mathfrak{i} \in \{0, 1\}$.

Proof. Recall Definition 3.10 that $v \in \mathsf{DEF}^{\mathfrak{i}}$ if and only if $v \in \mathsf{BSP}(D^{\kappa_*,\mathfrak{i}} \cap B_{R'/2}(v); B_{R'/2}(v))$. Let $\mathscr{T}' \equiv B_{R'/2}(T)$, that is, the union of the $(R'/2)$-neighborhoods of all variables in T. Let $A_0 \equiv D^{\kappa_*,\mathfrak{i}} \cap \mathscr{T}'$, and for $t \geq 1$ let A_t be the union of A_{t-1} with all variables in $\mathscr{T}'$ that have at least two neighbors in A_{t-1}. Then let

$$A_\infty \equiv \bigcup_{t \geq 0} A_t = \mathsf{BSP}(D^{\kappa_*,\mathfrak{i}} \cap \mathscr{T}'; \mathscr{T}').$$

Let G_t be the subgraph of $\mathscr{T}'$ induced by A_t. We will construct a sequence of trees $T_s \equiv (V_s, E_s, F_s) \subseteq G_\infty$ (terminating at T'), as follows. Let $T_0 \equiv T$. For all variables $u \in A_\infty$, let $\tau(u) \equiv \min\{t \geq 0 : u \in A_t\}$. For $s \geq 0$, let

$$U(T_s) \equiv \left\{ u \in V_s \;\middle|\; \begin{array}{l} 1 \leq \tau(u) < \infty, \text{ and } u \text{ does not have at least} \\ \text{two neighboring variables in } T_s \text{ with smaller } \tau \end{array} \right\}.$$

If $U(T_s) = \varnothing$, we terminate the process and set $T_s = T_\infty$. Otherwise, take any $u \in U(T_s)$: it must have at least one neighboring variable u' that does not lie in T_s and has $\tau(u') < \tau(u)$. Among all such u', choose one with minimal τ. Let T_{s+1} be the graph induced by $V_s \cup \{u'\}$. In the resulting $T_\infty \equiv T'$, any vertex that did not belong to T will have degree at most 3, while any vertex that did belong to T will have its degree in T' at most two larger than its degree in T. It follows that $T' \in \Lambda_{\mathfrak{c}',\ell'}(\mathscr{T})$, with $\mathfrak{c}' = \mathfrak{c} + 2$ and $\ell \leq \ell' \leq \ell(\mathfrak{c}')^{R'}$. It is straightforward to check (e.g. by induction) that the above construction implies

$$(\mathsf{DEF}^{\mathfrak{i}} \cap V_T) \cup (V_{T'} \setminus V_T) \subseteq \mathsf{BSP}(D^{\kappa_*,\mathfrak{i}} \cap V_{T'}; T').$$

Combining with Lemma 5.3 gives

$$\frac{1}{2k}\Big(|\mathsf{DEF}^{\mathfrak{i}} \cap V_T| + |V_{T'} \setminus V_T|\Big) \leq \frac{\mathsf{BSP}(D^{\kappa_*,\mathfrak{i}} \cap V_{T'}; T')}{2k} \leq |D^{\kappa_*,\mathfrak{i}} \cap V_{T'}|,$$

as claimed. □

Proof of Proposition 5.1. It follows from Lemma 5.4 that

$$\mathbf{PGW}\Big(|\mathsf{DEF}^{\mathfrak{i}} \cap V_T| \geq \ell\epsilon \text{ for some } T \in \Lambda_{\mathfrak{c},\ell}\Big)$$

$$\leq \sum_{\ell'=\ell}^{\lceil \ell(\mathfrak{c}')^{R'} \rceil} \mathbf{PGW}\Big(|D^{\kappa_*,\mathfrak{i}} \cap V_{T'}| \geq \frac{\ell\epsilon + (\ell' - \ell)}{2k} \text{ for some } T' \in \Lambda_{\mathfrak{c}',\ell'}\Big).$$

It follows from Lemma 5.2 that the last expression is upper bounded by

$$\sum_{\ell'=\ell}^{\lceil \ell(\mathfrak{c}')^{R'} \rceil} \frac{\exp(k^2[\ell + (\ell' - \ell)])}{\exp(2^{k/4}[\ell\epsilon + (\ell' - \ell)]/(2k))} \leq \frac{O(1)\exp(k^2\ell)}{\exp(\ell\epsilon 2^{k/4}/(2k))} \leq \frac{\exp(k^2\ell)}{\exp(\ell\epsilon 2^{k/4}/k^2)},$$

concluding the proof. □

Up to this point we have proved bounds for the $\mathbf{PGW}(T)$ measures (for instance, Proposition 4.15) by direct analysis — the basic intuition being that, when T is a sparse tree, the $\mathbf{PGW}(T)$ measure is not so different from the $\mathbf{PGW}$ measure. We now make this more precise by proving a general bound which allows us to more easily transfer bounds from $\mathbf{PGW}$ to $\mathbf{PGW}(T)$:

Lemma 5.5. *Let T be any fixed variable-rooted tree of maximum degree $\mathfrak{c}$, where $\mathfrak{c}$ is an absolute constant. Then the Radon–Nikodyn derivative between the measures $\mathbf{PGW}(T)$ and $\mathbf{PGW}$ satisfies the second moment bound*

$$\mathbb{E}\left[\left(\frac{d\mathbf{PGW}(T)}{d\mathbf{PGW}}\right)^2\right] \leq 2\,,$$

where $\mathbb{E}$ denotes expectation under $\mathbf{PGW}$. (The bound holds for $k \geq k_0$ where k_0 depends only on $\mathfrak{c}$.)

Proof. According to the original definitions, in a random tree $\mathscr{T} \sim \mathbf{PGW}$ or $\mathscr{T} \sim \mathbf{PGW}(T)$, only the root variable is distinguished, while all other variables are unlabelled. For the purposes of this proof, however, we now instead consider both $\mathscr{T} \sim \mathbf{PGW}$ and $\mathscr{T} \sim \mathbf{PGW}(T)$ as rooted *labelled* trees, where the children of each vertex are ordered uniformly at random. Separately, we also assume that T comes with a fixed labelling of its vertices, so that we may speak of embeddings $T \hookrightarrow \mathscr{T}$.

Step 1. Martingale of Radon–Nikodym derivatives. If $\mathbf{PGW}(T) \ll \mathbf{PGW}$, then the labellings discussed above do not affect the value of the Radon–Nikodym derivative, which we hereafter write as

$$\mathfrak{X} \equiv \frac{d\mathbf{PGW}(T)}{d\mathbf{PGW}}(\mathscr{T}) \equiv \mathfrak{X}(\mathscr{T};T)\,.$$

To show that $\mathbf{PGW}(T) \ll \mathbf{PGW}$ (so that $\mathfrak{X}$ is well-defined), and to obtain the bound on $\mathfrak{X}$ claimed in the statement of the lemma, we shall take the limit of depth $\ell \to \infty$. Let $T_\ell \equiv B_\ell(o;T)$; and let $\mathscr{T}_\ell \equiv B_\ell(v_{\mathrm{RT}};\mathscr{T})$ where v_{RT} is the root variable of $\mathscr{T}$. Let $\mathscr{F}_\ell$ be the σ-field generated by $\mathscr{T}_\ell$. Define the restricted measures

$$\mathbf{PGW} \equiv \mathbf{PGW}\Big|_{\mathscr{F}_\ell}\,, \quad \mathbf{PGW}_\ell(T) \equiv \mathbf{PGW}(T)\Big|_{\mathscr{F}_\ell} = \mathbf{PGW}_\ell(T_\ell)\,.$$

For finite ℓ it is clear that $\mathbf{PGW}_\ell(T) \ll \mathbf{PGW}_\ell$, so we can define the Radon–Nikodym derivative

$$\mathfrak{X}_\ell \equiv \frac{d\mathbf{PGW}_\ell(T_\ell)}{d\mathbf{PGW}_\ell} \equiv \mathfrak{X}_\ell(\mathscr{T}_\ell;T_\ell)\,,$$

which is a nonnegative martingale. We will show inductively that for all $\ell \geq 0$ we have

$$M(\ell) \equiv \max_{T_\ell} \mathbb{E}\left[\left(\mathfrak{X}_\ell(\mathscr{T}_\ell;T_\ell)\right)^2\right] \leq 2\,, \tag{5.7}$$

where the maximum is taken over all depth-ℓ trees T_ℓ of maximum degree $\mathfrak{c}$. This will imply that the martingale $\mathfrak{X}_\ell$ is bounded in L^2, so that (by the L^2 martingale convergence theorem) it converges almost surely to a finite limit $\mathfrak{X} \equiv \mathfrak{X}_\infty$ satisfying the same L^2 bound.

Step 2. L^2 bound on martingale. We now prove (5.7). The base case $\ell = 0$ holds trivially since $\mathfrak{X}_0 \equiv 1$, so let us suppose inductively that the bound holds up to depth ℓ. Suppose T has root degree t, meaning that the variable o has child clauses $a_1, \ldots, a_t$. Suppose a_i has $t(i)$ children in T, which we denote $v_{i,1}, \ldots, v_{i,t(i)}$. Let $\mathrm{emb}(T, \mathscr{T})$ denote the set of embeddings $T \hookrightarrow \mathscr{T}$ that map root to root, and note that at depth one we have

$$|\mathrm{emb}(T_1, \mathscr{T}_1)| = (d)_t \prod_{i=1}^{t}(k-1)_{t(i)}$$

where d is the root degree of $\mathscr{T}$. Let $\mathbf{E}$ denote expectation over $\mathscr{T}$ and over a uniformly random element ζ from $\mathrm{emb}(T_1, \mathscr{T}_1)$. We then have the recursion

$$\mathfrak{X}_{\ell+1}(\mathscr{T}_{\ell_1};T_{\ell+1}) = \frac{\mathrm{pois}_{\alpha k}(d-t)}{\mathrm{pois}_{\alpha k}(d)}\mathbf{E}\left\{\prod_{\substack{(i,j):1\leq i\leq t,\\ 1\leq j\leq t(i)}} \mathfrak{X}_\ell\Big(\mathscr{T}_\ell(\zeta(v_{ij}));T_{\ell+1}(v_{ij})\Big)\right\},$$

for a rooted tree U we write $U(x)$ for the subtree of U descended from vertex x, and $U_\ell(x) \equiv B_{\ell-1}(x; U)$ (we view $U(x)$ and $U_\ell(x)$ as being rooted at x). We can then express the second moment as

$$\mathbb{E}\Big[(\mathfrak{X}_{\ell+1})^2\Big] = \mathbb{E}\Bigg[\Big(\frac{(d)_t}{(\alpha k)^t}\Big)^2 \mathbf{E}\Big\{ \prod_{\substack{(i,j):1\leq i\leq t,\\ 1\leq j\leq t(i)}} \mathfrak{X}_\ell\Big(\mathscr{T}_\ell(\zeta(v_{ij})); T_{\ell+1}(v_{ij})\Big) \mathfrak{X}_\ell\Big(\mathscr{T}_\ell(\xi(v_{ij})); T_{\ell+1}(v_{ij})\Big) \Big| \mathscr{F}_1 \Big\}\Bigg]$$

where $\mathbf{E}$ now refers to expectation over $\mathscr{T}$ and over a uniformly random pair of elements ζ, ξ from $\mathrm{emb}(T_1, \mathscr{T}_1)$. Let $\jmath(\zeta,\xi)$ count the number of pairs (i,j) for which $\zeta(v_{ij}) = \xi(v_{ij})$, and note that

$$\mathbf{E}\Big\{ \prod_{\substack{(i,j):1\leq i\leq t,\\ 1\leq j\leq t(i)}} \mathfrak{X}_\ell\Big(\mathscr{T}_\ell(\zeta(v_{ij})); T_{\ell+1}(v_{ij})\Big) \mathfrak{X}_\ell\Big(\mathscr{T}_\ell(\xi(v_{ij})); T_{\ell+1}(v_{ij})\Big) \Big| \mathscr{F}_1 \Big\} \leq \mathbf{E}\Big[M(\ell)^{\jmath(\zeta,\xi)} \Big],$$

for $M(\ell)$ as in (5.7). Since T has maximum degree $\mathfrak{c}$, we can rather crudely bound $\imath(\zeta,\xi) \leq \mathfrak{c}a(\zeta,\xi)$ where $a(\zeta,\xi)$ counts the number of indices i for which $\zeta(a_i) = \xi(a_i)$. For any fixed ζ,

$$\frac{|\{\xi : a(\zeta,\xi) = s\}|}{|\mathrm{emb}(T_1, \mathscr{T}_1)|} = \frac{(d-t)_{t-s}(t)_s}{(d)_t}.$$

It follows from this that

$$\mathbb{E}\Big[(\mathfrak{X}_{\ell+1})^2\Big] \leq \mathbb{E}\Big\{\Big(\frac{(d)_t}{(\alpha k)^t}\Big)^2 \sum_{s=0}^{t} \frac{(d-t)_{t-s}(t)_s}{(d)_t} \mathbb{E}\big[(\mathfrak{X}_\ell)^2\big]^{\mathfrak{c}s}\Big\} = \frac{1}{(\alpha k)^{2t}} \sum_{s=0}^{t} (t)_s \mathbb{E}[(d)_{2t-s}] M(\ell)^{\mathfrak{c}s}.$$

Combining with the inductive hypothesis (5.7) gives

$$\mathbb{E}\Big[(\mathfrak{X}_{\ell+1})^2\Big] = \sum_{s=0}^{t} \frac{(t)_s}{(\alpha k)^s} M(\ell)^{\mathfrak{c}s} \leq \sum_{s\geq 0} \Big(\frac{2\mathfrak{c}}{\alpha k}\Big)^s \leq 2,$$

where the last bound holds for $k \geq k_0$, and verifies the induction. As noted above, it follows from the L^2 martingale convergence theorem that $\mathbf{PGW}(T) \ll \mathbf{PGW}$, with Radon–Nikodym derivative $\mathfrak{X}$ satisfying the claimed bound. □

5.2. **Orderliness and containment.** As above, let $\mathscr{T}$ denote a sample from the measure $\mathbf{PGW}$. This subsection is primarily occupied with the proofs of the following two propositions:

Proposition 5.6. *Let* $\mathfrak{i} \in \{0,1\}$, *and let* $\mathcal{S}^\mathfrak{i} \equiv \mathcal{S}^\mathfrak{i}(\mathscr{T})$ *denote the set of all variables in* $\mathscr{T}$ *that are not* $\mathfrak{i}$-*self-contained. There is an absolute constant* k_0 *(depending only on the absolute constant* δ_* *which appears in Definition 3.11) such that for all* $k \geq k_0$ *we have*

$$\mathbf{PGW}\left(\begin{array}{c} \mathscr{T} \text{ contains a path } P \text{ of } \ell \text{ variables,} \\ \text{emanating from } v_{\mathrm{RT}}, \text{ with } |\mathcal{S}^\mathfrak{i} \cap V_P| \geq \ell\epsilon \end{array} \right) \leq \frac{\exp(k^2 \ell)}{\exp\{2^{k\delta_*/3}\ell\epsilon\}},$$

for all $\epsilon \geq 0$ *and integer* $\ell \geq 0$.

Proposition 5.7. *Let* $\mathfrak{i} \in \{0,1\}$, *and let* $\mathcal{O}^\mathfrak{i} \equiv \mathcal{O}^\mathfrak{i}(\mathscr{T})$ *denote the set of all variables in* $\mathscr{T}$ *that are not* $\mathfrak{i}$-*orderly. There is an absolute constant* k_0 *(depending only on the absolute constant* δ_* *which appears in Definition 3.12) such that for all* $k \geq k_0$ *we have*

$$\mathbf{PGW}\left(\begin{array}{c} \mathscr{T} \text{ contains a path } P \text{ of } \ell \text{ variables,} \\ \text{emanating from } v_{\mathrm{RT}}, \text{ with } |\mathcal{O}^\mathfrak{i} \cap V_P| \geq \ell\epsilon \end{array} \right) \leq \frac{\exp(k^2 \ell)}{\exp(2^{k/4}\ell\epsilon/k^3)},$$

for all $\epsilon \geq 0$ *and integer* $\ell \geq 0$.

From these propositions, it will be fairly straightforward to deduce the next two corollaries, which are the main consequences from the analysis in this subsection:

Corollary 5.8. *For* $\mathfrak{i} \in \{0,1\}$, *recall what it means for a variable to be* $\mathfrak{i}$-*fair (Definition 3.13). For* δ_* *a sufficiently small positive absolute constant, we have for all* $k \geq k_0$ *(where* k_0 *is an absolute constant depending only on* δ_*) *that*

$$\mathbf{PGW}\Big(v_{\mathrm{RT}} \text{ not } \mathfrak{i}\text{-fair}\Big) \leq \frac{1}{\exp(2^{k\delta_*/3}R/k)}$$

for all R *as in* (3.1), *with* $r \geq 1$.

Corollary 5.9. *Recall what it means for a variable to be proper or improper (Definition 3.1). Further, for $\mathfrak{i} \in \{0, 1\}$, recall what it means for a variable to be $\mathfrak{i}$-excellent (Definition 3.14). We have*

$$\mathbf{PGW}\left(v_{\mathrm{RT}} \text{ is improper or not } \mathfrak{i}\text{-excellent}\right) \leq \frac{1}{\exp(2^{k\delta_*/4}R)}$$

for all $k \geq k_0$, where k_0 is an absolute constant (depending only on the absolute constant δ_*).

Proposition 5.7 is a straightforward consequence of the definitions together with Proposition 5.1, and we give its proof next. We then give the proof of Proposition 5.6, which is slightly more involved. Finally, at the end of this subsection we give the proofs for Corollaries 5.8 and 5.9.

Proof of Proposition 5.7. If $\ell\epsilon = 0$ the bound is vacuous, so we assume without loss that $\ell\epsilon \geq 1$. Suppose $\mathcal{T}$ contains a path P of ℓ variables emanating from v_{RT}, such that $S_0 \equiv \varnothing^{\mathfrak{i}} \cap V_P$ has size $|S_0| \geq \ell\epsilon$. For any $u \in S_0$ there is a path $\gamma(u) \subseteq \mathcal{T}$ that emanates from u, along which more than $(\delta_*)^3$ fraction of the variables are $\mathfrak{i}$-defective. We will define a sequence $S_0 \supseteq S_1 \supseteq \ldots$ as follows: as long as $S_{s-1} \neq \varnothing$, take

$$u_s \in \arg\max\left\{|\gamma(u)| : u \in S_{s-1}\right\}, \quad S_s \equiv S_{s-1} \setminus B_{2|\gamma(u_s)|+1}(u_s).$$

Eventually this terminates at $S_{s_{\max}} = \varnothing$, and the resulting sequence of paths $\gamma(u_s)$ will be mutually disjoint. Since S_0 was a subset of the path P, we have

$$\left|S_{s-1} \setminus S_s\right| \leq \left|V_P \cap B_{2|\gamma(u_s)|+1}(u_s)\right| \leq 4|V_{\gamma(u_s)}|.$$

Let U be the union of the paths $\gamma(u_s)$: then

$$|V_U| = \sum_{s=1}^{s_{\max}} |V_{\gamma(u_s)}| \geq \frac{1}{4}\sum_{s=1}^{s_{\max}} \left(|S_{s-1}| - |S_s|\right) = \frac{|S_0|}{4} \geq \frac{\ell\epsilon}{4}.$$

Let $T = P \cup U$. Then $T \in \Lambda_{3,\ell'}(\mathcal{T})$ for some $\ell' \geq \ell$, and $|\mathsf{DEF}^{\mathfrak{i}} \cap V_T| \geq |\mathsf{DEF}^{\mathfrak{i}} \cap V_U| \geq |V_U|(\delta_*)^3$. Note that since $|V_U| \geq \ell\epsilon/4$ and also $|V_T| \leq |V_P| + |V_U|$, we have

$$|V_U| \geq \frac{\ell\epsilon}{8} + \frac{|V_T| - |V_P|}{2} = \frac{\ell\epsilon}{8} + \frac{|V_T| - |V_P|}{2} = \frac{\ell\epsilon}{8} + \frac{\ell' - \ell}{2}.$$

It follows that

$$\mathbf{PGW}\left(\begin{array}{c} \mathcal{T} \text{ contains a path } P \text{ of } \ell \text{ variables,} \\ \text{emanating from } v_{\mathrm{RT}}, \text{ with } |\varnothing^{\mathfrak{i}} \cap V_P| \geq \ell\epsilon \end{array}\right)$$

$$\leq \sum_{\ell' \geq \ell} \mathbf{PGW}\left(\text{some } T \in \Lambda_{3,\ell'}(\mathcal{T}) \text{ has } |\mathsf{DEF}^{\mathfrak{i}} \cap V_T| \geq \left[\frac{\ell\epsilon}{8} + \frac{\ell' - \ell}{2}\right](\delta_*)^3\right).$$

It follows from Proposition 5.1 that the last expression is upper bounded by

$$\sum_{\ell' \geq \ell} \frac{\exp(k^2[\ell + (\ell' - \ell)])}{\exp((\delta_*)^3 2^{k/4}[\ell\epsilon/8 + (\ell' - \ell)/2]/k^2)} \leq \frac{O(1)\exp(k^2\ell)}{\exp((\delta_*)^3 2^{k/4}\ell\epsilon/(8k^2))},$$

and the result follows. $\qquad\square$

We now turn to the proof of Proposition 5.6. For the reader's convenience, we repeat (3.18) here:

$$\mathfrak{R}^{\mathfrak{i}}(v, t) \equiv \sum_{u : t \leq d(u,v) < 2R'} \frac{\exp\{k(\delta_*)^{-1}\mathfrak{B}^{\mathfrak{i}}(u, v)\}}{\exp\{(k\ln 2)(1 + \delta_*)d(u, v)\}},$$

for $1 \leq t \leq 2R'$. For any variable v, let

$$\Gamma^{\mathfrak{i}}(v) \equiv \left\{\begin{array}{c} \text{paths } P \text{ emanating from } v \text{ of length} \\ |P| \leq 2R' - 1 \text{ with } |\mathsf{DEF}^{\mathfrak{i}} \cap V_P| \geq |V_P|(\delta_*)^3 \end{array}\right\}. \tag{5.8}$$

(Recall that for a path P of length $|P| = \ell - 1$, the number of variables is $|V_P| = \ell$.) Let

$$\gamma^{\mathfrak{i}}(v) \equiv 1 + \max\left\{|P| : P \in \Gamma^{\mathfrak{i}}(v)\right\}.$$

By definition, every path in $\Gamma^{\mathrm{i}}(v)$ has length at most $2R' - 1$. If $\Gamma^{\mathrm{i}}(v)$ does not contain any (non-null) path, then we define $\gamma^{\mathrm{i}}(v) \equiv 1$. Therefore we always have $1 \leq \gamma^{\mathrm{i}}(v) \leq 2R'$. If $\gamma^{\mathrm{i}}(v) \leq d(u,v) \leq 2R' - 1$, then the shortest path between u and v must have less than $(\delta_*)^3$ fraction of i-defective variables. Define also

$$\bar{\gamma}(v) \equiv \min\left\{ t \geq 1 : |B_\ell(v)| \leq \exp\left\{ (k \ln 2)\left(1 + \frac{\delta_*}{2}\right)\ell \right\} \text{ for all } t \leq \ell < 2R' \right\}, \tag{5.9}$$

and note that $1 \leq \bar{\gamma}(v) \leq 2R'$. Recall the definition (3.19) of $\mathrm{rad}^{\mathrm{i}}(v)$. We claim that $\mathrm{rad}^{\mathrm{i}}(v)$ is upper bounded by

$$\chi^{\mathrm{i}}(v) \equiv \max\left\{ \gamma^{\mathrm{i}}(v), \bar{\gamma}(v) \right\}, \tag{5.10}$$

where $1 \leq \chi^{\mathrm{i}}(v) \leq 2R'$. Indeed, if $t \geq \chi^{\mathrm{i}}(v)$, then (by taking $k \geq k_0$ with k_0 an absolute constant) we have

$$\mathfrak{R}^{\mathrm{i}}(v, t) \leq \sum_{\ell=t}^{2R'-1} \frac{|B_\ell(v)| \exp\{ k(\delta_*)^{-1}(\delta_*)^3(1+\ell) \}}{\exp\{ (k \ln 2)(1 + \delta_*)\ell \}}$$

$$\leq \sum_{\ell=t}^{2R'-1} \frac{\exp\{ k(\delta_*)^2(1+\ell) \}}{\exp\{ (k \ln 2)(\delta_*/2)\ell \}} \leq \frac{O(1)\exp\{ 2k(\delta_*)^2 \}}{\exp\{ (k \ln 2)(\delta_*/2) \}} \leq \frac{1}{\exp(k\delta_*/3)} \leq \frac{1}{4} \, .$$

This shows that $\mathrm{rad}^{\mathrm{i}}(v) \leq \chi^{\mathrm{i}}(v)$, and we now turn to controlling $\chi^{\mathrm{i}}(v)$. The sketch for the proof of Proposition 5.6 is as follows: in Lemma 5.11 we show that if $\mathscr{T} \sim \mathbf{PGW}$, then it is very unlikely to have a sparse subtree $T \subseteq \mathscr{T}$ where many variables have a large value of $\bar{\gamma}(u)$. In Lemma 5.12 we show that if $\mathscr{T}$ has a path P emanating from v_{RT} with many variables that are not i-self-contained, then there is a sparse subtree $P \subseteq T \subseteq \mathscr{T}$ such that either (i) T has many variables with a large value of $\bar{\gamma}(u)$, or (ii) T has many i-defective variables. The probability of case (i) is bounded by Lemma 5.11 while that of case (ii) is bounded by Proposition 5.1, and the result of Proposition 5.6 follows. The remainder of this subsection gives the details of this argument.

Definition 5.10. For $T \in \mathbf{\Lambda}_{\mathfrak{c},\ell}(\mathscr{T})$ and $S \subseteq V_T$, we say that the pair (T, S) is ϵ-**thick** (in $\mathscr{T}$) if the neighborhoods $B_{\bar{\gamma}(u)}(u)$ for $u \in S$ are mutually disjoint with

$$\sum_{u \in S}\left(\bar{\gamma}(u) - 1 \right) \geq \ell\epsilon \, .$$

(Note that variables u with $\bar{\gamma}(u) = 1$ give no contribution to the above sum.) We then say that T is ϵ-**thick** if (T, S) is ϵ-**thick** for some $S \subseteq V_T$.

Lemma 5.11. *For $k \geq k_0$ (an absolute constant depending only on the absolute constant δ_*), we have*

$$\mathbf{PGW}\left(\text{some } T \in \mathbf{\Lambda}_{3,\ell}(\mathscr{T}) \text{ is } \epsilon\text{-thick in } \mathscr{T} \right) \leq \frac{\exp(k^2\ell)}{\exp\{ 2^{k\delta_*/2}\ell\epsilon/k^5 \}}$$

for all $\epsilon \geq 0$ and integer $\ell \geq 0$.

Proof. We shall bound the expected number of ϵ-thick subtrees of $\mathscr{T}$. To this end, first let us fix a sparse tree $T \in \mathbf{\Lambda}_{3,\ell}$ and a subset of variables $S \subseteq V_T$. Similarly as in the proof of Proposition 5.1 (cf. (5.3) and (5.5)), we use Markov's inequality to bound

$$\mathbf{PGW}\left(\text{there exists an embedding } \zeta : T \hookrightarrow \mathscr{T} \text{ such that } (\zeta(T), \zeta(S)) \text{ is } \epsilon\text{-thick in } \mathscr{T} \right)$$

$$\leq \int \left| \left\{ \text{embeddings } \zeta : T \hookrightarrow \mathscr{T} \text{ such that } (\zeta(T), \zeta(S)) \text{ is } \epsilon\text{-thick in } \mathscr{T} \right\} \right| \, d\mathbf{PGW}(\mathscr{T})$$

$$\leq (\alpha k^2)^{3\ell} \cdot \left[\mathbf{PGW}(T) \right]\left((T, S) \text{ is } \epsilon\text{-thick in } \mathscr{T} \right), \tag{5.11}$$

where in the last line $(\alpha k^2)^{3\ell}$ is an upper bound on $\mathrm{emb}(T)$ (similar to (5.4)). Now let λ denote any tuple of integers $(\lambda(u) : u \in S)$ such that $\lambda(u) \geq 1$ for all $u \in S$, and

$$|\lambda| \equiv \sum_{u \in S} \lambda(u) \geq \ell\epsilon \, .$$

Recalling the definition (5.9) of $\bar{\gamma}$, we see that if $\bar{\gamma}(u) - 1 \geq \lambda(u) \geq 1$ then the ball $|B_{\lambda(u)}(u)| \subseteq \mathcal{T}$ must be large. Therefore, writing T_u for the tree $T \cap B_{\lambda(u)}(u)$ rerooted at u, we have

$$\left[\mathbf{PGW}(T)\right]\left((T, S) \text{ is } \epsilon\text{-thick in } \mathcal{T}\right)$$

$$\leq \sum_\lambda \left[\mathbf{PGW}(T)\right]\left(\begin{array}{c} |B_{\lambda(u)}(u)| \geq \exp\{(k \ln 2)(1 + \delta_*/2)\lambda(u)\} \\ \text{for all } u \in S, \text{ and the } B_{\lambda(u)}(u) \text{ are mutually disjoint} \end{array}\right)$$

$$\leq \sum_\lambda \prod_{u \in S} \left[\mathbf{PGW}(T_u)\right]\left(\left|B_{\lambda(u)}(v_{\mathrm{RT}})\right| \geq \exp\left\{(k \ln 2)\left(1 + \frac{\delta_*}{2}\right)\lambda(u)\right\}\right)$$

$$\leq \sum_\lambda \prod_{u \in S} \left\{2\left(\frac{e}{\exp\{(2^{k\delta_*/2}/k^3)^{\lambda(u)}\}}\right)^{1/2}\right\},$$

where the last inequality follows by Lemma 4.7, Lemma 5.5, and the Cauchy–Schwarz inequality. Note that for all $A \geq e$ and $x \geq 1$ we have $\ln x \leq x - 1 \leq (\ln A)(x - 1)$, and rearranging gives $A^x \geq Ax$. It follows that

$$\left[\mathbf{PGW}(T)\right]\left((T, S) \text{ is } \epsilon\text{-thick in } \mathcal{T}\right) \leq \sum_\lambda \frac{\exp\{O(1)|S|\}}{\exp\{(2^{k\delta_*/2}/k^3)|\lambda|/2\}}$$

(where, as always, $O(1)$ refers to an absolute constant). Note that $|S| \leq \ell$, and the number of distinct tuples λ with $|\lambda| = L$ is crudely upper bounded by

$$\binom{L - 1}{|S| - 1} \leq 2^L.$$

Substituting these bounds into the preceding calculation gives

$$\left[\mathbf{PGW}(T)\right]\left((T, S) \text{ is } \epsilon\text{-thick in } \mathcal{T}\right) \leq \sum_{L \geq \ell\epsilon} \frac{2^L \exp(O(\ell))}{\exp\{(2^{k\delta_*/2}/k^3)L/2\}} \leq \frac{\exp(O(\ell))}{\exp\{2^{k\delta_*/2}\ell\epsilon/k^4\}}.$$

The result follows by substituting this into (5.11), and then summing over all (T, S). (The size of $\Lambda_{3,\ell}$ is bounded by $\exp(O(\ell))$; and for any given $T \in \Lambda_{3,\ell}$, the number of subsets $S \subseteq V_T$ is clearly at most 2^ℓ.) □

Lemma 5.12. *Let $\mathcal{T}$ be any bipartite factor tree rooted at variable v_{RT}. Let $\mathcal{S}^{\mathrm{i}} \equiv \mathcal{S}^{\mathrm{i}}(\mathcal{T})$ denote the subset of variables in $\mathcal{T}$ that are not i-self-contained. For $\epsilon > 0$ and $\ell \geq 1$, if $\mathcal{T}$ contains a path P emanating from v_{RT} with $|V_P| = \ell$ and $|\mathcal{S}^{\mathrm{i}} \cap V_P| \geq \epsilon\ell$, then at least one of the following must occur:*

(i) There is a tree $T \in \Lambda_{3,\ell'}(\mathcal{T})$ (with $\ell' \geq \ell$) which is ϵ'-thick for ϵ' defined by

$$\ell'\epsilon' = \frac{1}{4}\left[(\ell' - \ell) + \frac{\ell\epsilon}{9}\right].$$

(ii) There is a tree $T \in \Lambda_{4,\ell'}(\mathcal{T})$ (with $\ell' \geq \ell$) with

$$|\mathrm{DEF}^{\mathrm{i}} \cap V_T| \geq \frac{(\delta_*)^3}{4}\left[\frac{\ell' - \ell}{2} + \frac{\ell\epsilon}{9}\right].$$

Proof. Let $S_0 \equiv \mathcal{S}^{\mathrm{i}} \cap V_P$. For each $u \in S_0$, by definition there exists $g(u)$ such that

$$1 \leq d(u, g(u)) \leq R',$$
$$1 \leq d(u, g(u)) < \mathrm{rad}^{\mathrm{i}}(g(u)) \leq \chi^{\mathrm{i}}(g(u)) \leq 2R'. \tag{5.12}$$

We will first build a tree T', with $P \subseteq T' \subseteq \mathcal{T}$, by combining P with paths between u and $g(u)$ for a subset of u from S_0, as follows. We will keep track of

$$\mathcal{S}^{\mathrm{i}} \cap V_P = S_0 \supseteq S_1 \supseteq \ldots \supseteq S_{s_{\max}} = \varnothing,$$
$$P = T_0 \subseteq T_1 \subseteq \ldots \subseteq T_{s_{\max}} = T'.$$

As long as $S_{s-1} \neq \varnothing$, we can choose u_s from S_{s-1} such that $g(u_s)$ achieves

$$\chi^{\mathsf{i}}(g(u_s)) = \chi_s \equiv \max \left\{ \chi^{\mathsf{i}}(g(u)) : u \in S_{s-1} \right\}.$$

Let P_s be the path joining u_s to $g(u_s)$; it has length $|P_s| = d(u_s, g(u_s)) \leq \mathrm{rad}^{\mathsf{i}}(g(u_s)) - 1 \leq \chi_s - 1$. Let

$$T_s \equiv T_{s-1} \cup P_s,$$
$$S_s \equiv S_{s-1} \setminus B_{4(\chi_s - 1)}(g(u_s)).$$

Then the paths P_s will be mutually disjoint: indeed, for any $s < t$, by definition the variable u_t must lie outside the ball $B_{4(\chi_s-1)}(g(u_s))$, so the distance between $g(u_s)$ and u_t must be at least $4(\chi_s - 1) + 1$. Since $g(u_s) \in P_s$ and $u_t \in P_t$, we conclude that the minimum distance between P_s and P_t must satisfy

$$d(P_s, P_t) \geq d(g(u_s), u_t) - |P_s| - |P_t| \geq \left(4(\chi_s - 1) + 1 \right) - (\chi_s - 1) - (\chi_t - 1) \geq 2\chi_s - 1,$$

where the last step uses that $\chi_s \geq \chi_t$. We then continue the procedure until we reach $S_{s_{\max}} = \varnothing$, at which point we define $T' \equiv T_{s_{\max}}$. Note

$$|V_{T'}| \leq \ell + \sum_{s=1}^{s_{\max}} d(u_s, g(u_s)) \leq \ell + \sum_{s=1}^{s_{\max}} (\chi_s - 1). \tag{5.13}$$

The intersection between $B_{4(\chi_s-1)}(g(u_s))$ with (any) path P can contain at most $8(\chi_s - 1) + 1 \leq 9(\chi_s - 1)$ variables, where the last inequality holds since $\chi_s \geq 2$ from (5.12). Therefore $|S_s| \geq |S_{s-1}| - 9(\chi_s - 1)$, which implies

$$\sum_{s=1}^{s_{\max}} (\chi_s - 1) \geq \sum_{s=1}^{s_{\max}} \frac{|S_{s-1}| - |S_s|}{9} = \frac{|S_0|}{9} \geq \frac{\ell \epsilon}{9}. \tag{5.14}$$

Recall (5.10) that $\chi^{\mathsf{i}}(v) = \max\{\gamma^{\mathsf{i}}(v), \bar{\gamma}(v)\}$, so one of the following must hold:

$$\sum_{s=1}^{s_{\max}} \left(\gamma^{\mathsf{i}}(g(u_s)) - 1 \right) \geq \frac{1}{2} \sum_{s=1}^{s_{\max}} \left(\chi^{\mathsf{i}}(g(u_s)) - 1 \right), \tag{5.15}$$

$$\sum_{s=1}^{s_{\max}} \left(\bar{\gamma}(g(u_s)) - 1 \right) \geq \frac{1}{2} \sum_{s=1}^{s_{\max}} \left(\chi^{\mathsf{i}}(g(u_s)) - 1 \right) \tag{5.16}$$

We consider separately the two cases:

(i) If (5.16) occurs, then we let $T = T'$. Combining with (5.13) and (5.14) gives

$$\sum_{s=1}^{s_{\max}} \left(\bar{\gamma}(g(u_s)) - 1 \right) \geq \frac{1}{2} \sum_{s=1}^{s_{\max}} (\chi_s - 1) \geq \frac{1}{4} \left[(\ell' - \ell) + \frac{\ell \epsilon}{9} \right].$$

Therefore the tree $T = T'$ is ϵ'-thick for ϵ' as defined in the statement of the lemma.

(ii) If (5.15) occurs, then for each s we can find a path Q_s emanating from $g(u_s)$ with $\gamma^{\mathsf{i}}(g(u_s))$ variables, of which at least $(\delta_*)^3$ fraction are i-defective. Let T be the union of T' with all the paths Q_s, for $1 \leq s \leq s_{\max}$. The Q_s are mutually disjoint. Recalling (5.13), we have

$$\ell' = |V_T| \leq \ell + \sum_{s=1}^{s_{\max}} (\chi_s - 1) + \sum_{s=1}^{s_{\max}} \left(\gamma^{\mathsf{i}}(g(u_s)) - 1 \right) \leq \ell + 2 \sum_{s=1}^{s_{\max}} (\chi_s - 1). \tag{5.17}$$

Combining (5.14), (5.15), and (5.17) gives

$$|\mathsf{DEF}^{\mathsf{i}} \cap V_T| \geq \sum_{s=1}^{s_{\max}} |\mathsf{DEF}^{\mathsf{i}} \cap Q_s| \geq (\delta_*)^3 \sum_{s=1}^{s_{\max}} \gamma^{\mathsf{i}}(g(u_s)) \geq \frac{(\delta_*)^3}{2} \sum_{s=1}^{s_{\max}} (\chi_s - 1)$$

$$\geq \frac{(\delta_*)^3}{4} \left[\frac{\ell' - \ell}{2} + \frac{\ell \epsilon}{9} \right].$$

Combining the two cases gives the claim. $\qquad \square$

Proof of Proposition 5.6. It follows from Lemma 5.12 that

$$\mathbf{PGW}\left(\begin{array}{c}\mathscr{T} \text{ contains a path } P \text{ of } \ell \text{ variables,}\\ \text{emanating from } v_{\text{RT}}, \text{ with } |\mathcal{S}^{\mathfrak{i}} \cap V_P| \geq \ell\epsilon\end{array}\right) \leq P_{(\mathrm{i})} + P_{(ii)}$$

where $P_{(\mathrm{i})}$ and $P_{(ii)}$ refer to the two cases of Lemma 5.12. For case (i), we use Lemma 5.11 to bound

$$P_{(\mathrm{i})} \leq \sum_{\ell' \geq \ell} \frac{\exp\{k^2[\ell + (\ell' - \ell)]\}}{\exp\{2^{k\delta_*/2}[\ell\epsilon/36 + (\ell' - \ell)/4]/k^5\}} \leq \frac{\exp(k^2\ell)}{\exp\{2^{k\delta_*/2}\ell\epsilon/k^6\}} .$$

For case (ii), we use Proposition 5.1 to bound

$$P_{(ii)} \leq \sum_{\ell' \geq \ell} \frac{\exp\{k^2[\ell + (\ell' - \ell)]\}}{\exp\{(\delta_*)^3 2^{k/4}[\ell\epsilon/36 + (\ell' - \ell)/8]/k^2\}} \leq \frac{\exp(k^2\ell)}{\exp\{2^{k/4}\ell\epsilon/k^3\}} .$$

Combining gives the result. □

We conclude this subsection with the proofs of Corollaries 5.8 and 5.9, which are easy consequences of what was proved above.

Proof of Corollary 5.8. Recall from Definition 3.13 that an acyclic variable is termed $\mathfrak{i}$-fair if (i) it is $\mathfrak{i}$-stable; (ii) its $5R'$-neighborhood contains no more than $\exp\{k^2(5R')\}$ variables; and (iii) every length-R' path emanating from it contains at least one $\mathfrak{i}$-perfect variable, where $\mathfrak{i}$-perfect means both $\mathfrak{i}$-orderly and $\mathfrak{i}$-self-contained. For condition (i), it follows from Proposition 4.17 that

$$\mathbf{PGW}\left(v_{\text{RT}} \text{ not } \mathfrak{i}\text{-stable}\right) \leq \frac{1}{\exp(2^{k/20}R)} .$$

For condition (ii), it follows from Lemma 4.7 and Markov's inequality that

$$\mathbf{PGW}\left(|B_{5R'}(v_{\text{RT}})| \geq \exp\{k^2(5R')\}\right) \leq \frac{e}{\exp\{\exp\{k^2(5R')\}/(\alpha k^2)^{5R'}\}} \leq \frac{1}{\exp(\exp(kR))} .$$

For condition (iii), it follows by combining Propositions 5.6 and 5.7 that

$$\mathbf{PGW}\left(\begin{array}{c}\mathscr{T} \text{ contains a path } P \text{ of } R' \text{ variables,}\\ \text{emanating from } v_{\text{RT}}, \text{ with } V_P \subseteq \mathcal{S}^{\mathfrak{i}} \cup \mathcal{O}^{\mathfrak{i}}\end{array}\right) \leq \frac{O(1)\exp(k^2 R')}{\exp(\Omega(2^{k\delta_*/3}R'))} .$$

Combining these bounds gives the result. □

Proof of Corollary 5.9. Recall from Definition 3.14 that an acyclic variable v is termed $\mathfrak{i}$-excellent if its neighborhood $T = B_{10R'}(v)$ satisfies condition (3.20), which we repeat here for convenience:

$$p_{\text{ex},\mathfrak{i}}(T) \equiv \mathbf{PGW}\left(\begin{array}{c}B_{20R'}(u) \text{ contains any}\\ \text{variable which is not } \mathfrak{i}\text{-fair}\end{array}\bigg| B_{10R'}(u) \cong T\right) \leq \frac{1}{\exp\{k^3 R'\}} .$$

By Markov's inequality and iterated expectations,

$$p \equiv \mathbf{PGW}\left(v_{\text{RT}} \text{ not } \mathfrak{i}\text{-excellent}\right) \leq \int \frac{p_{\text{ex},\mathfrak{i}}(B_{10R'}(v_{\text{RT}}; \mathscr{T}))}{1/\exp\{k^3 R'\}} \, d\mathbf{PGW}(\mathscr{T})$$

$$= \exp\{k^3 R'\}\, \mathbf{PGW}\left(B_{20R'}(v_{\text{RT}}; \mathscr{T}) \text{ contains any variable which is not } \mathfrak{i}\text{-fair}\right) .$$

By another application of Markov's inequality, together with the unimodularity property (4.1), we find

$$p \leq \exp\{k^3 R'\} \int \sum_{u \in B_{20R'}(v_{\text{RT}}; \mathscr{T})} \mathbf{1}\left\{u \text{ is not } \mathfrak{i}\text{-fair}\right\} d\mathbf{PGW}(\mathscr{T})$$

$$= \exp\{k^3 R'\} \int |B_{20R'}(v_{\text{RT}}; \mathscr{T})| \mathbf{1}\left\{v_{\text{RT}} \text{ not } \mathfrak{i}\text{-fair}\right\} d\mathbf{PGW}(\mathscr{T}) \leq p_{\text{small}} + p_{\text{large}} ,$$

where p_{small} is the contribution from the event $|B_{20R'}(v_{\text{RT}}; \mathscr{T})| \leq \exp\{k^2 R\}$, and p_{large} is the contribution from the complementary event. It follows from Corollary 5.8 that

$$p_{\text{small}} = \exp\{k^3 R' + k^2 R\}\mathbf{PGW}\left(v_{\text{RT}} \text{ not } \mathfrak{i}\text{-fair}\right) \leq \frac{1}{\exp(2^{k\delta_*/3}R/k^2)} .$$

The contribution from the complementary event is

$$\frac{p_{\text{large}}}{\exp\{k^3 R'\}} = \int |B_{20R'}(v_{\text{RT}}; \mathscr{T})| \mathbf{1}\left\{|B_{20R'}(v_{\text{RT}}; \mathscr{T})| \geq \exp\{k^2 R\}\right\} d\mathbf{PGW}(\mathscr{T})$$

$$\leq \exp\{k^2 R\}\mathbf{PGW}\left(|B_{20R'}(v_{\text{RT}}; \mathscr{T})| \geq \exp\{k^2 R\}\right) + \int_{\exp\{k^2 R\}}^{\infty} \mathbf{PGW}\left(|B_{20R'}(v_{\text{RT}}; \mathscr{T})| \geq s\right) ds.$$

It follows from Lemma 4.7 that this is very small: $p_{\text{large}} \leq 1/\exp(\exp(kR))$, which is negligible compared with the bound on p_{small}. Combining these gives the claimed bound (it is easy to show, using Markov's inequality, that the probability for v_{RT} to be improper is negligible). $\qquad\square$

5.3. Combinatorial analysis of preprocessing.

In this section we analyze the preprocessing algorithm described by Definition 3.15, which maps the original k-SAT instance $\mathscr{G}$ to its pruned version $\text{pr}\,\mathscr{G}$. Recall that the procedure starts from an initial set $A \subseteq V$, which is the set of all variables in $\mathscr{G}$ that are improper (Definition 3.1) or not 1-good (Definition 3.14). It then iteratively produces a sequence $\mathscr{G} \supseteq {}_0\mathscr{G}_A \supseteq {}_1\mathscr{G}_A \supseteq \ldots$, terminating in $\text{pr}\,\mathscr{G} = \mathscr{G} \setminus \text{BSP}'(A; \mathscr{G})$. It is a straightforward consequence of Corollary 5.9 that the fraction of variables in the initial set A is $o_R(1)$, so the main challenge is to bound the effect of the BSP′ procedure. We do this in two parts:

a. The current subsection (§5.3) is devoted to the proof of a structural result, Proposition 5.14, which says roughly that if $\text{BSP}'(A; \mathscr{G})$ has a large connected component, then $\mathscr{G}$ must contain a certain kind of subgraph (either a "bicycle" or a "sparse corrupted subtree") of comparable size. This statement holds deterministically.

b. In the next subsection (§5.4) we bound the probability for such subgraphs to occur (Lemma 5.21). This allows us to control the typical size of $\text{BSP}'(A; \mathscr{G})$, and we conclude §5.4 with the proof of Proposition 3.22.

Turning to the task of this subsection, we note that BSP′ is highly analogous to the BSP procedure (3.17), which determined the $\mathfrak{i}$-defective variables based on the initial set $D^{\kappa_*,\mathfrak{i}}$ (see Definition 3.10). Recall from §5.1 the result Lemma 5.4, which roughly says that if $\mathscr{T}$ has a sparse subtree T that has a large intersection with the $\mathfrak{i}$-defective set, then it must also have a sparse subtree T' ($T \subseteq T' \subseteq \mathscr{T}$) that has a large intersection with the initial set $D^{\kappa_*,\mathfrak{i}}$. We now prove an analogous result for BSP′, given by Proposition 5.14 below. To state the result, we introduce the following definition:

Definition 5.13. Let $\mathscr{G} \equiv (V, F, E)$ be a bipartite factor graph, and fix any $A \subseteq V$. We say that a subtree $T \subseteq \mathscr{G}$ is ϵ-**corrupted with respect to** $(A, \mathscr{G})$ if there is a subset of vertices $B \subseteq V_T \cap A$ such that $|B| \geq \epsilon |V_T|$, the minimum pairwise distance between vertices in B exceeds $2R$, and the graph

$$T \cup B_R(B; \mathscr{G})$$

is acyclic. (We may say simply "ϵ-corrupted" if $(A, \mathscr{G})$ is unambiguous.)

Proposition 5.14. Let $\mathscr{G} = (V, F, E)$ be a finite bipartite factor graph in which all clauses have degree k, and let A be any subset of V. If $\text{BSP}'(A; \mathscr{G})$ contains a connected component $\mathscr{B}$ of diameter at least $5L$ where $L \geq 400R$, then either (i) there is a subgraph $B' \subseteq \mathscr{B}$ with $\text{diam}\, B' \leq 11L$ that contains at least two cycles; or (ii) there is an $1/(80R)$-corrupted subtree $T \subseteq \mathscr{B}$ of maximum degree at most four with $|V_T| \geq L$ and $L \leq \text{diam}\, T \leq 11L$.

The remainder of this subsection is devoted to the proof of Proposition 5.14.

Lemma 5.15. Let $\mathscr{T} = (V, F, E)$ be a finite bipartite factor tree in which all clauses have degree k. Assume that $\mathscr{T}$ has diameter $L \geq 8R$. If A is a subset of V such that $\text{BSP}'(A; \mathscr{T}) = \mathscr{T}$, then there is a $1/(80R)$-corrupted subtree $T \subseteq \mathscr{T}$ of diameter L and maximum degree at most four.

Proof. Since $\text{diam}\, \mathscr{T} = L$, it must contain some path P of length L. Let v_{RT} denote one of the endpoints of this path. From now on we regard $\mathscr{T}$ as being rooted at v_{RT}.

Step 1. Simplified bootstrap percolation of marked blocks. We now define a much simpler bootstrap percolation process, which we will show (in subsequent steps) to dominate $\text{BSP}'(A; \mathscr{T})$ in an appropriate sense. For integers $j \geq 0$, let $\mathscr{T}[j]$ denote the subgraph of $\mathscr{T}$ induced by variables whose distance to the root v_{RT} lies between $4Rj$ and $4R(j + 1)$; we then consider each connected component of $\mathscr{T}[j]$ as a "block." Note that the blocks themselves have a tree-like (i.e., hierarchical) structure: if U is a block rooted at depth $4R(j + 1)$, then its root is a leaf of a block U' rooted at depth $4Rj$, and we say that U is a "child block" of U'. Let $\mathfrak{B}$ be the tree structure of blocks. Let $\mathscr{A}_0$ be the

set of all blocks that intersect A. For $t \geq 1$ we will say that a block belongs to $\mathcal{A}_t$ if either it belongs to $\mathcal{A}_{t-1}$, or has at least two child blocks in $\mathcal{A}_{t-1}$. Iterate this to define the set of all "marked" blocks,

$$\mathsf{mark}(\mathcal{A}_0; \mathfrak{B}) \equiv \bigcup_{t \geq 0} \mathcal{A}_t \ .$$

Note that this "marking" is also a bootstrap-percolation-type process, but is much simpler than BSP$'$.

Step 2. A general property of BSP$'(A; \mathcal{T})$. We next state and prove a useful property of the BSP$'(A; \mathcal{T})$ process: if S is any subset of V which does not intersect A, then in the first round of BSP$'(A; \mathcal{T})$ in which *any* variable is deleted from S, it must be the case that some variable on its internal boundary

$$\partial^{\mathrm{int}} S \equiv \left\{ u \in S : u \in B_1(w; \mathcal{T}) \text{ for some } w \in V \setminus S \right\}$$

is also deleted. To see that this property holds, say the first deletion from S occurs at round $t+1$. This means that after round t, the subgraph remaining in the BSP$'(A; \mathcal{T})$ process is $\mathcal{T}_t = (V_t, F_t, E_t) \subseteq \mathcal{T}$ with $V_t \supseteq S$ (i.e., no variable has yet been deleted from S). The subgraph remaining after the next round $t+1$ is $\mathcal{T}_{t+1} = (V_{t+1}, F_{t+1}, E_{t+1}) \subseteq \mathcal{T}_t$ where $V_t \setminus V_{t+1}$ intersects S. By the definition of BSP$'$,

$$\mathcal{T}_{t+1} = \mathcal{T}_t \setminus B_R(A_t; \mathcal{T}_t)$$

where $A_{-1} = A$ and $A_t = \mathsf{Act}(\mathcal{T}_t)$ for $t \geq 1$. Now suppose for contradiction that $V_t \setminus V_{t+1}$ does not intersect $\partial^{\mathrm{int}} S$. For any variable v, its neighborhood $B_R(v; \mathcal{T}_t)$ is a connected subgraph of $\mathcal{T}_t$ (hence also of $\mathcal{T}$). As a result it must be that for all $v \in A_t$, the neighborhood $B_R(v; \mathcal{T}_t)$ is either disjoint from S, or contained in $S \setminus \partial^{\mathrm{int}} S$. Since $V_t \setminus V_{t+1}$ intersects S, there must be at least one $v \in A_t$ with $B_R(v; \mathcal{T}_t) \subseteq S \setminus \partial^{\mathrm{int}} S$. Since S does not intersect A, it must be that $t \geq 0$ and $v \in A_t = \mathsf{Act}(\mathcal{T}_t)$. By definition of $\mathsf{Act}(\mathcal{T}_t)$, this means that some clause a in $B_R(v; \mathcal{T}_t)$ must have degree less than k. Since all clauses in $\mathcal{T}$ were assumed to be of degree k, this means that in $\mathcal{T}$ there was a variable $u \in \partial a$ which was deleted by the end of round t. However u must also belong to S, contradicting the hypothesis that no deletion occurred from S by the end of round t. This proves the claim.

Step 3. Comparison of marking and BSP$'$. We now argue that the marking process "dominates" BSP$'(A; \mathcal{T})$ in the following sense. Let us say that a block U is a "tall block" if it has depth at least $2R$; otherwise we call U a "shallow block." We claim that any tall block must be marked. Suppose for contradiction that this is not the case, then let U be any block of maximal depth among all the unmarked tall blocks. This means that all the child blocks of U must be marked, or be shallow blocks. Since U itself is unmarked, the definition of the marking process implies that at most one child block of U can be marked. If such a block exists we will denote it U'. We let W denote any child block of U which is unmarked; this means that W must be a shallow block.

We now argue that throughout the BSP$'(A; \mathcal{T})$ process, W has no influence on U. To see this, apply the claim from Step 2 with $S = W$. Since W is a shallow block, it follows that at the first round during BSP$'(A; \mathcal{T})$ in which any variable is removed from the upper half of W, the root of W must also be removed. After this, any variables remaining in W will be disconnected from U. Thus, throughout BSP$'(A; \mathcal{T})$, there is no time at which any variable $u \in U$ has in its $(3R/10)$-neighborhood a clause a of degree less than k that lies in W. This shows that W has no influence on the evolution of U under BSP$'(A; \mathcal{T})$.

Let ρ denote the root of U, and let ρ' denote the root of U' (if it exists). Let τ be the first time during BSP$'(A; \mathcal{T})$ that any variable is deleted from U; note that τ must be finite by the assumption that U is a tall block. Let τ_ρ be the time that ρ is deleted, and let $\tau_{\rho'}$ be the time that ρ' is deleted. It follows from the preceding discussion that $\tau = \min\{\tau_\rho, \tau_{\rho'}\}$. Let us suppose first $\tau = \tau_\rho < \tau_{\rho'}$. At time τ, the BSP$'(A; \mathcal{T})$ process deletes a connected component U_ρ containing ρ. The definition of BSP$'$ implies that variables in U_ρ can lie at depth at most $13R/10$ below ρ, and hence at distance at least $27R/10$ from ρ'. As a result, no other variables will be removed from U until time $\tau_{\rho'}$, when the BSP$'(A; \mathcal{T})$ process deletes a connected component $U_{\rho'}$ containing ρ'. Variables in $U_{\rho'}$ can lie at distance at most $13R/10$ from ρ', hence at distance at least $7R/10$ from U_ρ. No other variables will be removed from U after time $\tau_{\rho'}$. In particular, this contradicts the assumption that U is a tall block and BSP$'(A; \mathcal{T}) = \mathcal{T}$. Very similar arguments give the desired contradiction in the cases $\tau = \tau_{\rho'} < \tau_\rho$ and $\tau = \tau_{\rho'} = \tau_\rho$. This proves our claim that any tall block must also be marked.

Step 4. Extraction of sparse subtree. Now recall that the tree $\mathcal{T}$ is rooted at a variable v_{RT}, from which there emanates a path P of length $L \geq 8R$. Let $\ell \equiv \ell/(8R)$ and note that $\lfloor L/(4R) \rfloor \geq L/(4R) - 1 \geq \ell$. It follows that there is a path of blocks $\mathfrak{P} = (U_1, \ldots, U_\ell)$, where U_1 is rooted at v_{RT} and U_i is a child block of U_{i-1} for each $2 \leq i \leq \ell$. By the claim proved in the previous step, each U_i must be a marked block: that is, each U_i either intersects A or has at least two marked child blocks. It follows from the definition of the marking process that the path $\mathfrak{P}$ can be covered by a disjoint union

$$\mathfrak{T} = \bigsqcup_{j=1}^{s} \mathfrak{T}_j$$

where each $\mathfrak{T}_j$ is a nonempty full binary tree of marked blocks, rooted at a block $U_{i(j)} \in \mathfrak{P}$, such that the leaves $\mathfrak{L}_j$ of $\mathfrak{T}_j$ are all blocks intersecting A.[19] Let $\mathfrak{T}$ be the tree of blocks given by the union of $\mathfrak{T}_1, \ldots, \mathfrak{T}_s$; it has maximum degree at most four. We then construct a subtree $T \subseteq \mathcal{T}$ as follows:

a. For $1 \leq j \leq s$ let $B_j \subseteq A$ be defined by taking one variable from $U \cap A$ for each $U \in \mathfrak{L}_j$.
b. For $1 \leq j \leq s$ let $T_j \subseteq \mathcal{T}$ be defined by taking the union of all paths between B_j and the root of $U_{i(j)}$.
c. For $1 \leq j \leq s-1$, let Q_j be the union of T_j with a path joining the roots of $U_{i(j+1)-1}$ and $U_{i(j+1)}$. Let Q_s be the union of T_s with a path joining the root of U_ℓ with a variable of maximal depth in U_ℓ. (The precise definition of Q_s is not so important; we choose this one as it will guarantee $\mathrm{diam}\, T = L$.)
d. Let $T \subseteq \mathcal{T}$ be the union of $Q_1, \ldots, Q_s$. Let B' be the (disjoint) union of $B_1, \ldots, B_s$.

Thus T is a subtree of $\mathcal{T}$ that includes the root v_{RT} and has maximum degree at most four. We now argue that T is $1/(80R)$-corrupted. The intersection of T_j with any block U is given by a union of at most two paths, so contains at most $8R$ variables. This implies

$$|V_{T_j}| \leq \sum_{U \in \mathfrak{T}_j} |V_{T_j} \cap U| \leq 8R|\mathfrak{T}_j|$$

where $|\mathfrak{T}_j|$ denotes the number of blocks in $\mathfrak{T}_j$. We then have $|V_{Q_j}| \leq |V_{T_j}| + 4R \leq 8R(|\mathfrak{T}_j| + 1)$, so

$$|V_T| \leq \sum_{j=1}^{s} |V_{Q_j}| \leq 8R \sum_{j=1}^{s} \left(|\mathfrak{T}_j| + 1 \right). \tag{5.18}$$

On the other hand, in any nonempty full binary tree, the number of leaf nodes is exactly one plus the number of internal nodes, so

$$|V_T \cap A| \geq |B'| = \sum_{j=1}^{s} |B_j| \geq \sum_{j=1}^{s} |\mathfrak{L}_j| = \sum_{j=1}^{s} \frac{|\mathfrak{T}_j| + 1}{2}. \tag{5.19}$$

Since $\mathfrak{T}$ has maximum degree at most four, we can extract $B \subseteq B'$ with $|B| \geq |B'|/5$ such that variables in B lie at pairwise distance greater than $2R$. Combining with (5.18) and (5.19) gives

$$\frac{|B|}{|V_T|} \geq \frac{|B'|}{5|V_T|} \geq \frac{1}{80R},$$

so T is $1/(80R)$-corrupted as claimed. (Note in this case that the condition that $T \cup B_R(B; \mathcal{T})$ be acyclic is trivially satisfied, since $\mathcal{T}$ is a tree.) □

In fact, by an essentially identical proof, we have the following slight generalization of Lemma 5.15, which will be used in the analysis that follows.

Corollary 5.16. *Suppose $\mathcal{G} = (V, F, E)$ is a finite bipartite factor graph in which all clauses have degree k, and which can be expressed as*

$$\mathcal{G} = \mathcal{T} \cup \mathcal{G}'$$

where $\mathcal{G}'$ is an arbitrary graph, and $\mathcal{T}$ is a tree that intersects $\mathcal{G}'$ at a single variable v, such that $\mathcal{T}$ has depth L when rooted at v. If A is a subset of V such that $\mathrm{BSP}'(A; \mathcal{G}) = \mathcal{G}$, then there is a $1/(80R)$-corrupted subtree $T \subseteq \mathcal{T}$ of maximum degree at most four, with $L \leq \mathrm{diam}\, T \leq \mathrm{diam}\, \mathcal{T} \leq 2L$.

---

[19]We will explain the $\mathfrak{T}_j$ by an example. The root of $\mathfrak{T}_1$ is always U_1, so $i(1) = 1$. Suppose that U_1 does not intersect A, in which case it must have two marked child blocks, say U_2 and $\tilde{U}_2$. Suppose that these do intersect A. Then $\mathfrak{T}_1$ consists of U_1, U_2, and $\tilde{U}_2$. We then take $\mathfrak{T}_2$ to be rooted at U_3, so in this case $i(2) = 3$. Suppose U_3 also intersects A; then $\mathfrak{T}_2$ consists only of U_3. We take $\mathfrak{T}_3$ to be rooted at U_4, and so on.

Proof. The proof of Lemma 5.15 applies to $\mathscr{T}$; the only difference is that we fix v to be the root of $\mathscr{T}$. In the final step of extracting B from B', it is easy to arrange that none of the variables in B lie in the topmost block of $\mathscr{T}$. By construction, the tree T has depth exactly L, hence diameter between L and $2L$. Then the requirement that $T \cup B_R(B; \mathscr{G})$ be acyclic is satisfied, since this will be a subgraph of the tree $\mathscr{T}$. $\qquad\square$

For the proof of Proposition 5.14, the main challenge remaining is to reduce to the case of Corollary 5.16 (or to its special case Lemma 5.15). To this end, it is useful to consider a slight variant of BSP′ (Definition 5.17) and prove a self-consistency property thereof (Lemma 5.18 below).

Definition 5.17 (slowed removal process BSP″). Let $\mathscr{G} = (V, F, E)$ be any bipartite factor graph (finite or infinite), and let A be a finite subset of V. Recall from Definition 3.15 that the removal process BSP′$(A; \mathscr{G})$ goes from $\mathscr{G}_t$ to $\mathscr{G}_{t+1}$ (for $t \geq -1$) by removing the R-neighborhoods of all the variables in A_t, where A_{-1} is the initial set A, and A_t is defined inductively using (3.21) as Act$(\mathscr{G}_t)$. We now define the process BSP″$(A; \mathscr{G})$ which is equivalent to BSP′$(A; \mathscr{G})$ except that it removes one R-neighborhood at a time. That is to say, for each $t \geq -1$, we arbitrarily order the variables in A_t as

$$\left(v_{t,i} : 1 \leq i \leq |A_t| \right),$$

and then remove first $B_R(v_{t,1}; \mathscr{G}_t)$, then $B_R(v_{t,2}; \mathscr{G}_t)$, and so on. We say that $v_{t,i}$ is **visited** at the i-th step of this process (even if it may have been deleted at an earlier step). Thus BSP″ reaches the graph $\mathscr{G}_{t+1}$ in

$$\sum_{s=-1}^{t} |A_s|$$

steps. The point of BSP″ is that each step cannot increase the maximum component diameter (of the removed subgraph) by too much: if the maximum component diameter before a BSP″ step is ℓ, the maximum component diameter after the BSP″ step is at most $2(\ell + R + 1)$. We let BSP^s denote the BSP″ process stopped after s steps.

Lemma 5.18. *Let $\mathscr{G} = (V, F, E)$ be any bipartite graph (finite or infinite) in which all clauses have degree k. Let A be a finite subset of V. If for some finite s we have BSP$^s(A; \mathscr{G}) = \mathscr{H}$, then BSP′$(A_{\mathscr{H}}; \mathscr{H}) = \mathscr{H}$ where $A_{\mathscr{H}}$ denotes the restriction of A to $\mathscr{H}$.*

Proof. First we note that $A_{\mathscr{H}}$ is a strict subset of A if and only if BSP^s stops before the initial round of BSP′$(A; \mathscr{G})$ is finished, i.e., if and only if $s < |A_{-1}|$. In this case the result is straightforward: let $U_{-1} \subseteq A_{-1}$ denote the first s variables in A_{-1}, so $\mathscr{H} = \text{BSP}^s(A; \mathscr{G}) = B_R(U_{-1}; \mathscr{G})$. This implies $U_{-1} \subseteq A_{\mathscr{H}}$ and $B_R(U_{-1}; \mathscr{G}) = B_R(U_{-1}; \mathscr{H})$. From this we obtain the chain of relations

$$\mathscr{H} \supseteq \text{BSP}'(A_{\mathscr{H}}; \mathscr{H}) \supseteq B_R(A_{\mathscr{H}}; \mathscr{H}) \supseteq B_R(U_{-1}; \mathscr{H}) = B_R(U_{-1}; \mathscr{G}) = \mathscr{H},$$

and so we have BSP′$(A_{\mathscr{H}}; \mathscr{H}) = \mathscr{H}$ as desired. We therefore assume from now on that $s \geq |A_{-1}|$. This means that BSP$^s(A; \mathscr{G})$ completes the initial round of BSP′$(A; \mathscr{G})$, producing $\mathscr{G}_0 = \mathscr{G} \setminus B_R(A; \mathscr{G})$. It also means that $A_{\mathscr{H}} = A$, and BSP′$(A; \mathscr{H})$ completes its initial round to produce $\mathscr{H}_0 = \mathscr{H} \setminus B_R(A; \mathscr{H}) = \mathscr{H} \setminus B_R(A; \mathscr{G})$. We hereafter index this initial round as the "zeroth round."

We now suppose inductively, for $i \geq 0$, the following hypothesis: provided that BSP$^s(A; \mathscr{G})$ fully completes the first i rounds of BSP′$(A; \mathscr{G})$ (again, i starts from zero), the following hold:

  (i) $A_t = \text{Act}(\mathscr{G}_t) = \text{Act}(\mathscr{H}_t)$ for all $t \leq i - 1$; and

  (ii) The subgraph $\mathscr{R}_i$ removed from $\mathscr{G}$ within the first i rounds of BSP′$(A; \mathscr{G})$ agrees with the subgraph removed from $\mathscr{H}$ within the first i rounds of BSP′$(A; \mathscr{H})$, i.e.,

$$\mathscr{R}_i \equiv \bigcup_{t=-1}^{i-1} B_R(A_t; \mathscr{G}_t) = \bigcup_{t=-1}^{i-1} B_R(A_t; \mathscr{H}_t). \tag{5.20}$$

The base case $i = 0$ follows from the preceding argument. Let $\mathscr{G}_i$ be the graph remaining after the first i rounds of BSP′$(A; \mathscr{G})$, and let $\mathscr{H}_i$ be the graph remaining after the first i rounds of BSP′$(A; \mathscr{H})$. By induction, $\mathscr{G}_i = \mathscr{G} \setminus \mathscr{R}_i$ and

$\mathcal{H}_i = \mathcal{H} \setminus \mathcal{R}_i = \mathcal{G}_i \cap \mathcal{H}$. We will say that a subgraph is "lacking" if it contains at least two clauses of degree $k-1$, or at least one clause of degree $\leq k-2$. Then, recalling (3.21), for the next round we must consider

$$\mathrm{Act}(\mathcal{G}_i) = \left\{ \text{variables } v \text{ in } \mathcal{G}_i \text{ such that } B_{3R/10}(v;\mathcal{G}_i) \text{ is lacking} \right\},$$

$$\mathrm{Act}(\mathcal{H}_i) = \left\{ \text{variables } v \text{ in } \mathcal{H}_i \text{ such that } B_{3R/10}(v;\mathcal{H}_i) \text{ is lacking} \right\}.$$

The inductive hypothesis implies $\mathcal{H}_i \subseteq \mathcal{G}_i$, so $B_R(v;\mathcal{H}_i) \subseteq B_R(v;\mathcal{G}_i)$ for any variable v. On the other hand, let U_i be the vertices in $\mathrm{Act}(\mathcal{G}_i)$ that are visited by $\mathrm{BSP}^s(A;\mathcal{G})$ (for "visited" in the sense of Definition 5.17). If $v \in U_i$ then the $\mathrm{BSP}^s(A;\mathcal{G})$ process removes $B_R(v;\mathcal{G}_i)$, which means it lies inside $\mathcal{H}$. This implies $B_R(v;\mathcal{G}_i) \subseteq \mathcal{G}_i \cap \mathcal{H} = \mathcal{H}_i$, where the last equality uses the inductive hypothesis. This proves that $B_R(v;\mathcal{G}_i) = B_R(v;\mathcal{H}_i)$ for all $v \in U_i$, hence also $B_{3R/10}(v;\mathcal{G}_i) = B_{3R/10}(v;\mathcal{H}_i)$ for all $v \in U_i$. Moreover, all variables in U_i are removed, so U_i must be contained in $\mathcal{G}_i \cap \mathcal{H} = \mathcal{H}_i$ (using the inductive hypothesis again). It follows from this that

$$U_i = \left\{ v \in U_i : B_{3R/10}(v;\mathcal{G}_i) \text{ is lacking} \right\} \subseteq \left\{ v \in \mathcal{H}_i : B_{3R/10}(v;\mathcal{H}_i) \text{ is lacking} \right\} = \mathrm{Act}(\mathcal{H}_i).$$

Conversely, suppose $v \in \mathrm{Act}(\mathcal{H}_i)$, so $B_{3R/10}(v;\mathcal{H}_i)$ is lacking: either there are two clauses $a_1, a_2 \in B_{3R/10}(v;\mathcal{H}_i)$ of degree $k-1$, or there is a single clause $a \in B_{3R/10}(v;\mathcal{H}_i)$ of degree $\leq k-2$. We have by induction $\mathcal{H}_i \subseteq \mathcal{G}_i$, therefore $B_{3R/10}(v;\mathcal{H}_i) \subseteq B_{3R/10}(v;\mathcal{G}_i)$ which means these clauses are also present in $B_{3R/10}(v;\mathcal{G}_i)$. All clauses have degree k in $\mathcal{G}$ (and hence also in $\mathcal{H}$), so if a clause $a \in B_{3R/10}(v;\mathcal{H}_i)$ has degree $k-j$ relative to $\mathcal{H}_i$, there must be j edges joining that clause to variables in $\mathcal{R}_i$. It then follows from the inductive hypothesis (5.20) that the clause has degree $\leq k-j$ in $\mathcal{G}_i$. This implies that $B_{3R/10}(v;\mathcal{G}_i)$ must also be lacking, so $\mathrm{Act}(\mathcal{H}_i) \subseteq \mathrm{Act}(\mathcal{G}_i)$. If the $(i+1)$-st round is completed, then we obtain $A_i = \mathrm{Act}(\mathcal{G}_i) = U_i = \mathrm{Act}(\mathcal{H}_i)$. Since we proved above that $B_R(v;\mathcal{G}_i) = B_R(v;\mathcal{H}_i)$ for all $v \in U_i = A_i$, we conclude $\mathcal{R}_{i+1} = \mathcal{R}_i \cup B_R(A_i;\mathcal{G}_i) = B_R(A_i;\mathcal{H}_i)$. This verifies the inductive hypothesis.

If $\mathrm{BSP}^s(A;\mathcal{G})$ completes exactly the first i rounds of $\mathrm{BSP}'(A;\mathcal{G})$ (and no more), then the above induction implies $\mathcal{H} = \mathcal{R}_i = \mathrm{BSP}'(A;\mathcal{H})$, as desired. It remains finally to consider the case that $\mathrm{BSP}^s(A;\mathcal{G})$ completes the first i rounds, and does not complete the $(i+1)$-st round. In this case, $\mathcal{H} = \mathrm{BSP}^s(A;\mathcal{G}) = \mathcal{R}_i \cup B_R(U_i;\mathcal{G}_i)$. The above argument also gives $U_i \subseteq \mathrm{Act}(\mathcal{H}_i)$. Altogether we obtain the chain of relations

$$\mathcal{H} \supseteq \mathrm{BSP}'(A;\mathcal{H}) \supseteq \mathcal{R}_i \cup B_R(\mathrm{Act}(\mathcal{H}_i);\mathcal{H}_i) \supseteq \mathcal{R}_i \cup B_R(U_i;\mathcal{H}_i) = \mathcal{H}.$$

This gives $\mathrm{BSP}'(A;\mathcal{H}) = \mathcal{H}$ in the case where $\mathrm{BSP}^s(A;\mathcal{G})$ stops partway through some round of $\mathrm{BSP}'(A;\mathcal{G})$, and concludes the proof of the lemma.   □

We are now finally prepared to prove the main result of this subsection:

Proof of Proposition 5.14. Instead of $\mathrm{BSP}'(A;\mathcal{G})$, we follow the slowed removal process $\mathrm{BSP}''(A;\mathcal{G})$ (Definition 5.17) where we remove one R-neighborhood at a time. Recall that if the maximum component diameter before a BSP'' step is ℓ, then the maximum component diameter after the BSP'' step is at most $2(\ell + R + 1)$. Let $t_\star$ be the first time that $\mathrm{BSP}''(A;\mathcal{G})$ creates any component $\mathcal{H} \subseteq \mathcal{B}$ of diameter at least $5L$. The maximum component diameter at time $t_\star - 1$ is at most $5L - 1$, so $\mathcal{H}$ can have diameter at most $2[(5L-1) + R + 1] \leq 11L$. In particular, if $\mathcal{H}$ contains more than one cycle, we are in scenario (i) and the conclusion follows.

If $\mathcal{H}$ is a tree, then Lemma 5.18 gives $\mathrm{BSP}'(A_{\mathcal{H}};\mathcal{H}) = \mathcal{H}$ where $A_{\mathcal{H}}$ is the restriction of A to $\mathcal{H}$. It follows from Lemma 5.15 (with $\mathcal{T} = \mathcal{H}$) that there is a subtree $T \subseteq \mathcal{H}$ with maximum degree at most four and

$$\mathrm{diam}\, T = \mathrm{diam}\, \mathcal{H} \in [5L, 11L],$$

which is $1/(80R)$-corrupted with respect to $(A_{\mathcal{H}}, \mathcal{H})$ — in particular, this gives that $T \cup B_R(A_{\mathcal{H}};\mathcal{H})$ is acyclic. Note however that the definition of the removal process implies $B_R(A_{\mathcal{H}};\mathcal{G}) \subseteq \mathcal{H}$, and so $T \cup B_R(A_{\mathcal{H}};\mathcal{H})$ is in fact the same as $T \cup B_R(A_{\mathcal{H}};\mathcal{G})$. It follows that T is also $1/(80R)$-corrupted with respect to $(A_{\mathcal{H}}, \mathcal{G})$, and hence also with respect to $(A, \mathcal{G})$. This proves that scenario (ii) holds if $\mathcal{H}$ is a tree.

Next we consider the case that $\mathcal{H}$ contains a single cycle C with $\mathrm{diam}\, C \geq 5L/2$. Let t' be the first time that the $\mathrm{BSP}''(A;\mathcal{G})$ process created any component $\mathcal{H}' \subseteq \mathcal{H}$ of diameter at least L. Then $t' < t$, and $\mathcal{H}'$ has diameter at most $2[(L-1) + R + 1] < 5L/2$. It follows that $\mathcal{H}'$ must be a tree. Lemma 5.18 gives $\mathrm{BSP}'(A_{\mathcal{H}'};\mathcal{H}') = \mathcal{H}'$. Arguing

as for the previous case, we apply Lemma 5.15 (now $\mathcal{T} = \mathcal{H}'$) to see that there is a subtree $T \subseteq \mathcal{H}'$ with maximum degree at most four and

$$\operatorname{diam} T = \operatorname{diam} \mathcal{H}' \in \left[L, \frac{5L}{2} \right],$$

which is $1/(80R)$-corrupted with respect to $(A_{\mathcal{H}'}, \mathcal{H}')$, and hence also with respect to $(A, \mathcal{G})$. This proves that scenario (ii) also holds in this case.

Finally, in the case that $\mathcal{H}$ contains a single cycle C with $\operatorname{diam} C < 5L/2$, we must be able to decompose

$$\mathcal{H} = \mathcal{T}' \cup \mathcal{H}''$$

where $C \subseteq \mathcal{H}''$, and $\mathcal{T}'$ is a tree that intersects $\mathcal{H}''$ at a single variable v, such that $\mathcal{T}'$ has depth exactly L. The conditions of Corollary 5.16 are then satisfied, so we conclude that there is a $1/(80R)$-corrupted subtree $T \subseteq \mathcal{T}'$ of maximum degree at most four, with $L \leq \operatorname{diam} T \leq \operatorname{diam} \mathcal{T}' \leq 2L$. Thus scenario (ii) again holds. □

5.4. **Probabilistic analysis of preprocessing.** First we use Lemma 5.5 to transfer the result of Corollary 5.9 (from §5.2) from the **PGW** measure to the **PGW**(T) measure of Definition 4.5, where T is any fixed sparse tree:

Corollary 5.19. *Let T be any fixed variable-rooted tree of maximum degree $\mathfrak{c}$, where $\mathfrak{c}$ is an absolute constant. Then*

$$\left[\mathbf{PGW}(T) \right] \left(v_{\mathrm{RT}} \text{ is improper or not } \mathfrak{i}\text{-excellent} \right) \leq \frac{1}{\exp(2^{k\delta_*/4} R/k)}$$

for $\mathfrak{i} \in \{0, 1\}$ and $k \geq k_0$, where k_0 is an absolute constant depending only on δ_* and $\mathfrak{c}$.

Proof. As in the proof of Lemma 5.5, let $\mathfrak{X}$ denote the Radon–Nikodym derivative of **PGW**(T) with respect to **PGW**. Let $\mathbb{E}$ denote expectation with respect to **PGW**. For any event E we can trivially bound

$$\mathbb{E}\left[\mathbf{1}_E \mathfrak{X} ; \mathfrak{X} \leq \frac{1}{\mathbf{PGW}(E)^{2/3}} \right] \leq \mathbf{PGW}(E)^{1/3}.$$

On the other hand we can use Lemma 5.5 together with Markov's inequality to bound

$$\mathbb{E}\left[\mathfrak{X} ; \mathfrak{X} > \frac{1}{\mathbf{PGW}(E)^{2/3}} \right]^2 \leq \mathbb{E}\left[\mathfrak{X}^2 \right] \mathbf{PGW}\left(\mathfrak{X} \geq \frac{1}{\mathbf{PGW}(E)^{2/3}} \right) \leq 2(\mathbb{E}\mathfrak{X}) \, \mathbf{PGW}(E)^{2/3} = 2\,\mathbf{PGW}(E)^{2/3}.$$

Combining the bounds gives

$$\left[\mathbf{PGW}(T) \right](E) = \mathbb{E}\left[\mathbf{1}_E \mathfrak{X} \right] \leq \mathbb{E}\left[\mathbf{1}_E \mathfrak{X} ; \mathfrak{X} \leq \frac{1}{\mathbf{PGW}(E)^{2/3}} \right] + \mathbb{E}\left[\mathfrak{X} ; \mathfrak{X} > \frac{1}{\mathbf{PGW}(E)^{2/3}} \right] \leq 3\,\mathbf{PGW}(E)^{1/3}.$$

The result then follows by combining with Corollary 5.9. □

We can use the above corollary to bound the probability, under the **PGW** measure, that the root lies in a sparse tree that is corrupt (see Definition 5.13) with respect to the set of non-excellent vertices. Recall (5.1) that $\Lambda_{\mathfrak{c},s}$ refers to $\mathfrak{c}$-sparse rooted trees having s variables. Also recall (5.2) that $\Lambda_{\mathfrak{c},s}(\mathcal{T})$ refers to $\mathfrak{c}$-sparse subtrees $v_{\mathrm{RT}} \in T \subseteq \mathcal{T}$ having s variables.

Lemma 5.20. *Let $\mathfrak{c}$ be an absolute constant. For $\mathcal{T} \sim \mathbf{PGW}$ let $A(\mathcal{T})$ denote the set of all variables in $\mathcal{T}$ that are improper or not $\mathfrak{i}$-excellent. It holds for all $s \geq 100R$ that*

$$\mathbf{PGW}\left(\begin{array}{c} \text{some } T \in \Lambda_{\mathfrak{c},s} \text{ is } 1/(80R)\text{-corrupt} \\ \text{with respect to } (\mathcal{T}, A(\mathcal{T})) \end{array} \right) \leq \frac{1}{\exp(2^{k\delta_*/4} s/k^2)}$$

provided $k \geq k_0$, where k_0 is an absolute constant depending only on δ_* and $\mathfrak{c}$.

Proof. This argument is similar to (but simpler than) part of the proof of Lemma 5.2. First we fix a tree $T' \in \Lambda_{\mathfrak{c},s}$, as well as a subset $B' \subseteq V_T$ with $|B'| \geq |V_T|/(80R)$ such that all variables in B' lie at pairwise distance greater than

$2R$. For this fixed pair (T', B') we bound (cf. (5.3), (5.5), and (5.6))

$$P(T', B') \equiv \mathbf{PGW}\left(\text{there exists an embedding } \zeta : T' \hookrightarrow \mathscr{T} \text{ such that } \zeta(B') \subseteq A(\mathscr{T})\right)$$

$$\leq \int \left|\left\{\text{embeddings } \zeta : T' \hookrightarrow \mathscr{T} \text{ such that } \zeta(B') \subseteq A(\mathscr{T})\right\}\right| d\mathbf{PGW}(\mathscr{T})$$

$$= \text{emb}(T') \cdot \left[\mathbf{PGW}(T')\right]\left(B' \subseteq A(\mathscr{T})\right) \leq (\alpha k^2)^{cs} \cdot \prod_{u \in B'} \left\{\left[\mathbf{PGW}(T_u)\right]\left(\text{root of } \mathscr{T}' \text{ is in } A(\mathscr{T}')\right)\right\},$$

where T_u now refers to the tree T' rerooted at u, $\mathscr{T}'$ refers to a sample from the measure $\mathbf{PGW}(T_u)$, and the probability factorizes over $u \in B'$ since "proper" and "i-excellent" are properties of the R-neighborhood, and we have assumed that variables in B' lie at pairwise distance greater than $2R$. Applying Corollary 5.19 (for the measures $\mathbf{PGW}(T_u)$) and summing over all possibilities of (T', B') gives

$$\sum_{T', B'} P(T', B') \leq \frac{e^{O(cks)}}{\exp(2^{k\delta_*/4} s/(80k))} \leq \frac{1}{\exp(2^{k\delta_*/4} s/k^2)} .$$

This implies the assertion of the lemma. $\square$

We next transfer the bound of Lemma 5.20 from the tree measure $\mathbf{PGW} \equiv \mathbf{PGW}^\alpha$ to the random k-SAT measure $\mathbb{P} \equiv \mathbb{P}^{n,\alpha}$ (as given in Definition 2.3, and including the random marking l_R of (3.2)). For $\mathscr{G} = (V, F, E) \sim \mathbb{P}$ and any $v \in V$, let $\mathbf{\Lambda}_{c,s}(\mathscr{G}, v)$ denote the set of subtrees $v \in T \subseteq \mathscr{G}$ having s variables and maximum degree bounded by c.

Lemma 5.21. *Let c be an absolute constant. For $\mathscr{G} = (V, F, E) \sim \mathbb{P} \equiv \mathbb{P}^{n,\alpha}$ let $A(\mathscr{G})$ denote the set of all variables in $\mathscr{G}$ that are improper or not i-excellent, for $i \in \{0, 1\}$. For any $v \in V$, and for all $100R \leq s \leq n^{1/10}$,*

$$\mathbb{P}\left(\begin{array}{c}\text{some } T \in \mathbf{\Lambda}_{c,s}(\mathscr{G}, v) \text{ is } 1/(80R)\text{-corrupt} \\ \text{with respect to } (\mathscr{G}, A(\mathscr{G}))\end{array}\right) \leq \frac{e^{O(cks)}}{\exp(2^{k\delta_*/4} s/(80k))} .$$

Proof. Fix a tree $T' \in \mathbf{\Lambda}_{c,s}$, as well as a subset $B' \subseteq V_{T'}$ with $|B'| \geq |V_{T'}|/(80R)$ such that all variables in B' lie at pairwise distance greater than $2R$. We consider the event

$$E_n(T', B') \equiv \left\{\begin{array}{c}\text{there exists an embedding } \zeta : T' \hookrightarrow \mathscr{G} \text{ such that} \\ \zeta(B') \subseteq A(\mathscr{G}) \text{ and } \zeta(T') \cup B_R(\zeta(B'); \mathscr{G}) \text{ is acyclic}\end{array}\right\} .$$

Write $P_n(T', B') \equiv \mathbb{P}(E_n(T', B'))$ and $P_{n,m}(T', B') \equiv \mathbb{P}_{n,m}(E_n(T', B'))$. Recalling (2.2), we can bound

$$P_n(T', B') \leq \sum_{m \geq 1} \mathbf{1}\left\{\frac{|m - n\alpha|}{n^{1/2} \ln n} \leq 1\right\} P_{n,m}(T', B') + \frac{1}{\exp\{(\ln n)^{3/2}\}}$$

— the last term accounts for the probability under $\mathbb{P}^{n,\alpha}$ that the total number of clauses in $\mathscr{G}$ (a $\text{Pois}(n\alpha)$ random variable) deviates from $n\alpha$ by more than $n^{1/2} \ln n$. We now fix any m satisfying

$$|m - n\alpha| \leq n^{1/2} \ln n ,$$

and consider $\mathscr{G}$ sampled from $\mathbb{P}_{n,m}$. Let ζ' denote any mapping that sends $V_{T'} \hookrightarrow [n]$, with $\zeta'(v_{\text{RT}}) = v$, and $F_{T'} \hookrightarrow [m]$; and let j denote a mapping $E_{T'} \to [k]$. We then write $\{\zeta' : T' \hookrightarrow_j \mathscr{G}\}$ for the event that the pair (ζ', j) is consistent with an actual embedding of T' into $\mathscr{G}$: this means that for every edge $e = (av) \in E_{T'}$, the edge $\zeta'(e) \equiv (\zeta'(a)\zeta'(v))$ is present in $E_{\mathscr{G}}$, and the index of this edge in the clause $\zeta'(a)$ is given by $j(e) \in [k]$. With this notation we can bound

$$P_{n,m}(T', B') \leq \sum_{\zeta', j} \mathbb{P}_{n,m}\left(\zeta' : T' \hookrightarrow_j \mathscr{G}\right) \mathbb{P}_{n,m}\left(\begin{array}{c}\zeta'(B') \subseteq A(\mathscr{G}) \text{ and} \\ \zeta'(T') \cup B_R(\zeta'(B'); \mathscr{G}) \text{ is acyclic}\end{array}\Big| \zeta' : T' \hookrightarrow_j \mathscr{G}\right). \tag{5.21}$$

In words, we interpret the above decomposition as follows. To sample $\mathscr{G} \sim \mathbb{P}_{n,m}$, we start with n variables and m clauses where each clause is equipped with k outgoing edges (indexed $j = 1, \ldots, k$). Each outgoing edge matches to a uniformly random variable, independently of all other edges — this means that we can sample $\mathscr{G}$ in a *sequential* way, revealing one edge at a time. In particular, when we fix a pair (ζ', j) and condition on the event $\{\zeta' : T' \hookrightarrow_j \mathscr{G}\}$,

it is equivalent to say that for all $a \in F_{T'}$ and all $e = (av) \in \delta a$, we reveal that the $j(e)$-th edge incident to clause $\zeta'(a)$ matches to the variable $\zeta'(v)$. Thus

$$\mathbb{P}_{n,m}\left(\zeta' : T' \hookrightarrow_j \mathscr{G}\right) = \frac{1}{n^{|E_{T'}|}},$$

and conditioning on the event $\{\zeta' : T' \hookrightarrow_j \mathscr{G}\}$ reveals nothing about the remaining $mk - |E_{T'}|$ edges in the graph $\mathscr{G}$. Note also that the total number of pairs (ζ', j') is upper bounded by $n^{|V_{T'}|-1} m^{|F_{T'}|} k^{|E_{T'}|}$.

Now fix (ζ', j) and condition on the event $\{\zeta' : T' \hookrightarrow_j \mathscr{G}\}$. We next want to explore the R-neighborhoods of the variables in $\zeta'(B')$. Note that this can also be done sequentially, in a "breadth-first search" manner: first take any $u \in \zeta'(B')$, and reveal all the edges incident to it. The degree of u is then

$$|\delta u| \sim |\delta_{\zeta'(T')}u| + \text{Bin}\left(mk - |E_{T'}|, \frac{1}{n}\right),$$

where $|\delta_{\zeta'(T')}u|$ is the degree of u in the subgraph $\zeta'(T')$. We then proceed to reveal the edges incident to the neighboring clauses of u, and so on until we have explored the entire subgraph

$$\mathscr{H} \equiv \zeta'(T') \cup B_R(\zeta'(B')).$$

If at any point in the exploration we reveal a cycle inside $\mathscr{H}$, we can simply stop because this means the event of interest does not occur. It is extremely unlikely to reveal any variable degree larger than $(\ln n)^2$, so if this occurs we also simply halt the exploration. Restricted to the event that no cycle is formed and all revealed variable degrees are at most $(\ln n)^2$, the law of $\mathscr{H}$ can be bounded in terms of the $\mathbf{PGW}^\alpha$ law: more precisely, we claim that

$$\mathbb{P}_{n,m}\left(\left\{\zeta'(B') \subseteq A(\mathscr{G})\right\} \cap \left\{\begin{array}{c} \text{exploration of } \mathscr{H} \text{ completes} \\ \text{without revealing any cycle or} \\ \text{variable of degree larger than } (\ln n)^2 \end{array}\right\}\right)$$

$$\leq 2 \prod_{u \in B'} \left\{\left[\mathbf{PGW}(T_u)\right]\left(\text{root of } \mathscr{T}' \text{ is in } A(\mathscr{T}')\right)\right\}, \tag{5.22}$$

where T_u is the tree T' rerooted at u. The bound (5.22) can be justified by noting that for all integers m and j with $|m - n\alpha| \leq n^{1/2} \ln n$ and all $0 \leq j \leq (\ln n)^2$, we have

$$\frac{\mathbb{P}(\text{Bin}(mk, 1/n) = j)}{\mathbb{P}(\text{Pois}(\alpha k) = j)} = \binom{mk}{j} \frac{1}{n^j}\left(1 - \frac{1}{n}\right)^{mk-j} \frac{e^{\alpha k} j!}{(\alpha k)^j} \leq \left(\frac{n}{n-1}\right)^j = \exp\left\{O\left(\frac{(\ln n)^2}{n}\right)\right\}.$$

If the exploration of $\mathscr{H}$ completes without revealing any variable of degree larger than $(\ln n)^2$, then the total number of variables explored is upper bounded by $n^{1/9}$ (for n large enough), so the joint distribution of degrees will be close to i.i.d. $\text{Pois}(\alpha k)$ random variables, and this gives (5.22). Applying Corollary 5.19 and summing over (ζ', j) gives

$$P_{n,m}(T', B') \leq \frac{n^{|V_{T'}|-1} m^{|F_{T'}|} k^{|E_{T'}|}}{n^{|E_{T'}|} \exp(2^{k\delta_*/4} s/(80k))} = \frac{\alpha^{|F_{T'}|} k^{|E_{T'}|}}{\exp(2^{k\delta_*/4} s/(80k))} \leq \frac{e^{O(cks)}}{\exp(2^{k\delta_*/4} s/(80k))}.$$

Finally, summing over all (T', B') proves the claim. $\qquad \square$

Lemma 5.22. *If* $1 \leq L \leq (\ln n)/k^2$ *and* $\gamma \leq \exp(-k^2 L)/k$, *then*

$$\mathbb{P}\left(\begin{array}{c} \text{there exists a subset } S \subseteq V \text{ with} \\ |S| = n\gamma \text{ and } |B_L(S)| \geq n\gamma \exp(k^2 L) \end{array}\right) \leq \frac{1}{\exp(ne^{kL}\gamma)}.$$

Proof. As in the proof of Lemma 5.21, we fix m satisfying (5.21) and consider $\mathscr{G} \sim \mathbb{P} \equiv \mathbb{P}_{n,m}$. We then fix a subset $S \subseteq V$ of size $|S| = n\gamma$, and explore its neighborhood by breadth-first search. Let $Z_\ell(S)$ denote the number of variables at distance exactly ℓ from S. Note that

$$\mathbb{E}\left(\exp\left\{\frac{Z_\ell(S)}{(\alpha k^2)^\ell}\right\} \Big| B_{\ell-1}(S)\right) \leq \mathbb{E}\left(\exp\left\{\frac{\text{Bin}\left(mk, \frac{Z_{\ell-1}(S)}{n - |B_{\ell-1}(S)|}\right)}{(\alpha k^2)^\ell}\right\} \Big| B_{\ell-1}(S)\right). \tag{5.23}$$

Let F_ℓ be the event that $|Z_\ell(S)| \le n\gamma \exp(k^{3/2}L)$, and let $E_\ell \equiv F_1 \cap \cdots \cap F_\ell$. On the event $E_{\ell-1}$ the neighborhood $B_{\ell-1}(S)$ cannot be too large (crudely, by the assumption on γ, it contains less than $1/k$ fraction of V), and it follows from (5.23) that

$$\mathbb{E}\left(\exp\left\{ \frac{Z_\ell(S)}{(\alpha k^2)^\ell} \right\}; E_{\ell-1} \right) \le \mathbb{E}\left(\exp\left\{ \frac{Z_{\ell-1}(S)O(1/k)}{(\alpha k^2)^{\ell-1}} \right\}; E_{\ell-1} \right) \le \mathbb{E}\left(\exp\left\{ \frac{Z_{\ell-1}(S)}{(\alpha k^2)^{\ell-1}} \right\}; E_{\ell-1} \right) \le e,$$

where the last bound holds by induction on ℓ. It then follows by Markov's inequality that

$$\mathbb{P}\left(E_{\ell-1} \setminus F_\ell \right) \le \mathbb{P}\left(\exp\left\{ \frac{Z_\ell(S)}{(\alpha k^2)^\ell} \right\} \ge \exp\left\{ \frac{n\gamma \exp(k^{3/2}L)}{(\alpha k^2)^\ell} \right\}; E_{\ell-1} \right) \le \frac{e}{\exp(n\gamma \exp(k^{3/2}L/2))}.$$

Summing the last bound over $1 \le \ell \le L$ gives

$$\mathbb{P}\left((E_L)^c \right) = \sum_{\ell=1}^{L} \mathbb{P}\left(E_{\ell-1} \setminus F_\ell \right) \le \frac{Le}{\exp(n\gamma \exp(k^{3/2}L/2))} \le \frac{1}{\exp(n\gamma \exp(k^{3/2}L/3))}.$$

The claim follows by enumerating over at most $2^{n\gamma}$ choices for the subset $S \subseteq V$. $\square$

Proof of Proposition 3.22. Fix m satisfying (5.21) and consider $\mathcal{G} \sim \mathbb{P} \equiv \mathbb{P}_{n,m}$. For $v \in V$ we let Y_v be the indicator that v lies within distance $10R$ of a variable that is removed during processing, i.e.,

$$Y_v = \mathbf{1}\left\{ v \in B_{10R}\left(\mathrm{BSP}'(A; \mathcal{G}); \mathcal{G} \right) \right\}$$

where A is the set of all variables that are improper or not 1-good in $\mathcal{G}$ (Definition 3.15). Let

$$\bar{Y}_v = \mathbf{1}\left\{ \begin{array}{c} v \text{ lies within distance } 10R \text{ of a connected component of} \\ \mathrm{BSP}'(A; \mathcal{G}) \text{ having diameter at most } L_{\mathrm{BSP}'} \equiv (\ln n)/2^{k\delta_*/5} \end{array} \right\} \le Y_v.$$

We first argue that, with high probability, $Y_v = \bar{Y}_v$ for all $v \in V$. Indeed, Proposition 5.14 implies that if $\mathrm{BSP}'(A; \mathcal{G})$ has a connected component of diameter more than $L_{\mathrm{BSP}'}$, then either (i) there is a connected subgraph $B' \subseteq \mathcal{G}$ with $\mathrm{diam}\, B' \le 11 L_{\mathrm{BSP}'}/5$ that contains more than one cycle, or (ii) there is a $1/(80R)$-corrupted subtree $T' \subseteq \mathcal{G}$ that has maximum degree at most four and $\mathrm{diam}\, T \ge L_{\mathrm{BSP}'}/5$. Thus

$$\mathbb{P}\left(\begin{array}{c} Y_v \ne \bar{Y}_v \\ \text{for any } v \in V \end{array} \right) \le \sum_{v \in V} \left\{ \mathbb{P}\left(\begin{array}{c} B_{11 L_{\mathrm{BSP}'}/5}(v) \text{ contains} \\ \text{more than one cycle} \end{array} \right) + \mathbb{P}\left(\begin{array}{c} \text{some } T \in \mathbf{\Lambda}_{4, L_{\mathrm{BSP}'}/5}(\mathcal{G}, v) \\ \text{is } 1/(80R)\text{-corrupt} \end{array} \right) \right\}$$

$$\overset{\odot}{\le} n\left\{ \left(\frac{\exp(5 L_{\mathrm{BSP}'} k^2)}{n} \right)^2 + \frac{1}{\exp\{2^{k\delta_*/4} L_{\mathrm{BSP}'}/k^2\}} \right\} = o_n(1),$$

where the bound marked $\odot$ follows from Lemma 5.21 together with very crude bounds on the chance to see more than one cycle in a neighborhood of v of radius $5 L_{\mathrm{BSP}'}$. This proves that $Y = \bar{Y}$ with high probability where

$$Y \equiv \sum_{v \in V} Y_v, \quad \bar{Y} \equiv \sum_{v \in V} \bar{Y}_v.$$

We now turn to estimating $\bar{Y}$. Let $B'(\ell)$ be the union of all connected components of $\mathrm{BSP}'(A; \mathcal{G})$ of diameter ℓ that contain more than one cycle; and let $T'(\ell)$ be the union of all subtrees of $\mathrm{BSP}'(A; \mathcal{G})$ of diameter at least ℓ that are $1/(80R)$-corrupted. Let $A'(\ell) \equiv B'(\ell) \cup T'(\ell/5)$. Proposition 5.14 implies that if $\bar{Y}_v = 1$, then either v lies within distance $2010R$ from A, or v lies within distance $\ell + 10R$ from $A'(\ell)$ for some $2000R \le \ell \le L_{\mathrm{BSP}'}$. Therefore

$$\mathbb{E}\bar{Y} \le \mathbb{E}\left| B_{2010R}(A) \right| + \sum_{\ell=2000R}^{L_{\mathrm{BSP}'}} \mathbb{E}\left| B_{\ell+10R}(A'(\ell)) \right|. \tag{5.24}$$

Let $\gamma = \exp(-2^{k\delta_*/4} R/(2k))$, and note that Corollary 5.19 together with Markov's inequality gives

$$\mathbb{P}\left(|A| \ge n\gamma \right) \le \frac{\mathbb{E}A}{n\gamma} = \frac{\mathbb{P}(v \in A)}{\gamma} \le \frac{\exp(2^{k\delta_*/4} R/(2k))}{\exp(2^{k\delta_*/4} R/k)} = \gamma.$$

Combining with Lemma 5.22 gives, for $L = 2010R$ and $\gamma = \exp(-2^{k\delta_*/4}R/(2k))$,

$$\mathbb{E}\left|B_{2010R}(A)\right| \le n\gamma \exp(k^2 L) + n\mathbb{P}\left(|A| \ge n\gamma\right) + n\mathbb{P}\left(\begin{array}{c} \text{there exists } S \subseteq V \text{ with } |S| = n\gamma \\ \text{and } |B_L(S)| \ge n\gamma\exp(k^2 L) \end{array}\right)$$

$$\le \frac{n\exp(2010k^2R)}{\exp(2^{k\delta_*/4}R/(2k))} + \frac{n}{\exp(2^{k\delta_*/4}R/(2k))} + \frac{n}{\exp(ne^{kL}\gamma)} \le \frac{n}{\exp(2^{k\delta_*/4}R/(3k))} \, . \tag{5.25}$$

Similarly, if we take $\gamma_\ell = \exp(-2^{k\delta_*/4}\ell/k^2)$, then Lemma 5.21 together with Markov's inequality gives

$$\mathbb{P}(|A'(\ell)| \ge n\gamma_\ell) \le \frac{\mathbb{E}|A'(\ell)|}{n\gamma_\ell} \le \frac{\mathbb{P}(v \in A'(\ell))}{\gamma_\ell} \le o_n(1) + \frac{\exp(2^{k\delta_*/4}\ell/k^2)}{\exp(2^{k\delta_*/4}\Omega(\ell/k))} \le \gamma_\ell \, ,$$

where the $o_n(1)$ error term accounts for the probability that $B_\ell(v)$ contains more than one cycle. Combining with Lemma 5.22 gives

$$\mathbb{E}\left|B_{\ell+10R}(A'(\ell))\right| \le n\gamma_\ell \exp(k^2(\ell + 10R)) + n\mathbb{P}\left(|A'(\ell)| \ge n\gamma_\ell\right) + n\mathbb{P}\left(\begin{array}{c} \text{there exists } S \subseteq V \text{ with } |S| = n\gamma_\ell \text{ and} \\ |B_{\ell+10R}(S)| \ge n\gamma_\ell\exp(k^2(\ell + 10R)) \end{array}\right)$$

$$\le \frac{n\exp(k^2(\ell + 10R))}{\exp(2^{k\delta_*/4}\ell/k^2)} + \frac{n}{\exp(2^{k\delta_*/4}\ell/k^2)} + \frac{n}{\exp(ne^{k(\ell+10R)}\gamma_\ell)} \le \frac{n}{\exp(2^{k\delta_*/4}\ell/k^3)} \, , \tag{5.26}$$

for all $\ell \ge 2000R$. Substituting (5.25) and (5.26) into (5.24) gives

$$\mathbb{E}\bar{Y} \le \frac{n}{\exp(2^{k\delta_*/4}R/k^4)} \, .$$

Next note that under $\mathbb{P}_{n,m}$, for any $u \ne v$ the radius-$(L_{\mathrm{BSP}'} + 10R)$ neighborhoods of u and v intersect with chance $o_n(1)$, and so are nearly independent of one another. If follows that $\mathrm{Cov}(\bar{Y}_u, \bar{Y}_v) = o_n(1)$. Therefore

$$\mathrm{Var}\,\bar{Y} \le n + \sum_{u \ne v} \mathrm{Cov}(\bar{Y}_u, \bar{Y}_v) = o_n(n^2) \, .$$

We can then use Chebychev's inequality to conclude that

$$\mathbb{P}\left(\bar{Y} \ge \frac{n}{\exp(2^{k\delta_*/5}R)}\right) \le \mathbb{P}\left(\bar{Y} \ge 2\mathbb{E}\bar{Y}\right) \le \frac{\mathrm{Var}\,\bar{Y}}{(\mathbb{E}\bar{Y})^2} = o_n(1) \, .$$

Since we already argued above that $\mathbb{P}(Y \ne \bar{Y}) = o_n(1)$, the result follows. $\qquad \square$

5.5. Combinatorial analysis for positive type fractions.

We now turn to the proof of Proposition 3.23. In this subsection we prove a deterministic result, Corollary 5.27, which says essentially that if the local neighborhood of a clause a in $\mathscr{G}$ satisfies certain properties (to be detailed below), then we can guarantee that a has a certain total type in the processed graph $\mathrm{pr}\,\mathscr{G}$.

In preparation for this result, we recall some definitions and notations. As in Definition 3.19, we denote a clause total type as $\boldsymbol{L} \equiv (L_0, L_\infty)$ where L_0 is the initial clause type and L_∞ is the final clause type. The initial type L_0 corresponds to the $(R + 1/2)$-neighborhood of the clause. If the clause does not lie in a compound enclosure (Definition 3.16), then the final type L_∞ corresponds to the $(R + 1/2)$-neighborhood of the clause in the processed graph. If the clause does lie in a compound enclosure U, then L_∞ instead encodes the structure of $B_R(U)$. Recall also from Definition 3.16 that any compound enclosure has diameter at most $R/100$, so in any case L_∞ has depth at most $R(1 + 1/100)$. Of course, we restrict our attention to types $\boldsymbol{L}$ that can actually occur, as formalized by the following:

Definition 5.23 (feasible types). A clause total type $\boldsymbol{L} \equiv (L_0, L_\infty)$ is termed **feasible** if there exists some bipartite factor graph $\mathscr{H}$, with girth larger than $8R$, such that $\mathbb{P}_{n',n'\alpha'}(\mathscr{H}) > 0$ for some n', α', and some clause $a_\star \in \mathrm{pr}\,\mathscr{H}$ has total type $\boldsymbol{L}$. Thus L_0 and L_∞ are both trees, which we regard as being rooted at $a_\star$.

Given a feasible type $\boldsymbol{L}$, we now construct the tree $\mathscr{T}(\boldsymbol{L})$ (see Figure 5 for a schematic depiction):

Definition 5.24 (tree based on feasible type). Given a feasible clause total type $\boldsymbol{L}$, we now construct a corresponding tree $\mathscr{T}(\boldsymbol{L})$ as follows. Among all pairs $(\mathscr{H}, a_\star)$ for which the conditions of Definition 5.23 hold, fix one such that $\mathscr{H}$ has minimal size. Let $\mathscr{T}_\star$ be the $(4R)$-neighborhood of $a_\star$ in $\mathscr{H}$, which we regard as a tree rooted at $a_\star$. We regard L_0, L_∞ as subtrees of $\mathscr{T}_\star$, and make the following definitions:

a. Let V_1 denote the set of variables of $L_0 \setminus L_\infty$ which neighbor some clause in L_∞. For each $v \in V_1$, let $\mathscr{T}_\star(v, \ell)$ be the subtree of $\mathscr{T}_\star$ induced by all the descendants of v that lie within distance ℓ of v. Let $Y_{\text{out}}(v)$ be the variables in $\mathscr{T}_\star(v)$ that lie at depth exactly R below v — we will argue in the proof of Corollary 5.27 below that $Y_{\text{out}}(v)$ is nonempty for all $v \in V_1$. We take a union over V_1 to define

$$\mathscr{T}_{\star,\text{out}} \equiv \bigcup_{v \in V_1} \mathscr{T}_\star(v, 2R), \quad \mathscr{N}_{\text{out}} \equiv \bigcup_{v \in V_1} \mathscr{T}_\star(v, R-1), \quad Y_{\text{out}} \equiv \bigcup_{v \in V_1} Y_{\text{out}}(v). \tag{5.27}$$

For each $y \in Y_{\text{out}}$, let $Z_{\text{out}}(y)$ be the variables that lie at depth exactly R below y. If $Z_{\text{out}}(y)$ is nonempty, then we fix an arbitrary $w \in Z_{\text{out}}(v)$, and modify w by redefining $l(w)$ to be the same as $l(v)$ where v is the ancestor of y in V_1. We let $\mathscr{T}_{\text{out}}$ denote the resulting modification of $\mathscr{T}_{\star,\text{out}}$ (the two graphs can differ only in the markings l at the last level).

b. Now let $\mathscr{T}_{\star\star}$ be the component of $\mathscr{T}_\star \setminus \mathscr{T}_{\text{out}}$ that contains the root clause $a_\star$. Let W denote the variables at the boundary of L_∞. From the definition of the processing algorithm, each variable left in pr $\mathscr{H}$ is 1-good, which means (Definition 3.14) that any length-$(2R/5)$ path emanating from the variable must contain at least one variable that is 1-excellent. This in turn means we can find a set X of 1-excellent variables that lie below W, at distance at most $2R/5$ from W, such that X forms a cutset in $\mathscr{T}_{\star\star}$ that encloses L_∞. Let Y be the variables in $\mathscr{T}_{\star\star}$ that lie at depth exactly $R/10$ below X. Let $\mathscr{T}_{\text{in}}$ be the subtree of $\mathscr{T}_{\star\star}$ enclosed by Y.

Let $\mathscr{T}(L)$ be the subtree of $\mathscr{T}_\star$ induced by the union of $\mathscr{T}_{\text{in}}$ and $\mathscr{T}_{\text{out}}$.
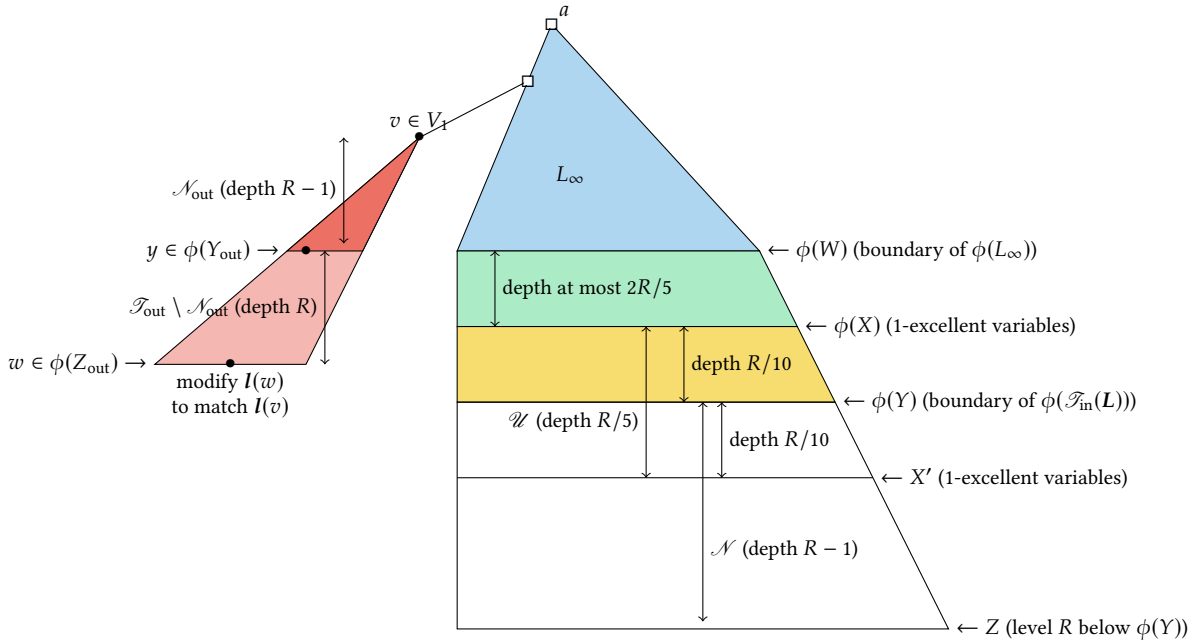
FIGURE 5. Depiction of the events $E_a(L)$ and $G_a(L)$ (Definition 5.25). The shaded region shows $\phi(\mathscr{T}(L))$ where L is a feasible type (Definition 5.23), $\mathscr{T}(L)$ is the tree based on L (Definition 5.24), and ϕ embeds $\mathscr{T}(L)$ into a neighborhood of a on the event $E_a(L)$. The unshaded region points out the main features of the event $G_a(L)$.

We next define an event E_a which says, roughly, that the neighborhood of a in $\mathscr{G}$ looks like $\mathscr{T}(L)$ and does not contain cycles of length $\leq 8R$. We then define events G_a and K_a which capture additional desirable properties.

Definition 5.25 (events E_a, G_a, K_a). Fix a feasible type L, and let $\mathscr{T}(L)$ be as given by Definition 5.24, rooted at clause $a_\star$. Let $\mathscr{G} = (V, F, E)$ be sampled from $\mathbb{P}_{n,m}$. For any clause $a \in F$, we let $E_a \equiv E_a(L)$ be the event that $\mathscr{T}' \equiv B_{4R}(a; \mathscr{G})$ is a tree, and there is an embedding $\phi : \mathscr{T}(L) \hookrightarrow \mathscr{T}'$ that maps $a_\star \mapsto a$, and satisfies the following:

(i) For any variable $v \in \phi(\mathscr{T}(L)) \subseteq \mathscr{T}'$, either all or none of its child variables (with respect to $\mathscr{T}'$) lie in $\phi(\mathscr{T}(L))$.

(ii) For any variable $v \in V_1$, the subtree of v in $\mathscr{T}(L)$ agrees with the depth-$2R$ subtree of $\phi(v)$ in $\mathscr{T}'$.

If E_a occurs, let X' be the cutset of variables in $\mathscr{T}'$ lying at distance exactly $R/10$ below $\phi(Y)$. Let Z be the cutset of variables in $\mathscr{T}'$ lying at distance exactly R below $\phi(Y)$. Let $\mathscr{U}$ be the subgraph of $\mathscr{T}'$ that lies sandwiched between $\phi(X)$ and X' (inclusive). Let $\mathscr{N}$ be the subgraph of $\mathscr{T}'$ that lies sandwiched between $\phi(Y)$ and Z, including $\phi(Y)$ but not including Z. Let $G_a \equiv G_a(L)$ denote the event that $\mathscr{G} \in E_a$, and moreover satisfies the following:

(I) All variables in $\phi(\mathscr{T}_{\text{in}}) \cup \mathscr{N}$ are proper;
(II) $l(u) \neq l(w)$ for any $u \in \phi(\mathscr{T}_{\text{out}})$ and $w \in \phi(\mathscr{T}_{\text{in}}) \cup \mathscr{N}$;
(III) All variables in $\mathscr{U}$ are 1-fair;
(IV) All variables in X' are 1-excellent.

If E_a occurs, then we also define the subgraph

$$\mathscr{G}' \equiv \mathscr{G} \setminus \phi\left(\mathscr{T}_{\text{in}} \setminus Y\right). \tag{5.28}$$

We then let $K_a \equiv K_a(L)$ be the event that $\mathscr{G} \in E_a$, and every variable in $\mathscr{N} \cup Z$ survives in $\operatorname{pr}\mathscr{G}'$.

Lemma 5.26. *Suppose that L is a feasible clause type in the sense of Definition 5.23. Suppose $\mathscr{G}$ has girth greater than $8R$ and belongs to the event $G_a \cap K_a$, and let $\mathscr{G}'$ be as defined by (5.28). Let A, A' be the variables that are improper or not 1-good with respect to $\mathscr{G}, \mathscr{G}'$ respectively. Then $A \subseteq A' \subseteq \mathscr{G}' \setminus \mathscr{N}$ and $A' \setminus A = \phi(\mathscr{N}_{\text{out}}) \cap A'$.*

Proof. Throughout the proof we fix $\mathscr{H}$ as in Definition 5.23, and let $\mathscr{T}(L)$ be the tree given by Definition 5.24. Let $A_{\mathscr{H}}$ denote the set of all variables that are improper or not 1-good with respect to $\mathscr{H}$. Then, with $\mathscr{N}_{\text{out}}$ as defined by (5.27), we first observe that

$$A_{\mathscr{H}} \cap \left(\mathscr{T}_{\text{in}} \cup \mathscr{N}_{\text{out}}\right) = \varnothing \tag{5.29}$$

— this is because the R-neighborhood of each variable in $A_{\mathscr{H}}$ is removed in the first step of processing on $\mathscr{H}$, but $\mathscr{T}_{\text{in}}$ is assumed to survive in $\operatorname{pr}\mathscr{H}$.

We now turn to the comparison of A and A'. To begin, recall that one way for a variable to be improper is that its R-neighborhood contains a cycle — however, by the girth assumption, this never happens in $\mathscr{G}$ or $\mathscr{G}'$. It remains to consider all the ways for an acyclic variable to belong to A or A':

- Let $A(R)$ denote the variables $v \in A$ for which $B_R(v; \mathscr{G})$ contains a repeat marking — i.e., two variables $u \neq w$ with the same marking $l(u) = l(w)$. Define similarly $A'(R) \subseteq A'$, and note $A'(R) \subseteq A(R)$ since the R-neighborhood of any variable relative to $\mathscr{G}'$ is a subgraph of its R-neighborhood relative to $\mathscr{G}$. Let

$$\mathscr{G}'' \equiv \mathscr{G}' \setminus \left(\mathscr{N} \cup \phi(\mathscr{N}_{\text{out}})\right),$$

that is to say, $\mathscr{G}''$ is the subgraph of $\mathscr{G}$ induced by variables lying at distance at least R from $\phi(\mathscr{T}_{\text{in}})$. If variable v lies in $\mathscr{G}''$, then $B_R(v; \mathscr{G}) = B_R(v; \mathscr{G}')$, and so

$$A(R) \cap \mathscr{G}'' = A'(R) \cap \mathscr{G}''.$$

If v lies in $\mathscr{G} \setminus \mathscr{G}''$, we distinguish two cases:

○ If $v \in \phi(\mathscr{T}_{\text{in}}) \cup \mathscr{N}$, property (I) implies that v must be proper in $\mathscr{G}$, so $v \notin A(R)$.
○ If $v \in \phi(\mathscr{N}_{\text{out}})$, it follows from (5.29) that $v' \equiv \phi^{-1}(v)$ must be proper with respect to $\mathscr{H}$, so $B_R(v'; \mathscr{H})$ has no repeat marking. We claim that the same holds for $B_R(v; \mathscr{G})$. Indeed, suppose for contradiction that in $B_R(v; \mathscr{G})$ there are two variables $u \neq w$ with $l(u) = l(w)$.
  – If both u, w lie in $\phi(\mathscr{T}_{\text{out}})$, then we have two variables $\phi^{-1}(u) \neq \phi^{-1}(w)$ inside $B_R(v'; \mathscr{H})$ with the same marking, contradicting the above observation.
  – If both u, w lie in $\phi(\mathscr{T}_{\text{in}}) \cup \mathscr{N}$, then it must be possible to join them by a path *inside* $\phi(\mathscr{T}_{\text{in}}) \cup \mathscr{N}$ of length at most $2R - 1$. It follows that $u, w \in B_R(\bar{v}; \mathscr{G})$ for $\bar{v} \in \phi(\mathscr{T}_{\text{in}}) \cup \mathscr{N}$, which contradicts property (I).
  – If $u \in \phi(\mathscr{T}_{\text{out}})$ while $w \in \phi(\mathscr{T}_{\text{in}}) \cup \mathscr{N}$, we must have $l(u) \neq l(w)$ by property (II).

The above shows that $A(R)$ cannot intersect $\mathscr{G} \setminus \mathscr{G}''$. Since $A'(R) \subseteq A(R)$, we obtain that $A'(R)$ also cannot intersect $\mathscr{G} \setminus \mathscr{G}''$. Combining with the earlier observation gives

$$A(R) = A'(R) \subseteq \mathscr{G}''. \tag{5.30}$$

- Let $A(1), A'(1)$ denote the variables that fail to be 1-good with respect to $\mathscr{G}, \mathscr{G}'$ respectively. Recall that any variable in $\mathscr{G}$ or $\mathscr{G}'$ is acyclic, so whether it is 1-good depends only on its $(2R/5)$-neighborhood. If v lies in $\mathscr{G}$ at distance at least $2R/5$ from $\phi(\mathscr{T}_{\text{in}})$, then it has the same $(2R/5)$-neighborhood in both $\mathscr{G}$ and $\mathscr{G}'$, so $v \in A(1)$ if and only if $v \in A'(1)$. For v lying at distance less than $2R/5$ from $\phi(\mathscr{T}_{\text{in}})$, we distinguish three cases:
  - First suppose that v lies either in $\phi(\mathscr{N}_{\text{out}})$, or in the part of $\phi(\mathscr{T}_{\text{in}})$ above $\phi(X)$. Then it follows from (5.29) that $v' \equiv \phi^{-1}(v)$ must be 1-good (hence 1-fair) in $\mathscr{H}$. Since being 1-fair is a property of the $(R/10)$-neighborhood, and $B_{R/10}(v; \mathscr{G}) \cong B_{R/10}(v'; \mathscr{H})$, we see that v is 1-fair in $\mathscr{G}$. Now consider a path γ of length $2R/5$ emanating from v:
    - Suppose γ never intersects $\phi(X)$. Since $v' = \phi^{-1}(v)$ is 1-good in $\mathscr{H}$, the path $\phi^{-1}(\gamma)$ must contain a variable $u' = \phi^{-1}(u)$ which is 1-excellent in $\mathscr{H}$. Since being 1-excellent is a property of the $(R/10)$-neighborhood, and $B_{R/10}(u; \mathscr{G}) \cong B_{R/10}(u'; \mathscr{H})$, we see that $u \in \gamma$ is 1-excellent in $\mathscr{G}$.
    - Otherwise, γ contains a variable $u \in \phi(X)$. From the construction (Definition 5.24), the variable $u' = \phi^{-1}(u)$ is 1-excellent in $\mathscr{H}$. By the same reasoning as in the last case, u must then be 1-excellent in $\mathscr{G}$.
    This proves that v is 1-good in $\mathscr{G}$, i.e., $v \notin A(1)$.
  - Next suppose v lies in $\mathscr{U}$ (i.e., between $\phi(X)$ and X', inclusive). Then v is 1-fair in $\mathscr{G}$ simply by property (III). A path of length $2R/5$ emanating from v must contain a variable u from either $\phi(X)$ or X'. If $u \in \phi(X)$, then it is 1-excellent in $\mathscr{G}$ as argued above. If $u \in X'$, then it is 1-excellent in $\mathscr{G}$ by property (IV). Thus, $v \notin A(1)$.
  - It remains to consider the case that v lies in $\mathscr{N} \setminus \mathscr{U}$. For any variable in $\mathscr{G}$ at distance at least $R/10$ from $\phi(\mathscr{T}_{\text{in}})$, the $(R/10)$-neighborhoods in $\mathscr{G}, \mathscr{G}'$ are the same, so that variable is 1-fair in $\mathscr{G}$ if and only if it is 1-fair in $\mathscr{G}'$; the same applies to the 1-excellent property. In particular, this tells us that $v \in \mathscr{N} \setminus \mathscr{U}$ is 1-fair in $\mathscr{G}$ if and only if it is 1-fair in $\mathscr{G}'$. If a path of length $2R/5$ emanates from v and never intersects X', then it must stay at least distance $R/10$ away from $\phi(\mathscr{T}_{\text{in}})$, so the path contains a 1-excellent variable of $\mathscr{G}$ if and only if it contains a 1-excellent variable of $\mathscr{G}'$. Otherwise, the path contains a variable from X' which is 1-excellent in both $\mathscr{G}$ and $\mathscr{G}'$ by property (IV). This proves that $A(1) \cap (\mathscr{N} \setminus \mathscr{U}) = A'(1) \cap (\mathscr{N} \setminus \mathscr{U})$.
  It follows from the above that $A(1) \subseteq A'(1)$ and

$$A(1) \cap \phi(\mathscr{N}_{\text{out}}) = \varnothing . \tag{5.31}$$

Now note that A' is by definition a subset of $\mathscr{G}'$, and on the event K_a it must not intersect $\mathscr{N}$. Thus, on the event K_a, we have $A(1) \subseteq A'(1) \subseteq \mathscr{G}' \setminus \mathscr{N}$. We also noted that variables at distance at least $2R/5$ from $\phi(\mathscr{T}_{\text{in}})$ belong in $A(1)$ if and only if they belong in $A'(1)$, so in particular $A(1) \cap \mathscr{G}'' = A'(1) \cap \mathscr{G}''$. It follows that

$$A'(1) \setminus A(1) \subseteq (\mathscr{G}' \setminus \mathscr{N}) \Big\backslash \mathscr{G}'' = \phi(\mathscr{N}_{\text{out}}) .$$

We already saw in (5.30) (without appealing to K_a) that $A(R) = A'(R) \subseteq \mathscr{G}''$. Combining with (5.31) finishes the proof of the lemma, since $A = A(R) \cup A(1)$ and $A' = A'(R) \cup A'(1)$. $\qquad \square$

Corollary 5.27. *Suppose that $L \equiv (L_0, L_\infty)$ is a feasible clause type in the sense of Definition 5.23. Let G_a, K_a be the events from Definition 5.25. If $\mathscr{G}$ has girth greater than $8R$ and belongs to the event $G_a \cap K_a$, then the clause a has total type $L_a = L$.*

Proof. As before, we fix $\mathscr{H}$ as in Definition 5.23, and let $\mathscr{T}(L)$ be the tree given by Definition 5.24. We first argue that (using the same notation as in Definition 5.24) we have

$$Y_{\text{out}}(v) \cap A \neq \varnothing \quad \text{for all } v \in V_1 . \tag{5.32}$$

We separate this into two cases:

- If $Z_{\text{out}}(y) \neq \varnothing$ for some $y \in Y_{\text{out}}(v)$, then recall from Definition 5.24 that we choose some $w \in Z_{\text{out}}(y)$ and set $l(w) = l(y)$, which ensures $y \in A(R) \subseteq A$.
- Now suppose that we do not have $Z_{\text{out}}(y) \neq \varnothing$ for any $y \in Y_{\text{out}}(v)$. Let a' be the parent clause of v in $\mathscr{T}(L)$, so a' lies in $\mathscr{T}_{\text{in}}$. Since the preprocessing algorithm on $\mathscr{H}$ removes v but leaves a' behind, it must be that at some stage of the algorithm a removal is triggered by a vertex u which lies at depth exactly R below v. (In particular, $u \in Y_{\text{out}}(v)$ which proves that $Y_{\text{out}}(v)$ must be nonempty.) We now argue that in fact there must be some $u \in Y_{\text{out}}(v) \cap A$. If not, then u triggers a removal *after* the initial stage of $\mathrm{BSP}'(A; \mathscr{G})$ — that is to say, at some stage, the R-neighborhood of u must contain either one clause of degree $\leq k - 2$, or two clauses of degree $k - 1$. Let u' be the parent variable of u: by the assumption that $Z_{\text{out}}(v) = \varnothing$, the R-neighborhood of u' contains the R-neighborhood of u, which

means that u' should trigger a removal at the same stage of $\text{BSP}'(A;\mathscr{G})$. This is a contradiction, since removing the R-neighborhood of u' would remove the parent clause a' of v.

This concludes the proof of (5.32), which immediately implies that $V_1 \subseteq B_R(A';\mathscr{G}')$. For the remainder of the proof, we partition $Y_{\text{out}} = Y_+ \sqcup Y_\varnothing$ where

$$Y_+ \equiv \left\{ y \in Y_{\text{out}} : Z_{\text{out}}(y) \neq \varnothing \right\}.$$

Note that, by the construction from Definition 5.24, $Y_+ \subseteq A(R)$. Let $\mathscr{T}_\varnothing$ denote the subgraph of $\mathscr{T}_{\text{out}}$ induced by all descendants of variables in $Y_\varnothing$.

Now assume that $\mathscr{G}$ has girth greater than $8R$ and belongs to the event $\boldsymbol{G}_a \cap \boldsymbol{K}_a$. This means that the clause $a \in \mathscr{G}$ has initial type L_0. We next argue that

$$B_R(A';\mathscr{G}') \backslash \phi(\mathscr{N}_{\text{out}}) = \left\{ B_R(A;\mathscr{G}) \backslash \phi(\mathscr{N}_{\text{out}}) \right\} \sqcup K_\varnothing \tag{5.33}$$

for some $K_\varnothing \subseteq \mathscr{T}_\varnothing$. Indeed, recall that if $v \in \mathscr{G}''$ then v has the same R-neighborhood in $\mathscr{G}$ as in $\mathscr{G}'$, and Lemma 5.26 implies that $A \subseteq \mathscr{G}''$, $A \cap \mathscr{G}'' = A' \cap \mathscr{G}''$, and $A' \backslash A = \phi(\mathscr{N}_{\text{out}}) \cap A'$. As a result we can express

$$B_R(A';\mathscr{G}') \backslash \phi(\mathscr{N}_{\text{out}}) = \left\{ \bigcup_{v \in A} \left(B_R(v;\mathscr{G}) \backslash \phi(\mathscr{N}_{\text{out}}) \right) \right\} \cup \left\{ \bigcup_{v \in A' \backslash A} \left(B_R(v;\mathscr{G}') \backslash \phi(\mathscr{N}_{\text{out}}) \right) \right\}$$

$$= \left\{ B_R(A;\mathscr{G}) \backslash \phi(\mathscr{N}_{\text{out}}) \right\} \cup \left\{ \bigcup_{v \in \phi(\mathscr{N}_{\text{out}}) \cap A'} \left(B_R(v;\mathscr{G}') \backslash \phi(\mathscr{N}_{\text{out}}) \right) \right\}.$$

Now consider $v \in \phi(\mathscr{N}_{\text{out}}) \cap A'$ and $u \in B_R(v;\mathscr{G}') \backslash \phi(\mathscr{N}_{\text{out}})$. The path between u and v must intersect exactly one variable $y \in Y_{\text{out}}$. If this y belongs to Y_+, then (since $Y_+ \subseteq A$) we have

$$u \in B_R(y;\mathscr{G}) \backslash \phi(\mathscr{N}_{\text{out}}) \subseteq B_R(A;\mathscr{G}) \backslash \phi(\mathscr{N}_{\text{out}}).$$

If instead y belongs to $Y_\varnothing$, then $u \in K_\varnothing$. Combining these observations proves (5.33).

Next, we also observe that $B_R(Y_{\text{out}};\mathscr{G}) \cap \phi(\mathscr{N}_{\text{out}})$ is the same as $B_R(A;\mathscr{G}) \cap \phi(\mathscr{N}_{\text{out}})$, and this must be a subset of $B_R(A';\mathscr{G}') \cap \phi(\mathscr{N}_{\text{out}})$. Denote

$$K \equiv \phi(\mathscr{N}_{\text{out}}) \backslash B_R(A;\mathscr{G}) = \phi(\mathscr{N}_{\text{out}}) \backslash B_R(Y_{\text{out}};\mathscr{G}), \quad K' \equiv \phi(\mathscr{N}_{\text{out}}) \backslash B_R(A';\mathscr{G}') \subseteq K.$$

It follows using (5.33) that

$$\mathscr{G} \backslash B_R(A;\mathscr{G}) = \left(\mathscr{G}'' \cup \mathscr{N} \cup \phi(\mathscr{N}_{\text{out}}) \right) \backslash B_R(A;\mathscr{G}) = \left\{ \left(\mathscr{G}'' \cup \mathscr{N} \right) \backslash B_R(A';\mathscr{G}') \right\} \cup (K \cup K_\varnothing),$$

$$\mathscr{G}'(0) \equiv \mathscr{G}' \backslash B_R(A';\mathscr{G}') = \left(\mathscr{G}'' \cup \mathscr{N} \cup \phi(\mathscr{N}_{\text{out}}) \right) \backslash B_R(A';\mathscr{G}') = \left\{ \left(\mathscr{G}'' \cup \mathscr{N} \right) \backslash B_R(A';\mathscr{G}') \right\} \cup K', \tag{5.34}$$

where $\mathscr{G}'(0)$ is what is left of $\mathscr{G}'$ after the initial stage of $\text{BSP}'(A';\mathscr{G}')$. For comparison, if $\mathscr{G}(0)$ denotes what is left of $\mathscr{G}$ after the initial stage of $\text{BSP}'(A;\mathscr{G})$, then

$$\mathscr{G}(0) \equiv \mathscr{G} \backslash B_R(A;\mathscr{G}) = \left\{ \left(\mathscr{G}'' \cup \mathscr{N} \right) \backslash B_R(A';\mathscr{G}') \right\} \cup \phi(\mathscr{T}_{\text{in}} \backslash Y) \cup (K \cup K_\varnothing). \tag{5.35}$$

Recall that (5.32) implies that $V_1 \subseteq B_R(A';\mathscr{G}')$. Consequently, in (5.34) the subgraph induced by K' is disconnected from the rest, and likewise in (5.35) the subgraph induced by $K \cup K_\varnothing$ is disconnected from the rest. It follows by comparing (5.34) with (5.35) that

$$\mathscr{G}(0) \backslash (K \cup K_\varnothing) = \left(\mathscr{G}'(0) \cup \phi(\mathscr{T}_{\text{in}} \backslash Y) \right) \backslash K'. \tag{5.36}$$

Let $\mathscr{G}(t)$ and $\mathscr{G}'(t)$ be the graphs $\mathscr{G}, \mathscr{G}'$ after the t-th stages of $\text{BSP}'(A;\mathscr{G})$ and $\text{BSP}'(A';\mathscr{G}')$ respectively. We will argue by induction (with (5.36) being the base case) that

$$\mathscr{G}(t) \backslash (K \cup K_\varnothing) = \left(\mathscr{G}'(t) \cup \phi(\mathscr{T}_{\text{in}} \backslash Y) \right) \backslash K'. \tag{5.37}$$

Indeed, suppose inductively that (5.37) holds up to stage $t \geq 0$. Recall the notation (3.21). It is clear that

$$A'(t) \equiv \mathrm{Act}\Big(\mathscr{G}'(t) \setminus K'\Big) \subseteq \mathrm{Act}\Big(\mathscr{G}(t) \setminus K\Big) \equiv A(t),$$

where we have discarded the disconnected components induced by $K \cup K_\varnothing$ and by K' in defining $A(t), A'(t)$. In order for v to belong to $A(t)$, there are two possibilities:

- In $\mathscr{G}(t)$ there is a path γ of length at most $3R/10$ that joins v to a clause a' of degree $\leq k-2$. If γ is contained in $\mathscr{G}'(t)$, then clearly v must belong to $A'(t)$ as well. Otherwise, if γ is not contained in $\mathscr{G}'(t)$, then using the inductive hypothesis (5.37) it must be that γ intersects $\phi(\mathscr{T}_{\mathrm{in}} \setminus Y)$. On the other hand, the clause a' cannot be in $\phi(\mathscr{T}_{\mathrm{in}} \setminus Y)$ (which cannot contain any clauses of degree $\leq k-2$, since it was constructed from a processed graph $\mathrm{pr}\,\mathscr{H}$). Moreover, on the event K_a, the clause a' cannot be in $\mathscr{N}$. This is a contradiction, since there is no path of length $3R/10$ that intersects both $\phi(\mathscr{T}_{\mathrm{in}} \setminus Y)$ and $\mathscr{G}(t) \setminus [\phi(\mathscr{T}_{\mathrm{in}} \setminus Y) \cup \mathscr{N}]$.
- In $\mathscr{G}(t)$ there are two paths γ_1, γ_2 of length at most $3R/10$ that join v to clauses $a_1 \neq a_2$ of degree $k-1$. If the γ_i are both contained in $\mathscr{G}'(t)$, then v must belong to $A'(t)$ as well. Otherwise, by the same argument as above, at least one of the paths must intersect $\phi(\mathscr{T}_{\mathrm{in}} \setminus Y)$. On the event K_a, the clauses a_i cannot be in $\mathscr{N}$, so they must both be $\phi(\mathscr{T}_{\mathrm{in}} \setminus Y)$. This contradicts the construction of $\phi(\mathscr{T}_{\mathrm{in}} \setminus Y)$ which was based on the processed graph $\mathrm{pr}\,\mathscr{H}$.

The above shows that $A(t) = A'(t) \subseteq \mathscr{G}'(t) \setminus \mathscr{N}$, from which it follows that

$$B_R\Big(A(t); \mathscr{G}(t)\Big) = B_R\Big(A'(t); \mathscr{G}'(t)\Big) \subseteq \mathscr{G}'(t).$$

This implies that (5.37) holds at the next stage $t+1$. It follows that

$$\mathrm{pr}\,\mathscr{G} \Big\backslash (K \cup K_\varnothing) = \Big(\mathrm{pr}\,\mathscr{G}' \cup \phi(\mathscr{T}_{\mathrm{in}} \setminus Y)\Big) \Big\backslash K',$$

so that the clause a in graph $\mathscr{G}$ has final type L_∞. It follows that clause a has total type $L_a = (L_0, L_\infty)$, as required. □

5.6. **Probabilistic analysis for positive type fractions.** We now conclude the proof of Proposition 3.23. In view of Corollary 5.27 from the preceding subsection, it suffices to show that for any feasible L, conditional on $\mathscr{G}$ having girth greater than $8R$, with high probability the events $G_a(L)$ and $K_a(L)$ (Definition 5.25) will occur for a positive fraction of clauses $a \in F$. We argue this in a few steps, below. The general idea is that all these events are fairly local in nature, so they should occur a linear number of times in the random graph $\mathscr{G}$.

Lemma 5.28. *Let* $\mathbb{P} = \mathbb{P}_{n,m}$ *for* $|m - n\alpha| \leq n^{1/2} \ln n$, *and let* $\mathscr{G}$ *denote a sample from* $\mathbb{P}$. *There is a positive constant* $c_0(k, R)$ *such that for all n large enough, we have*

$$\mathbb{P}\Big(\mathscr{G} \in E_a(L)\Big) \geq c_0(k, R).$$

for all $a \in [m]$ and all clause total types L that are feasible in the sense of Definition 5.23.

Proof. We first argue that there is a finite constant $C(k, R)$ such that

$$\left| \left\{ \begin{array}{c} \text{feasible clause total} \\ \text{types } L \equiv (L_0, L_\infty) \end{array} \right\} \right| \leq C_0(k, R). \tag{5.38}$$

To this end, let $L \equiv (L_0, L_\infty)$ be a feasible type, with $\mathscr{H}, a_\star$ as in Definition 5.23. Recall that the initial type L_0 encodes the $(R + 1/2)$-neighborhood of $a_\star$ in $\mathscr{H}$. If any variable in the $(R - 1/2)$-neighborhood of $a_\star$ in $\mathscr{H}$ has degree more than $\exp\{k^2(5R')\}$, then one of the variables $u \in a_\star$ will fail to be 1-fair in $\mathscr{H}$ (since it will violate property (ii) in Definition 3.13). But then the R-neighborhood of u will be removed in the initial stage of processing on $\mathscr{H}$, contradicting the assumption that $a_\star \in \mathrm{pr}\,\mathscr{H}$. This proves that all variables in L_0 must have degree at most $\exp\{k^2(5R')\}$. All variables in the final type L_∞ must be 1-good, hence also 1-fair, so they must also have degree at most $\exp\{k^2(5R')\}$. Finally L_0 is a tree of depth at most R, while L_∞ is a tree of depth at most $R(1 + 1/100)$. This proves (5.38). Since we fix a mapping from L to $\mathscr{T}(L)$ (see Definition 5.24), it immediately follows from (5.38) that the number of distinct $\mathscr{T}(L)$ is also upper bounded by $C_0(k, R)$.

Now take $\mathbb{P} = \mathbb{P}_{n,m}$ as in the statement of the lemma, fix $a \in [m]$, and let $\mathscr{G} \sim \mathbb{P}$. By revealing the neighborhood of of a in $\mathscr{G}$ in breadth-first fashion, it is easy to see that $\mathbb{P}(E_a(L))$ is lower bounded by a constant which depends only on $\mathscr{T}(L)$. It then follows from the above that in fact there is a constant $c_0(k, R)$ such that (for n large enough) we have $\mathbb{P}(E_a(L)) \geq c_0(k, R)$ for all feasible L. This proves the lemma. □

Lemma 5.29. *Let* $\mathbb{P} = \mathbb{P}_{n,m}$ *for* $|m - n\alpha| \le n^{1/2} \ln n$, *and* $\mathscr{G} \sim \mathbb{P}$. *Then, for all n large enough, we have*

$$\mathbb{P}\Big(\mathscr{G} \in G_a(L) \,\Big|\, \mathscr{G} \in E_a(L)\Big) \ge \frac{7}{8}\,.$$

for all $a \in [m]$ and all clause total types L that are feasible in the sense of Definition 5.23.

Proof. Throughout the proof, L is fixed and often suppressed from the notation. Let P_a be the event that E_a holds, and that $\mathscr{G}$ satisfies properties (I) and (II) of Definition 5.25. Then $E_a \subseteq P_a \subseteq K_a$, and it is clear that

$$\mathbb{P}\Big(\mathscr{G} \in P_a \,\Big|\, \mathscr{G} \in E_a\Big) \ge \frac{24}{25}\,. \tag{5.39}$$

This is simply because the markings l are chosen uniformly at random from a very large set (see (3.2)), so if we consider all the variables in $\mathscr{G} \setminus \phi(\mathscr{T}(L))$ within distance $2R$ of $\phi(\mathscr{T}(L))$, it holds with probability at least $9/10$ that all their markings are distinct from one another and from the markings on $\phi(\mathscr{T}(L))$.

Next we let F_a be the event that E_a holds, and that $\mathscr{G}$ satisfies property (III) of Definition 5.25. For $x \in \phi(X)$, let $\mathscr{U}(x)$ denote the subtree of $\mathscr{U}$ descended from x. Let $F_a(x)$ be the event that E_a holds, and that every $u \in \mathscr{U}(x)$ is 1-fair (with respect to $\mathscr{G}$). Then

$$F_a = \bigcap_{x \in \phi(X)} F_a(x)\,.$$

For $x \in \phi(X)$, let $\mathscr{U}_+(x)$ denote the subtree of $\mathscr{U} \cup \mathscr{N}$ induced by descendants of x that lie within distance $3R/10$ of x, so that $\mathscr{U}(x) \subseteq \mathscr{U}_+(x) \subseteq \mathscr{U} \cup \mathscr{N}$. Since $F_a \subseteq E_a$ where E_a imposes that $B_{4R}(a; \mathscr{G})$ is a tree, for any variable in $B_{(4-1/10)R}(a; \mathscr{G})$ we can determine whether it is 1-fair based on its $(R/10)$-neighborhood only. It follows that the event $F_a(x)$ can be determined from

$$B_{R/10}(x; \mathscr{G}) \cup \mathscr{U}_+(x)\,.$$

An important point is that the above does not see any part of $\phi(\mathscr{T}(L))$ at distance more than $R/10$ from x, because

$$\phi(\mathscr{T}(L)) \cap \Big\{ B_{R/10}(x; \mathscr{G}) \cup \mathscr{U}_+(x) \Big\} = \phi(\mathscr{T}(L)) \cap B_{R/10}(x; \mathscr{G})\,. \tag{5.40}$$

Indeed, when we condition on E_a, we in fact reveal information about the neighborhood of x in $\mathscr{G}$ beyond depth $R/10$ (simply because $\phi(\mathscr{T}(L))$ contains vertices at distance more than $R/10$ from x), but (5.40) allows us to disregard the additional information. It follows from the 1-excellence condition (3.20) that

$$\mathbb{P}\Big(\mathscr{G} \in F_a(x) \,\Big|\, \mathscr{G} \in E_a\Big) \ge 1 - \frac{1}{\exp(k^3 R)} + o_n(1)\,, \tag{5.41}$$

where the $o_n(1)$ comes from the discrepancy between the Galton–Watson law and the breadth-first exploration in $\mathscr{G}$. Since all variables in $\mathscr{T}_{\text{in}}$ must be 1-fair, it follows from property (ii) in Definition 3.13 that $\mathscr{T}_{\text{in}}$ has cardinality at most $\exp(O(k^2 R))$. Therefore we conclude

$$\mathbb{P}\Big(\mathscr{G} \in F_a \,\Big|\, \mathscr{G} \in E_a\Big) \ge \frac{24}{25} \tag{5.42}$$

by taking a union bound of (5.41) over $x \in \phi(X)$.

Finally, let D_a be the event that E_a holds, and that $\mathscr{G}$ satisfies property (IV) of Definition 5.25. For $y \in \phi(Y)$, let $X'(y)$ denote the subset of variables in X' that lie at depth $R/10$ below y. Let $X''(y)$ denote the variables in $X'(y)$ that fail to be 1-excellent in $\mathscr{G}$. Again, on the event E_a which imposes that $B_{4R}(a; \mathscr{G})$ is acyclic, for any variable in $B_{(4-1/10)R}(a; \mathscr{G})$ we can determine whether it is 1-excellent based on its $(R/10)$-neighborhood only. Consequently, by similar considerations as for (5.41), we have

$$\mathbb{E}\Big(X'(y) \,\Big|\, \mathscr{G} \in E_a\Big) = o_n(1) + \int \left| \left\{ \begin{matrix} x \in \mathscr{T} : d(v_{\text{RT}}, x) = R/10 \\ \text{and } x \text{ is not 1-excellent} \end{matrix} \right\} \right| d\mathbf{PGW}(\mathscr{T})$$

$$= o_n(1) + \int \mathbf{1}\left\{ \begin{matrix} \text{the root } v_{\text{RT}} \text{ of } \mathscr{T} \text{ is} \\ \text{not 1-excellent} \end{matrix} \right\} Z_{R/10}(\mathscr{T}) \, d\mathbf{PGW}(\mathscr{T})\,, \tag{5.43}$$

where $Z_{R/10}(\mathscr{T})$ denotes the number of variables at depth $R/10$ in $\mathscr{T}$, and the last equality is by the unimodularity property (4.1). Now recall from Corollary 5.9 that

$$\mathbf{PGW}\Big(v_{\text{RT}} \text{ not 1-excellent}\Big) \le \frac{1}{\exp(2^{k\delta_*/4}R)}\,.$$

On the other hand, by a similar argument as in the proof of Lemma 5.5, we can bound

$$\int \left(Z_{R/10}(\mathscr{T}) \right)^2 d\mathbf{PGW}(\mathscr{T}) \leq \exp(O(kR)).$$

Combining these and applying Cauchy–Schwarz in (5.43) gives

$$\mathbb{E}\left(X'(y) \,\middle|\, \mathscr{G} \in E_a \right) \leq \frac{\exp(O(kR))}{\exp(2^{k\delta_*/4}R)}.$$

Since we noted above that $\mathscr{T}_{\text{in}}$ has cardinality at most $\exp(O(k^2 R))$, it follows by Markov's inequality and a union bound over $y \in \phi(Y)$ that

$$\mathbb{P}\left((D_a)^c \,\middle|\, \mathscr{G} \in E_a \right) \leq \sum_{y \in \phi(Y)} \mathbb{E}\left(X'(y) \,\middle|\, \mathscr{G} \in E_a \right) \leq \frac{1}{25}. \tag{5.44}$$

The lemma follows by combining (5.39), (5.42), and (5.44), since $G_a = P_a \cap F_a \cap D_a$.   □

Lemma 5.30. *Let $\mathbb{P} = \mathbb{P}_{n,m}$ for $|m - n\alpha| \leq n^{1/2} \ln n$, and $\mathscr{G} \sim \mathbb{P}$. Then, for all n large enough, we have*

$$\mathbb{P}\left(K_a(L) \,\middle|\, \mathscr{G} \in E_a(L) \right) \geq \frac{7}{8}.$$

for all $a \in [m]$ and all clause total types L that are feasible in the sense of Definition 5.23.

Proof. Again, throughout the proof L is fixed and often suppressed from the notation. Let $\mathscr{G} \sim \mathbb{P} = \mathbb{P}_{n,m}$. Recalling Definition 5.25, let C_a be the event that there is an embedding $\phi : \mathscr{T}(L) \hookrightarrow \mathscr{G}$ satisfying properties (i) and (ii) of Definition 5.25. Then E_a is C_a with the additional restriction that $B_{4R}(a; \mathscr{G})$ is a tree. Provided that $\mathscr{G} \in C_a$, we define, as in (5.28),

$$\mathscr{G}' \equiv \mathscr{G} \,\backslash\, \phi\left(\mathscr{T}_{\text{in}} \backslash Y \right).$$

Recalling Definition 5.25, we can express K_a as the intersection of E_a with the event that no variable in $\phi(Y)$ lies within distance R of any variable removed in the processing of $\mathscr{G}'$, equivalently,

$$K_a = E_a \cap \left\{ \phi(Y) \text{ does not intersect } S(\mathscr{G}') \equiv B_R\left(\mathsf{BSP}'(A'; \mathscr{G}'); \mathscr{G}' \right) \right\}. \tag{5.45}$$

Now consider the graph $\mathscr{G}'$ where only $\phi(\mathscr{T}_{\text{out}})$ is labelled (in particular, we ignore $\phi(Y)$ for the moment). Let n' denote n minus the number of internal variables in $\mathscr{T}_{\text{in}} \backslash Y$. Likewise, let m' denote m minus the number of clauses in $\mathscr{T}_{\text{in}} \backslash Y$. Under the measure $\mathbb{P}_{n,m}(\cdot \,|\, \mathscr{G} \in C_a)$, the induced law of $\mathscr{G}'$ is equivalent (up to graph isomorphism) to the law $\mathbb{P}_{n',m'}(\cdot \,|\, \mathscr{G}' \in C')$ where C' is the event that there is an (arbitrary, fixed) embedding $\phi_{\text{out}} : \mathscr{T}_{\text{out}} \hookrightarrow \mathscr{G}'$. This is only a minor modification of the original measure $\mathbb{P}_{n,m}$, and a trivial extension of Proposition 3.22 gives

$$\mathbb{P}_{n',m'}\left(|S(\mathscr{G}')| \geq \frac{n'}{\exp(2^{ck}R)} \,\middle|\, \mathscr{G}' \in C' \right) = o_n(1). \tag{5.46}$$

Since we can choose a fixed embedding ϕ_{out}, we will suppose that it maps the variables in $\mathscr{T}_{\text{out}}$ to the last variables in $\mathscr{G}'$, so that the variables left in $\mathscr{G}' \backslash \phi_{\text{out}}(\mathscr{T}_{\text{out}})$ can be written as $[n''] = \{1, \ldots, n''\}$. We now return to $Y' \equiv \phi(Y)$, which we regard as an element of

$$[n'']_y = \left\{ (v_1, \ldots, v_y) \in [n'']^y : v_i \neq v_j \text{ for all } i \neq j \right\},$$

where $y \equiv |Y|$. Given $(\mathscr{G}', Y')$ where $\mathscr{G}' \in C'$ and $Y' \in [n'']_y$, we can uniquely recover the original graph $\mathscr{G}$ (modulo isomorphism) by gluing back the tree $\phi(\mathscr{T}_{\text{in}})$, so we write $\mathscr{G} = \mathscr{G}(\mathscr{G}', Y')$ (in this graph, $\phi(\mathscr{T}(L))$ is labelled). We then denote

$$J \equiv \left\{ (\mathscr{G}', Y') \in C' \times [n'']_y : \text{in } \mathscr{G} = \mathscr{G}(\mathscr{G}', Y'), \text{ the neighborhood } B_{4R}(a; \mathscr{G}) \text{ is a tree} \right\}. \tag{5.47}$$

Recall from above that under $\mathbb{P}_{n,m}(\cdot \mid \mathscr{G} \in C_a)$, the induced law of $\mathscr{G}'$ is equivalent (modulo graph isomorphism) to $\mathbb{P}_{n',m'}(\cdot \mid \mathscr{G}' \in C')$. Recall also that E_a is the same as C_a with the added restriction that $B_{4R}(a;\mathscr{G})$ is a tree. It follows that if K is any event that is invariant under graph isomorphism, then we have

$$\mathbb{P}\Big((\mathscr{G}',Y') \in K \,\Big|\, \mathscr{G} \in E_a\Big) = \mathbb{P}_{n',m'}\Big((\mathscr{G}',Y') \in K \,\Big|\, J\Big). \tag{5.48}$$

Conditional on C', it is clear that Y' is a uniformly random element of $[n'']_y$, which is independent of the structure of $\mathscr{G}'$, and hence independent of $S(\mathscr{G}')$: for any $Y_0 \in [n'']_y$ and any subset $S_0 \subseteq [n']$,

$$\mathbb{P}_{n',m'}\Big(Y' = Y_0 \,\Big|\, \mathscr{G}' \in C', S(\mathscr{G}') = S_0\Big) = \frac{1}{(n'')_y}\,.$$

We now argue that this does not change much if we condition further on $B_{4R}(a;\mathscr{G})$ being a tree. Indeed, since all the random graphs we consider are locally tree-like, it is clear that the event J of (5.47) occupies $1 - o_n(1)$ fraction of $C' \times [n'']_y$. This implies that, for any $Y_0 \in [n'']_y$ and any $S_0 \subseteq [n']$, we have

$$p(Y_0|S_0) \equiv \mathbb{P}_{n',m'}\Big(Y' = Y_0 \,\Big|\, J, S(\mathscr{G}') = S_0\Big) = \frac{\mathbb{P}_{n',m'}(Y' = Y_0, J \mid \mathscr{G}' \in C', S(\mathscr{G}') = S_0)}{\mathbb{P}_{n',m'}(J \mid \mathscr{G}' \in C', S(\mathscr{G}') = S_0)}$$

$$\leq \frac{\mathbb{P}_{n',m'}(Y' = Y_0 \mid \mathscr{G}' \in C', S(\mathscr{G}') = S_0)}{1 - o_n(1)} = \frac{1 - o_n(1)}{(n'')_y}\,. \tag{5.49}$$

By a very similar argument, we can extend (5.46) to

$$\mathbb{P}_{n',m'}\bigg(|S(\mathscr{G}')| \geq \frac{n'}{\exp(2^{ck}R)} \,\bigg|\, J\bigg) = o_n(1)\,. \tag{5.50}$$

Recall that $|Y| \leq \exp(O(k^2R))$. It follows by combining (5.50) and (5.49) that

$$\mathbb{P}_{n',m'}\Big(Y' \cap S(\mathscr{G}') \neq \varnothing \,\Big|\, J\Big) \leq o_n(1) + \mathbb{P}_{n',m'}\bigg(Y' \cap S(\mathscr{G}') \neq \varnothing \,\bigg|\, J, |S(\mathscr{G}')| < \frac{n'}{\exp(2^{ck}R)}\bigg)$$

$$\leq o_n(1) + \frac{\exp(O(k^2R))}{\exp(2^{ck}R)} \leq \frac{1}{8}\,.$$

The lemma follows by the equivalence (5.48). □

Proof of Proposition 3.23. Let $\mathscr{G} = (V,F,E) \sim \mathbb{P} = \mathbb{P}_{n,m}$. Let L be any feasible type, in the sense of Definition 5.23. For any fixed L we consider the events $E_a(L)$, $G_a(L)$, $F_a(L)$ as in Definition 5.25. Let

$$J(L) \equiv \sum_{a \in F} J_a(L) \equiv \sum_{a \in F} \mathbf{1}\Big\{E_a(L) \cap G_a(L) \cap K_a(L)\Big\}\,.$$

It follows by combining Lemmas 5.28–5.30 that for all feasible L we have

$$\mathbb{E}J(L) \geq \frac{3nc_0(k,R)}{4}\,.$$

We will argue that each $J(L)$ is sufficiently concentrated around its mean, such that

$$\mathbb{P}\bigg(J(L) \geq \frac{nc_0(k,R)}{2}\bigg) = o_n(1)\,. \tag{5.51}$$

Note that (5.51) implies the result: as discussed in the proof of Lemma 5.28, the total number of feasible types L is upper bounded by some $C_0(k,R)$, so (5.51) (together with the fact that $\mathscr{G}$ has girth greater than $8R$ with probability at least $c_0(k,R)$) implies

$$\mathbb{P}\bigg(J(L) \geq \frac{nc_0(k,R)}{2} \text{ for all feasible } L \,\bigg|\, \text{girth}(\mathscr{G}) > 8R\bigg) = o_n(1)\,.$$

In light of Corollary 5.27, this directly implies the result. It therefore remains only to prove the concentration result (5.51). Let $\mathscr{T}(L)$ be the tree based on L that is given by Definition 5.24, and recall from Definition 5.25 that $E_a(L)$ is the event that there is a certain embedding $\phi_a : \mathscr{T}(L) \hookrightarrow \mathscr{G}$ mapping $a_\star \mapsto a$, together with some local girth

condition. The events $E_a(L)$ and $G_a(L)$ are clearly local — they do not depend on more than $B_{4R}(a; \mathscr{G})$. By contrast, the event $K_a(L)$ concerns preprocessing on the graph

$$\mathscr{G}'(a) \equiv \mathscr{G} \setminus \phi_a\left(\mathscr{T}_{\text{in}} \setminus Y\right),$$

which (*a priori*) cannot be locally determined. This can be addressed with a similar argument as for Proposition 3.22: as we saw in that proof, preprocessing is in fact fairly localized with high probability. In particular, we can define

$$\mathscr{G}^\circ(a) \equiv \left\{ B_{(\ln n)/2^{k\delta_*/6}}(a; \mathscr{G}) \right\} \setminus \phi_a\left(\mathscr{T}_{\text{in}} \setminus Y\right) \subseteq \mathscr{G}'(a).$$

Let $A^\circ(a)$ be the variables $u \in \mathscr{G}^\circ(a)$ with $B_R(u; \mathscr{G}'(a)) \subseteq \mathscr{G}^\circ(a)$, that fail to proper or 1-good with respect to $\mathscr{G}^\circ(a)$, and let (cf. (5.45))

$$\overline{K}_a(L) = E_a(L) \cap \left\{ \phi_a(Y) \text{ does not intersect } B_R\left(\text{BSP}'\left(A^\circ(a); \mathscr{G}^\circ(a)\right); \mathscr{G}^\circ(a)\right) \right\}$$

Let $\overline{J}(L)$ be defined as $J(L)$, but with $\overline{K}_a(L)$ in place of $K_a(L)$. It follows from the proof of Proposition 3.22 that $J(L)$ and $\overline{J}(L)$ agree with high probability, and so

$$\mathbb{E}\overline{J}(L) = \mathbb{E}J(L) - o(n) \geq \frac{3nc_0(k, R)}{4} - o(n).$$

For any pair of distinct clauses $a \neq b$, their $(\ln n)/2^{k\delta_*/6}$-neighborhoods do not intersect with high probability, and so are roughly independent. It follows that $\overline{J}(L)$ has variance $o(n^2)$, and so by Chebyshev's inequality

$$\mathbb{P}\left(\overline{J}(L) < \frac{nc_0(k, R)}{2}\right) \leq \mathbb{P}\left(\left|\overline{J}(L) - \mathbb{E}\overline{J}(L)\right| \geq \frac{nc_0(k, R)}{5}\right) \leq o_n(1).$$

Since $\mathbb{P}(J(L) \neq \overline{J}(L)) = o_n(1)$, this proves (5.51), and the result follows as argued above. $\square$

5.7. Uniformity of processed graph. We conclude this section with the proof of Proposition 3.24.

Proposition 5.31. *Fix any k-SAT instance $\mathscr{G} = (V, F, E)$ that can arise under the measure $\mathbb{P} = \mathbb{P}^{n,\alpha}$. In the processed graph $\text{pr}\,\mathscr{G}$, choose any two edges $e_i = (a_i u_i)$ ($i = 1, 2$) having the same total type (Definition 3.18). Let sw denote the switching operation in which we cut the edges e_i, and reconnect the resulting half-edges as $f_1 = (a_1 u_2)$ and $f_2 = (a_2 u_1)$. Then the switching and processing operations commute:*

$$\text{pr}\left(\text{sw}\,\mathscr{G}\right) = \text{sw}\left(\text{pr}\,\mathscr{G}\right).$$

In particular, this implies that f_i survive in $\text{pr}(\text{sw}\,\mathscr{G})$.

We begin with a preliminary lemma:

Lemma 5.32. *In the setting of Proposition 5.31, the initial set (of variables that are improper or not 1-good) in $\mathscr{G}$ is the same as in $\text{sw}\,\mathscr{G}$.*

Proof. Denote the initial set (of variables that are improper or not 1-good) as A_{-1} in $\mathscr{G}$, and S_{-1} in $\text{sw}\,\mathscr{G}$. Membership of a variable in A_{-1} (resp. S_{-1}) is a property of its R-neighborhood relative to $\mathscr{G}$ (resp. $\text{sw}\,\mathscr{G}$). If $B_R(v; \mathscr{G})$ is the same as $B_R(v; \text{sw}\,\mathscr{G})$, then v belongs either to both sets A_{-1} and S_{-1}, or to neither.

If $B_R(v; \mathscr{G})$ is not the same as $B_R(v; \text{sw}\,\mathscr{G})$, then it must be that $B_R(v; \mathscr{G})$ contains at least one of the e_i, and so $B_R(v; \text{sw}\,\mathscr{G})$ contains at least one of the f_i. In this situation, we have two observations:

(I) We must have $v \notin A_{-1}$ by the assumption that the e_i survive in $\text{pr}\,\mathscr{G}$ — in particular, this implies that $B_R(v; \mathscr{G})$ must be proper (acyclic, with no repeated l markings). It follows that $B_R(v; \mathscr{G})$ must contain exactly one of the e_i while being disjoint from the other — otherwise, it would contain two variables with the same marking l, making v improper.

(II) Similarly, $B_R(v; \text{sw}\,\mathscr{G})$ must contain exactly one of the f_i while being disjoint from the other — otherwise, there is a path $\gamma_\star \subseteq \mathscr{G} \cap (\text{sw}\,\mathscr{G})$ of length at most $2R - 1$ that joins the f_i in $\text{sw}\,\mathscr{G}$. In the graph $\mathscr{G}$, either $\gamma_\star$ forms a cycle with one of the e_i, or it joins e_1 to e_2. In both cases $\gamma_\star$ will contain an improper variable that will cause one of the e_i to be removed during processing on $\mathscr{G}$.

Since we only switch two edges $e_i = (a_i u_i)$ of the same type, if $B_R(v; \mathrm{sw}\, \mathcal{G})$ is also acyclic then it must be isomorphic to $B_R(v; \mathcal{G})$, so in this case $v \notin S_{-1}$. The last possibility is that $B_R(v; \mathrm{sw}\, \mathcal{G})$ contains a cycle C. By observation (II), we can suppose without loss that $B_R(v; \mathrm{sw}\, \mathcal{G})$ contains f_1 but not f_2. If f_1 is disjoint from C, then C will also appear in $B_R(u_i; \mathcal{G})$ for one of the u_i. If f_1 lies on C, then $C \setminus f_1$ is a path in $\mathcal{G}$ joining e_1 to e_2. In both cases C will contain an improper variable that will cause one of the e_i to be removed during processing on $\mathcal{G}$. Altogether this proves that $A_{-1} = S_{-1}$, as claimed. $\qquad\square$

Proof of Proposition 5.31. Denote the initial set (of variables that are improper or not 1-good) as A_{-1} in $\mathcal{G}$, and S_{-1} in $\mathrm{sw}\, \mathcal{G}$. We showed in Lemma 5.32 that $A_{-1} = S_{-1}$. Note also that $B_R(A_{-1}; \mathcal{G})$ cannot include either of the e_i, again by the assumption that the e_i survive in $\mathrm{pr}\, \mathcal{G}$. It follows that $B_R(A_{-1}; \mathcal{G}) = B_R(A_{-1}; \mathrm{sw}\, \mathcal{G})$, and consequently

$$\mathrm{pr}_0 \left(\mathrm{sw}\, \mathcal{G} \right) = \left(\mathrm{sw}\, \mathcal{G} \right) \setminus B_R \left(A_{-1}; \mathrm{sw}\, \mathcal{G} \right) = \mathrm{sw} \left(\mathcal{G} \setminus B_R \left(A_{-1}; \mathcal{G} \right) \right) = \mathrm{sw} \left(\mathrm{pr}_0\, \mathcal{G} \right), \tag{5.52}$$

that is to say, switching commutes with the initial preprocessing step. For $t \geq 0$ let us abbreviate

$$A_t \equiv \mathrm{Act}(\mathrm{pr}_t\, \mathcal{G}),$$
$$S_t \equiv \mathrm{Act}(\mathrm{pr}_t\, \mathrm{sw}\, \mathcal{G}).$$

If $A_\ell = S_\ell$ for all $0 \leq \ell < t$, then the same logic that led to (5.52) gives

$$\mathrm{pr}_t \left(\mathrm{sw}\, \mathcal{G} \right) = \mathrm{sw} \left(\mathrm{pr}_t\, \mathcal{G} \right). \tag{5.53}$$

Thus, let t be the first time that $A_t \neq S_t$; we will argue that this leads to a contradiction. Abbreviate $\mathcal{H} \equiv \mathrm{pr}_t\, \mathcal{G}$.

Case 1. $A_t \setminus S_t \neq \varnothing$. Recall from the definition (3.21) that membership of a variable in A_t or S_t is a property of its $(3R/10)$-neighborhood. Consequently, if there is any variable $v \in A_t \setminus S_t$, then $B_{3R/10}(v; \mathcal{H})$ must contain at least one of the e_i. But the very next step of processing on $\mathcal{G}$ will remove $B_R(v; \mathcal{H})$, contradicting the assumption that both edges e_i survive in $\mathrm{pr}\, \mathcal{G}$. This proves that $A_t \subseteq S_t$.

Case 2. $S_t \setminus A_t \neq \varnothing$. Suppose $v \in S_t \setminus A_t$. From the above definitions and the relation (5.53), we have $A_t = \mathrm{Act}(\mathcal{H})$ and $S_t = \mathrm{Act}(\mathrm{sw}\, \mathcal{H})$. Thus, since $v \in S_t$, it must be that $B_{3R/10}(v; \mathrm{sw}\, \mathcal{H})$ contains at least one clause of degree $\leq k-2$, or two clauses of degree $k-1$. On the other hand, since $v \notin A_t$, the same statement must not hold for $B_{3R/10}(v; \mathcal{H})$. It follows that $B_{3R/10}(v; \mathrm{sw}\, \mathcal{H})$ must contain at least one of the switched edges f_i. In fact, by observation (II) in the proof of Lemma 5.32, it must contain exactly one of the f_i, say f_1.

(a) If $B_{3R/10}(v; \mathrm{sw}\, \mathcal{H})$ contains any clause $b_\star$ of degree $\leq k-2$, then the path joining v to $b_\star$ must pass through at least one of the f_i, so the distance between $b_\star$ and $\{f_1, f_2\}$ in $\mathrm{sw}\, \mathcal{H}$ is less than $3R/10$. It follows that the distance between $b_\star$ and $\{e_1, e_2\}$ in $\mathcal{H}$ is also less than $3R/10$, which again yields a contradiction since it means that at least one of the e_i will be deleted in the next preprocessing step.

(b) It remains to consider the case that $B_{3R/10}(v; \mathrm{sw}\, \mathcal{H})$ contains two clauses $b_1 \neq b_2$, each of degree $k-1$. If π_i is the path joining v to b_i in $B_{3R/10}(v; \mathrm{sw}\, \mathcal{H})$, then at least one of the π_i, say π_1, must pass through a switched edge. It follows that the distance between b_1 and $\{f_1, f_2\}$ in $\mathrm{sw}\, \mathcal{H}$ is less than $3R/10$; and so the distance between b_1 and $\{e_1, e_2\}$ in $\mathcal{H}$ is also less than $3R/10$. We now consider the path γ in $B_{3R/10}(v; \mathrm{sw}\, \mathcal{H})$ that joins b_1 to b_2, and distinguish two cases:

   (i) If γ does not pass through either f_i, then it is also a path in $\mathcal{H}$. It has length at most $6R/10$ and joins b_1 to b_2, and we noted above that the distance between b_1 and $\{e_1, e_2\}$ in $\mathcal{H}$ is less than $3R/10$. It follows that the next preprocessing step in $\mathcal{G}$ will remove the R-neighborhood of some variable on γ, and thereby also remove one of the e_i. This gives a contradiction.

   (ii) Now consider the case that γ passes through f_1. This situation is shown in Figure 6. We label $f_1 = (xz)$ and $f_2 = (wy)$ where $\{x, w\}$ is the same as either $\{a_1, a_2\}$ or $\{u_1, u_2\}$. Then $e_1 = (xy)$ and $e_2 = (wz)$. Let ρ_1 denote the path between b_1 and x, and ρ_2 the path between b_2 and z: these paths are the same in $\mathcal{H}$ as in $\mathrm{sw}\, \mathcal{H}$, and must satisfy

$$\mathrm{len}(\rho_1) + \mathrm{len}(\rho_2) + \frac{1}{2} = \mathrm{len}(\gamma) \leq \frac{6R}{10}$$

where we use len to denote path length. Recall from Remark 2.4 that an edge $e = (av)$ has length $1/2$, so a path joining two neighboring variables has length one. On the left-hand side above, the $1/2$ term accounts for the length of the edge $f_1 = (xz)$. In the graph $\mathcal{H}$, the path ρ_1 joins b_1 to x without passing through y.

Since we assume that $e_1 = (xy)$ survives in $\mathrm{pr}\,\mathscr{G}$, any further processing on $\mathscr{G}$ cannot remove b_1 without leaving behind a clause on ρ_1 of degree less than k that lies even closer to x. It follows that the final graph $\mathrm{pr}\,\mathscr{G}$ contains a path $\bar{\rho}_1 \subseteq \rho_1$ joining x to a clause $\bar{b}_1$ of degree less than k. Likewise, $\mathrm{pr}\,\mathscr{G}$ must contain a path $\bar{\rho}_2 \subseteq \rho_2$ joining z to a clause $\bar{b}_2$ of degree less than k. Now note that x and z lie on opposite sides of the bipartite graph (one is a clause while the other is a variable). On the other hand, by the assumption that the e_i have the same total type, there must be an isomorphism

$$\phi : B_R\Big(x; \mathrm{pr}\,\mathscr{G}\Big) \to B_R\Big(w; \mathrm{pr}\,\mathscr{G}\Big), \quad \phi(x) = w, \quad \phi(y) = z\,.$$

It follows that $B_R(w; \mathrm{pr}\,\mathscr{G})$ must contain a path $\phi(\bar{\rho}_1)$ that joins w to a clause $\phi(\bar{b}_1)$ of degree less than k. Then in the graph $\mathrm{pr}\,\mathscr{G}$ we have

$$d\Big(\bar{b}_2, \phi(\bar{\rho}_1); \mathrm{pr}\,\mathscr{G}\Big) = \mathrm{len}(\phi(\bar{\rho}_1)) + \mathrm{len}(\bar{\rho}_2) + \frac{1}{2} \leq \mathrm{len}(\rho_1) + \mathrm{len}(\rho_2) + \frac{1}{2} \leq \frac{6R}{10}\,.$$

This contradicts the fact that, by definition, $\mathrm{Act}(\mathrm{pr}\,\mathscr{G})$ must be empty.

Conclusion. The above argument shows that in fact $A_t = S_t$ for all t, so that the same logic leading to (5.52) and (5.53) gives the desired conclusion, that $\mathrm{pr}(\mathrm{sw}\,\mathscr{G}) = \mathrm{sw}(\mathrm{pr}\,\mathscr{G})$. $\qquad\square$
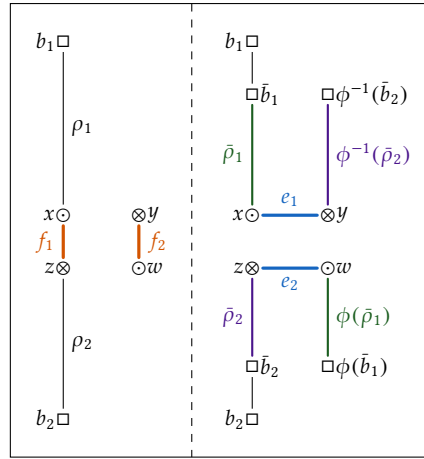

FIGURE 6. Case (ii) in the proof of Proposition 5.31. We let t be the first time that $A_t \neq S_t$ and $\mathscr{H} = \mathrm{pr}_t\,\mathscr{G}$, so that $\mathrm{pr}_t(\mathrm{sw}\,\mathscr{G}) = \mathrm{sw}(\mathrm{pr}_t\,\mathscr{G}) = \mathrm{sw}\,\mathscr{H}$ by (5.53). In each of the panels (A)–(C), the left side of the dashed line depicts part of $\mathrm{sw}\,\mathscr{H}$ (including the switched edges f_i) while the right side depicts part of $\mathscr{H}$ (including the original edges e_i). We use the symbols $\odot$ and $\otimes$ to indicate vertices from the two sides of the bipartite graph — either $\odot$ indicates clause and $\otimes$ indicates variable, or the other way around (it does not matter which).

Proof of Proposition 3.24. Let $\mathscr{G} \sim \mathbb{P} \equiv \mathbb{P}_{n,m}$. As in Proposition 5.31, let e_1, e_2 be two edges in the processed graph $\mathrm{pr}\,\mathscr{G}$ having the same total type, and define the switching operation sw. For any $H \in \mathbf{CM}(\mathcal{D})$,

$$\mathbb{P}(\mathrm{pr}\,\mathscr{G} = H) = \sum_G \mathbb{P}(\mathscr{G} = G)\mathbf{1}\{\mathrm{pr}\,G = H\} \overset{(a)}{=} \sum_G \mathbb{P}(\mathscr{G} = G)\mathbf{1}\{\mathrm{pr}(\mathrm{sw}\,G) = H\}$$

$$\overset{(b)}{=} \sum_G \mathbb{P}(\mathscr{G} = G)\mathbf{1}\{\mathrm{sw}(\mathrm{pr}\,\mathscr{G}) = H\} = \mathbb{P}(\mathrm{pr}\,\mathscr{G} = \mathrm{sw}\,H).$$

In the above, the step marked (a) follows from the fact the law of the original graph $\mathscr{G}$ is invariant under the switching operation: $\mathbb{P} \circ \mathrm{sw}^{-1} = \mathbb{P}$. Equality (b) holds by Proposition 5.31, and equality (c) holds since the switching operation is involutive. Returning to the definition of $\mathbf{CM}(\mathcal{D})$ (Remark 3.21), the above proves $\mathbb{P}(\mathrm{pr}\,\mathscr{G} = H) = \mathbb{P}(\mathrm{pr}\,\mathscr{G} = H')$ for any $H, H' \in \mathbf{CM}(\mathcal{D})$. Finally, it is clear that the law of $\mathrm{pr}\,\mathscr{G}$ is invariant under any permutation of the ordering among the variables or among the clauses, so if we condition on $\mathcal{D}_{\mathrm{pr}\,\mathscr{G}} = \mathcal{D}$ then the law of $\mathcal{D}_{\mathrm{pr}\,\mathscr{G}}$ is uniformly random among all $\mathcal{D} \sim \mathscr{D}$. This concludes the proof. $\qquad\square$

6. Extendibility and separability

In this section we prove Proposition 3.30 and Proposition 3.31. The section is organized as follows:

– In §6.1 we return to the issue that if a clause a receives incoming messages $\star\dot{q}_e$ ($e \in \delta a$), the resulting marginals on its incident edges does not generally agree with the canonical measures $\star\pi_e$ from Definition 3.4. We previously saw in §3.7 (specifically, Corollary 3.55) that if the clause is **coherent**, then it can be reweighted to achieve marginals $\star\pi_e$. In §6.1 we show that if the clause is **nice**, then we can explicitly construct and estimate these weights. This result will be used in the remainder of the current section, as well as in later sections.

– In §6.2 we introduce the **planted measure** $\mathbb{Q}_{\mathscr{D}}$. This is a well-known idea, which in the context of our problem gives a way to view the quantitites $\mathbb{E}_{\mathscr{D}}Z$, $\mathbb{E}_{\mathscr{D}}Z_{\mathrm{ext}}$, and $\mathbb{E}_{\mathscr{D}}Z_{\mathrm{sep}}$ in terms of a certain modified configuration model (Remark 6.3) that is tractable to analyze.

– In §6.3 we prove Proposition 3.30, which can be reinterpreted as saying that most colorings are extendible under the planted measure. The idea of the proof is to show that, under the planted measure, the subgraph of dependent free variables is with high probability a disjoint union of trees and unicyclic components, which can be completed to produce a valid satisfying assignment. The analysis of the free subgraph relies on the containment property (3.24) of compound enclosures.

– In §6.4 we prove Proposition 3.31. The idea of the proof is to note that if a (processed) k-sat instance $\mathscr{G} = (V, F, E)$ admits two judicious colorings that disagree on subset of variables V' with $1 \ll |V'| \ll |V|$, then $\mathscr{G}$ must admit a certain combinatorial structure (Lemma 6.11) which we then show is unlikely to occur. The extraction of the compound structure relies on the fact that compound enclosures are bounded by variables that are perfect, hence orderly (Definition 3.12).

6.1. **Explicit Lagrange multipliers for nice clauses.** Recall from §3.7 (in particular, see Corollary 3.55) that whenever a clause a is strictly coherent (Definition 3.6), there exists a set of weights such that the associated Gibbs measure on colorings of a has edge marginals $\star\pi$ (Definition 3.4). In this subsection, under the assumption that the clause is sufficiently nice, we give a direct construction and error estimate for these weights. This result will be used in the proofs that appear later in the current section. Moreover, the analysis in this subsection is a simplified version of the analysis of Section 7.

Through this subsection we are concerned with the reweighting of the edges around a single clause. Recall from (3.5) that we introduced the composite color $\mathtt{c} \equiv \mathtt{cyan} \equiv \{\mathtt{green}, \mathtt{blue}\}$, for the reason that the clause factor (2.20) does not distinguish between $\mathtt{green}$ and $\mathtt{blue}$. Therefore in this subsection we can work on the reduced alphabet $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$. Given a variable-to-clause message $\dot{q}$ which is a probability measure over $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$, we now abuse notation and write $\dot{q}$ for the measure on $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$ where $\mathtt{c}$ takes the combined weight of $\{\mathtt{g}, \mathtt{b}\}$. On the other hand, given a clause-to-variable message $\hat{q}$ which is a probability measure over $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$ such that $\hat{q}(\mathtt{b}) = \hat{q}(\mathtt{g})$, we define a probability measure $\tilde{q}$ on $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$ with weights

$$\left(\tilde{q}(\mathtt{r}), \tilde{q}(\mathtt{y}), \tilde{q}(\mathtt{c})\right) = \left(\frac{\hat{q}(\mathtt{r})}{1 - \hat{q}(\mathtt{b})}, \frac{\hat{q}(\mathtt{y})}{1 - \hat{q}(\mathtt{b})}, \frac{\hat{q}(\mathtt{b})}{1 - \hat{q}(\mathtt{b})}\right). \tag{6.1}$$

For the rest of the subsection we work with the measures $\dot{q}, \tilde{q}$ over $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$.

As we will soon see, the bounds that we can obtain for the BP recursion are different for $\mathtt{red}$ versus the other colors. For this reason, given a variable-to-clause message $\dot{q} \equiv \dot{q}_{va}$, it will be useful to define a reweighted version

$$Q_{va}(\sigma) = \frac{\dot{q}(\sigma)}{(2^{|\partial a|-1})^{\mathbf{1}\{\sigma=\mathtt{r}\}}} \Bigg/ \left\{\frac{\dot{q}(\mathtt{r})}{2^{|\partial a|-1}} + \dot{q}(\mathtt{c}) + \dot{q}(\mathtt{y})\right\}. \tag{6.2}$$

With this notation, we can now state and prove the main technical result of this subsection:

Lemma 6.1. *Let $\delta \in (0, 1]$ be a fixed constant. Suppose the clause a receives incoming variable-to-clause messages $\dot{q}_e$ ($e \in \delta a$) whose reweightings Q_e (as defined by (6.2)) satisfy the bounds*

$$\max\left\{Q_e(\mathtt{y}) - \frac{1}{2}, Q_e(\mathtt{r})\right\} \le \frac{1}{2^{k\delta}} \tag{6.3}$$

for all $e \in \delta a$. Suppose also we have outgoing messages $\tilde{q}_e$ such that, for all $e \in \delta a$, we have

$$\left| \frac{[\text{BP}_e[\dot{q}]](\sigma)}{\tilde{q}_e(\sigma)} - 1 \right| \leq \begin{cases} \epsilon_e & \text{for } \sigma \in \{y, c\}, \\ \dot{\epsilon}_e & \text{for } \sigma = r, \end{cases} \tag{6.4}$$

where all the errors $\epsilon_e, \dot{\epsilon}_e$ are at most $1/k^4$. Then there exist edge weights $\Gamma \equiv (\gamma_e : e \in \delta a)$ with $\gamma_e(y) = 1$ such that

$$\max_{e \in \delta a} \left\{ \sum_{\sigma \in \{r,c\}} |\gamma_e(\sigma) - 1| \right\} \leq k \sum_{e \in \delta a} \left(\epsilon_e + \dot{\epsilon}_e \right), \tag{6.5}$$

and such that the Γ-weighted BP recursion at clause a maps $\dot{q}$ to $\tilde{q}$: that is, such that $\tilde{q}_e = \text{BP}_e[\dot{q}; \Gamma]$ for all $e \in \delta a$.

Proof. We will iteratively define a sequence of weights $\Gamma^t \equiv (\gamma_{e,t} : e \in \delta a)$, started from $\Gamma^0 \equiv 1$ and converging in the limit $t \to \infty$ to the desired weights Γ. We will maintain for all t that $\gamma_{e,t}(y) = 1$. Denote the output of the Γ^t-weighted recursion by

$$\tilde{q}_{e,t}(\sigma) \equiv \left(\text{BP}_e[\dot{q}, \Gamma^t] \right)(\sigma) \cong \gamma_{e,t}(\sigma) \left\{ \left(\text{BP}_e[\gamma^t \dot{q}] \right)(\sigma) \right\}. \tag{6.6}$$

Step 1. Definition of weights and errors between weights. In this step, we fix an edge e and abbreviate $\gamma \equiv \gamma_e$, $\tilde{q} \equiv \tilde{q}_e$, and so on. Given γ^t and $\tilde{q}^t$, we define the next weight γ^{t+1} by setting

$$\gamma^{t+1}(\sigma) \equiv \gamma^t(\sigma) \cdot \frac{\tilde{q}(\sigma)/\tilde{q}(y)}{\tilde{q}^t(\sigma)/\tilde{q}^t(y)}$$

for each $\sigma \in \{r, y, c\}$. Note this choice ensures $\gamma^{t+1}(y) = 1$. It remains to estimate the error between γ^t and γ^{t+1} on the other two colors $\{r, c\}$. To this end, let us define the error quantities

$$\epsilon^t(\sigma) \equiv \left| \frac{\gamma^{t+1}(\sigma)}{\gamma^t(\sigma)} - 1 \right| = \left| \frac{\tilde{q}(\sigma)/\tilde{q}(y)}{\tilde{q}^t(\sigma)/\tilde{q}^t y} - 1 \right|. \tag{6.7}$$

for $\sigma \in \{r, c\}$. The error at the next iteration is then given (after a short algebraic manipulation) by

$$\epsilon^{t+1}(\sigma) = \left| \underbrace{\left\{ \frac{\gamma^t(\sigma)}{\gamma^{t+1}(\sigma)} \frac{\tilde{q}(\sigma)/\tilde{q}(y)}{\tilde{q}^t(\sigma)/\tilde{q}^t(y)} \right\}}_{\text{equals one}} \frac{\tilde{q}^t(\sigma)/(\gamma^t(\sigma)\tilde{q}^t(y))}{\tilde{q}^{t+1}(\sigma)/(\gamma^{t+1}(\sigma)\tilde{q}^{t+1}(y))} - 1 \right|. \tag{6.8}$$

For the rest of this proof we use the shorthand $\epsilon^t \equiv \epsilon^t(c)$ and $\dot{\epsilon}^t \equiv \epsilon^t(r)$.

Step 2. Errors for variable-to-clause quantities. In this step we continue to consider a single edge $e \in \delta a$, which we suppress from the notation. As in (6.2), let Q^t be the reweighted version of $\gamma^t \dot{q}$ defined by

$$Q^t(\sigma) \equiv \frac{\gamma^t(\sigma)\dot{q}(\sigma)}{(2^{|\partial a|-1})^{\mathbf{1}\{\sigma=r\}}} \left/ \left\{ \frac{\gamma^t(r)\dot{q}(r)}{2^{|\partial a|-1}} + \dot{q}(y) + \gamma^t(c)\dot{q}(c) \right\} \right. . \tag{6.9}$$

Recall the assumption (6.3) that $Q(r) = Q_0(r) \leq 2^{-k\delta}$. It will follow from the inductive analysis below that all the weights γ^t are of constant order, so we will also have $Q^t(r) \leq O(2^{-k\delta})$. It follows that

$$\frac{\gamma^{t+1}(r)\dot{q}(r)}{2^{|\partial a|-1}} + \dot{q}(y) + \gamma^{t+1}(c)\dot{q}(c) = \frac{\gamma^t(r)\dot{q}(r)}{2^{|\partial a|-1}}\left(1 + O(\dot{\epsilon}^t)\right) + \dot{q}(y) + \gamma^t(c)\dot{q}(c)\left(1 + O(\epsilon^t)\right)$$

$$= \left\{ \frac{\gamma^t(r)\dot{q}(r)}{2^{|\partial a|-1}} + \dot{q}(y) + \gamma^t(c)\dot{q}(c) \right\}\left\{ 1 + O\left(\epsilon^t + \frac{\dot{\epsilon}^t}{2^{k\delta}} \right) \right\},$$

for $\dot{\epsilon}^t$ and ϵ^t as in (6.7). Combining this with (6.9) gives

$$\left| \frac{Q^{t+1}(\sigma)}{Q^t(\sigma)} - 1 \right| \leq O(1) \begin{cases} \epsilon^t + \dot{\epsilon}^t/2^{k\delta} & \text{for } \sigma \in \{y, c\}, \\ \epsilon^t + \dot{\epsilon}^t & \text{for } \sigma = r. \end{cases} \tag{6.10}$$

Step 3. Errors output by clause recursion. Recall from (6.6) that $\tilde{q}_{e,t}/\gamma_{e,t} \cong \mathrm{BP}_e[\gamma^t \dot{q}]$, and recall from (6.9) that Q^t is a reweighted version of $\gamma^t \dot{q}$. We therefore let $z_{e,t}$ denote the normalizing constant such that

$$\frac{\tilde{q}_{e,t}(\mathtt{r})}{\gamma_{e,t}(\mathtt{r})} = \frac{1}{z_{e,t}} \prod_{e' \in \delta a \setminus e} Q_{e',t}(\mathtt{y}).$$

It the follows from the first bound in (6.10) that for $\sigma = \mathtt{red}$ we have

$$\frac{z_{e,t+1}\tilde{q}_{e,t+1}(\mathtt{r})}{\gamma_{e,t+1}(\mathtt{r})} = \prod_{e' \in \delta a \setminus e} Q_{e'}^{t+1}(\mathtt{y}) = \frac{z_{e,t}\tilde{q}_{e,t+1}(\mathtt{r})}{\gamma_{e,t}(\mathtt{r})} \left\{ 1 + O\left(\sum_{e' \in \delta a} \left(\epsilon_{e',t} + \frac{\dot{\epsilon}_{e',t}}{2^{k\delta}} \right) \right) \right\}.$$

Next, by (6.10) together with the assumption (6.3), we have for $\sigma = \mathtt{cyan}$ that

$$\frac{z_{e,t+1}\tilde{q}_{e,t+1}(\mathtt{c})}{\gamma_{e,t+1}(\mathtt{c})} = \prod_{e' \in \delta a \setminus e} \left(1 - Q_{e'}^{t+1}(\mathtt{r}) \right) - \prod_{e' \in \delta a \setminus e} Q_{e'}^{t+1}(\mathtt{y})$$

$$= \prod_{e' \in \delta a \setminus e} \left(1 - Q_{e',t}(\mathtt{r})\left\{ 1 + O\left(\epsilon_{e',t} + \dot{\epsilon}_{e',t} \right) \right\} \right) - \prod_{e' \in \delta a \setminus e} \left(Q_{e',t}(\mathtt{y})\left\{ 1 + O\left(\epsilon_{e',t} + \frac{\dot{\epsilon}_{e',t}}{2^{k\delta}} \right) \right\} \right)$$

$$= \frac{z_{e,t}\tilde{q}_{e,t}(\mathtt{c})}{\gamma_{e,t}(\mathtt{c})} \left\{ 1 + O\left(\sum_{e' \in \delta a} \left(\frac{\epsilon_{e',t} + \dot{\epsilon}_{e',t}}{2^{k\delta}} \right) \right) \right\}.$$

Lastly, for $\sigma = \mathtt{yellow}$, it follows from (6.3) and (6.10) that

$$z_{e,t+1}\tilde{q}_{e,t+1}(\mathtt{y}) = \prod_{e' \in \delta a \setminus e} \left(1 - Q_{e'}^{t+1}(\mathtt{r}) \right) - \prod_{e' \in \delta a \setminus e} Q_{e'}^{t+1}(\mathtt{y}) + \sum_{e' \in \delta a \setminus e} \left(2^{|\partial a|-1} Q_{e'}^{t+1}(\mathtt{r}) - Q_{e'}^{t+1}(\mathtt{c}) \right) \prod_{e'' \in \delta a \setminus \{e,e'\}} Q_{e''}^{t+1}(\mathtt{y})$$

$$= z_{e,t}\tilde{q}_{e,t}(\mathtt{y}) \left\{ 1 + O\left(k \sum_{e' \in \delta a} \left(\frac{\epsilon_{e',t} + \dot{\epsilon}_{e',t}}{2^{k\delta}} \right) \right) \right\}.$$

Now substitute these estimates into the quantity appearing on the right-hand side of (6.8): it gives

$$\epsilon_{e,t+1}(\mathtt{c}) = \left| \frac{\tilde{q}_{e,t}(\mathtt{c})/(\gamma_{e,t}(\mathtt{c})\tilde{q}_{e,t}(\mathtt{y}))}{\tilde{q}_{e,t+1}(\mathtt{c})/(\gamma_{e,t+1}(\mathtt{c})\tilde{q}_{e,t+1}(\mathtt{y}))} - 1 \right| = O\left(k \sum_{e' \in \delta a} \left(\frac{\epsilon_{e',t} + \dot{\epsilon}_{e',t}}{2^{k\delta}} \right) \right),$$

$$\epsilon_{e,t+1}(\mathtt{r}) = \left| \frac{\tilde{q}_{e,t}(\mathtt{r})/(\gamma_{e,t}(\mathtt{r})\tilde{q}_{e,t}(\mathtt{y}))}{\tilde{q}_{e,t+1}(\mathtt{r})/(\gamma_{e,t+1}(\mathtt{r})\tilde{q}_{e,t+1}(\mathtt{y}))} - 1 \right| = O\left(k \sum_{e' \in \delta a} \left(\epsilon_{e',t} + \frac{\dot{\epsilon}_{e',t}}{2^{k\delta}} \right) \right).$$

Consequently, if we aggregate all the error terms at time t as

$$E(t) \equiv \sum_{e \in \delta a} \left(\epsilon_{e,t}(\mathtt{c}) + \frac{\epsilon_{e,t}(\mathtt{r})}{2^{k\delta/2}} \right),$$

then we shall obtain at the next step

$$E(t+1) \le O(k^2) \sum_{e' \in \delta a} \left(\frac{\epsilon_{e',t} + \dot{\epsilon}_{e',t}}{2^{k\delta}} \right) + \frac{O(k^2)}{2^{k\delta/2}} \sum_{e' \in \delta a} \left(\epsilon_{e',t} + \frac{\dot{\epsilon}_{e',t}}{2^{k\delta}} \right) \le \frac{O(k^2)}{2^{k\delta/2}} E(t).$$

It follows that $E(t)$ decays exponentially in t, so the sequence Γ^t converges to a limit Γ. Summing over t gives the claimed error bound on the weights γ_e. $\qquad\square$

An application of Lemma 6.1 is a direct construction, with an error estimate, for the weights of Corollary 3.55 and Corollary 3.56 in the case that the edges are nice:

Corollary 6.2. *In the setting of Corollaries 3.55 and 3.56, suppose all the edges in U are stable and nice. Then, for every clause a of U, the canonical messages $_\star\dot{q}_e$ and $_\star\hat{q}_e$ (for $e \in \delta a$) satisfy the conditions of Lemma 6.1, with $\delta = 1/11$ in (6.3) and $\epsilon_e, \dot{\epsilon}_e \le O(1/k^r)$ in (6.4). Consequently, Lemma 6.1 guarantees, for each a in U, the existence of edge weights $_\star\Gamma_a = (\gamma_e : e \in \delta a)$ satisfying (cf. (6.5)) $|\gamma_e(\sigma) - 1| \le 1/k^{r/2}$ for all σ, and $_\star\hat{q}_e = \mathrm{BP}_e(_\star\dot{q}, {}_\star\Gamma_a)$ for all $e \in \delta a$. As a consequence, the weighted Gibbs measure (3.90) has all edge marginals and messages agreeing with the canonical ones $_\star\pi, {}_\star\dot{q}, {}_\star\hat{q}$. Redistributing the weights as in (2.30) (and as in Corollary 3.56) produces a system $_\star\Lambda$ of variable weights*

such that the $_\star\Lambda$-weighted Gibbs measure again has edge marginals $_\star\pi$. (However, since the weights were shifted from clauses to variables, the BP *messages will be given by $_\bullet\dot{q}, _\bullet\hat{q}$ from* (3.92) *rather than by $_\star\dot{q}, _\star\hat{q}$).*

Proof. We need only to check that the conditions of Lemma 6.1 are satisfied. Indeed, for each edge $e \in U$, if we take $_\star\dot{q}_e$ and use (6.2) to define its reweighted version $_\star Q_e$, then $_\star Q_e$ will satisfy the error bound (6.3) with $\delta = 1/11$, because the edges in U are assumed to be nice (Definition 3.8). Moreover, we will have the error bounds (6.4) with $\epsilon_e, \dot{\epsilon}_e \leq O(1/k^r)$ because the edges in U are assumed to be stable (Definition 3.7). The claimed result then follows directly from Lemma 6.1.   □

6.2. The planted measure for judicious colorings.

In preparation for the proofs of Propositions 3.30 and 3.31, we introduce the planted measure in this subsection. Given a processed neighborhood sequence $\mathcal{D}$ (as in Definition 3.20), let $\mathbb{Q}_\mathcal{D}$ denote the uniform measure over all pairs $(\mathscr{G}, \underline{\sigma})$ such that $\mathscr{G} \equiv (V, F, E)$ is consistent with the sequence $\mathcal{D}$, and $\underline{\sigma}$ is a valid judicious coloring of $\mathscr{G}$. Then, recalling Definition 3.28 and (3.32), we have

$$\mathbb{Q}_\mathcal{D}\left(\left\{(\mathscr{G}, \underline{\sigma}) : \underline{\sigma} \text{ is extendible}\right\}\right) = \frac{\mathbb{E}_\mathcal{D} Z_{\text{ext}}}{\mathbb{E}_\mathcal{D} Z} = \frac{\mathbb{E}_\mathscr{D} Z_{\text{ext}}}{\mathbb{E}_\mathscr{D} Z}, \tag{6.11}$$

where $\mathscr{D}$ is the unordered version of $\mathcal{D}$ (see Definition 3.20). Thus, to prove Proposition 3.30, it suffices to show that $\underline{\sigma}$ is extendible with high probability under $\mathbb{Q}_\mathcal{D}$. Similarly, to prove Proposition 3.31, it suffices to show that $\underline{\sigma}$ is separable with high probability under $\mathbb{Q}_\mathcal{D}$. Following standard convention, we call $\mathbb{Q}_\mathcal{D}$ the **planted measure**.

Remark 6.3 (sampling from the planted measure). Recall from Proposition 3.24 that $\mathbb{P}_\mathcal{D}$ coincides with the uniform measure over $\mathbf{CM}(\mathcal{D})$, and can be sampled as a configuration model, as discussed in Remark 3.21. It is well known that one can also sample from $\mathbb{Q}_\mathcal{D}$ by a configuration-model-type procedure, as we now describe. As in Remark 3.21, we fix a set of variables V and a set of clauses F, equipped with incident half-edges δV and δF, all labelled with types according to $\mathcal{D}$. Then a graph $\mathscr{G}$ corresponds to a matching $\mathfrak{M}$ of δV to δF that respects the edge types. The number of all such matchings is given by (3.27). Let σ be any coloring of $\delta V \sqcup \delta F$ that gives a valid coloring of δx for every $x \in V \sqcup F$. We say that a valid coloring σ is **judicious**, abbreviated $\sigma \in \mathcal{J}$, if the empirical measure of $\sigma : \delta V \to \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$ agrees with the canonical marginal $_\star\pi$, and the empirical measure of $\sigma : \delta F \to \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$ agrees with $_\star\omega$ (up to rounding, cf. Definition 3.26). We write $\mathfrak{M} \sim \sigma$ if the matching $\mathfrak{M}$ also respects the edge colors specified by σ. Then a pair $(\mathscr{G}, \underline{\sigma})$ is equivalent to a pair $(\mathfrak{M}, \sigma)$ such that $\mathfrak{M} \sim \sigma$. Thus we can regard $\mathbb{Q}_\mathcal{D}$ as the measure over pairs $(\mathfrak{M}, \sigma)$ given by

$$\mathbb{Q}_\mathcal{D}(\mathfrak{M}, \sigma) = \frac{\mathbf{1}\{\sigma \in \mathcal{J}, \mathfrak{M} \sim \sigma\}}{\mathcal{Z}},$$

where $\mathcal{Z}$ is the normalizing constant. The marginal probability of σ under $\mathbb{Q}_\mathcal{D}$ is given by

$$\mathbb{Q}_\mathcal{D}(\sigma) = \frac{\mathbf{1}\{\sigma \in \mathcal{J}\}}{\mathcal{Z}} \cdot \left|\left\{\mathfrak{M} : \mathfrak{M} \sim \sigma\right\}\right| = \frac{\mathbf{1}\{\sigma \in \mathcal{J}\}}{\mathcal{Z}} \cdot \prod_{t,\sigma}(n_t \cdot {_\star\pi_t}(\sigma))!,$$

which we emphasize is constant over all judicious σ. This implies

$$\mathbb{Q}_\mathcal{D}(\sigma) = \frac{\mathbf{1}\{\sigma \in \mathcal{J}\}}{|\mathcal{J}|},$$

i.e., the marginal law of σ under $\mathbb{Q}_\mathcal{D}$ is simply uniform over $\mathcal{J}$. It further implies that for any $\sigma \in \mathcal{J}$,

$$\mathbb{Q}_\mathcal{D}(\mathfrak{M} \mid \sigma) = \frac{\mathbf{1}\{\mathfrak{M} \sim \sigma\}}{\mathcal{Z}/|\mathcal{J}|} = \frac{\mathbf{1}\{\mathfrak{M} \sim \sigma\}}{|\{\mathfrak{M} : \mathfrak{M} \sim \sigma\}|},$$

i.e., under $\mathbb{Q}_\mathcal{D}$, the law of $\mathfrak{M}$ conditional on σ is uniform among the matchings compatible with σ. In conclusion, to generate a sample from $\mathbb{Q}_\mathcal{D}$, we can first sample a uniformly random coloring $\sigma \in \mathcal{J}$, then sample a uniformly matching $\mathfrak{M}$ that satisfies $\mathfrak{M} \sim \sigma$.

Remark 6.4 (empirical measure of colorings under the planted measure). Given a coloring $\sigma \in \mathcal{J}$, we associate the variable empirical measure $\dot{\nu}$ as in (3.40), and the clause empirical measure $\hat{\nu}$ as in (3.41). Abbreviate $\nu \equiv (\dot{\nu}, \hat{\nu}) \equiv \nu(\sigma)$. Then, for any ν that is consistent with $_\star\pi$ and $_\star\omega$, we have

$$\left|\left\{\sigma \in \mathcal{J} : \nu(\sigma)\right\}\right| = \left\{\prod_T \binom{n_T}{n_T \hat{\nu}_T}\right\}\left\{\prod_L \binom{m_L}{m_L \hat{\nu}_L}\right\}.$$

By comparing with (3.48) and (3.49), we see that if σ is sampled uniformly at random from $\mathcal{J}$, then its empirical measure $\nu(\sigma)$ is very close to $\nu^{op}[\star\omega]$ with high probability: more precisely, as long as $\mathscr{D}$ is bounded away from zero in the sense of (3.45) (as guaranteed by Proposition 3.23 with high probability), then

$$\mathbb{Q}_{\mathcal{D}}\left(\left\|\nu - \nu^{op}[\star\omega]\right\|_{\infty} \geq \frac{\ln n}{n^{1/2}}\right) \leq \exp\left\{-C(k, R)(\ln n)^2\right\}.$$

We recall that $\nu^{op}[\star\omega] = (\dot{\nu}^{op}[\star\omega], \hat{\nu}^{op}[\star\omega])$ where $\dot{\nu} = \dot{\nu}^{op}[\star\omega]$ is given by Lemma 3.38, while $\hat{\nu} = \hat{\nu}^{op}[\star\omega]$ is given by a reweighted measure: if a denotes a clause of type $\mathbf{L}$, then

$$\hat{\nu}_{\mathbf{L}}(\underline{\sigma}_{\delta a}) = \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{e \in \delta a} \left\{\star\dot{q}_e(\sigma_e)\gamma_e(\sigma_e)\right\}$$

where the weights γ_e are given by Corollary 3.55, since all clauses in the processed graph must be coherent. If all the edges in the clause are nice, then the weights γ_e are estimated by Lemma 6.2. We will use this observation in the proofs that follow.

We conclude this subsection with a simple lemma regarding the matching of edges near defects under the planted measure. It will be used in the analysis of §6.4 below. We will say simply "defect" to refer to a 0-defect in the graph $\mathscr{G}$.

Remark 6.5. Recall from Definition 3.10 that if a variable is non-defective, then it has at most one defective variable in its depth-one neighborhood. The clauses $a \in F$ therefore can be divided into three categories:

(i) The clause a neighbors only defective variables, in which case we say that a is a **defective clause**.
(ii) The clause a has only one defective variable among its neighbors.
(iii) The clause a has no defective variables among its neighbors, in which case we say that a is a **strongly non-defective clause**. (The strongly non-defective property will be used in the proof of Proposition 9.17.)

In both cases (ii) and (iii) we say that a is a **non-defective clause**. If v is a non-defective variable, then ∂v contains only non-defective clauses; moreover, at most one clause in ∂v can fail to be strongly non-defective. We say that v is **strongly non-defective** if every clause in ∂v is strongly non-defective. We say that an edge e is **strongly non-defective** if and only if $e \in \delta v$ for a strongly non-defective variable v. Lastly, we say that an edge e is **internal** to a defect if it is incident to some internal clause of the defect.

Lemma 6.6. *As in Remark 6.3, fix V, F, δV, δF labelled with total types according to $\mathcal{D}$. From the notion of compound type (Definition 3.18), for each half-edge $e \in \delta V \sqcup \delta F$, we can deduce from its type t_e whether, in the final graph $\mathscr{G}$, the half-edge e will participate in an edge that is internal to a defect. For any matching $\mathfrak{m}$ of these half-edges, we have $\mathbb{Q}_{\mathcal{D}}(\mathfrak{m} \subseteq \mathscr{G}) = \mathbb{P}_{\mathcal{D}}(\mathfrak{m} \subseteq \mathscr{G})$.*

Proof. As in the statement of the lemma, fix V, F, δV, δF labelled with total types according to $\mathcal{D}$. It is clear from Definition 3.18 that for a half-edge e, the total type t_e encodes whether e will participate in an edge that is internal to a defect. That is to say, there is a set of edge types $\mathfrak{T}$ such that $t_e \in \mathfrak{T}$ if and only if e participates in the internal edge of a defect.

Let $\mathscr{G}$, $\mathscr{G}'$ be any two graphs with neighborhood sequence $\mathcal{D}$. Let $\mathscr{H}$, $\mathscr{H}'$ be the corresponding subgraphs induced by their defects; these are encoded by matchings $\mathfrak{m}$, $\mathfrak{m}'$ as described in the statement of the lemma. The matchings involve precisely the half-edges with total types in $\mathfrak{T}$. Since defects are contained in enclosures which are encoded by compound types, it follows that $\mathscr{H}$ and $\mathscr{H}'$ are isomorphic — equivalently, that there is an isomorphism ι which takes $\mathfrak{m} \mapsto \mathfrak{m}'$. Let $\mathscr{G}$ be any graph in which $\mathfrak{m}$ appears. We can extend ι to $\mathscr{G}$ by applying the identity map to edges not in $\mathfrak{m}$; then $\mathfrak{m}'$ appears in $\iota(\mathscr{G})$. If $\underline{\sigma}$ is a valid coloring of $\mathscr{G}$, then a valid coloring of $\iota(\mathscr{G})$ is given by

$(\underline{\sigma} \circ \iota^{-1})_e \equiv \sigma_{\iota^{-1}(e)}$. If $\mathbb{E}_\mathcal{D}$ denotes expectation under $\mathbb{P}_\mathcal{D}$, then

$$\mathbb{E}_\mathcal{D}\left(\mathbf{Z} \,\middle|\, \mathfrak{m} \subseteq \mathcal{G}\right) = \frac{1}{\mathbb{P}_\mathcal{D}(\mathfrak{m} \subseteq \mathcal{G})} \sum_{G:\mathfrak{m} \subseteq G} \mathbb{P}_\mathcal{D}(G) \sum_{\underline{\sigma}} \mathbf{1}\{\underline{\sigma} \text{ is a judicious coloring of } G\}$$

$$= \frac{1}{\mathbb{P}_\mathcal{D}(\mathfrak{m}' \subseteq \mathcal{G})} \sum_{G:\mathfrak{m} \subseteq G} \mathbb{P}_\mathcal{D}(\iota(G)) \sum_{\underline{\sigma}} \mathbf{1}\{\underline{\sigma} \circ \iota^{-1} \text{ is a judicious coloring of } \iota(G)\}$$

$$= \frac{1}{\mathbb{P}_\mathcal{D}(\mathfrak{m}' \subseteq \mathcal{G})} \sum_{G':\mathfrak{m}' \subseteq G'} \mathbb{P}_\mathcal{D}(G') \sum_{\underline{\sigma}'} \mathbf{1}\{\underline{\sigma}' \text{ is a judicious coloring of } G'\} = \mathbb{E}_\mathcal{D}\left(\mathbf{Z} \,\middle|\, \mathfrak{m}' \subseteq \mathcal{G}\right).$$

It follows from the definition of $\mathbb{Q}_\mathcal{D}$ that

$$\mathbb{Q}_\mathcal{D}(\mathfrak{m} \subseteq \mathcal{G}) = \frac{1}{\mathbb{E}_\mathcal{D}\mathbf{Z}} \sum_G \mathbf{1}\{\mathfrak{m} \subseteq G\}\mathbb{P}_\mathcal{D}(G)\mathbf{Z}(\mathcal{G}) = \mathbb{P}_\mathcal{D}(\mathfrak{m} \subseteq \mathcal{G})\frac{\mathbb{E}_\mathcal{D}(\mathbf{Z} \,|\, \mathfrak{m} \subseteq \mathcal{G})}{\mathbb{E}_\mathcal{D}\mathbf{Z}} = \mathbb{P}_\mathcal{D}(\mathfrak{m} \subseteq \mathcal{G}),$$

as claimed. $\square$

6.3. **Extendibility for judicious colorings.** In this subsection we prove Proposition 3.30, which says that the first moment $\mathbb{E}_\mathcal{D}\mathbf{Z}$ of judicious colorings is dominated by extendible colorings (Definition 3.28), with high probability over $\mathcal{D}$.

Definition 6.7 (free subgraph). Given a valid coloring $\underline{\sigma}$ of $\mathcal{G} = (V, F, E)$, we define a subgraph $\mathfrak{F} \subseteq \mathcal{G}$ as follows. Let $V_\mathfrak{F}$ be the variables in V that are incident to only green edges under $\underline{\sigma}$. Let $F_\mathfrak{F}$ be the clauses in F that are incident to only green or yellow edges under $\underline{\sigma}$. Note that, by the rules of the coloring model, each $a \in F_\mathfrak{F}$ must be incident to at least two variables $u, v \in V_\mathfrak{F}$. Let $E_\mathfrak{F}$ be the edges between $V_\mathfrak{F}$ and $F_\mathfrak{F}$. We shall call $\mathfrak{F} \equiv (V_\mathfrak{F}, F_\mathfrak{F}, E_\mathfrak{F})$ the **free subgraph of $\mathcal{G}$ induced by $\underline{\sigma}$**.

Lemma 6.8. *In the setting of Proposition 3.30 we have*

$$\mathbb{Q}_\mathcal{D}\left(\left\{(\mathcal{G}, \underline{\sigma}) : \text{any connected component of } \mathfrak{F} \text{ contains a bicycle}\right\}\right) = o_n(1),$$

with high probability over $\mathcal{D}$.

Proof. We assume throughout the proof that $\mathcal{D}$ is fixed, such that its unordered version $\mathscr{D}$ is bounded away from zero in the sense of (3.45) (as guaranteed by Proposition 3.23 with high probability). Then, as in Remark 6.3, we first sample a uniformly random coloring $\sigma \in \mathcal{J}$, followed by a uniformly random matching $\mathfrak{M}$ such that $\mathfrak{M} \sim \sigma$. The resulting pair $(\mathfrak{M}, \sigma)$ is equivalent to a sample $(\mathcal{G}, \underline{\sigma}) \sim \mathbb{Q}_\mathcal{D}$. Let $\nu \equiv (\dot{\nu}, \hat{\nu})$ be the empirical measure of $\underline{\sigma}$. We can assume that

$$\left\|\nu - \nu^{\mathrm{op}}[\star\omega]\right\|_\infty \le \frac{\ln n}{n^{1/2}}, \tag{6.12}$$

since this event holds with high probability under $\mathbb{Q}_\mathcal{D}$ by Remark 6.4. Moreover, given σ, the random matching $\mathfrak{M}$ can be explored in breadth-first manner, as we analyze next.

Fix an initial variable $v \in V$ of type T, and suppose under σ that it is incident to only green edges. Let $\mathfrak{F}(v)$ be the connected component of $\mathfrak{F}$ containing v. Under the randomness of the matching $\mathfrak{M}$, the expected number of variables $w \in \mathfrak{F}(v)$ at unit distance from v is

$$f(T) \equiv \sum_{e \in \delta v} \sum_L \pi(L \,|\, t_e) \underbrace{\sum_{\underline{\sigma}_{\delta a} \in \{\mathtt{g},\mathtt{y}\}^{\delta a}} \hat{\nu}_L(\underline{\sigma}_{\delta a} \,|\, \sigma_e = \mathtt{g}) \sum_{e' \in \delta a \setminus e} \mathbf{1}\{\sigma_{e'} = \mathtt{g}\}}_{\text{denote this } g(t, L)}. \tag{6.13}$$

In the above $\pi(L \,|\, t)$ is the chance that an edge $e \in \delta v$ of type t is matched with an edge $e' \in \delta a$ such that $L_a = L$. Recall from Definition 2.1 that the type t includes the position $j(t)$ that the edge takes in the clause, so

$$\pi(L \,|\, t) = \hat{\mathscr{D}}(L) \,\middle/\, \left(\sum_{L'} \mathbf{1}\{(L')_{j(t)} = t\}\hat{\mathscr{D}}(L')\right). \tag{6.14}$$

Consider the quantity $g(t, L)$ defined by (6.13): it is given explicitly by

$$g(t, L) = \frac{\displaystyle\sum_{e' \in \delta a \setminus e} \star\dot{q}_{e'}(\mathbf{g})\gamma_{e'}(\mathbf{g}) \prod_{e'' \in \delta a \setminus \{e, e'\}} \left(\sum_{\sigma \in \{\mathbf{y}, \mathbf{g}\}} \star\dot{q}_{e''}(\sigma)\gamma_{e''}(\sigma)\right)}{\displaystyle\prod_{e'' \in \delta a \setminus e} \left(\sum_{\sigma \in \{\mathbf{y}, \mathbf{g}, \mathbf{b}\}} \star\dot{q}_{e''}(\sigma)\gamma_{e''}(\sigma)\right) - \prod_{e'' \in \delta a \setminus e} \left(\star\dot{q}_{e''}(\mathbf{y})\gamma_{e''}(\mathbf{y})\right)} \left[1 + O\left(\frac{\ln n}{n^{1/2}}\right)\right], \qquad (6.15)$$

where the error estimate comes from (6.12). If L is a nice clause type, then we can use Definition 3.8 together with Corollary 6.2 to bound $g(t, L) \le O(k/4^k)$.

If the variable v is of a non-defective type T, (Definition 3.10), then it can only neighbor nice clause types L, and must have degree $|\delta v| = O(k2^k)$. Thus we see from (6.13) that $f(T) \le O(k^2/2^k)$. If T is a defective type, then we recall from (3.24) that v must lie in a compound enclosure U of diameter at most R', such that the containment radius $s = \mathrm{rad}(v)$ is upper bounded by the distance between v and the boundary of U. Moreover, by Definition 3.18, the total type T encodes the isomorphism class of U, the position of v within U, and the simple total type of every edge in U. It follows that T encodes the isomorphism class of $B_s(v)$, and the simple total type of every clause in $B_s(v)$. As before, suppose that v is incident under σ to only green edges. Then, under the randomness of the matching $\mathfrak{M}$, the expected number of variables in $\mathfrak{F}(v) \cap B_s(v)$ is

$$f(T) \equiv \sum_{\ell=1}^{s} \sum_{v_0 a_1 v_1 \cdots a_\ell v_\ell} \prod_{j=1}^{\ell} \underbrace{\left(\sum_{\underline{\sigma}_{\delta a_j}} \hat{v}_{a_j}\left(\underline{\sigma}_{\delta a_j} \mid \sigma_{v_{j-1} a_j} = \mathbf{g}\right) \mathbf{1}\left\{\underline{\sigma}_{\delta a_j} \in \{\mathbf{y}, \mathbf{g}\}^{\delta a_j}, \sigma_{a_j v_j} = \mathbf{g}\right\}\right)}_{g(v_{j-1}, a_j, v_j)},$$

where the second summation is over all paths $v = v_0 a_1 v_1 \cdots a_\ell v_\ell$ emanating from v. If a_j is a nice clause, then we have $g(v_{j-1}, a_j, v_j) \le O(1/4^k)$ by a similar calculation as for (6.15). In any case we always have $g(v_{j-1}, a_j, v_j) \le 1$. Recall from Definition 3.11 that $\mathfrak{B}(v, w)$ counts the number of defective variables on the shortest path between v and w (including the endpoints). By Definition 3.10, any clause neighboring to a non-defective variable must be nice, so the number of nice clauses on the shortest path between v and w must be at least $d(v, w) - \mathfrak{B}(v, w)$. It follows that

$$f(T) \le \sum_{w \in B_s(v)} \left(\frac{O(1)}{4^k}\right)^{d(v,w) - \mathfrak{B}(v,w)} \le \frac{1}{4},$$

where the last inequality holds by the definition (3.19) of the containment radius, since we set $s = \mathrm{rad}(v)$.

Now consider exploration of $\mathfrak{F}(v)$ by the following modified breadth-first search procedure. We maintain a queue (a first-in first-out list) of variables, starting from $Q_0 = (v)$. Then, at each time step $\ell \ge 1$, we remove the first element w_ℓ of $Q_{\ell-1}$ to produce $(Q_{\ell-1})'$. We then explore the neighborhood of w_ℓ to depth s_ℓ where $s_\ell = 1$ if w_ℓ is non-defective, and $s_\ell = \mathrm{rad}(w_\ell)$ otherwise. Let $\mathfrak{F}(v)_\ell$ be the variables of $\mathfrak{F}(v)$ that are *newly* discovered in this exploration. We then take the ones at the boundary of $B_{s_\ell}(w_\ell)$ and append them to the queue:

$$Q_\ell = \left((Q_{\ell-1})', \mathfrak{F}(v)_\ell \cap \partial_\circ B_{s_\ell}(w_\ell)\right).$$

The exploration continues until the first time $\tau(v)$ that $Q_{\tau(v)} = \varnothing$. Let $\zeta_\ell \equiv |Q_\ell| - |Q_{\ell-1}|$. Under the randomness of the matching $\mathfrak{M}$, as long as $\ell/n = o_n(1)$ and $Q_{\ell-1} \ne \varnothing$, it follows from the preceding bounds on $f(T)$ that

$$\mathbb{E}\zeta_\ell \le -1 + \frac{1}{4}\left(1 + o_n(1)\right).$$

Once some vertices have already been explored, (6.14) is no longer an exact expression for the conditional law of subsequent clause types in the exploration — however, since we assume that $\ell/n = o_n(1)$ and that $\mathscr{D}$ is bounded away from zero, there remains a linear number of unexplored vertices of each type, so (6.14) is correct up to a multiplicative factor $1 + o_n(1)$. Thus, as long as $\ell/n = o_n(1)$ and $Q_{\ell-1} \ne \varnothing$, we have $\mathbb{E}\zeta_\ell \le -1/2$. Since all variables in $\mathscr{G}$ must be fair (Definition 3.13), we have $0 \le 1 + \zeta_\ell \le \exp(k^2 R)$ with probability one. It follows by the Azuma–Hoeffding

inequality that for a large enough constant $C = C(k, R)$,

$$\mathbb{P}\Big(\tau(v) \geq C \ln n\Big) \leq \mathbb{P}\Big(|Q_{C \ln n}| - \mathbb{E}|Q_{C \ln n}| \geq \frac{C \ln n}{2}\Big) \ll \frac{1}{n}.$$

On the other hand, again using that $\mathscr{D}$ is bounded away from zero, we see that the chance for the exploration to close more than two cycles is at most $(\ln n)^{O(1)}/n^2 \ll 1/n$. Taking a union bound over all $v \in V$ gives

$$\mathbb{Q}_{\mathscr{D}}\Big(\Big\{(\mathscr{G}, \underline{\sigma}) : \text{any connected component of } \mathfrak{F} \text{ contains a bicycle}\Big\}\Big) = o_n(1),$$

as long as $\mathscr{D}$ is bounded away from zero in the sense of (3.45). The result follows by appealing to Proposition 3.23. $\quad\square$

Proof of Proposition 3.30. Let $\mathscr{G}' \sim \mathbb{P} \equiv \mathbb{P}_{n,m}$ for $|m - n\alpha| \leq n^{1/2} \ln n$. Let $\mathscr{G} \equiv \operatorname{pr} \mathscr{G}'$ be the processed graph given by Definition 3.15. Let $\underline{\sigma}$ be any valid coloring of $\mathscr{G}$, and let $\underline{x}$ be its corresponding frozen configuration: as long as the free subgraph of $\mathscr{G}$ induced by $\underline{\sigma}$ does not contain a bicycle, $\underline{x}$ can be extended to a satisfying assignment of $\mathscr{G}$, that is to say, $\underline{\sigma}$ is extendible (Definition 3.28). Recalling (6.11), it follows that

$$\frac{\mathbb{E}_{\mathscr{D}} Z_{\text{ext}}}{\mathbb{E}_{\mathscr{D}} Z} = \mathbb{Q}_{\mathscr{D}}\Big(\Big\{(\mathscr{G}, \underline{\sigma}) : \underline{\sigma} \text{ is extendible}\Big\}\Big)$$

$$\geq \mathbb{Q}_{\mathscr{D}}\left(\left\{\begin{array}{c}(\mathscr{G}, \underline{\sigma}) : \underline{\sigma} \text{ is a judicious coloring of } \mathscr{G}, \text{ and no connected} \\ \text{component of its free subgraph contains a bicycle}\end{array}\right\}\right) = 1 - o_n(1)$$

with high probability, where the last inequality is by Lemma 6.8. This concludes the proof. $\quad\square$

6.4. Separability for judicious colorings.

In this subsection we prove Proposition 3.31, which says that the first moment $\mathbb{E}_{\mathscr{D}} Z$ of judicious colorings is dominated by separable colorings (Definition 3.27), with high probability over $\mathscr{D}$. We again let $\mathscr{G} = (V, F, E)$ be the processed graph (with neighborhood sequence $\mathscr{D}$), so that $\mathscr{G} = \operatorname{pr} \mathscr{G}'$ where $\mathscr{G}'$ is the original k-SAT instance. We first show by a direct second moment calculation that the original instance $\mathscr{G}'$ is very unlikely to have pairs of satisfying assignments of "intermediate" overlap:

Lemma 6.9. *Let $\mathscr{G}' \sim \mathbb{P}_{n',n'\alpha}$ for α in the regime (1.7). Recall (2.8) that $Z^2[z]$ counts the number of pairs $(\underline{x}^1, \underline{x}^2)$ of satisfying assignments of $\mathscr{G}'$ with overlap z. Then, for any positive absolute constant c, it holds for all*

$$z \in \left[0, \frac{1}{2}\Big(1 - \frac{ck}{2^{k/2}}\Big)\right] \cup \left[\frac{1}{2}\Big(1 - \frac{ck}{2^{k/2}}\Big), 1 - \frac{k^2}{2^k}\right],$$

we have $\mathbb{E}_{n',n'\alpha} Z^2[z] \leq \exp\{-\Omega(n'k^2/2^k)\}$.

Proof. Let $\mathbb{E}$ denote expectation under $\mathbb{E}_{n',n'\alpha}$. Recall from (2.9) that

$$\mathbb{E}Z^2[z] = \frac{\exp\{n' \mathsf{f}_{\text{SAT},2}(z)\}}{(n')^{O(1)}} = \frac{1}{(n')^{O(1)}} \exp\left\{n'\left[\ln 2 + \mathcal{H}(z) + \alpha \ln\Big(1 - \frac{2 - z^k}{2^k}\Big)\right]\right\}.$$

We use the inequality $\ln(1 - x) \leq -x$ to bound

$$\mathsf{f}_{\text{SAT},2}(z) \leq \varphi(z) \equiv \ln 2 + \mathcal{H}(z) - \alpha\Big(\frac{2 - z^k}{2^k}\Big).$$

It suffices to prove $\varphi(z) \leq -\Omega(k^2/2^k)$ for all z in the claimed interval. By the restriction (1.7) on α, we have

$$\varphi(z) \leq \mathcal{H}(z) - \ln 2 + z^k \ln 2 + O\Big(\frac{1}{2^k}\Big).$$

Next recall that $\mathcal{H}''(z) \leq -4$ for all $z \in [0, 1]$, so expanding around $z = 1/2$ gives

$$\varphi(z) \leq -2\Big(z - \frac{1}{2}\Big)^2 + z^k \ln 2.$$

This readily implies $\varphi(z) \leq -\Omega(k^2/2^k)$ for

$$z \in \left[0, \frac{1}{2}\Big(1 - \frac{k}{2^{k/2}}\Big)\right] \cup \left[\frac{1}{2}\Big(1 + \frac{k}{2^{k/2}}\Big), \frac{1}{2}\Big(1 + \frac{\ln k}{k}\Big)\right].$$

It also implies $\varphi(z) \leq -\Omega((\ln k)^2/k^2)$ for

$$z \in \left[\frac{1}{2}\left(1 + \frac{\ln k}{k}\right), 1 - \frac{2\ln k}{k}\right],$$

as well as $\varphi(z) \leq -\Omega(1/k^{1/2})$ for

$$z \in \left[1 - \frac{2\ln k}{k}, , 1 - \frac{1}{k^{3/2}}\right].$$

Finally, a straightforward Taylor expansion near $z = 1$ gives $\varphi(z) \leq -\Omega(k^2(\ln k)/2^k)$ for

$$\epsilon \in \left[1 - \frac{1}{k^{3/2}}, 1 - \frac{k^2}{2^k}\right],$$

concluding the proof. $\qquad\square$

Lemma 6.9 controls pairs of satisfying assignments of the original instance $\mathscr{G}'$. The next result transfers this to a bound on pairs of judicious colorings of the processed instance $\mathscr{G} = \mathrm{pr}\,\mathscr{G}'$.

Corollary 6.10. *Recall from* (3.34) *that* $\mathbf{Z}^2[z]$ *counts the number of pairs* $(\underline{\sigma}^1, \underline{\sigma}^2)$ *of judicious colorings of* $\mathscr{G}$, *such that their corresponding frozen configurations* $\underline{x}(\underline{\sigma}^i)$ *have overlap* z. *Let*

$$I_1 \equiv \left[0, \frac{1}{2}\left(1 - \frac{k}{2^{k/2}}\right)\right] \cup \left[\frac{1}{2}\left(1 + \frac{k}{2^{k/2}}\right), 1 - \frac{k^2}{2^k}\right],$$

and let $\mathbf{Z}^2[I_1]$ *be the sum of* $\mathbf{Z}^2[z]$ *over* $z \in I_1$. *Then* $\mathbb{E}_{\mathscr{D}}[\mathbf{Z}^2[I_1]] \leq \exp\{-\Omega(nk^2/2^k)\}$ *with high probability over* $\mathscr{D}$.

Proof. Let $\mathscr{G}' = (V', F', E')$ with $|V'| = n'$, and $\mathscr{G} = (V, F, E) = \mathrm{pr}\,\mathscr{G}'$ with $|V| = n$. It follows from Proposition 3.22 that $n = n'(1 - o_R(1))$ with high probability. Moreover, with high probability over $\mathscr{D}$, in any judicious coloring the fraction of variables set to free is at most $4/2^k$. Given any frozen configuration $\underline{x}$ of $\mathscr{G}$, we can extend it to $\underline{\acute{x}} \in \{\text{-},\text{+},\text{f}\}^{V'}$ by setting $\acute{x}_v = \text{f}$ for all $v \in V' \setminus V$. The resulting $\underline{\acute{x}}$ is an "almost-SAT assignment" in the sense that any clause in $\mathscr{G}$ that does not neighbor a free variable must be satisfied. As long as $n = n'(1 - o_R(1))$, the number of free variables under $\underline{\acute{x}}$ must be (crudely) at most $n'(5/2^k)$.

Suppose $(\underline{\sigma}^1, \underline{\sigma}^2)$ is a pair of judicious colorings of $\mathscr{G}$, such that their corresponding frozen configurations $(\underline{x}^1, \underline{x}^2)$ agree on nz variables. Define the extended configurations $\underline{\acute{x}}^i \in \{\text{-},\text{+},\text{f}\}^{V'}$, and note that

$$V'(\text{f}) \equiv \left\{v \in V' : (\acute{x}^1)_v = \text{f or } (\acute{x}^2)_v = \text{f}\right\}$$

has size $n'\pi \equiv |V'(\text{f})| \leq n'(10/2^k)$. Moreover, by definition we have $V'(\text{f}) \supseteq V' \setminus V$. Let

$$V_= \equiv \left\{v \in V' \setminus V'(\text{f}) : (\acute{x}^1)_v = (\acute{x}^2)_v\right\},$$

and note that $n'(1 - \pi)y \equiv |V_=| \in [nz - n'\pi, nz]$. Thus, with high probability over $\mathscr{D}$, we have

$$\mathbf{Z}^2[z] \leq \sum_{\pi \leq 10/2^k} \sum_{y:|y-z|\leq 15/2^k} A^2[\pi, y] \tag{6.16}$$

where $A^2[\pi, y]$ counts the number of pairs $(\underline{x}^1, \underline{x}^2)$ of almost-SAT assignments with $|V'(\text{f})| = n'\pi$ and $|V_=| = n'y$. We then calculate

$$\mathbb{E}_{n',n'\alpha}A^2[\pi, y] \leq \binom{n'}{n'\pi}\binom{n'(1 - \pi)}{n'(1 - \pi)y}5^{n'\pi}2^{n'(1-\pi)}\left(1 - (1 - \pi)^k\left(\frac{2 - y^k}{2^k}\right)\right)^{n'\alpha}$$

$$\leq \exp\left\{n'\left[f_{\text{SAT},2}(y) + O\left(\frac{k}{2^k}\right)\right]\right\} \leq \exp\left\{-\Omega\left(\frac{n'k^2}{2^k}\right)\right\},$$

where the last inequality follows by Lemma 6.9 for all π, y in the range specified by (6.16), when $z \in I_1$. It follows that $\mathbb{E}_{n',n'\alpha}\mathbf{Z}^2[z] \leq \exp\{-\Omega(n'k^2/2^k)\}$, and then applying Markov's inequality gives $\mathbb{E}_{\mathscr{D}}[\mathbf{Z}^2[I_1]] \leq \exp\{-\Omega(nk^2/2^k)\}$ with high probability over $\mathscr{D}$, as claimed. $\qquad\square$

To prove Proposition 3.31 it remains to address pairs $(\underline{x}^1, \underline{x}^2)$ of frozen configurations with overlap in $[1-k^2/2^k, 1]$, since the rest of $[0, 1]$ is covered by I_1 from Corollary (6.10), or by I_0 from (3.31). We can also ignore overlaps very close to one, since for any $\underline{x}^1$, the total number of configurations $\underline{x}^2$ within Hamming distance $(\ln n)^4$ is much smaller than $\exp\{(\ln n)^5\}$. We thus restrict our attention to pairs $(\underline{x}^1, \underline{x}^2)$ of frozen configurations on $\mathcal{G}$ with overlap in

$$I_2 \equiv \left[1 - \frac{k^2}{2^k}, 1 - \frac{(\ln n)^4}{n}\right]. \tag{6.17}$$

We begin with a combinatorial lemma which says that if such a pair exists, then the graph $\mathcal{G}$ must (deterministically) contain a particular structure. In the remainder of this section we show that such structures are very unlikely to exist, from which Proposition 3.31 will follow.

Lemma 6.11. *Let $\mathcal{G} = (V, F, E) = \mathrm{pr}\,\mathcal{G}'$ be a processed k-SAT instance, with $|V| = n$. Suppose that $(\underline{x}^1, \underline{x}^2)$ is a pair of frozen configurations on $\mathcal{G}$, each corresponding to a judicious coloring of $\mathcal{G}$, with $\{v \in V : (x^1)_v = (x^2)_v\} = nz$ for some $z \in I_2$, as defined by (6.17). Then there exists a subset $S \subseteq V$,*

$$\frac{(\ln n)^3}{n} \le \frac{|S|}{|V|} \le \frac{k^2}{2^k}, \tag{6.18}$$

and for each $v \in S$ a set $C(v)$ of directed paths $u \to a \to v$ inside $\mathcal{G}$, such that the following hold:

a. Let $\epsilon_* \equiv (\delta_*)^3$ (cf. Definition 3.12). Let D be the defective variables in S, and $B \equiv S \setminus D$. Then $|D| \le 2\epsilon_*|S|$.
b. For each $v \in S$, every element of $C(v)$ is of the form $u \to a \to v$ where a is forcing to v under $\underline{x}^1$, and $u \in S \setminus \{v\}$.
c. For each $v \in D$, the set $C(v)$ contains exactly one element $u \to a \to v$.
d. For each $v \in B$, every clause a that forces to v under $\underline{x}^1$ appears in exactly one element of $C(v)$.

Let C_S denote the set of all paths appearing in the sets $C(v)$ for $v \in S$.

Proof. We divide the proof into a few steps:

Step 1. Use $(\underline{x}^1, \underline{x}^2)$ to define an "internally forced" subset $X \subseteq V$. Since both the $\underline{x}^i$ correspond to judicious colorings $\underline{\sigma}^i$, they must each contain the same number of free variables, from which it follows that

$$\left|\left\{v \in V : (x^1)_v \ne \mathtt{f}, (x^2)_v = \mathtt{f}\right\}\right| = \left|\left\{v \in V : (x^1)_v = \mathtt{f}, (x^2)_v \ne \mathtt{f}\right\}\right|.$$

Consequently, if $d_H(\underline{x}^1, \underline{x}^2) = |\{v \in V : (x^1)_v \ne (x^2)_v\}|$, the set

$$X \equiv \left\{v \in V : (x^1)_v \ne (x^2)_v, (x^1)_v \ne \mathtt{f}\right\}$$

must have cardinality $|X| \ge d_H(\underline{x}^1, \underline{x}^2)/2$. The assumed overlap $z \in I_2$ then implies

$$\frac{(\ln n)^4}{2n} \le \frac{|X|}{|V|} \le \frac{k^2}{2^k}. \tag{6.19}$$

Since the $\underline{x}^i$ differ on X and $\underline{x}^1$ is not $\mathtt{free}$ on X, it must be that X is **internally forced** with respect to $\underline{x}^1$ — that is to say, each $v \in X$ is forced, but only by clauses involving at least one other variable from X. This implies that for any $v \in X$ we can find a path

$$v = u_0 \leftarrow a_1 \leftarrow u_1 \leftarrow a_2 \leftarrow u_2 \leftarrow \ldots \tag{6.20}$$

where a_i is forcing to u_{i-1}, and $u_i \in X$ with $u_i \ne u_{i-1}$. Since the graph is finite, any such path must eventually close on itself to form a directed cycle within X. It follows that every maximal connected component of X must contain at least one cycle.

Step 2. Extract $S \subseteq X$ which has a small fraction of defects, and is internally forced. If X does not intersect any compound enclosure (Definition 3.16) in the processed graph $\mathcal{G}$, then X does not contain any defective variable, and we simply take $S = X$.

If $U = U^\circ \sqcup \partial_\circ U$ is a compound enclosure and $X \cap U \ne \varnothing$, then we must have $X \cap \partial_\circ U \ne \varnothing$ — indeed, if it were not the case, then $X \cap U = X \cap U^\circ$ would be a collection of maximal connected components of X, and each of these components would be a tree (since U itself must be a tree). This contradicts our earlier observation that every maximal connected component of X must contain at least one cycle.

We shall define a subset $S_U \subseteq X \cap U$ as follows. As we just saw, each connected component of $X \cap U$ induces a tree $\mathscr{T} \subseteq \mathscr{G}$ that intersects $\partial_\circ U$. Since X is internally forced, the tree $\mathscr{T}$ must be covered by forcing paths (6.20); this makes $\mathscr{T}$ into a DAG (directed acyclic graph). Consider the variables of the DAG with out-degree zero (meaning that they do not participate in any clause that forces another variable): if these all lie in $\partial_\circ U$, then we simply put all the variables of $\mathscr{T}_U$ into S_U. Otherwise, suppose the DAG contains a variable $v \notin \partial_\circ U$ with out-degree zero: then we can replace X by $X \setminus \{v\}$, which is also internally forced. We continue pruning variables in this way until we arrive at $X' \subseteq X$ such that in the DAG corresponding to $X' \cap U$, all variables of out-degree zero lie in $\partial_\circ U$. We then set $S_U = X' \cap U$: by construction, S_U contains $X \cap \partial_\circ U$. Moreover, if we take S to be the union of all the perfect variables in X together with S_U for all compound enclosures U, then S is internally forced.

Each variable in $S_U \cap U^\circ$ has positive out-degree (by the above construction) and positive in-degree (since S is internally forced). This means that if we view S_U as an undirected subgraph of $\mathscr{G}$, each of its connected components must be a tree all of whose leaves lie in $\partial_\circ U$. It follows that S_U can be covered by a disjoint union of (undirected) paths γ where each γ has a variable in $\partial_\circ U$ at one (or both) of its endpoints. Each γ has length at most R', the maximum diameter of U. By the construction of compound enclosures (Definition 3.16), each variable in $v \in \partial_\circ U$ is perfect, hence orderly (Definition 3.12), so each γ has at most ϵ_* fraction defective variables. This implies that S_U has at most ϵ_* fraction of defective variables. Moreover, if γ contains any defective variable, then it must contain at least $1/\epsilon_*$ variables. It follows that the fraction of defective variables in $\gamma \setminus \partial_\circ U$ is at most

$$\frac{\epsilon_* \cdot (\text{number of variables in } \gamma)}{\text{number of variables in } (\gamma \setminus \partial_\circ U)} \leq \frac{\epsilon_*}{1 - 2\epsilon_*} \leq 2\epsilon_* \,.$$

It follows that $S_U \setminus \partial_\circ U$ has at most $2\epsilon_*$ fraction of defective variables. Since every defective variable in S must be contained in some $S_U \setminus \partial_\circ U$, and the sets $S_U \setminus \partial_\circ U$ are pairwise disjoint, we conclude that S has at most $2\epsilon_*$ fraction of defective variables.

Step 3. Choose paths $C(v)$ **for** $v \in S$. Recall that $S = B \sqcup D$ where D is the subset of defective variables in S. For $v \in D$, choose a single path $u \to a \to v$ such that a is forcing to v and $u \in S$, and let $C(v)$ consist of this path alone. For $v \in B$, for every clause a that is forcing to v, choose a path $u \to a \to v$ with $u \in S$, and add this path to $C(v)$. Thus, for $v \in B$, every clause forcing to v is covered by exactly one path in $C(v)$. This finishes the construction. Finally, since $S \subseteq X$ and X satisfies (6.19), it is clear that S satisfies the upper bound in (6.18), so it remains only to verify the lower bound. To this end, partition $X = X' \sqcup X''$ where X' is the subset of all perfect variables in X. We can define a mapping $f : X'' \to X'$ where we map $u \in X''$ to $v \in X'$ such that v is on the boundary of the compound enclosure containing u. Recall that the maximum size of a compound enclosure is bounded by a constant depending only on k and R. Moreover, a perfect variable must be nice, so its degree is $O(k2^k)$. This implies that for any $v \in X'$, its preimage $f^{-1}(v) \subseteq X''$ has size at most $C(k, R)$. It follows that $|X''| \leq C(k, R)|X|$. Since $S \supseteq X'$ by construction, we conclude using the lower bound in (6.19) that (for large n) we have

$$\frac{|S|}{|V|} \geq \frac{|X'|}{|V|} \geq \frac{(\ln n)^4}{C(k, R) \cdot 2n} \geq \frac{(\ln n)^3}{n} \,,$$

as claimed. This concludes the proof. $\qquad\square$

In the remainder of this subsection, we show that the structure described by Lemma 6.11 is unlikely to occur under the planted measure. We make the following definitions based on the lemma:

Definition 6.12 (permissible tuples)**.** As in Remarks 3.21 and 6.3, we fix a set of variables V and a set of clauses F, equipped with incident half-edges δV and δF, all labelled with types according to $\mathcal{D}$. We do not, as yet, take any matching of δV to δF. We say that the tuple (u, j_u, a, j_v, v) is **permissible** if u and v are distinct variables, a is a clause, and $j_u \neq j_v$ are indices such that $L_a(j_u) \in T_u$ and $L_a(j_v) \in T_v$. There is a unique half-edge $e_u \in \delta u$ of type $t_{e_u} = L_a(j_u)$; we call this the **initial half-edge** of the tuple. Likewise there is a unique half-edge $e_v \in \delta v$ of type $t_{e_v} = L_a(j_v)$; we call this the **final half-edge** of the tuple.

Let $\boldsymbol{\Pi}(\mathcal{D})$ denote the collection of all pairs (S, P_S) such that S is a subset of V satisfying (6.18), and P_S is a collection of permissible tuples (u, j_u, a, j_v, v) such that $u, v \in S$; every variable in S appears as the final (i.e., fifth) entry of at least one tuple in P_S; and no clause appears in more than one tuple in P_S. Moreover we require that for every $v \in S$, all half-edges in $P(v)$ have the same sign L, where $P(v)$ denotes the set of all half-edges $e_v \in \delta v$ that arise as the final half-edge of a tuple in P_S.

Note that, given $\mathcal{D}$, we can partition $S = B \sqcup D$ where D denotes the defective variables in S. Let $\mathbf{\Pi}_*(\mathcal{D})$ denote the subset of elements $(S, P_S) \in \mathbf{\Pi}(\mathcal{D})$ such that $|D| \le 2\epsilon_*|S|$; and each $v \in D$ appears as the final element of exactly one tuple in P_S.

Definition 6.13 (events based on permissible tuples). Again fix $V, F, \delta V, \delta F$, labelled with types according to $\mathcal{D}$. Take $(S, P_S) \in \mathbf{\Pi}(\mathcal{D})$ and partition $S = B \sqcup D$ as above. Let F_S denote all the clauses appearing in P_S, and note that given $\mathcal{D}$ we can also partition $F_S = F_B \sqcup F_D$ where F_D denotes all the clauses in F_S that are internal to defects, and $F_B \equiv F_S \setminus F_B$. As in Remark 6.3, a pair $(\mathfrak{M}, \sigma)$ with $\mathfrak{M} \sim \sigma$ is equivalent to a pair $(\mathcal{G}, \underline{\sigma})$ where $\mathcal{G}$ is a graph with neighborhood sequence $\mathcal{D}$, and $\underline{\sigma}$ is a judicious coloring on $\mathcal{G}$. We define the following events on the space of all pairs $(\mathfrak{M}, \sigma)$. First recall that each clause $a \in F_S$ appears in a unique tuple $(u, j_u, a, j_v, v) \in P_S$. Let

$$Y_a = \left\{ (\mathfrak{M}, \sigma) : \sigma_{aj_v} = \mathtt{r} \text{ and } \sigma_{a\ell} = \mathtt{y} \text{ for all } \ell \ne j_v \right\}, \quad Y_B \equiv \bigcap_{a \in F_B} Y_a .$$

Membership of $(\mathfrak{M}, \sigma)$ in Y_a depends only on $\sigma_{\delta a}$. Let

$$K_a = \left\{ \begin{array}{c} (\mathfrak{M}, \sigma) : \mathfrak{M} \text{ contains a path} \\ \text{passing through } (u, a, v) \end{array} \right\}, \quad K_B \equiv \bigcap_{a \in F_B} K_a , \quad K_D \equiv \bigcap_{a \in F_D} K_a .$$

Let $K_S \equiv K_B \cap K_D$. Membership in K_S depends only on the matching $\mathfrak{M}$, or equivalently the graph $\mathcal{G}$, so we sometimes abuse notation and write simply $\mathcal{G} \in K_S$. Next recall that for each $v \in S$, all the half-edges $P(v) \subseteq \delta v$ have the same sign $\mathtt{L}$. Let

$$X_v \equiv \left\{ (\mathfrak{M}, \sigma) : \left\{ \begin{array}{l} \sigma_{\delta v} \text{ is } \mathtt{red} \text{ on all half-edges in } P(v), \\ \mathtt{blue} \text{ on all edges in } \delta v(\mathtt{+L}) \setminus P(v), \\ \mathtt{yellow} \text{ on all edges in } \delta v(\mathtt{-L}) \end{array} \right\} \right\}, \quad X_B \equiv \bigcap_{v \in B} X_v .$$

Membership of $(\mathfrak{M}, \sigma)$ in X_v depends only on $\sigma_{\delta v}$.

It follows immediately from Lemma 6.6 that $\mathbb{Q}_\mathcal{D}(K_D) = \mathbb{P}_\mathcal{D}(K_D)$.

Lemma 6.14. *Fix a pair $(S, P_S) \in \mathbf{\Pi}_*(\mathcal{D})$ as in Definition 6.12. For the events of Definition 6.13, we have*

$$\mathbb{Q}_\mathcal{D}\left(X_B, Y_B \,\middle|\, K_D\right) = \mathbb{Q}_\mathcal{D}(X_B)\mathbb{Q}_\mathcal{D}(Y_B) \le \left(\frac{O(1)}{2^k}\right)^{|B| + 2|F_B|} .$$

Proof. Let $V, F, \delta V, \delta F$ be fixed as in Remark 6.3. The event K_D depends on the matching of half-edges internal to defects. The event X_B concerns the coloring σ on variable-incident half-edges that will not be internal to defects. The event Y_B concerns the coloring σ on clause-incident half-edges that will not be internal to defects. It follows that the three events are mutually independent, therefore

$$\mathbb{Q}_\mathcal{D}\left(X_B, Y_B \,\middle|\, K_D\right) = \mathbb{Q}_\mathcal{D}(X_B)\mathbb{Q}_\mathcal{D}(Y_B) . \tag{6.21}$$

It follows from Remark 6.3 that $\mathbb{Q}_\mathcal{D}(X_B)$ is the probability that a uniformly random element σ from $\mathcal{J}$ satisfies the conditions of X_B. Suppose we instead sample $\sigma : \delta V \sqcup \delta F \to \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$ according to the measure

$$\check{\mathbb{Q}}_\mathcal{D}(\sigma) = \left\{ \prod_{v \in V} \dot{v}_v(\sigma_{\delta v}) \right\}\left\{ \prod_{a \in F} \hat{v}_a(\sigma_{\delta a}) \right\},$$

where $(\dot{v}, \hat{v}) \equiv v \equiv v^{\mathrm{op}}[\star\omega]$. It follows from Lemma 3.38 and Corollary 3.55 that

$$\check{\mathbb{Q}}_\mathcal{D}(\sigma) = \left\{ \prod_{v \in V} \left(\frac{\varphi_v(\sigma_{\delta v})}{\dot{z}_v} \prod_{e \in \delta v} {}_\star \hat{q}_e(\sigma_e) \right) \right\}\left\{ \prod_{a \in F} \left(\frac{\hat{\varphi}_a(\sigma_{\delta a})}{\hat{z}_a} \prod_{f \in \delta a} \left(\gamma_f(\sigma_f)_\star \dot{q}_f(\sigma_f) \right) \right) \right\} .$$

Moreover, if a is a clause of type L and e is the j-th edge in δa, then γ_e depends only on (L, j). Thus each σ in $\mathcal{J}$ receives precisely the same weight under the measure $\check{\mathbb{Q}}_\mathcal{D}$. Moreover, if we sample $\sigma \sim \check{\mathbb{Q}}_\mathcal{D}$, then the *expected*

empirical measure of σ is $\nu = \nu^{\mathrm{op}}$. Since the total number of types is at most $C(k, R)$, it follows by the local central limit theorem that $\check{\mathbb{Q}}_{\mathcal{D}}(\sigma \in \mathcal{J}) = 1/n^{O(1)}$, so

$$\mathbb{Q}_{\mathcal{D}}(X_B) \leq n^{O(1)}\check{\mathbb{Q}}_{\mathcal{D}}(X_B) = n^{O(1)}\prod_{v \in B} \dot{\nu}_v(X_v).$$

A variable $v \in B$ must be non-defective, hence nice (Definition 3.8). It follows that

$$\dot{\nu}_v(X_v) = \frac{1}{\dot{z}_v}\prod_{e \in \delta v \cap P(v)} \star\hat{q}_e(\mathtt{r})\prod_{e \in \delta v(\mathtt{L})\backslash P(v)} \star\hat{q}_e(\mathtt{b})\prod_{e \in \delta v(-\mathtt{L})} \star\hat{q}_e(\mathtt{y})$$

$$\leq O(1)\prod_{e \in \delta v \cap P(v)} \frac{\star\hat{q}_e(\mathtt{r})}{\star\hat{q}_e(\mathtt{b})}\prod_{e \in \delta v(\mathtt{L})} \frac{\star\hat{q}_e(\mathtt{b})}{\star\hat{q}_e(\mathtt{r},\mathtt{b})} \leq \frac{O(1)}{(2^k)^{1+|P(v)|}}.$$

Substituting into the previous calculation gives

$$\mathbb{Q}_{\mathcal{D}}(X_B) \leq n^{O(1)}\left(\frac{O(1)}{2^k}\right)^{|B|+|F_B|}. \tag{6.22}$$

Similarly, a clause $a \in F_B$ can neighbor at most one defective variable, so it follows from Definition 3.10 that all the variables in the clause must be nice. It then follows using Corollary 6.2 that

$$\hat{\nu}_a(Y_a) = \frac{1}{\hat{z}_a}\gamma_j(\mathtt{r})\star\dot{q}_j(\mathtt{r})\prod_{\ell \neq j}\left\{\gamma_\ell(\mathtt{y})\star\dot{q}_\ell(\mathtt{y})\right\} \leq O(1)\frac{\star\dot{q}_j(\mathtt{r})}{\star\dot{q}_j(\{\mathtt{y},\mathtt{g},\mathtt{b}\})}\prod_{\ell \neq j} \frac{\star\dot{q}_\ell(\mathtt{y})}{\star\dot{q}_\ell(\{\mathtt{y},\mathtt{g},\mathtt{b}\})} \leq \frac{O(1)}{2^k}.$$

It follows by a similar argument as for X_B that

$$\mathbb{Q}_{\mathcal{D}}(Y_B) \leq n^{O(1)}\check{\mathbb{Q}}_{\mathcal{D}}(Y_B) \leq n^{O(1)}\left(\frac{O(1)}{2^k}\right)^{|B|+|F_B|}. \tag{6.23}$$

The result follows by combining (6.21), (6.22), and (6.23). □

Lemma 6.15. *Fix a pair* $(S, P_S) \in \Pi_*(\mathcal{D})$ *as in Definition 6.12. For the events of Definition 6.13, we have*

$$\frac{\mathbb{Q}_{\mathcal{D}}(K_B \,|\, X_B, Y_B, K_D)}{\mathbb{P}_{\mathcal{D}}(K_B \,|\, K_D)} = \frac{\mathbb{Q}_{\mathcal{D}}(K_B \,|\, X_B, Y_B)}{\mathbb{P}_{\mathcal{D}}(K_B)} \leq \left(O(1)2^k\right)^{|F_B|}.$$

Proof. Recall from Definition 6.12 that each tuple $(u, j_u, a, j_v, v) \in P_S$ distinguishes an initial half-edge $e_u \in \delta u$ and a final half-edge $e_v \in \delta v$. For each edge type t, let $a_t(\mathtt{y})$ count the number of tuples in P_S where the initial half-edge has type t. Let $a_t(\mathtt{r})$ count the number of tuples in P_S where the final half-edge has type t. Let $a_t \equiv a_t(\mathtt{r}) + a_t(\mathtt{y})$. Then note that

$$\sum_t a_t(\mathtt{y}) = \sum_t a_t(\mathtt{r}) = \frac{1}{2}\sum_t a_t = |F_B|.$$

Since we restricted to clauses in F_B, any type t with $a_t > 0$ must be nice. The events K_B and K_D involve edges of distinct types, so they are independent under $\mathbb{P}_{\mathcal{D}}$. It follows that

$$\mathbb{P}_{\mathcal{D}}(K_B \,|\, K_D) = \mathbb{P}_{\mathcal{D}}(K_B) = \prod_t\left\{\prod_{i=0}^{a_t-1}\frac{1}{n_t - i}\right\} = \prod_t \frac{1}{(n_t)_{a_t}}$$

(using the standard notation for the falling factorial). On the other hand, under the measure $\mathbb{Q}_{\mathcal{D}}$, if we condition on the events X_B and Y_B, then the event K_B is more likely to occur because we have conditioned the edges involved to have compatible colorings:

$$\mathbb{Q}_{\mathcal{D}}\left(K_B \,\Big|\, X_B, Y_B, K_D\right) = \mathbb{Q}_{\mathcal{D}}\left(K_B \,\Big|\, X_B, Y_B\right) = \prod_t \frac{1}{(n_t \cdot \star\pi_t(\mathtt{y}))_{a_t(\mathtt{y})}(n_t \cdot \star\pi_t(\mathtt{r}))_{a_t(\mathtt{r})}},$$

which is clearly larger than the quantity $\mathbb{P}_{\mathcal{D}}(K_B \,|\, K_D)$ calculated just above. For any integers $0 < a < b$ we have using Stirling's formula that

$$b^a \geq (b)_a = \frac{b!}{(b-a)!} \asymp \frac{b^a}{e^a}\left(\frac{b}{b-a}\right)^{1/2+b-a} \geq \frac{b^a}{e^a}.$$

It follows by combining the above that

$$\frac{\mathbb{Q}_{\mathcal{D}}(K_B \mid X_B, Y_B, K_D)}{\mathbb{P}_{\mathcal{D}}(K_B \mid K_D)} = \frac{\mathbb{Q}_{\mathcal{D}}(K_B \mid X_B, Y_B)}{\mathbb{P}_{\mathcal{D}}(K_B)} \leq \prod_t \frac{e^{a_t}}{(\star \pi_t(\mathbf{y}))^{a_t(\mathbf{y})}(\star \pi_t(\mathbf{r}))^{a_t(\mathbf{r})}} \leq \left(O(1)2^k\right)^{|F_B|},$$

as claimed.   □

Corollary 6.16. *Fix a pair* $(S, P_S) \in \Pi_*(\mathcal{D})$ *as in Definition 6.12. For the events of Definition 6.13, we have*

$$\frac{\mathbb{Q}_{\mathcal{D}}(X_B, Y_B, K_S)}{\mathbb{P}_{\mathcal{D}}(K_S)} \leq \left(\frac{O(1)}{2^k}\right)^{|B|+|F_B|} \leq \left(\frac{O(1)}{2^k}\right)^{(1-2\epsilon_*)(|S|+|P_S|)}$$

Proof. The first inequality follows directly by combining Lemmas 6.6, 6.14, and 6.15. Next, recall the assumption from Definition 6.12 that $|B| \geq (1 - 2\epsilon_*)|S|$. It follows that

$$\frac{|F_B|}{|F_S|} \geq \frac{\sum\limits_{v \in B} |P(v)|}{|F_S|} = \frac{\sum\limits_{v \in B} |P(v)|}{\sum\limits_{v \in B} |P(v)| + |D|} \geq \frac{|B|}{|B| + |D|} = \frac{|B|}{|S|} \geq 1 - 2\epsilon_*,$$

which implies the second inequality since $|F_S| = |P_S|$.   □

Proof of Proposition 3.31. Take I_1 as in Corollary 6.10, and I_2 as in (6.17). Then, recalling Definition 3.27 and (3.32),

$$\mathbf{Z}_{\text{sep}} \leq \mathbf{Z}^2[I_1] + \underbrace{\sum_{\underline{\sigma}} \mathbf{1}\left\{\left|\left\{\underline{\sigma}' : \frac{|v \in V : x(\underline{\sigma})_v = x(\underline{\sigma}')_v|}{|V|} \in I_2\right\}\right| \geq \frac{\exp\{(\ln n)^5\}}{2}\right\}}_{\text{denote this } (\mathbf{Z}_{\text{sep}})'}.$$

We have a bound on $\mathbb{E}_{\mathcal{D}}\mathbf{Z}^2[I_1] = \mathbb{E}_{\mathcal{D}}\mathbf{Z}^2[I_1]$ from Corollary 6.10. It follows from the definition of $\mathbb{Q}_{\mathcal{D}}$ that

$$r'(\mathcal{D}) \equiv \frac{\mathbb{E}_{\mathcal{D}}(\mathbf{Z}_{\text{sep}})'}{\mathbb{E}_{\mathcal{D}}\mathbf{Z}_{\text{sep}}} = \mathbb{Q}_{\mathcal{D}}\left(\left\{(\mathcal{G}, \underline{\sigma}) : \left|\left\{\underline{\sigma}' : \frac{|v \in V : x(\underline{\sigma})_v = x(\underline{\sigma}')_v|}{|V|} \in I_2\right\}\right| \geq \frac{\exp\{(\ln n)^5\}}{2}\right\}\right).$$

Recall that the events of Definition 6.13 all depend on the choice of (S, P_S). To make this explicit, we now write $K[S, P_S] \equiv K_S$ and $T[S, P_S] \equiv X_B \cap Y_B \cap K_S$. It then follows from Lemma 6.11 and Corollary 6.16 that

$$r'(\mathcal{D}) \leq \sum_{(S, P_S) \in \Pi_*(\mathcal{D})} \mathbb{Q}_{\mathcal{D}}(T[S, P_S]) \leq \sum_{(S, P_S) \in \Pi_*(\mathcal{D})} \frac{\mathbb{P}_{\mathcal{D}}(K[S, P_S])}{2^{k(1-3\epsilon_*)(|S|+|F_S|)}}.$$

We now wish to take expectation over the law $\mathbb{P} = \mathbb{P}_{n',m'}$ of the original graph $\mathcal{G}' = (V', F', E')$, of which $\mathcal{G} = \text{pr}\,\mathcal{G}'$ is the processed version. To this end, let Π' denote the set of all pairs (S, P_S) where $S \subseteq V$ satisfies (6.18), and P_S is a collection of tuples (u, j_u, a, j_v, v) where $u \neq v$ in S, $a \in F'$, every $v \in S$ appears as the last entry of at least one element of P_S, and no clause appears more than once in P_S. Note that if $(S, P_S) \in \Pi'$ and $\mathbb{P}_{\mathcal{D}}(K[S, P_S])$ is positive, then in fact $(S, P_S) \in \Pi(\mathcal{D})$. Thus, writing $\mathbb{E}$ for expectation under $\mathbb{P}$, we have (by the tower property of conditional expectation)

$$\mathbb{E}r'(\mathcal{D}) \leq \sum_{(S, P_S) \in \Pi'} \mathbb{E}\left(\mathbf{1}\{(S, P_S) \in \Pi(\mathcal{D})\} \frac{\mathbb{P}_{\mathcal{D}}(\mathcal{G} \in K[S, P_S])}{2^{k(1-3\epsilon_*)(|S|+|F_S|)}}\right) = \sum_{(S, P_S) \in \Pi'} \mathbb{E}\left(\frac{\mathbb{P}_{\mathcal{D}}(\mathcal{G} \in K[S, P_S])}{2^{k(1-3\epsilon_*)(|S|+|F_S|)}}\right)$$

$$= \sum_{(S, P_S) \in \Pi'} \frac{\mathbb{P}(\mathcal{G} \in K[S, P_S])}{2^{k(1-3\epsilon_*)(|S|+|F_S|)}} \leq \sum_{(S, P_S) \in \Pi'} \frac{\mathbb{P}(\mathcal{G}' \in K[S, P_S])}{2^{k(1-3\epsilon_*)(|S|+|F_S|)}},$$

where the last step uses that if $K[S, P_S]$ occurs for the processed graph $\mathcal{G}$, then it also occurs for the original graph $\mathcal{G}'$. Taking into account the number of choices for S and P_S, as well as the probability for the edges to be present under $\mathbb{P}$, we can bound the above as

$$\mathbb{E}r'(\mathcal{D}) \leq o_n(1) + \sum_s \mathbf{1}\left\{\frac{(\ln n')^2}{n'} \leq s \leq \frac{k^2}{2^k}\right\} \binom{n'}{n's} \frac{1}{2^{k(1-4\epsilon_*)n's}} \left(\sum_{\ell \geq 1} \left(\frac{n's \cdot k^2 \cdot m'/(n')^2}{2^{k(1-4\epsilon_*)}}\right)^\ell\right)^{n's}.$$

In the above, the $o_n(1)$ error accounts for the probability that more than $o_R(1)$ fraction of variables are removed during preprocessing, which is controlled by Proposition 3.22. On the complementary event, the number of variables n in $\mathscr{G}$ is very close to the original number of variables n', so the restriction on $s = |S|/n'$ follows from (6.18). The inner sum over $\ell \geq 1$ is for the possible sizes of the sets $P(v)$. Simplifying the above gives

$$\mathbb{E}r'(\mathcal{D}) \leq \sum_s \mathbf{1}\left\{\frac{(\ln n')^2}{n'} \leq s \leq \frac{k^2}{2^k}\right\}\binom{n'}{n's}\frac{1}{2^{k(1-3\epsilon_*)n's}}\left(\sum_{\ell \geq 1}(2^{4\epsilon_*k}s)^\ell\right)^{n's}$$

$$\leq \sum_s \mathbf{1}\left\{\frac{(\ln n')^2}{n'} \leq s \leq \frac{k^2}{2^k}\right\}\binom{n}{ns}\frac{s^{ns}}{2^{k(1-8\epsilon_*)ns}} \leq \frac{1}{\exp\{\Omega((\ln n)^2)\}}\ .$$

It follows that $r'(\mathcal{D}) = o_n(1)$ with high probability over $\mathcal{D}$, and the result follows. $\qquad\square$

7. Contraction estimates

In this section we state and prove three key technical results:

– Proposition 7.3 considers a **nice** (Definition 3.8) subtree T of a compound region U, and analyzes the maximal-entropy judicious measure on T subject to edge marginals $\omega_{\delta T}$ on the boundary edges δT.
– Proposition 7.4 considers the depth-one neighborhood of a non-compound variable U, and analyzes the maximal-entropy judicious measure on U subject to edge marginals $\omega_{\delta U}$ on the boundary edges δU.
– Proposition 7.6 shows how to reweight a clause to achieve a desired set of outgoing BP messages.

The significance of these results in the proof outline is as follows:

– Proposition 7.4 solves the optimization problem for non-compound variables that was derived in Proposition 3.46. The main technical difficulty of this result is that the clause types around a non-compound variable are not fixed.
– On the other hand, in Section 8 we will use Propositions 7.3 and 7.6 to prove Proposition 8.1, which analyzes the maximal-entropy judicious measure on all of U subject to boundary conditions $\omega_{\delta U}$. This completes the solution of the optimization for compound regions that was derived in Proposition 3.44. The main technical difficulty of this result is to deal with **non-nice** regions of the compound enclosure, which are not covered by Proposition 7.3. At the same time, we are helped in the analysis by the notion of compound type (Definition 3.18), which ensures that in the interior of a compound enclosure, the clause types around a variable are fixed.

A more detailed outline of this section is given below, before the start of §7.1. We begin by specifying the form of the subtrees T that we will consider for Proposition 7.3.

Definition 7.1 (entropy maximization on rooted trees). Let T be a finite tree rooted at a clause a_{RT} with exactly one child, which we refer to as the root variable v_{RT}. We use $\mathcal{L}T$ to denote the set of all leaf vertices of T other than a_{RT}, and we assume $\mathcal{L}T$ consists of variables only. We use δT to denote the edges incident to $\mathcal{L}T$. From now on we will refer to $\mathcal{L}T$ and δT respectively as the "boundary variables" and "boundary edges" of T. A small example appears in Figure 7. In the applications of this definition in the analysis of Section 8, we will take T to be a nice (Definition 3.8) subtree of a compound enclosure U. As a result, the notations that follow are purposefully similar to those of Definition 3.43. Let

$$\Delta(T) \equiv \left\{\begin{array}{l} \text{probability measures } \nu \text{ over} \\ \text{pair colorings } (\underline{\sigma}^1, \underline{\sigma}^2) \text{ of } T \end{array}\right\}\ .$$

We say that $\nu \in \Delta(T)$ is **judicious** if all of its edge marginals match the canonical marginals of Definition 3.4: that is, for all edges e in T and for both $j = 1, 2$, we have $\nu((\sigma_e)^j = \sigma) = {}_\star\pi_e(\sigma)$ for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$. Let $\omega_{\delta T}$ denote a tuple $(\omega_e)_{e \in \delta T}$ where each ω_e is a judicious probability measure on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$. We then let

$$J(T; \omega_{\delta T}) \equiv \left\{\begin{array}{l} \nu \in \Delta(T) : \nu \text{ is judicious,} \\ \text{and } \nu_e = \omega_e \text{ for all } e \in \delta T \end{array}\right\}\ ,$$

where ν_e denotes the marginal of ν on edge e. Define the constrained optimizer

$$\nu^{\mathrm{op}}(T; \omega_{\delta T}) \equiv \arg\max_\nu\left\{\mathcal{H}(\nu) : \nu \in J(T, \omega_{\delta T})\right\}\ . \tag{7.1}$$

Given weights Λ on T, we also define the unconstrained optimizer

$$\nu[T;\Lambda] \equiv \arg\max_{\nu} \left\{ \mathcal{H}(\nu) + \langle \ln\Lambda, \nu \rangle : \nu \in \Delta(T) \right\}, \tag{7.2}$$

where $\langle \ln\Lambda, \nu \rangle$ denotes the expected value of $\ln\Lambda(\underline{\sigma})$ with $\underline{\sigma}$ distributed according to ν. Note that by elementary calculus, the solution $\nu = \nu[T;\Lambda]$ of (7.2) is given explicitly by

$$\nu(\underline{\sigma}) = \frac{\mathbf{1}\{\underline{\sigma} \text{ is a valid coloring of } T\}\Lambda(\underline{\sigma})}{z[T;\Lambda]} \tag{7.3}$$

where $z[T;\Lambda]$ is the normalizing constant. Finally, we say that Λ are **Lagrangian weights** (for the constrained optimization problem $\nu^{\mathrm{op}}(T;\omega_{\delta T}$ in (7.1)) if $\langle \ln\Lambda, \nu \rangle$ is constant over $\nu \in J(T;\omega_{\delta T})$. We will parametrize Lagrangian weights in a particular way, described in Definition 7.8 below.

Definition 7.2 (error notation for edge distributions). On any edge e, denote $_*\omega_e \equiv {}_\star\pi_e \otimes {}_\star\pi_e$ where $_\star\pi_e$ is the canonical marginal on edge e from Definition 3.4. For any other probability measure ω_e on $\{\mathtt{r},\mathtt{y},\mathtt{g},\mathtt{b}\}^2$, we define the **discrepancy** of ω_e (relative to $_*\omega_e$) by

$$\mathrm{disc}(\omega_e) \equiv \mathrm{disc}_e(\omega) \equiv \sum_{\sigma \in \{\mathtt{r},\mathtt{y},\mathtt{g},\mathtt{b}\}^2} (\vartheta_*)^{\mathbf{1}\{\mathtt{r}\in\{\sigma^1,\sigma^2\}\}} \left| \frac{\omega_e(\sigma)}{_*\omega_e(\sigma)} - 1 \right|. \tag{7.4}$$

where $\vartheta_* \equiv 2^{-k\zeta/6}$ for an absolute constant $0 < \zeta \leq 1/20$. We shall ultimately take $\zeta = \zeta_*/4$ where ζ_* is the absolute constant in Proposition 8.4 below.
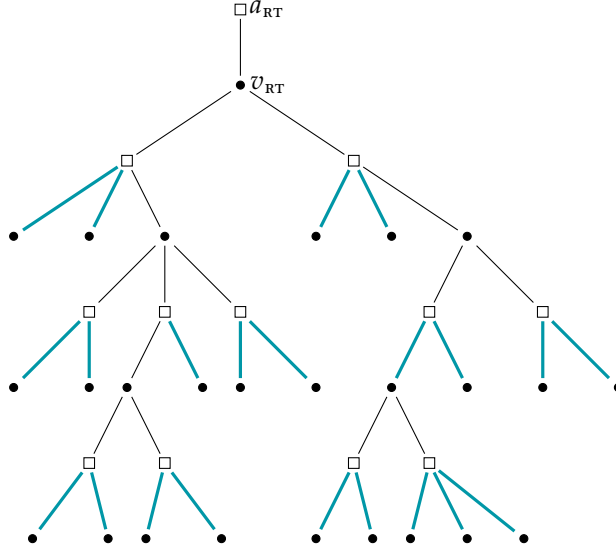

FIGURE 7. A small example of a tree T as in Definition 7.1. The edges of δT are shown as thick colored lines. In the applications of this definition in Section 8, T will be a nice subtree of a compound enclosure U.

Proposition 7.3 (contraction result inside compound enclosures). *Assume that $R \geq k$, where R is the neighborhood radius in (3.1). Let U be a compound enclosure (Definition 3.16). Take $T \subseteq U$ of the form described in Definition 7.1, and suppose T contains no defective variables. Consider the constrained optimization problem $\nu^{\mathrm{op}}(T;\omega_{\delta T})$ from (7.1), where we assume that $\omega_{\delta T}$ satisfies, for an absolute constant $0 < \zeta \leq 1/20$, the bounds*

$$\max_{\sigma \in \{\mathtt{r},\mathtt{y},\mathtt{g},\mathtt{b}\}^2} \left\{ \frac{1}{(2^k)^{\mathbf{1}\{\sigma=\mathtt{rr}\}}} \left| \frac{\omega_e(\sigma)}{_*\omega_e(\sigma)} - 1 \right| \right\} \leq \frac{1}{2^{2k\zeta}} \tag{7.5}$$

for all $e \in \delta T$. Then the following hold:

a. *We can explicitly construct Lagrangian weights $\Lambda \equiv \Lambda(T; \omega_{\delta T})$ such that the constrained optimizer $\nu = \nu^{\mathrm{op}}(T; \omega_{\delta T})$ of (7.1) coincides with the unconstrained optimizer $\nu = \nu[T; \Lambda]$ of (7.2) (where we recall that the latter is simply the Λ-weighted Gibbs measure on T).*

b. *Let q be the solution of the Λ-weighted BP recursions on T, where we fix the message "from the root" to be*

$$\left(\star \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}} \otimes \star \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}} \right)(\sigma) \equiv \prod_{j=1,2} \star \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}(\sigma^j) .$$

Then $\nu_e \cong \dot{q}_e \hat{q}_e$ for each edge e in T.

c. *For any edge (av) in T, the discrepancy of ν_{av} (defined by (7.4)) satisfies the bound*

$$\mathrm{disc}_{av}(\nu) \leq k^{O(1)} \underbrace{\sum_{e \in \delta T} \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{\mathfrak{b}_T(e,a)} \mathrm{disc}_e(\omega)}_{\xi_a(T; \omega_{\delta T})}, \tag{7.6}$$

where $\mathfrak{b}_T(e, a)$ denotes the number of variables between e and a on the unique path in T that joins e to a.

The explicit parametrization of Λ is given in Definition 7.8 below, and the construction is given in Definition 7.9.

Proposition 7.4 (contraction result for non-compound variables). *Assume that $R \geq k$, where R is the neighborhood radius in (3.1). In the same setting as Proposition 3.46, consider the constrained optimization problem*

$$\nu^{\mathrm{op}}(U; \omega_{\delta U}) = \arg\max_{\nu} \left\{ \mathcal{H}(\nu) : \nu \in J_{\mathscr{D}}(U, \omega_{\delta U}) \right\},$$

where U is the depth-one tree from Definition 3.45, and we assume that (cf. (7.5))

$$\max_{L} \max_{\sigma \in \{\mathrm{r,y,g,b}\}^2} \left\{ \frac{1}{(2^k)^{\mathbf{1}\{\sigma = \mathrm{rr}\}}} \left| \frac{\omega_{L,j(t_e)}(\sigma)}{\star \omega_{L,j(t_e)}(\sigma)} - 1 \right| \right\} \leq \frac{1}{2^{2k\zeta}} \tag{7.7}$$

for all $e \in \delta U$. Then the following hold:

a. *We can explicitly construct Lagrangian weights $\Psi \equiv \Psi_{\mathscr{D}}(U; \omega_{\delta U})$ such that the solution $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ of the above coincides with $\nu = \nu[U; \Psi]$, i.e., the Ψ-weighted Gibbs measure on U for the augmented pair coloring model.*

b. *Let q be the solution of the Ψ-weighted BP recursions on U. Then $\nu_e \cong \dot{q}_e \hat{q}_e$ for each e in U, where ν_e, $\dot{q}_e$, and $\hat{q}_e$ are probability measures on pairs (σ, L) with $\sigma \in \{\mathrm{r, y, g, b}\}^2$.*

c. *For the solution $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ of the above, denote (cf. (7.4))*

$$\mathrm{disc}_e(\nu) \equiv \max_{L} \left\{ \sum_{\sigma \in \{\mathrm{r,y,g,b}\}^2} (\vartheta_*)^{\mathbf{1}\{\mathrm{r} \in \{\sigma^1, \sigma^2\}\}} \left| \frac{\nu_e(\sigma \mid L)}{\star \omega_e(\sigma)} - 1 \right| \right\} \tag{7.8}$$

Then, for each edge (av) incident to the root variable v of U, we have (cf. (7.6))

$$\mathrm{disc}_{av}(\nu) \leq k^{O(1)} \underbrace{\sum_{e \in \delta U} \frac{(\vartheta_*)^{1/4}}{2^k} \mathrm{disc}_e(\omega_e)}_{\xi_a(U; \omega_{\delta U})} . \tag{7.9}$$

The explicit parametrization of Ψ is given in Definition 7.24 below, and the construction is given in Definition 7.25.

Definition 7.5 (error notation for clause-to-variable messages). *In the pair model, given any two functions $g, h : \{\mathrm{r, y, c}\}^2 \to (0, \infty)$, we write*

$$\widehat{\mathrm{ERR}}(g, h) = \begin{pmatrix} \max \left\{ \left| \frac{h(\sigma)}{g(\sigma)} - 1 \right| : \mathrm{r}[\sigma] = 0 \right\} \\ \max \left\{ \left| \frac{h(\sigma)}{g(\sigma)} - 1 \right| : \mathrm{r}[\sigma] = 1 \right\} \\ \max \left\{ \left| \frac{h(\sigma)}{g(\sigma)} - 1 \right| : \mathrm{r}[\sigma] = 2 \right\} \end{pmatrix} \in \mathbb{R}^3.$$

We write $\widehat{\mathrm{ERR}}(g, h) \leq \underline{s}$ to mean that $\widehat{\mathrm{ERR}}(g, h)$ is coordinatewise upper bounded by $\underline{s}$ in $\mathbb{R}^3$. In most cases we use this notation when both g and h are clause-to-variable messages in the pair model.

Recall that in the single-copy model, given a variable-to-clause message $\dot{q}_{va}$ (a measure on colors $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$), we defined a reweighted version Q_{va} by (6.2). We make the analogous definition in the pair model: given a variable-to-clause message $\dot{q} \equiv \dot{q}_{va}$ (a measure on pairs of colors $(\sigma^1, \sigma^2) \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$), define its reweighted version

$$Q_{va}(\sigma) = \frac{\dot{q}(\sigma)}{(2^{|\partial a|-1})\mathrm{r}[\sigma]} \Bigg/ \left\{ \dot{q}(\{\mathtt{y}, \mathtt{c}\}^2) + \frac{\dot{q}(\{\mathtt{r}\} \times \{\mathtt{y}, \mathtt{c}\}) + \dot{q}(\{\mathtt{y}, \mathtt{c}\} \times \{\mathtt{r}\})}{2^{|\partial a|-1}} + \frac{\dot{q}(\mathtt{rr})}{4^{|\partial a|-1}} \right\}, \tag{7.10}$$

where we denote $\mathrm{r}[\sigma] \equiv \mathbf{1}\{\sigma^1 = \mathtt{r}\} + \mathbf{1}\{\sigma^2 = \mathtt{r}\}$ for $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$.

Proposition 7.6 (pair version of Lemma 6.1). *Suppose in the pair model that the clause a receives incoming messages $\dot{q}_e$ ($e \in \delta a$) whose reweighted versions Q_e (defined by (7.10)) satisfy, for some absolute constant $0 < \zeta \leq 1/20$,*

$$\max_{\sigma \in \{\mathtt{y}, \mathtt{c}\}^2} \left\{ \left| Q_e(\sigma) - \frac{1}{4} \right| \right\} \leq \frac{O(1)}{2^{k\zeta}}, \quad Q_e(\sigma) = \frac{1 + O(2^{-k\zeta})}{4 \cdot 2^{|\partial a|-1}} \text{ if } \mathrm{r}[\sigma] = 1, \quad Q_e(\mathtt{rr}) \leq \frac{O(1)}{2^{k(1+\zeta)}}, \tag{7.11}$$

Let $h \equiv \mathrm{BP}[\dot{q}]$. Meanwhile, let $\hat{q}_{e,\infty}$ ($e \in \delta a$) denote a set of desired outgoing messages, and suppose that

$$\widehat{\mathrm{ERR}}(\hat{q}_{e,\infty}, h_e) \leq \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \ddot{\epsilon}_e \end{pmatrix} \leq \frac{1}{k^4} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

for all $e \in \delta a$ (in the notation of Definition 7.5). Then there exist clause weights $\Gamma = (\gamma_e)_{e \in \delta a}$ such that the weighted BP recursion outputs $\hat{q}_{av,\infty} = \mathrm{BP}_{av}[\dot{q}; \Gamma]$ for all $v \in \partial a$, and these weights satisfy the bounds

$$\widehat{\mathrm{ERR}}(1, \gamma_e) \leq O(1) \begin{pmatrix} \epsilon_e \\ \epsilon_e + \dot{\epsilon}_e \\ \epsilon_e + \ddot{\epsilon}_e \end{pmatrix} + O(k^6) \begin{pmatrix} 2^{-k} & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 2^{-k(1-\zeta)} & 2^{-k} \end{pmatrix} \sum_{e' \in \delta a} \begin{pmatrix} \epsilon_{e'} \\ \dot{\epsilon}_{e'} \\ \ddot{\epsilon}_{e'} \end{pmatrix}$$

for all $e \in \delta a$.

Organization of the remainder of this section:

– In §7.1 we give the precise construction (see Definition 7.9) of the weights $\Lambda \equiv \Lambda(T; \omega_{\delta T})$ of Proposition 7.3. The construction iterates between a series of clause and variable updates, which we then proceed to analyze in the subsequent subsections, as follows:
– In §7.2 we analyze the clause updates in Definition 7.9. As a byproduct of the analysis we prove Proposition 7.6.
– In §7.3 we analyze the variable updates in Definition 7.9.
– In §7.4 we combine the analysis of §7.2 and §7.3 to prove Proposition 7.3.
– In §7.5 we give the analysis in the non-compound case to prove Proposition 7.4.

7.1. Lagrange multipliers for subtrees of compound enclosures.

Fix T as in Definition 7.1. For the constrained optimization problem $\nu^{\mathrm{op}}(T; \omega_{\delta T})$ of (7.1), we will always set boundary conditions $\omega_{\delta T}$ such that there is a unique solution in the interior of the feasible domain of measures ν. It then follows by general theory that there is a set of weights $\Lambda \equiv \Lambda(T; \omega_{\delta T})$ that are Lagrangian (recalling Definition 7.1, this means that $\langle \ln \Lambda, \nu \rangle$ is constant over $\nu \in J(T; \omega_{\delta T})$), and such that the solutions of (7.1) and (7.2) coincide — i.e., such that

$$\arg\max_{\nu} \left\{ \mathcal{H}(\nu) : \nu \in J(T, \omega_{\delta T}) \right\} = \arg\max_{\nu} \left\{ \mathcal{H}(\nu) + \langle \ln \Lambda, \nu \rangle : \nu \in \Delta(T) \right\}. \tag{7.12}$$

In this subsection, we describe the explicit construction of these weights in the setting of Proposition 7.3. We begin by remarking on a single-copy analogue which was obtained in previous sections.

Remark 7.7. Recall that in the single-copy model, if U is a tree whose leaves are all variables and whose clauses are all strictly coherent, then Corollary 3.56 guarantees the existence of weights $_\star\Lambda_U$ such that the $_\star\Lambda_U$-weighted Gibbs measure on U has edge marginals $_\star\pi$. If U is moreover nice, then the weights are constructed explicitly by Corollary 6.2. Recall that the corresponding BP solution is not given by $_\star q$, but rather by $_\bullet q$ as defined by (3.92). It follows from Corollary 6.2 that

$$\left\| \frac{_\bullet\dot{q}_e}{_\star\dot{q}_e} - 1 \right\|_\infty + \left\| \frac{_\bullet\hat{q}_e}{_\star\hat{q}_e} - 1 \right\|_\infty \leq \frac{O(1)}{k^{r/2}}$$

for all edges e of T. If $R \geq k$ (as assumed in Proposition 7.3), then this error is very small. In particular, on nice edges it implies that $_\bullet q$ will satisfy the same estimates as $_\star q$ from Definition 3.8. For this reason, we assume $R \geq k$ in the rest of the section, even when not explicitly stated. From now on we denote $_* q \equiv {_\bullet q} \otimes {_\bullet q}$. In the setting of Proposition 7.3, we have a tree T as in Definition 7.1 or Figure 7. Let $T \setminus a_{\mathrm{RT}}$ denote the tree T with a_{RT} removed: then all the leaves of $T \setminus a_{\mathrm{RT}}$ are variables, so $_\star \Lambda_{T \setminus a_{\mathrm{RT}}}$ is defined by Corollaries 3.56 and 6.2. Then, for a single-copy coloring $\underline{\tau} = \underline{\tau}_T$ on T, we define

$$_\star \Lambda_T(\underline{\tau}) \equiv {_\bullet \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}}(\tau_{a_{\mathrm{RT}} v_{\mathrm{RT}}}) \cdot {_\star \Lambda_{T \setminus a_{\mathrm{RT}}}}(\underline{\tau}_{T \setminus a_{\mathrm{RT}}}), \tag{7.13}$$

that is, we put an additional weight $_\bullet \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}$ on the root clause a_{RT}. As a consequence, in the $_\star \Lambda_T$-weighted model on T, the BP message from a_{RT} to v_{RT} will be precisely $_\bullet \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}$.

Definition 7.8 (parametrization of Lagrangian weights Λ on a subtree of a compound enclosure). Let T be as in Definition 7.1. We parametrize the Lagrangian weights Λ on T as follows. Given a pair $\underline{\sigma} \equiv (\underline{\sigma}^1, \underline{\sigma}^2)$ of valid colorings of T, let $\underline{x} \equiv (\underline{x}^1, \underline{x}^2) \equiv \underline{x}(\underline{\sigma})$ be the corresponding pair of frozen configurations on the variables of T. Recall from Remark 7.7 that $_* q = {_\bullet q} \otimes {_\bullet q}$. We then take the weights to be of form

$$\Lambda(\underline{\sigma}) = {_* \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}}(\sigma_{a_{\mathrm{RT}} v_{\mathrm{RT}}}) \left\{ \prod_{v \in V_T \setminus \mathcal{L}T} \Lambda_v(\underline{\sigma}_{\delta v}) \right\} \left\{ \prod_{e \in \delta T} \lambda_e(\sigma_e) \right\},$$

where for each internal variable $v \in V_T \setminus \mathcal{L}T$ the weight factorizes as

$$\Lambda_v(\underline{\sigma}_{\delta v}) \equiv \overbrace{\left(\prod_{j=1,2} (\lambda_v)^j((x_v)^j) \right)}^{\lambda_v(x_v)} \prod_{e \in \delta v} \overbrace{\left(\prod_{j=1,2} (\lambda_e)^j((\sigma_e)^j) \right)}^{\lambda_e(\sigma_e)}. \tag{7.14}$$

We assume that $(\lambda_v)^j(+) = 1$, and $(\lambda_e)^j((\sigma_e)^j) = 1$ for all $(\sigma_e)^j \neq \mathtt{r}$; we will often abbreviate $(\lambda_e)^j \equiv (\lambda_e)^j(\mathtt{r})$. If Λ is any weight on T of the functional form just described, then Λ is a Lagrangian weight in the sense of Definition 7.1, meaning that $\langle \ln \Lambda, \nu \rangle$ is constant over $\nu \in J(T; \omega_{\delta T})$. Conversely, it is easy to see that any Lagrangian weight can be expressed in this form. Next, recalling the discussion of §2.6, there is a unique solution $q = q_{\mathrm{BP}}(T; \Lambda)$ for the Λ-weighted BP recursions on T, obtained by recursing inwards from the leaves. In particular, for $e \in \delta T$ we have simply $\dot{q}_e = \lambda_e$, and at the root we have $\hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}} = {_* \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}}$. Analogously to (2.33), the marginals of $\nu = \nu[T; \Lambda]$ (the Λ-weighted measure, see (7.2) and (7.3)) can be expressed in terms of $q = q_{\mathrm{BP}}(T; \Lambda)$: for instance, the marginal law of $\underline{\sigma}_{\delta v}$ under ν is given by

$$\left(\nu_{\delta v}[\Lambda_v; \hat{q}] \right)(\underline{\sigma}_{\delta v}) \equiv \frac{1}{\dot{z}_v} \varphi_v(\underline{\sigma}_{\delta v}) \Lambda_v(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \hat{q}_e(\sigma_e), \tag{7.15}$$

where $\dot{z}_v$ is the normalization. (This holds even if $v \in \mathcal{L}T$, provided we define $\Lambda_v(\underline{\sigma}_{\delta v}) \equiv \lambda_e(\sigma_e)$ where e is the one edge in δv.) Likewise, the marginal law on any edge e of T is given by $\nu_e(\sigma) \cong \dot{q}_e(\sigma) \hat{q}_e(\sigma)$ for $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$.

Definition 7.9 (iterative construction of Lagrangian weights on a subtree of a compound enclosure). Continuing in the setting of Proposition 7.3, we now define a sequence of weights Λ_t (parametrized as in Definition 7.8 for each $t \geq 0$), which will be proved in this section to converge as $t \to \infty$ to the Lagrangian weights $\Lambda_\infty \equiv \Lambda(T; \omega_{\delta T})$ of Proposition 7.3. At the same time we will define messages q_t which will converge as $t \to \infty$ to the BP solution $q \equiv q_{\mathrm{BP}}(T; \Lambda_\infty)$ for the Λ_∞-weighted model on T, as discussed above. We initialize our construction at $t = 0$ with weights

$$\Lambda_0(\underline{\sigma}) \equiv \left({_\star \Lambda_T} \otimes {_\star \Lambda_T} \right)(\underline{\sigma}) \equiv \prod_{j=1,2} {_\star \Lambda_T}(\underline{\sigma}^j)$$

where $_\star \Lambda_T$ are the single-copy weights defined by (7.13) in Remark 7.7. More explicitly, if $\underline{\tau}$ denotes a single-copy coloring of T, then its weight under $_\star \Lambda_T$ is given by

$$_\star \Lambda_T(\underline{\tau}) = {_\bullet \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}}}(\sigma_{a_{\mathrm{RT}} v_{\mathrm{RT}}}) \left\{ \prod_{v \in V_T \setminus \mathcal{L}T} \overbrace{\left({_\star \lambda_v}(x_v) \prod_{e \in \delta v} {_\star \lambda_e}(\tau_e) \right)}^{_\star \Lambda_v(\underline{\tau}_{\delta v})} \right\} \left\{ \prod_{e \in \delta T} {_\star \lambda_e}(\tau_e) \right\}, \tag{7.16}$$

where $x_v \in \{\text{-}, \text{+}, \text{f}\}$ denotes the frozen spin corresponding to $\underline{\tau}_{\delta v}$, and we recall from Corollary 3.56 that for $e \in \delta T$ we take $_\star \lambda_e = {}_\bullet \dot{q}_e$ as defined by (3.92). Note that Λ_0 is indeed of the functional form specified in Definition 7.8. Let $q_0 \equiv q_{\text{BP}}(T; \Lambda_0)$; explicitly, we have $q_0 = {}_\star q \equiv {}_\bullet q \otimes {}_\bullet q$ (see Remark 7.7). For $t \geq 1$, given Λ_{t-1} we define updated weights Λ_t (also of the functional form from Definition 7.8) by making the following series of updates, started from the boundary δT and working up to the root, then working back down to the boundary:

I. **Boundary updates.** Recall from Definition 7.8 that for each boundary edge $e = (av) \in \delta T$, the weight λ_e is equivalent to the variable-to-clause message $\dot{q}_e$. We update these weights by setting

$$\lambda_{av,t}(\sigma) \equiv \dot{q}_{va,t}(\sigma) \equiv \frac{\omega_{av}(\sigma)}{\hat{q}_{av,t-1}(\sigma)} \bigg/ \left(\sum_\tau \frac{\omega_{av}(\tau)}{\hat{q}_{av,t-1}(\tau)} \right).$$

II. **Upward pass (from boundary to root).** Having updated the boundary messages, go up the tree, alternating steps (a) and (b), starting with step (a) at the clauses incident to δT:

(a) At a clause $a \neq a_{\text{RT}}$, suppose the upward messages $\dot{q}_{ua}$ from the child variables u have just been updated to their t-versions, while the downward message $\hat{q}_{va}$ from the parent variable v is still at its $(t-1)$-version. Then apply BP to update the upward message from the clause,

$$\hat{q}_{av,t} = \text{BP}_{av}[\dot{q}_t].$$

(Note that we have not yet defined $\dot{q}_t$ on all edges of T, but the right-hand side is well-defined since it depends only on the $\dot{q}_t$ messages coming from the child variables u of a.)

(b) At an internal variable v, suppose the upward messages $\hat{q}_{bv}$ from the child clauses b have just been updated to their t-versions; while the downward message $\hat{q}_{av}$ from the parent clause a, as well as the variable weight Λ_v, are still at their $(t-1)$-versions. Let $\Lambda_{v,t}$ be the unique choice of weights, of the functional form (7.14) from Definition 7.8, such that the measure

$$\nu_{\delta v,t}(\underline{\sigma}_{\delta v}) \equiv \frac{1}{\dot{z}_{v,t}} \varphi_v(\underline{\sigma}_{\delta v}) \Lambda_{v,t}(\underline{\sigma}_{\delta v}) \left\{ \hat{q}_{av,t-1}(\sigma_{av}) \prod_{b \in \partial v \setminus a} \hat{q}_{bv,t}(\sigma_{bv}) \right\} \tag{7.17}$$

is judicious (in the sense of Definition 7.1). For now we assume that the weights $\Lambda_{v,t}$ exist; we will explicitly construct them in Proposition 7.13 below. After obtaining $\Lambda_{v,t}$, we apply BP to update the message upward from variable v, that is to say, we let

$$\dot{q}_{va,t} = \text{BP}_{va}[\hat{q}_t; \Lambda_{v,t}].$$

(Note again that we have not yet defined $\hat{q}_t$ on all edges of T, but the right-hand side is well-defined since it depends only on the $\hat{q}_t$ messages coming from the child clauses b of v.)

The upward pass is completed once we have applied step (2b) at the root v_{RT}.

III. **Downward pass (from root to boundary).** Recall from Definition 7.8 that we always put weight $_\star \hat{q}_{a_{\text{RT}} v_{\text{RT}}}$ on the root clause a_{RT}, and this will also be the BP message from a_{RT} to v_{RT}. We therefore start the downward pass with the trivial update

$$\hat{q}_{a_{\text{RT}} v_{\text{RT}}, t} = \hat{q}_{a_{\text{RT}} v_{\text{RT}}, t-1} = \hat{q}_{a_{\text{RT}} v_{\text{RT}}, 0} = {}_\star \hat{q}_{a_{\text{RT}} v_{\text{RT}}}. \tag{7.18}$$

Then continue down the tree, alternating steps (a) and (b), starting from step (a) at v_{RT}:

(a) At an internal variable v, suppose that the downward message $\hat{q}_{bv}$ from the parent clause b has just been updated to its t-version. This means the upward messages $\hat{q}_{av}$ from the child clauses a, as well as the variable weight Λ_v, were already updated to their t-versions during the preceding upward pass. Now apply BP to update the downward messages from the variable,

$$\dot{q}_{va,t} = \text{BP}_{va}\left[\left(\hat{q}_{bv,t-1}; (\hat{q}_{cv,t} : c \in \partial v \setminus \{a, b\}) \right); \Lambda_{v,t} \right]$$

where we have chosen to use the $(t-1)$-version of the downward message $\hat{q}_{bv}$ rather than the t-version. This is merely convenient for our analysis, but will have no effect on the limit since the messages all converge.

(b) At a clause $a \neq a_{\mathrm{RT}}$, suppose the downward message $\dot{q}_{ua}$ from the parent variable u has just been updated to its t-version. This means the upward messages $\dot{q}_{va}$ from the child variables v were already updated to their t-versions during the preceding upward pass. Apply BP to update the downward messages from the clause, again using for convenience the $(t-1)$-version of the message from above:

$$\hat{q}_{av,t} = \mathrm{BP}_{av}\left[\left(\dot{q}_{ua,t-1}, (\dot{q}_{wa,t} : w \in \partial a \setminus \{u, v\})\right)\right].$$

The downward pass is completed once we have applied step (3b) at each of the clauses incident to the boundary edges δT. This completes the definition of Λ_t and q_t.

In summary, by iterating the above steps, we obtain (Λ_t, q_t) for all $t \geq 0$ on T. The next few subsections (§7.2–7.4) are devoted to analyzing this iteration under the assumptions of Proposition 7.3.

7.2. **Analysis of clause BP recursion in pair model.** In this subsection we analyze the clause BP recursion in the pair model. The main result of the subsection is Proposition 7.12, which will be applied to give estimates for Steps II(a) and III(b) in Definition 7.9. At the end of the subsection we also give the proof of Proposition 7.6. Recall that clauses do not distinguish between the green and blue colors, so in this subsection we shall work with the reduced alphabet $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$ with $\mathtt{c}$ as in (3.5).

Definition 7.10 (error notation for variable-to-clause messages). In the pair model, given two variable-to-clause messages p and q (both probability measures on $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$), let P and Q be their reweightings as defined by (7.10). For $j = 1, 2$ let P^j and Q^j be the marginals of P and Q on the j-th coordinate. We then write

$$\mathrm{E\ddot{R}R}(p, q) = \begin{pmatrix} \max\left\{\left|\dfrac{Q(\sigma)}{P(\sigma)} - 1\right| : \sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2, \mathtt{r}[\sigma] = 0\right\} \\[2mm] \max\left\{\left|\dfrac{Q(\sigma)}{P(\sigma)} - 1\right| : \sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2, \mathtt{r}[\sigma] = 1\right\} \\[2mm] \max\left\{\left|\dfrac{Q(\sigma)}{P(\sigma)} - 1\right| : \sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2, \mathtt{r}[\sigma] = 2\right\} \\[2mm] \max\left\{\displaystyle\sum_{j=1,2}\left|\dfrac{Q^j(\sigma)}{P^j(\sigma)} - 1\right| : \sigma \in \{\mathtt{y}, \mathtt{c}\}\right\} \\[2mm] \displaystyle\sum_{j=1,2}\left|\dfrac{Q^j(\mathtt{r})}{P^j(\mathtt{r})} - 1\right| \end{pmatrix} \in \mathbb{R}^5.$$

We write $\mathrm{E\ddot{R}R}(p, q) \leq \underline{s}$ to mean that $\mathrm{E\ddot{R}R}(p, q)$ is coordinatewise upper bounded by $\underline{s} \in \mathbb{R}^5$.

Definition 7.11 (subsets of clause colorings). Given a clause $a \in F$, we now fix some notation for various subsets of colorings of δa. Abbreviate $K = |\delta a| \in \{k, k-1\}$. Let $\mathtt{W} \equiv \{\mathtt{y}, \mathtt{c}\}^{K-1}$ and $\mathtt{Y} \equiv \{\mathtt{y}\}^{K-1}$. Then let

$$\mathtt{R} \equiv \left\{\underline{\sigma} \in \{\mathtt{r}, \mathtt{y}\}^{K-1} : \left|\left\{j \in [K] : \sigma_j = \mathtt{r}\right\}\right| = 1\right\},$$

$$\mathtt{C} \equiv \left\{\underline{\sigma} \in \{\mathtt{y}, \mathtt{c}\}^{K-1} : \left|\left\{j \in [K] : \sigma_j = \mathtt{c}\right\}\right| = 1\right\},$$

$$\mathtt{C}_{\geq j} \equiv \left\{\underline{\sigma} \in \{\mathtt{y}, \mathtt{c}\}^{K-1} : \left|\left\{j \in [K] : \sigma_j = \mathtt{c}\right\}\right| \geq j\right\}.$$

For any two subsets $\mathtt{A}$ and $\mathtt{B}$ of $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^{K-1}$, we will abbreviate $\mathtt{AB} \equiv \mathtt{A} \times \mathtt{B}$.

Proposition 7.12. *Suppose the clause $a \in F$ receives two sets of incoming messages p and q, with each message a probability measure over $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$ satisfying the estimates* (7.11), *such that, in the notation of Definition* 7.10, *we have*

$$\ddot{\mathrm{ERR}}(p_e, q_e) \leq \begin{pmatrix} \delta_e \\ \dot{\delta}_e \\ \ddot{\delta}_e \\ {}^{\mathrm{m}}\delta_e \\ {}^{\mathrm{m}}\dot{\delta}_e \end{pmatrix}$$

for each $e \in \delta a$. Then the outgoing clause-to-variable messages $g \equiv \mathrm{BP}[p]$ and $h \equiv \mathrm{BP}[q]$ satisfy, for all $e' \in \delta a$,

$$\widehat{\mathrm{ERR}}(g_{e'}, h_{e'}) \leq O(k^2) \begin{pmatrix} 2^{-k(1+\zeta)} & 4^{-k} & 2^{-k(1+\zeta)} & 2^{-k} & 2^{-k} \\ 2^{-k} & 2^{-k} & 2^{-k(1+\zeta)} & 1 & 2^{-k} \\ 1 & 4^{-k} & 2^{-k(1+\zeta)} & 2^{-k} & 2^{-k} \end{pmatrix} \sum_{e \in \delta a \backslash e'} \begin{pmatrix} \delta_e \\ \dot{\delta}_e \\ \min\{\ddot{\delta}_e, 1\} \\ {}^{\mathrm{m}}\delta_e \\ {}^{\mathrm{m}}\dot{\delta}_e \end{pmatrix}. \tag{7.19}$$

Proof. Throughout the proof, we fix an edge $e' = (av)$ on which to calculate the outgoing BP messages, and suppress e' from the notation when possible. We abbreviate

$$\begin{pmatrix} \delta \\ \dot{\delta} \\ \ddot{\delta} \\ {}^{\mathrm{m}}\delta \\ {}^{\mathrm{m}}\dot{\delta} \end{pmatrix} \equiv \sum_{e \in \delta a \backslash e'} \begin{pmatrix} \delta_e \\ \dot{\delta}_e \\ \min\{\ddot{\delta}_e, 1\} \\ {}^{\mathrm{m}}\delta_e \\ {}^{\mathrm{m}}\dot{\delta}_e \end{pmatrix}.$$

Step 1. Rewriting of BP equations. For $e \in \delta a$ let P_e, Q_e be the reweighted versions of p_e, q_e defined by (7.10). Let $\mathcal{X}$ and $\mathcal{Y}$ be the (non-normalized) measures over configurations $\underline{\sigma}_{\delta a \backslash e'}$ defined by

$$\mathcal{X}(\underline{\sigma}_{\delta a \backslash e'}) = \prod_{e \in \delta a \backslash e'} \left\{ (2^{|\partial a|-1})^{\mathtt{r}[\sigma_e]} P_e(\sigma_e) \right\},$$

$$\mathcal{Y}(\underline{\sigma}_{\delta a \backslash e'}) = \prod_{e \in \delta a \backslash e'} \left\{ (2^{|\partial a|-1})^{\mathtt{r}[\sigma_e]} Q_e(\sigma_e) \right\}.$$

Then the equations $g \equiv g_{av} \equiv \mathrm{BP}_{av}[p]$ and $h \equiv h_{av} \equiv \mathrm{BP}_{av}[q]$ can be rewritten as

$$g(\sigma) = \frac{1}{Z_g} \sum_{\underline{\sigma}_{\delta a} : \sigma_{e'} = \sigma} \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \mathcal{X}(\underline{\sigma}_{\delta a \backslash e'}),$$

$$h(\sigma) = \frac{1}{Z_h} \sum_{\underline{\sigma}_{\delta a} : \sigma_{e'} = \sigma} \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \mathcal{Y}(\underline{\sigma}_{\delta a \backslash e'}),$$

where Z_g and Z_h are the normalizing constants, and $\hat{\varphi}_a$ denotes the factor for the pair coloring model, as in (3.66). In the remainder of the proof, we first obtain error bounds between $\mathcal{X}$ and $\mathcal{Y}$, then use this to deduce error bounds between g and h.

Step 2. Error bounds between $\mathcal{X}$ and $\mathcal{Y}$. Using the notation of Definition 7.11, we have

$$\mathcal{X}(\mathtt{yy}) = \prod_{e \in \delta a \backslash e'} P_e(\mathtt{yy}) = \prod_{e \in \delta a \backslash e'} \left\{ Q_e(\mathtt{yy}) + O(\delta_e) \right\} = \mathcal{Y}(\mathtt{yy}) + O\left(\frac{\delta}{4^k}\right) = \Theta\left(\frac{1}{4^k}\right),$$

$$\mathcal{X}(\mathtt{ry}) = \sum_{e \in \delta a \backslash e'} \left\{ 2^{|\partial a|-1} Q_e(\mathtt{ry}) + O(\dot{\delta}_e) \right\} \prod_{e'' \in \delta a \backslash \{e, e'\}} \left\{ Q_{e''}(\mathtt{yy}) + O(\delta_{e''}) \right\} = \mathcal{Y}(\mathtt{ry}) + O\left(\frac{\dot{\delta} + k\delta}{4^k}\right) = \Theta\left(\frac{k}{4^k}\right),$$

$$\mathcal{X}(\mathtt{cy}) = \sum_{e \in \delta a \backslash e'} \left\{ Q_e(\mathtt{cy}) + O(\delta_e) \right\} \prod_{e'' \in \delta a \backslash \{e, e'\}} \left\{ Q_{e''}(\mathtt{yy}) + O(\delta_{e''}) \right\} = \mathcal{Y}(\mathtt{cy}) + O\left(\frac{k\delta}{4^k}\right) = \Theta\left(\frac{k}{4^k}\right), .$$

Recall for $j = 1, 2$ we use $(P_e)^j$ to denote the marginal of P_e on the j-th copy. Then

$$X(\text{YW}) = \prod_{e \in \delta a \setminus e'} \left\{ (P_e)^1(\text{y}) - P_e(\text{yr}) \right\} = \prod_{e \in \delta a \setminus e'} \left\{ (Q_e)^1(\text{y}) \Big(1 + O(^{\text{m}}\dot{\delta}_e) \Big) - Q_e(\text{yr}) \Big(1 + O(\dot{\delta}_e) \Big) \right\}$$

$$= \prod_{e \in \delta a \setminus e'} \left\{ (Q_e)^1(\text{y}) - Q_e(\text{yr}) + O\Big(^{\text{m}}\dot{\delta}_e + \frac{\dot{\delta}_e}{2^k}\Big) \right\} = \mathcal{Y}(\text{YW}) + O\Big(\frac{^{\text{m}}\dot{\delta}}{2^k} + \frac{\dot{\delta}}{4^k} \Big) = \Theta\Big(\frac{1}{2^k} \Big),$$

where the transition to the second line uses the estimates (7.11). By similar calculations, we have

$$X(\text{CW}) = \sum_{e \in \delta a \setminus e'} \left\{ (P_e)^1(\text{c}) - P_e(\text{cr}) \right\} \prod_{e'' \in \delta a \setminus \{e, e'\}} \left\{ (P_e)^1(\text{y}) - P_e(\text{yr}) \right\}$$

$$= \mathcal{Y}(\text{CW}) + O\Big(\frac{k^{\text{m}}\dot{\delta}}{2^k} + \frac{k\dot{\delta}}{4^k} \Big) = \Theta\Big(\frac{k}{2^k} \Big)$$

$$X(\text{RW}) = 2^{|\partial a|-1} \sum_{e \in \delta a \setminus e'} \left\{ (P_e)^1(\text{r}) - P_e(\text{rr}) \right\} \prod_{e'' \in \delta a \setminus \{e, e'\}} \left\{ (P_e)^1(\text{y}) - P_e(\text{yr}) \right\}$$

$$= \mathcal{Y}(\text{RW}) + O\Big(\frac{k^{\text{m}}\dot{\delta} + {}^{\text{m}}\ddot{\delta}}{2^k} + \frac{\ddot{\delta}}{2^{k(1+\zeta)}} + \frac{k\dot{\delta}}{4^k} \Big) = \Theta\Big(\frac{k}{2^k} \Big),$$

$$X(\text{WW}) = \prod_{e \in \delta a \setminus e'} \left\{ 1 - (Q_e)^1(\text{r}) - (Q_e)^2(\text{r}) + Q_e(\text{rr}) + O\Big(\frac{^{\text{m}}\ddot{\delta}_e}{2^k} + \frac{\min\{\ddot{\delta}_e, 1\}}{2^{k(1+\zeta)}} \Big) \right\}$$

$$= \mathcal{Y}(\text{WW}) + O\Big(\frac{^{\text{m}}\ddot{\delta}}{2^k} + \frac{\ddot{\delta}}{2^{k(1+\zeta)}} \Big) = \Theta(1).$$

In the last calculation, the term involving $\min\{\ddot{\delta}_e, 1\}$ arises because on the one hand the definition of $\ddot{\delta}_e$ implies

$$Q_e(\text{rr}) = P_e(\text{rr}) \left\{ 1 + O(\ddot{\delta}_e) \right\},$$

and on the other hand we assume that both p and q satisfy (7.11) which implies that

$$\max \left\{ P_e(\text{rr}), Q_e(\text{rr}) \right\} \le \frac{O(1)}{2^{k(1+\zeta)}}.$$

We also comment that we have kept track of explicit powers of k, but this is not an important point; one could also write the proof with factors $k^{O(1)}$ where $O(1)$ stays bounded throughout. Lastly we have

$$\mathcal{X}(\mathtt{cc}) = \sum_{e\in\delta a\backslash e'} P_e(\mathtt{cc}) \prod_{e''\in\delta a\backslash\{e,e'\}} P_{e''}(\mathtt{yy}) + \sum_{e\in\delta a\backslash e'}\sum_{e''\in\delta a\backslash\{e,e'\}} P_e(\mathtt{cy})P_{e''}(\mathtt{yc}) \prod_{e'''\in\delta a\backslash\{e,e',e''\}} P_{e'''}(\mathtt{yy})$$

$$= \mathcal{Y}(\mathtt{cc}) + O\left(\frac{k^2\delta}{4^k}\right) = \Theta\left(\frac{k^2}{4^k}\right),$$

$$\mathcal{X}(\mathtt{RC}) = 2^{|\partial a|-1}\left\{ \sum_{e\in\delta a\backslash e'} P_e(\mathtt{rc}) \prod_{e''\in\delta a\backslash\{e,e'\}} P_{e''}(\mathtt{yy}) + \sum_{e\in\delta a\backslash e'}\sum_{e''\in\delta a\backslash\{e,e'\}} P_e(\mathtt{ry})P_{e''}(\mathtt{yc}) \prod_{e'''\in\delta a\backslash\{e,e',e''\}} P_{e'''}(\mathtt{yy}) \right\}$$

$$= \mathcal{Y}(\mathtt{RC}) + O\left(\frac{k^2\delta + k\dot{\delta}}{4^k}\right) = \Theta\left(\frac{k^2}{4^k}\right),$$

$$\mathcal{X}(\mathtt{RR}) = 4^{|\partial a|-1}\left\{ \sum_{e\in\delta a\backslash e'} P_e(\mathtt{rr}) \prod_{e''\in\delta a\backslash\{e,e'\}} P_{e''}(\mathtt{yy}) + \sum_{e\in\delta a\backslash e'}\sum_{e''\in\delta a\backslash\{e,e'\}} P_e(\mathtt{ry})P_{e''}(\mathtt{yr}) \prod_{e'''\in\delta a\backslash\{e,e',e''\}} P_{e'''}(\mathtt{yy}) \right\}$$

$$= \mathcal{Y}(\mathtt{RR}) + O\left(\frac{k\delta + \ddot{\delta}}{2^{k(1+\zeta)}} + \frac{k\dot{\delta} + k^2\delta}{4^k}\right) = \mathcal{Y}(\mathtt{RR}) + O\left(\frac{k\delta + \ddot{\delta}}{2^{k(1+\zeta)}} + \frac{k\dot{\delta}}{4^k}\right) = O\left(\frac{k}{2^{k(1+\zeta)}}\right).$$

Step 3. Error bounds between g and h. Recall that in Step 1 we expressed g and h in terms of $\mathcal{X}$ and $\mathcal{Y}$. We now use the bounds from Step 2 on the error between $\mathcal{X}$ and $\mathcal{Y}$ to deduce bounds on the error between g and h. The two easiest bounds are

$$Z_g g(\mathtt{rr}) = \mathcal{X}(\mathtt{YY}) = \mathcal{Y}(\mathtt{YY})\left\{1 + O(\delta)\right\} = Z_h h(\mathtt{rr})\left\{1 + O(\delta)\right\},$$

$$Z_g g(\mathtt{rc}) = \mathcal{X}(\mathtt{YC}_{\geq 1}) = \mathcal{X}(\mathtt{YW}) - \mathcal{X}(\mathtt{YY}) = Z_h h(\mathtt{rc})\left\{1 + O\left({}^{\mathtt{m}}\delta + \frac{\dot{\delta} + \delta}{2^k}\right)\right\}.$$

The remaining estimates are similar, only slightly more complicated. We next have

$$Z_g g(\mathtt{ry}) = \mathcal{X}\left(\mathtt{Y} \times (\mathtt{R} \cup \mathtt{C}_{\geq 2})\right) = \mathcal{X}(\mathtt{YR}) + \mathcal{X}(\mathtt{YW}) - \mathcal{X}(\mathtt{YY}) - \mathcal{X}(\mathtt{YC}) = Z_h h(\mathtt{ry})\left\{1 + O\left({}^{\mathtt{m}}\delta + \frac{\dot{\delta} + k\delta}{2^k}\right)\right\}$$

The calculations for the spins in $\{\mathtt{y},\mathtt{c}\}^2$ are all similar:

$$Z_g g(\mathtt{cc}) = \mathcal{X}\left((\mathtt{C}_{\geq 1})^2\right) = \mathcal{X}(\mathtt{WW}) - \mathcal{X}(\mathtt{WY}) - \mathcal{X}(\mathtt{YW}) + \mathcal{X}(\mathtt{YY}) = Z_h h(\mathtt{cc})\left\{1 + O\left(\frac{{}^{\mathtt{m}}\delta + {}^{\mathtt{m}}\dot{\delta}}{2^k} + \frac{\ddot{\delta}}{2^{k(1+\zeta)}} + \frac{\delta + \dot{\delta}}{4^k}\right)\right\},$$

$$Z_g g(\mathtt{cy}) = \mathcal{X}\left(\mathtt{C}_{\geq 1} \times (\mathtt{R} \cup \mathtt{C}_{\geq 2})\right) = \left(\mathcal{X}(\mathtt{WR}) - \mathcal{X}(\mathtt{YR})\right) + \left(\mathcal{X}\left((\mathtt{C}_{\geq 1})^2\right) - \mathcal{X}(\mathtt{WC}) + \mathcal{X}(\mathtt{YC})\right)$$

$$= Z_h h(\mathtt{cy})\left\{1 + O\left(\frac{k{}^{\mathtt{m}}\delta + {}^{\mathtt{m}}\dot{\delta}}{2^k} + \frac{\ddot{\delta}}{2^{k(1+\zeta)}} + \frac{k^2\delta + k\dot{\delta}}{4^k}\right)\right\},$$

$$Z_g g(\mathtt{yy}) = \mathcal{X}\left((\mathtt{R} \cup \mathtt{C}_{\geq 2})^2\right) = \mathcal{X}(\mathtt{RR}) + \left(\mathcal{X}(\mathtt{RW}) - \mathcal{X}(\mathtt{RY}) - \mathcal{X}(\mathtt{RC})\right) + \left(\mathcal{X}(\mathtt{WR}) - \mathcal{X}(\mathtt{YR}) - \mathcal{X}(\mathtt{CR})\right)$$

$$+ \left(\mathcal{X}(\mathtt{C}_{\geq 1}\mathtt{C}_{\geq 1}) - (\mathcal{X}(\mathtt{CW}) - \mathcal{X}(\mathtt{CY})) - (\mathcal{X}(\mathtt{WC}) - \mathcal{X}(\mathtt{YC})) + \mathcal{X}(\mathtt{cc})\right)$$

$$= Z_h h(\mathtt{yy})\left\{1 + O\left(\frac{k{}^{\mathtt{m}}\delta + {}^{\mathtt{m}}\dot{\delta}}{2^k} + \frac{k\delta + \ddot{\delta}}{2^{k(1+\zeta)}} + \frac{k\dot{\delta}}{4^k}\right)\right\}.$$

Combining these estimates gives

$$Z_g = Z_h \left\{ 1 + O\left(\frac{k^{\mathfrak{m}}\delta + {}^{\mathfrak{m}}\dot{\delta}}{2^k} + \frac{k\delta + \ddot{\delta}}{2^{k(1+\zeta)}} + \frac{k\dot{\delta}}{4^k} \right) \right\}.$$

The claimed bounds directly follow. □

Let us briefly summarize where we are, with respect to the proof of Proposition 7.3. In §7.1 we specified the construction of the weights Λ (Definition 7.9). In the current subsection (§7.2) we made a general analysis of a **clause** update (Proposition 7.12) which can be applied to estimate the effects of both Step II(a) and Step III(b). In the next subsection §7.3 we will make a general analysis of a **variable** update, which will be applied to estimate the effects of both Step II(b) and Step III(a). We will combine these estimates in §7.4 to complete the proof of Proposition 7.3.

To conclude the current subsection, we momentarily digress from the plan just described, in order to complete the proof of Proposition 7.6, which which makes use of the preceding result Proposition 7.12.

Proof of Proposition 7.6. We first define a sequence of weights $\Gamma_t \equiv (\gamma_{e,t} : e \in \delta a)$, then show that they converge in the limit $t \to \infty$ to the desired weights $\Gamma_\infty = \Gamma$. We emphasize that this index t is **purely local to the proof of this proposition**, and is not the same as the t that indexes the up-and-down passes in Definition 7.9.

Step 1. Iterative definition of weights. Recall that the goal is to find weights Γ such that $\text{BP}_{av}[\dot{q}; \Gamma] = \hat{q}_{av,\infty}$ for all $v \in \partial a$. We initialize $\Gamma_0 \equiv 1$. For all $t \geq 0$ let $h_t \equiv \text{BP}[\dot{q}; \Gamma_t]$ and $g_t \equiv \text{BP}(\gamma_t \dot{q})$, so that h_t is proportional to $\gamma_t g_t$ — in this notation, the goal is to have in the limit $\gamma_\infty g_\infty \cong \hat{q}_\infty$. We then recursively set

$$\gamma_{t+1}(\sigma) = \gamma_t(\sigma) \cdot \frac{\hat{q}_\infty(\sigma)}{h_t(\sigma)} \left/ \left(\frac{\hat{q}_\infty(yy)}{h_t(yy)} \right) \right. \tag{7.20}$$

on all the edges of δa. The rationale for this choice is that it keeps $\gamma_t(yy) = 1$ for all $t \geq 0$, and gives us

$$\gamma_{t+1} g_t = \frac{\gamma_{t+1}}{\gamma_t} \gamma_t g_t \cong \frac{\hat{q}_\infty}{h_t} \gamma_t g_t \cong \frac{\hat{q}_\infty}{h_t} h_t \cong \hat{q}_\infty \,,$$

(which is, clearly, compatible with our eventual goal of $\gamma_\infty g_\infty \cong \hat{q}_\infty$).

Step 2. Analysis of weights. Suppose, in the notation of Definition 7.5, that for all $e \in \delta a$ we have

$$\widehat{\text{ERR}}(\gamma_{e,t}, \gamma_{e,t+1}) = \begin{pmatrix} \epsilon_e(t) \\ \dot{\epsilon}_e(t) \\ \ddot{\epsilon}_e(t) \end{pmatrix}. \tag{7.21}$$

Recall that $\gamma_0 \equiv 1$ and $h_0 \equiv \text{BP}[\dot{q}] \equiv h$, so at $t = 0$ the assumptions give

$$\widehat{\text{ERR}}(\gamma_{e,0}, \gamma_{e,1}) = \widehat{\text{ERR}}\left(1, \frac{\hat{q}_{e,\infty}}{h_e} \cdot \frac{h_e(yy)}{\hat{q}_{e,\infty}(yy)} \right) \leq O(1) \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \widehat{\text{ERR}}(\hat{q}_{e,\infty}, h_e) = O(1) \begin{pmatrix} \epsilon_e \\ \epsilon_e + \dot{\epsilon}_e \\ \epsilon_e + \ddot{\epsilon}_e \end{pmatrix}. \tag{7.22}$$

For all $t \geq 0$, it follows from (7.20) and the relation $h_t \cong \gamma_t g_t$ that

$$\frac{\gamma_{t+2}(\sigma)}{\gamma_{t+1}(\sigma)} = \frac{\hat{q}_\infty(\sigma)/\hat{q}_\infty(yy)}{h_{t+1}(\sigma)/h_{t+1}(yy)} = \overbrace{\left\{ \frac{\gamma_t(\sigma)\hat{q}_\infty(\sigma)/\hat{q}_\infty(yy)}{\gamma_{t+1}(\sigma)h_t(\sigma)/h_t(yy)} \right\}}^{\text{equals one by (7.20)}} \frac{\gamma_{t+1}(\sigma)h_t(\sigma)/h_t(yy)}{\gamma_t(\sigma)h_{t+1}(\sigma)/h_{t+1}(yy)}$$

$$= \frac{[h_t(\sigma)/\gamma_t(\sigma)]/[h_t(yy)/\gamma_t(yy)]}{[h_{t+1}(\sigma)/\gamma_{t+1}(\sigma)]/[h_{t+1}(yy)/\gamma_{t+1}(yy)]} = \frac{g_t(\sigma)/g_t(yy)}{g_{t+1}(\sigma)/g_{t+1}(yy)}. \tag{7.23}$$

(the second-to-last step uses that $\gamma_t(yy) \equiv 1$). We can bound the error between $g_t = \text{BP}[\gamma_t \dot{q}]$ and $g_{t+1} = \text{BP}[\gamma_{t+1}\dot{q}]$ by applying Proposition 7.12, as follows: let p_t denote the normalized measure corresponding to $\gamma_t \dot{q}$, so that, in the

notation of Definition 7.10,

$$\ddot{\text{ERR}}(p_{e,t}, p_{e,t+1}) \le O(1) \begin{pmatrix} \epsilon_e(t) \\ \dot{\epsilon}_e(t) \\ \ddot{\epsilon}_e(t) \\ \epsilon_e(t) + \dot{\epsilon}_e(t)/2^k \\ \dot{\epsilon}_e(t) + \ddot{\epsilon}_e(t)/2^{k\zeta} \end{pmatrix} + O\left(\epsilon_e(t) + \frac{\dot{\epsilon}_e(t)}{2^k} + \frac{\ddot{\epsilon}_e(t)}{2^{k(1+\zeta)}} \right) \mathbf{1}_5$$

— on the right-hand side, the first term comes from the error between the non-normalized measures $\gamma_t q$ and $\gamma_{t+1} q$, while the second term comes from the normalization. In matrix notation,

$$\ddot{\text{ERR}}(p_{e,t}, p_{e,t+1}) \le O(1) \begin{pmatrix} 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 1 & 2^{-k(1+\zeta)} \\ 1 & 2^{-k} & 1 \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 1 & 2^{-k\zeta} \end{pmatrix} \begin{pmatrix} \epsilon_e(t) \\ \dot{\epsilon}_e(t) \\ \ddot{\epsilon}_e(t) \end{pmatrix}.$$

The message $p_0 = \dot{q}$ satisfies the bounds (7.11) by assumption. It follows from (7.22) that the message p_1 (again, the normalization of $\gamma_1 \dot{q}$) also satisfies (7.11), simply because the error between γ_0 and γ_1 is negligible. We will argue by induction that p_t satisfies (7.11) for all $t \ge 0$. First we note that by the definition (7.21) and by (7.23), we have

$$\begin{pmatrix} \epsilon_e(t+1) \\ \dot{\epsilon}_e(t+1) \\ \ddot{\epsilon}_e(t+1) \end{pmatrix} = \widehat{\text{ERR}}(\gamma_{e,t+2}, \gamma_{e,t+1}) \le O(1) \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \widehat{\text{ERR}}(g_{t+1}, g_t),$$

where the 3×3 matrix in the last expression is the same as the one in (7.22). Combining with Proposition 7.12 (more precisely, taking the product of the 3×3 matrix in the last display, the 3×5 matrix in (7.19), and the 5×3 matrix above), we find

$$\begin{pmatrix} \epsilon_e(t+1) \\ \dot{\epsilon}_e(t+1) \\ \ddot{\epsilon}_e(t+1) \end{pmatrix} \le O(k^2) \begin{pmatrix} 2^{-k} & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \end{pmatrix} \sum_{e' \in \delta a \setminus e} \begin{pmatrix} \epsilon_{e'}(t) \\ \dot{\epsilon}_{e'}(t) \\ \ddot{\epsilon}_{e'}(t) \end{pmatrix}. \tag{7.24}$$

It is clear from this that p_t satisfies (7.11) for all $t \ge 0$, and moreover that the iteration defined in Step 1 converges as $t \to \infty$, with limiting weights $\gamma_\infty \equiv \gamma$. Quantitatively, it follows from (7.24) that

$$E(t+1) \equiv \begin{pmatrix} 1 & 2^{-k/2} & 2^{-k(1/2+\zeta)} \end{pmatrix} \sum_{e \in \delta a} \begin{pmatrix} \epsilon_e(t+1) \\ \dot{\epsilon}_e(t+1) \\ \ddot{\epsilon}_e(t+1) \end{pmatrix} \le \frac{O(k^3) E(t)}{2^{k/2}},$$

(the 3×1 row vector in the definition of $E(t)$ is chosen because it is an approximate left eigenvector of the 3×3 matrix in the last bound of (7.24)). We shall bound

$$\widehat{\text{ERR}}(1, \gamma_e) \le O(1) \sum_{t=0}^{\infty} \widehat{\text{ERR}}(\gamma_{e,t}, \gamma_{e,t+1}) \le O(1) \sum_{t=0}^{1} \begin{pmatrix} \epsilon_e(t) \\ \dot{\epsilon}_e(t) \\ \ddot{\epsilon}_e(t) \end{pmatrix} + O(1) \sum_{t=2}^{\infty} E(t) \begin{pmatrix} 1 \\ 2^{k/2} \\ 2^{k(1/2+\zeta)} \end{pmatrix}.$$

On the right-hand side, the $t = 0$ is bounded by (7.22), and the $t = 1$ term can be bounded by (7.24). Thanks to the exponential decay of $E(t)$ in t, the sum over $t \ge 2$ can be bounded by

$$\sum_{t=2}^{\infty} E(t) \begin{pmatrix} 1 \\ 2^{k/2} \\ 2^{k(1/2+\zeta)} \end{pmatrix} \le \frac{O(k^6)}{2^k} E(0) \begin{pmatrix} 1 \\ 2^{k/2} \\ 2^{k(1/2+\zeta)} \end{pmatrix} \le \frac{O(k^6)}{2^k} \begin{pmatrix} 1 \\ 2^{k/2} \\ 2^{k(1/2+\zeta)} \end{pmatrix} \begin{pmatrix} 1 & 2^{-k/2} & 2^{-k(1/2+\zeta)} \end{pmatrix} \sum_{e' \in \delta a} \begin{pmatrix} \epsilon_{e'} \\ \epsilon_{e'} + \dot{\epsilon}_{e'} \\ \epsilon_{e'} + \ddot{\epsilon}_{e'} \end{pmatrix}$$

$$= O(k^6) \begin{pmatrix} 2^{-k} & 2^{-3k/2} & 2^{-k(3/2+\zeta)} \\ 2^{-k/2} & 2^{-k} & 2^{-k(1+\zeta)} \\ 2^{-k(1/2-\zeta)} & 2^{-k(1-\zeta)} & 2^{-k} \end{pmatrix} \sum_{e' \in \delta a} \begin{pmatrix} \epsilon_{e'} \\ \epsilon_{e'} + \dot{\epsilon}_{e'} \\ \epsilon_{e'} + \ddot{\epsilon}_{e'} \end{pmatrix}$$

Combining these bounds gives altogether

$$\widehat{\text{ERR}}(1, \gamma_e) \le O(1) \begin{pmatrix} \epsilon_e \\ \epsilon_e + \dot{\epsilon}_e \\ \epsilon_e + \ddot{\epsilon}_e \end{pmatrix} + O(k^6) \begin{pmatrix} 2^{-k} & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 2^{-k(1-\zeta)} & 2^{-k} \end{pmatrix} \sum_{e' \in \delta a} \begin{pmatrix} \epsilon_{e'} \\ \epsilon_{e'} + \dot{\epsilon}_{e'} \\ \epsilon_{e'} + \ddot{\epsilon}_{e'} \end{pmatrix},$$

from which the claimed result follows. □

7.3. Analysis of variable BP recursion in pair model.

In this subsection we analyze the variable BP recursion in the pair model. The analysis will be applied to give estimates for Steps II(b) and III(a) in Definition 7.9. The main results of the current subsection are Proposition 7.13 and Corollary 7.14, which are stated next. The remaining results of the subsection are technical lemmas used in the proofs of Proposition 7.13 and Corollary 7.14, which we briefly outline here:

– Lemma 7.15 shows that, under the conditions of Proposition 7.13 on the messages in coming to variable v, the resulting law of the frozen spin $x_v \in \{-, +, \mathtt{f}\}^2$ is close to a product measure.
– Lemma 7.16 shows (using the result of Lemma 7.15) that that reweighting the frozen spin $(x_v)^1$ does not greatly affect the law of the other spin $(x_v)^2$, and vice versa.
– Lemma 7.17 shows that when we reweight all the edges $e \in \delta v$, the marginal law of $\sigma_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$ is mainly affected by the weight on edge e alone, with a much smaller effect from the edges $\delta v \setminus e$.
– Lemma 7.18 gives a bound on the marginal error in variable-to-clause messages under certain conditions.

The proof of Proposition 7.13 uses Lemmas 7.16 and 7.17. Corollary 7.14 follows from Proposition 7.13 together with Lemma 7.18. From this point on, rather than keeping track of explicit powers of k, we will simply write $k^{O(1)}$ to indicate k raised to powers bounded by an absolute constant.

Proposition 7.13. *Let v be a nice variable with incoming messages $\hat{p}$ (in the pair coloring model). Recalling the notation of Definition 7.5, assume that for all $e \in \delta v$ we have*

$$\widehat{\mathrm{ERR}}(_*\hat{q}_e, \hat{p}_e) \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix}. \tag{7.25}$$

Suppose v has weight $\Theta \equiv \Theta_v$, of the functional form (7.14) from Definition 7.8 (with θ in place of λ). Assume that the measure $\mu \equiv \nu_{\delta v}[\Theta; \hat{p}]$ (defined using the notation of (7.15)) is judicious in the sense of Definition 7.1, and that

$$\sum_{j=1,2} \sum_{x \in \{-,+,\mathtt{f}\}} \left| \theta^j(x) - 1 \right| \leq \frac{k^{O(1)}}{2^{k\zeta}}, \quad \max_{e \in \delta v} \left\{ \sum_{j=1,2} \left| (\theta_e)^j - 1 \right| \right\} \leq \frac{k^{O(1)}}{2^{k\zeta}}. \tag{7.26}$$

Let $\hat{q}$ be another set of incoming messages, also satisfying (7.25), such that for all $e \in \delta v$ we have

$$\widehat{\mathrm{ERR}}(\hat{p}_e, \hat{q}_e) \leq \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \ddot{\epsilon}_e \end{pmatrix} \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix}.$$

Then there exist weights $\Lambda \equiv \Lambda_v$ (also parametrized as (7.14) from Definition 7.8) such that

$$\sum_{j=1,2} \sum_{x \in \{-,\mathtt{f}\}} \left| \frac{\lambda^j(x)}{\theta^j(x)} - 1 \right| \leq k^{O(1)} \overbrace{\left(\sum_{e' \in \delta v} \left\{ \epsilon_{e'} + \frac{\dot{\epsilon}_{e'}}{2^k} + \frac{\min\{\ddot{\epsilon}_{e'}, 1\}}{2^{k(1+\zeta)}} \right\} \right)}^{\text{denote this } \mathrm{err}_v}, \tag{7.27}$$

$$\sum_{j=1,2} \left| \frac{(\lambda_e)^j}{(\theta_e)^j} - 1 \right| \leq k^{O(1)} \left(\dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} + \mathrm{err}_v \right)$$

and $\nu_{\delta v}[\Lambda; \hat{q}]$ is again judicious in the sense of Definition 7.1.

Corollary 7.14. *In the setting of Proposition 7.13, for $\dot{p} = \mathrm{BP}[\hat{p}; \Theta]$ and $\dot{q} = \mathrm{BP}[\hat{q}; \Lambda]$ we have*

$$\ddot{\mathrm{ERR}}(\dot{p}_e, \dot{q}_e) \leq k^{O(1)} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2^{-k\zeta} \\ 2^{-k\zeta} \end{pmatrix} \mathrm{err}_v + k^{O(1)} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \min\{\ddot{\epsilon}_e, 1\}/2^{k\zeta} \end{pmatrix},$$

for all $e \in \delta v$ (using the notation of Definition 7.10), and with err_v as defined by (7.27).

Lemma 7.15. *Let v be a nice variable with incoming messages $\hat{q}$ (in the pair model) satisfying condition (7.25). Let v be the resulting probability measure on the frozen spin $x_v \in \{-,+,\mathtt{f}\}^2$. Then v is close to a product measure:*

$$v(x) = \left\{1 + O\left(\frac{1}{2^{k\zeta}}\right)\right\} \cdot \begin{cases} 1/4 & \text{for } x \in \{-,+\}^2 \,, \\ 1/(4 \cdot 2^k) & \text{for } x \in \{\mathtt{-f},\mathtt{+f},\mathtt{f-},\mathtt{f+}\} \,, \\ 1/(4 \cdot 4^k) & \text{for } x = \mathtt{ff} \end{cases} \tag{7.28}$$

for an absolute constant $0 < \zeta \le 1/20$.

Proof. This is a fairly direct calculation. Recall that v is a nice variable, so by Remark 7.7, condition (7.25) also holds with $_{\star\star}\hat{q}_e \equiv {}_\star\hat{q}_e \otimes {}_\star\hat{q}_e$ in place of $_*\hat{q}_e \equiv {}_\bullet\hat{q}_e \otimes {}_\bullet\hat{q}_e$. The degrees $|\delta v(\pm)|$ are constrained by Definition 3.8, as are the canonical messages $_\star\hat{q}$; recall moreover that $_\star\hat{q}_e(\mathtt{y}) = {}_\star\hat{q}_e(\mathtt{g}) = {}_\star\hat{q}_e(\mathtt{b})$. We will use these facts repeatedly in what follows. Recall moreover from (3.5) that we defined $\mathtt{p} \equiv \texttt{purple} \equiv \{\mathtt{r},\mathtt{b}\}$.

Step 1. Estimates in single-copy model. Consider the variable v with incoming messages $_\star\hat{q}$, and let π denote the resulting law of the frozen spin $x_v \in \{-,+,\mathtt{f}\}$. Then, for a normalizing constant $z_\star$, we have

$$\pi(\mathtt{f}) = \frac{1}{z_\star}\left\{\prod_{e \in \delta v(+)} {}_\star\hat{q}_e(\mathtt{g})\right\}\left\{\prod_{e \in \delta v(-)} {}_\star\hat{q}_e(\mathtt{g})\right\} = \frac{1}{z_\star}\left\{\prod_{e \in \delta v(+)} {}_\star\hat{q}_e(\mathtt{b})\right\}\left\{\prod_{e \in \delta v(-)} {}_\star\hat{q}_e(\mathtt{b})\right\}$$

Then, using the conditions from Definition 3.8, we have

$$\frac{\pi(+)}{\pi(\mathtt{f})} = \left\{1 - O\left(\frac{1}{2^k}\right)\right\}\left\{\prod_{e \in \delta v(+)} \frac{{}_\star\hat{q}_e(\mathtt{p})}{{}_\star\hat{q}_e(\mathtt{g})}\right\}\left\{\prod_{e \in \delta v(-)} \frac{{}_\star\hat{q}_e(\mathtt{y})}{{}_\star\hat{q}_e(\mathtt{g})}\right\}$$

$$= \left\{1 - O\left(\frac{1}{2^k}\right)\right\}\prod_{e \in \delta v(+)}\left(1 + \frac{1 + O(2^{-k/10})}{2^{|\partial a(e)|-1}}\right) = 2^k\left\{1 + O\left(\frac{k}{2^{k/10}}\right)\right\},$$

where $a(e)$ refers to the clause incident to edge e. The same estimate holds for $\pi(-)$. It follows that

$$\pi(x) = \left\{1 + O\left(\frac{k}{2^{k/10}}\right)\right\} \times \begin{cases} 1/2 & \text{for } x \in \{-,+\}\,, \\ 1/2^{k+1} & \text{for } x = \mathtt{f}\,. \end{cases}$$

Step 2. Estimates in pair model. Let $v_\star \equiv \pi \otimes \pi$. Let $z, z_\star$ be the normalizing constant such that

$$v(\mathtt{ff}) = \frac{1}{z}\prod_{e \in \delta v} \hat{q}_e(\mathtt{gg})\,, \qquad v_\star(\mathtt{ff}) = \frac{1}{z_\star}\prod_{e \in \delta v} {}_{\star\star}\hat{q}_e(\mathtt{gg})\,.$$

Taking the ratio between the two and using the assumption (7.25) gives

$$\frac{z v(\mathtt{ff})}{z_\star v_\star(\mathtt{ff})} = \prod_{e \in \delta v}\left(1 + \frac{k^{O(1)}}{2^{k(1+\zeta)}}\right) = 1 + O\left(\frac{k^{O(1)}}{2^{k\zeta}}\right).$$

Similarly, we use (7.25) again to calculate that

$$z v(\mathtt{+f}) = \left\{\prod_{e \in \delta v(+)} \hat{q}_e(\mathtt{pg})\right\}\left\{\prod_{e \in \delta v(-)} \hat{q}_e(\mathtt{yg})\right\}\left\{1 - \prod_{e \in \delta v(+)}\left(1 - \frac{\hat{q}_e(\mathtt{rg})}{\hat{q}_e(\mathtt{pg})}\right)\right\}$$

$$= \left\{\prod_{e \in \delta v(+)} \hat{q}_e(\mathtt{pg})\right\}\left\{\prod_{e \in \delta v(-)} \hat{q}_e(\mathtt{yg})\right\}\left\{1 - \prod_{e \in \delta v(+)}\left(1 - \frac{1 + O(k^{O(1)}/2^{k\zeta})}{2^{|\partial a(e)|-1}}\right)\right\}$$

$$= \left\{\prod_{e \in \delta v(+)} \hat{q}_e(\mathtt{pg})\right\}\left\{\prod_{e \in \delta v(-)} \hat{q}_e(\mathtt{yg})\right\}\left\{1 + O\left(\frac{1}{2^k}\right)\right\} = z_\star v_\star(\mathtt{+f})\left\{1 + O\left(\frac{k^{O(1)}}{2^{k\zeta}}\right)\right\}.$$

The same estimate holds for the spins in $\{\texttt{f+},\texttt{-f},\texttt{f-}\}$. A similar calculation gives

$$z\nu(\texttt{++}) = \left\{\prod_{e\in\delta v(\texttt{+})}\hat{q}_e(\texttt{pp})\right\}\left\{\prod_{e\in\delta v(\texttt{-})}\hat{q}_e(\texttt{yy})\right\}\left\{1 - \prod_{e\in\delta v(\texttt{+})}\left(1 - \frac{\hat{q}_e(\texttt{rp})}{\hat{q}_e(\texttt{pp})}\right) - \prod_{e\in\delta v(\texttt{+})}\left(1 - \frac{\hat{q}_e(\texttt{pr})}{\hat{q}_e(\texttt{pp})}\right) + \prod_{e\in\delta v(\texttt{+})}\frac{\hat{q}_e(\texttt{bb})}{\hat{q}_e(\texttt{pp})}\right\}$$

$$= \left\{\prod_{e\in\delta v(\texttt{+})}\hat{q}_e(\texttt{pp})\right\}\left\{\prod_{e\in\delta v(\texttt{-})}\hat{q}_e(\texttt{yy})\right\}\left\{1 + O\left(\frac{k^{O(1)}}{2^{k\zeta}}\right)\right\} = z_\star\nu_\star(\texttt{++})\left\{1 + O\left(\frac{k^{O(1)}}{2^{k\zeta}}\right)\right\},$$

and the same estimate holds for the other spins in $\{\texttt{+-},\texttt{-+},\texttt{--}\}$. Therefore

$$\nu(x) = \pi(x^1)\pi(x^2)\left\{1 + O\left(\frac{k^{O(1)}}{2^{k\zeta}}\right)\right\}, \tag{7.29}$$

and the claim follows. $\qquad\square$

Lemma 7.16. *Let v be a nice variable with incoming messages $\hat{q}$ (in the pair model) satisfying (7.25). Consider two weight functions $\lambda = \lambda^1 \otimes \lambda^2$ and $\theta = \theta^1 \otimes \theta^2$ where $\lambda^j, \theta^j : \{\texttt{-},\texttt{+},\texttt{f}\} \to (0,\infty)$ with*

$$\sum_{j=1,2}\sum_{x\in\{\texttt{-},\texttt{+},\texttt{f}\}}\left(\left|\lambda^j(x) - 1\right| + \left|\theta^j(x) - 1\right|\right) \le \frac{k^{O(1)}}{2^{k\zeta}}. \tag{7.30}$$

Let v be the law of the frozen spin $x_v \in \{\texttt{-},\texttt{+},\texttt{f}\}^2$ that results from the messages $\hat{q}$ and the weights λ, and let μ be the law of x_v that results from $\hat{q}$ and θ. Then for the marginal on the first copy we have

$$v^1(x^1) = \left\{1 + O\left(\frac{k^{O(1)}\|\delta^2\|_\infty}{2^{k\zeta}}\right)\right\}\left(\frac{\lambda^1\mu^1}{\theta^1}\right)(x^1)\bigg/\left(\sum_{y^1}\left(\frac{\lambda^1\mu^1}{\theta^1}\right)(y^1)\right), \qquad \delta^j \equiv \frac{\lambda^j}{\theta^j} - 1.$$

That is to say, v^1 is approximately equal to the reweighting of μ^1 by λ^1/θ^1, and does not depend much on λ^2/θ^2.

Proof. Let $\bar{\pi}$ be the law of the frozen spin $x = x_v \in \{\texttt{-},\texttt{+},\texttt{f}\}^2$ that results from messages $\hat{q}$ and weights $\lambda^1 \otimes \theta^2$. Then

$$\bar{\pi}(x) \cong \frac{\lambda^1(x^1)}{\theta^1(x^1)}\mu(x),$$

from which it follows that the marginal of $\bar{\pi}$ on the first copy is exactly

$$\bar{\pi}^1(x^1) = \left(\frac{\lambda^1\mu^1}{\theta^1}\right)(x^1)\bigg/\left(\sum_{y^1}\left(\frac{\lambda^1\mu^1}{\theta^1}\right)(y^1)\right).$$

It remains to compare v^1 with $\bar{\pi}^1$. To this end, note it follows from Lemma 7.15 together with the assumption (7.30) that $\bar{\pi}$ also satisfies the estimates (7.28) or equivalently (7.29). We thus have $e : \{\texttt{-},\texttt{+},\texttt{f}\}^2 \to \mathbb{R}$ such that

$$\bar{\pi}(x) = \pi(x^1)\pi(x^2)\left(1 + e(x)\right), \qquad |e(x)| \le \frac{k^{O(1)}}{2^{k\zeta}}. \tag{7.31}$$

Recall that we defined $\delta^j \equiv \lambda^j/\theta^j - 1$. Then, with z a normalizing constant, we have

$$v(x) = \frac{1}{z}\frac{\lambda^2(x^2)}{\theta^2(x^2)}\bar{\pi}(x) = \frac{1}{z}\left(1 + \delta^2(x^2)\right)\bar{\pi}(x).$$

We can use (7.31) to estimate the normalizing constant as

$$z = \sum_x\left(1 + \delta^2(x^2)\right)\bar{\pi}(x) = 1 + \sum_x\delta^2(x^2)\bar{\pi}(x) = 1 + \sum_{x^1}\pi(x^1)\sum_{x^2}\pi(x^2)\delta^2(x^2)\left(1 + e(x)\right)$$

$$= \left\{1 + \sum_{x^2}\pi(x^2)\delta^2(x^2)\right\}\left\{1 + O\left(\frac{k^{O(1)}\|\delta^2\|_\infty}{2^{k\zeta}}\right)\right\}.$$

Similarly, we can also use (7.31) to estimate

$$\sum_{x^2} \left(1 + \delta^2(x^2)\right)\bar{\pi}(x) = \bar{\pi}^1(x^1) + \sum_{x^2} \delta^2(x^2)\bar{\pi}(x) = \bar{\pi}^1(x^1) + \pi(x^1)\sum_{x^2} \delta^2(x^2)\pi(x^2)\left(1 + e(x)\right)$$

$$= \bar{\pi}^1(x^1) + \pi(x^1)\sum_{x^2} \pi(x^2)\delta^2(x^2) + O\left(\frac{k^{O(1)}\|\delta^2\|_\infty}{2^{k\zeta}}\right)$$

$$= \bar{\pi}^1(x^1)\left\{1 + \sum_{x^2} \pi(x^2)\delta^2(x^2)\right\}\left\{1 + O\left(\frac{k^{O(1)}\|\delta^2\|_\infty}{2^{k\zeta}}\right)\right\}.$$

Taking the ratio between the last two quantities gives

$$\nu^1(x^1) = \bar{\pi}^1(x^1)\left\{1 + O\left(\frac{k^{O(1)}\|\delta^2\|_\infty}{2^{k\zeta}}\right)\right\}.$$

This proves the claim. □

Lemma 7.17. *Let v be a nice variable. Suppose we are given two sets of weights for v, Θ and Λ, both of the functional form (7.14) from Definition 7.8, satisfying the bound (7.30) from Lemma 7.16, and additionally such that*

$$\max_{e \in \delta v}\left\{\sum_{j=1,2}\left(\left|(\theta_e)^j - 1\right| + \left|(\lambda_e)^j - 1\right|\right)\right\} \le \frac{k^{O(1)}}{2^{k\zeta}}. \tag{7.32}$$

Assume that the weights on the frozen spin agree, i.e., $\lambda^j(x) = \theta^j(x)$ for $j = 1, 2$ and all $x \in \{\text{-}, \text{+}, \text{f}\}$, so that only the edge weights can differ between Θ and Λ. Denote the error in the edge weights by

$$(\rho_e)^j \equiv \frac{(\lambda_e)^j}{(\theta_e)^j} - 1 \equiv \frac{(\lambda_e)^j(\mathtt{r})}{(\theta_e)^j(\mathtt{r})} - 1.$$

Let v have incoming messages $\hat{q}$ (in the pair model) satisfying (7.25). Let $\mu = \nu_{\delta v}[\Theta; \hat{q}]$ and $\nu = \nu_{\delta v}[\Lambda; \hat{q}]$, with edge marginals μ_e and ν_e for $e \in \delta v$. Then it holds for all $e \in \delta v$ that

$$\frac{(\nu_e)^1(\mathtt{r})}{(\nu_e)^1(\mathtt{b})} = \frac{(\lambda_e)^1}{(\theta_e)^1} \cdot \frac{(\mu_e)^1(\mathtt{r})}{(\mu_e)^1(\mathtt{b})}\left\{1 + O\left(k^{O(1)}\left[\frac{(\rho_e)^2}{2^{k\zeta}} + \sum_{j=1,2}\sum_{e' \in \delta v \setminus e}\frac{|(\rho_{e'})^j|}{2^{k(1+\zeta)}}\right]\right)\right\}.$$

This says that the marginal red-to-blue *ratio in the first copy is changed by a factor of approximately $(\lambda_e)^1/(\theta_e)^1$ — i.e., for the first-copy marginal on e, changing the edge weights in both copies on all of δv has approximately the same effect as changing the edge weight in the first copy only, on e alone. The analogous bound holds exchanging the copy indices $j = 1, 2$.*

Proof. For $e \in \delta v$ let g_e, h_e be the non-normalized measures on $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$ defined by

$$g_e(\sigma) \equiv \hat{q}_e(\sigma)\theta_e(\sigma) = \hat{q}_e(\sigma)\left\{\prod_{j=1,2}((\theta_e)^j)^{\mathbf{1}\{\sigma^1=\mathtt{r}\}}\right\},$$

$$h_e(\sigma) \equiv \hat{q}_e(\sigma)\lambda_e(\sigma) = g_e(\sigma)\left\{\prod_{j=1,2}\left(1 + ((\rho_e)^j)^{\mathbf{1}\{\sigma^j=\mathtt{r}\}}\right)\right\}.$$

Similarly as in (2.22), let $\delta v(\pm e) \equiv \{e' \in \delta v \setminus e : \mathtt{L}_e = \pm\mathtt{L}_{e'}\}$. Then, for some normalizing constants $z_e, \bar{z}_e$ we have

$$\frac{\bar{z}_e\mu_e(\mathtt{rr})}{\theta^1(\mathtt{L}_e)\theta^2(\mathtt{L}_e)} = (\theta_e)^1(\theta_e)^2\hat{q}_e(\mathtt{rr})\left\{\prod_{e' \in \delta v(+e)} g_{e'}(\mathtt{pp})\right\}\left\{\prod_{e' \in \delta v(-e)} g_{e'}(\mathtt{yy})\right\},$$

$$\frac{z_e\nu_e(\mathtt{rr})}{\theta^1(\mathtt{L}_e)\theta^2(\mathtt{L}_e)} = (\lambda_e)^1(\lambda_e)^2\hat{q}_e(\mathtt{rr})\left\{\prod_{e' \in \delta v(+e)} h_{e'}(\mathtt{pp})\right\}\left\{\prod_{e' \in \delta v(-e)} h_{e'}(\mathtt{yy})\right\}.$$

Note that $g_{e'}(\mathtt{yy}) = \hat{q}_{e'}(\mathtt{yy}) = h_{e'}(\mathtt{yy})$ for all $e' \in \delta v$. On the other hand, we can estimate

$$\prod_{e' \in \delta v(+e)} \frac{h_{e'}(\mathtt{pp})}{g_{e'}(\mathtt{pp})} = \prod_{e' \in \delta v(+e)} \left\{ 1 + \frac{(\rho_{e'})^1 g_{e'}(\mathtt{rp}) + (\rho_{e'})^2 g_{e'}(\mathtt{pr}) + (\rho_{e'})^1 (\rho_{e'})^2 g_{e'}(\mathtt{rr})}{g_{e'}(\mathtt{pp})} \right\}$$

$$\overset{(7.32)}{=} \prod_{e' \in \delta v(+e)} \left\{ 1 + \left(1 + \frac{k^{O(1)}}{2^{k\zeta}} \right) \frac{(\rho_{e'})^1 \hat{q}_{e'}(\mathtt{rp}) + (\rho_{e'})^2 \hat{q}_{e'}(\mathtt{pr})}{\hat{q}_{e'}(\mathtt{pp})} + O\left(\frac{(\rho_{e'})^1 (\rho_{e'})^2}{2^{k(1+\zeta)}} \right) \right\}$$

$$\overset{(7.25)}{=} 1 + \sum_{j=1,2} \underbrace{\left(\sum_{e' \in \delta v(+e)} \frac{(\rho_{e'})^j {}_\star \hat{q}_{e'}(\mathtt{r})}{{}_\star \hat{q}_{e'}(\mathtt{b})} \right)}_{\text{denote this } A^j(+)} + O\underbrace{\left(\sum_{j=1,2} \sum_{e' \in \delta v \backslash e} \frac{k^{O(1)} |(\rho_{e'})^j|}{2^{k(1+\zeta)}} \right)}_{\text{denote this } R_0}$$

(having implicitly used Remark 7.7 together with (7.25)). It follows that

$$\frac{z_e \nu_e(\mathtt{rr})}{\bar{z}_e \mu_e(\mathtt{rr})} \left\{ \prod_{j=1,2} \frac{(\theta_e)^j}{(\lambda_e)^j} \right\} = 1 + A^1(+) + A^2(+) + O(R_0) .$$

By very similar calculations we obtain, for all $\sigma \in \{\mathtt{r}, \mathtt{b}\}^2$,

$$\frac{z_e \nu_e(\sigma)}{\bar{z}_e \mu_e(\sigma)} \left\{ \prod_{j=1,2} \left(\frac{(\theta_e)^j}{(\lambda_e)^j} \right)^{\mathbf{1}\{\sigma^j = \mathtt{r}\}} \right\} = 1 + A^1(+) + A^2(+) + O(R_0) .$$

Let $A^j(-)$ be define analogously to $A^j(+)$, except that we sum over $e' \in \delta v(-e)$ rather than $e' \in \delta v(+e)$. Then similar calculations as above give the estimates

$$\frac{z_e \nu_e(\mathtt{ry})}{\bar{z}_e \mu_e(\mathtt{ry})} \frac{(\theta_e)^1}{(\lambda_e)^1} = 1 + A^1(+) + A^2(-) + O(R_0) = \frac{z_e \nu_e(\mathtt{by})}{\bar{z}_e \mu_e(\mathtt{by})} ,$$

$$\frac{z_e \nu_e(\mathtt{rg})}{\bar{z}_e \mu_e(\mathtt{rg})} \frac{(\theta_e)^1}{(\lambda_e)^1} = 1 + A^1(+) + O(R_0) = \frac{z_e \nu_e(\mathtt{bg})}{\bar{z}_e \mu_e(\mathtt{bg})} .$$

Combining these estimates gives

$$\frac{(\nu_e)^1(\mathtt{r})}{(\nu_e)^1(\mathtt{b})} = \frac{(\lambda_e)^1}{(\theta_e)^1} \frac{(\mu_e)^1(\mathtt{r})[1 + A^1(+)] + \mu_e(\mathtt{rp})A^2(+) + \mu_e(\mathtt{ry})A^2(-) + O(\mu_e(\mathtt{rr})(\rho_e)^2)}{(\mu_e)^1(\mathtt{b})[1 + A^1(+)] + \mu_e(\mathtt{bp})A^2(+) + \mu_e(\mathtt{by})A^2(-) + O(\mu_e(\mathtt{br})(\rho_e)^2)} \left\{ 1 + O(R_0) \right\}$$

$$= \frac{(\lambda_e)^1}{(\theta_e)^1} \frac{(\mu_e)^1(\mathtt{r})}{(\mu_e)^1(\mathtt{b})} \frac{1 + \dfrac{\mu_e(\mathtt{rp})A^2(+)}{(\mu_e)^1(\mathtt{r})[1 + A^1(+)]} + \dfrac{\mu_e(\mathtt{ry})A^2(-)}{(\mu_e)^1(\mathtt{r})[1 + A^1(+)]} + O\left(\dfrac{k^{O(1)}(\rho_e)^2}{2^{k\zeta}} \right)}{1 + \dfrac{\mu_e(\mathtt{bp})A^2(+)}{(\mu_e)^1(\mathtt{b})[1 + A^1(+)]} + \dfrac{\mu_e(\mathtt{by})A^2(-)}{(\mu_e)^1(\mathtt{b})[1 + A^1(+)]} + O\left(\dfrac{(\rho_e)^2}{2^k} \right)} \left\{ 1 + O(R_0) \right\}$$

$$= \frac{(\lambda_e)^1}{(\theta_e)^1} \frac{(\mu_e)^1(\mathtt{r})}{(\mu_e)^1(\mathtt{b})} \left\{ 1 + O\left(\frac{k^{O(1)}(\rho_e)^2}{2^{k\zeta}} + R_0 \right) \right\} .$$

In the last step we used the assumptions (7.25), (7.30), and (7.32), which together guarantee that

$$\frac{\mu_e(\mathtt{rp})}{(\mu_e)^1(\mathtt{r})} = {}_\star \pi_e(\mathtt{p}) \left\{ 1 + O\left(\frac{k^{O(1)}}{2^{k\zeta}} \right) \right\} = \frac{\mu_e(\mathtt{bp})}{(\mu_e)^1(\mathtt{b})} ,$$

$$\frac{\mu_e(\mathtt{ry})}{(\mu_e)^1(\mathtt{r})} = {}_\star \pi_e(\mathtt{y}) \left\{ 1 + O\left(\frac{k^{O(1)}}{2^{k\zeta}} \right) \right\} = \frac{\mu_e(\mathtt{by})}{(\mu_e)^1(\mathtt{b})} .$$

The claim follows. $\qquad \square$

Proof of Proposition 7.13. We first define a sequence of weights Λ_t, then analyze the construction to show that they converge to the desired weights $\Lambda_\infty = \Lambda$. We emphasize that this index t is **purely local to the proof of this**

proposition, and is not the same as the t that indexes the up-and-down passes in Definition 7.9. The proof below is divided into a few numbered parts.

Part 1. Iterative definition of weights. Initialize $\Lambda_0 = \Theta$. For all $t \geq 0$, given the weights Λ_t, let $\nu_t \equiv \nu_{\delta v}[\Lambda_t; \hat{q}]$. Then, for all $e \in \delta v$, update the edge weights by setting, for both $j = 1, 2$,

$$(\lambda_{e,t+1})^j = (\lambda_{e,t})^j \cdot \frac{\star \pi_e(\mathtt{r})/\star \pi_e(\mathtt{b})}{(\nu_{e,t})^j(\mathtt{r})/(\nu_{e,t})^j(\mathtt{b})} . \tag{7.33}$$

Let $\nu_{t+1/2}$ be the measure that results from incoming messages $\hat{q}$, frozen spin weights λ_t, and edge weights $\lambda_{e,t+1}$ for all $e \in \delta v$. Let $\pi_{t+1/2}$ be the induced measure on the frozen spin $x \in \{-,+,\mathtt{f}\}^2$. Let π be the measure on the single spin $x^j \in \{-,+,\mathtt{f}\}$ that is induced by $\star \pi$. Then update the frozen spin weights by setting, for both $j = 1, 2$,

$$(\lambda_{t+1})^j(x^j) = (\lambda_t)^j(x^j) \cdot \frac{\pi(x^j)/\pi(+)}{(\pi_{t+1/2})^j(x^j)/(\pi_{t+1/2})^j(+)} . \tag{7.34}$$

This concludes the definition of the weights Λ_{t+1}.

Part 2. Definition of error quantities. Define the error quantities

$$(\delta_t)^j \equiv \sum_{x \in \{-,\mathtt{f}\}} \left| \frac{\pi(x^j)/\pi(+)}{(\pi_t)^j(x^j)/(\pi_t)^j(+)} - 1 \right|, \quad (\rho_{e,t})^j \equiv \left| \frac{\star \pi_e(\mathtt{r})/\star \pi_e(\mathtt{b})}{(\nu_{e,t})^j(\mathtt{r})/(\nu_{e,t})^j(\mathtt{b})} - 1 \right|.$$

We will keep track of the aggregate errors

$$\delta_t \equiv \sum_{j=1,2} (\delta_t)^j, \quad \rho_{e,t} \equiv \sum_{j=1,2} (\rho_{e,t})^j, \quad \rho_t \equiv \sum_{e \in \delta v} \rho_{e,t} .$$

For the initial measure $\nu_0 = \nu_{\delta v}[\Lambda_0; \hat{q}] = \nu_{\delta v}[\Theta; \hat{q}]$, it follows by straightforward calculations that

$$(\delta_0)^j \leq O\left(\sum_{e' \in \delta v} \left\{ \epsilon_{e'} + \frac{\dot{\epsilon}_{e'}}{2^k} + \frac{\min\{\ddot{\epsilon}_{e'}, 1\}}{2^{k(1+\zeta)}} \right\} \right) = O(\mathrm{err}_v) ,$$

$$(\rho_{e,0})^j \leq O\left(\dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} + \mathrm{err}_v \right)$$

for both $j = 1, 2$ — this makes use of the assumptions (7.25) and (7.26).

Part 3. Effect of update on edge weights (from t to $t + 1/2$). Recall that for each integer $t \geq 0$, the procedure described in Part 1 goes from t to $t + 1/2$ by the update (7.33) on all the edge weights, leaving the frozen spin weights unchanged. For this update, Lemma 7.17 gives

$$(\rho_{e,t+1/2})^1 \leq \frac{k^{O(1)}}{2^{k\zeta}} \left\{ (\rho_{e,t})^2 + \frac{\rho_t}{2^k} \right\}, \tag{7.35}$$

with the analogous bound if we exchange $j = 1, 2$. Summing (7.35) over $j = 1, 2$ and over all $e \in \delta v$ gives

$$\rho_{t+1/2} \leq \frac{k^{O(1)}}{2^{k\zeta}} \rho_t . \tag{7.36}$$

Meanwhile, the error in the frozen spin marginals can change by at most

$$\delta_{t+1/2} \leq \delta_t + \frac{k^{O(1)}}{2^{k(1+\zeta)}} \rho_t , \tag{7.37}$$

by a very similar calculation.

Part 4. Effect of update on frozen spin weights (from $t + 1/2$ to $t + 1$. Next recall that for each integer $t \geq 0$, the procedure of Part 1 goes from $t + 1/2$ to $t + 1$ by the update (7.34) on the frozen spin weights, leaving the edge weights unchanged. As a result the red-to-blue ratios are unaffected, so $(\rho_{e,t+1})^j = (\rho_{e,t})^j$. As for the frozen spin marginals, the result of Lemma 7.16 gives

$$\delta_{t+1} \leq \frac{k^{O(1)} \delta_{t+1/2}}{2^{k\zeta}} \leq \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \delta_t + \frac{\rho_t}{2^{k(1+\zeta)}} \right\} \tag{7.38}$$

where the last step is by (7.37).

Part 5. Conclusion. For each integer $t \geq 0$, when we go from t to $t + 1$ in the procedure of Part 1, it follows by combining (7.36) and (7.38) that

$$\begin{pmatrix} \boldsymbol{\delta}_{t+1} \\ \boldsymbol{\rho}_{t+1} \end{pmatrix} \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 1 & 2^{-k(1+\zeta)} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta}_t \\ \boldsymbol{\rho}_t \end{pmatrix}$$

(having used also that $\boldsymbol{\rho}_{t+1/2} = \boldsymbol{\rho}_t$). This implies

$$\Upsilon_{t+1} \equiv \boldsymbol{\delta}_{t+1} + \frac{\boldsymbol{\rho}_{t+1}}{2^k} \leq \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \boldsymbol{\delta}_t + \frac{\boldsymbol{\rho}_t}{2^k} \right\} = \frac{k^{O(1)}}{2^{k\zeta}} \Upsilon_t \, .$$

It follows that the iteration defined in Part 1 converges. The error on the frozen spin weights can be bounded as

$$\sum_{j=1,2} \sum_{x \in \{-,\mathtt{f}\}} \left| \frac{\lambda^j(x)}{\theta^j(x)} - 1 \right| \leq O\left(\sum_{t=0}^{\infty} \sum_{j=1,2} \sum_{x \in \{-,\mathtt{f}\}} \left| \frac{(\lambda_{t+1})^j(x)}{(\lambda_t)^j(x)} - 1 \right| \right)$$

$$\leq O\left(\sum_{t=0}^{\infty} \boldsymbol{\delta}_t \right) \leq O\left(\sum_{t=0}^{\infty} \Upsilon_t \right) \leq O(\Upsilon_0) \leq O(\mathrm{err}_v) \, .$$

To bound the error on the edge weights, let us first note that for any $e \in \delta v$, (7.35) implies

$$\sum_{t \geq 1} \rho_{e,t} \leq \frac{k^{O(1)}}{2^{k\zeta}} \sum_{t \geq 0} \left\{ \rho_{e,t} + \frac{\boldsymbol{\rho}_t}{2^k} \right\} \leq \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \sum_{t \geq 1} \rho_{e,t} + \left(\rho_{e,0} + \sum_{t \geq 0} \Upsilon_t \right) \right\} \, ,$$

and rearranging the inequality gives

$$\sum_{t \geq 1} \rho_{e,t} \leq \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \rho_{e,0} + \sum_{t \geq 0} \Upsilon_t \right\} \leq \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \rho_{e,0} + \mathrm{err}_v \right\} \, .$$

It follows from this that the total error on the edge weights can be bounded as

$$\sum_{j=1,2} \left| \frac{(\lambda_e)^j}{(\theta_e)^j} - 1 \right| \leq O\left(\sum_{t=0}^{\infty} \rho_{e,t} \right) \leq O\left(\rho_{e,0} + \frac{k^{O(1)}}{2^{k\zeta}} \mathrm{err}_v \right) \leq O\left(\dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} + \mathrm{err}_v \right) \, .$$

This concludes the proof. $\qquad \square$

Next, recall that for a variable-to-clause message $\dot{q}_{va}$ on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$, we defined a reweighted version Q_{va} by (7.10). Now, for a clause-to-variable message $h \equiv h_{av}$ on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$ we define the compensatory reweighting

$$H(\sigma) = \frac{W(\sigma)}{Z_h} = \frac{h(\sigma)(2^{|\partial a| - 1})^{\mathtt{r}[\sigma]}}{Z_h} \, , \tag{7.39}$$

where Z_h is the normalizing constant that makes H a probability measure over $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$, and we write W for the non-normalized version of H. The following lemma records an elementary bound which will be used in the proof of Corollary 7.14 at the end of this subsection:

Lemma 7.18. *Suppose at the edge $e = (av)$ we have variable-to-clause messages p and q (both going from v to a), and a clause-to-variable message h (going from a to v), all probability measures on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$. Let P, Q be the reweightings of p, q defined by (7.10). Let H be the reweighting of h, with non-normalized version W, as defined by (7.39). Assume*

$$\left\| W - \frac{1}{9} \right\|_{\infty} \leq \frac{1}{18} \, . \tag{7.40}$$

Define the corresponding edge marginals (probability measures on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$)

$$\pi_p(\sigma) \equiv \frac{P(\sigma)W(\sigma)}{z_p} \, , \quad \pi_q(\sigma) \equiv \frac{Q(\sigma)W(\sigma)}{z_q} \, ,$$

where z_p, z_q are the normalizing constants. If π_p and π_q have the same marginals on the first copy, $(\pi_p)^1 = (\pi_q)^1$, then the first-copy marginals of P and Q satisfy

$$\overbrace{\left|\frac{Q^1(\sigma^1)}{P^1(\sigma^1)} - 1\right| \le O(1)\left\|W - \frac{1}{9}\right\|_\infty \sum_\tau P(\tau)\left(1 + \frac{\mathbf{1}\{\tau^1 = \sigma^1\}}{P^1(\sigma^1)}\right)\left|\frac{Q(\sigma)}{P(\sigma)} - 1\right|}^{\text{this tighter bound is used in the proof of Lemma 7.21}} \le O(1)\left\|W - \frac{1}{9}\right\|_\infty\left\|\frac{P}{Q} - 1\right\|_\infty.$$

The analogous statement holds if we instead have $(\pi_p)^2 = (\pi_q)^2$.

Proof. For the purposes of the proof, for $\sigma \in \{\mathbf{r}, \mathbf{y}, \mathbf{c}\}^2$ denote

$$\delta(\sigma) \equiv \frac{Q(\sigma)}{P(\sigma)} - 1, \quad \beta(\sigma) \equiv W(\sigma) - \frac{1}{9}.$$

Note that, since P is a probability measure over $\{\mathbf{r}, \mathbf{y}, \mathbf{c}\}^2$, we have

$$z_p = \sum_\sigma P(\sigma)\left(W(\sigma) - \frac{1}{9}\right) + \frac{1}{9}\sum_\sigma P(\sigma) = \sum_\sigma P(\sigma)\beta(\sigma) + \frac{1}{9} \in \left[\frac{1}{18}, \frac{3}{18}\right],$$

where the last step uses the assumption (7.40). The same bound holds for z_q. Next we have

$$\frac{Q^1(\sigma^1)}{P^1(\sigma^1)} - 1 = \frac{z_q - z_p}{z_p} + \frac{z_q}{P^1(\sigma^1)}\left(\frac{Q^1(\sigma^1)}{z_q} - \frac{P^1(\sigma^1)}{z_p}\right);$$

we will bound separately the two terms on the right-hand side. For the first term we have

$$|z_q - z_p| = \left|\sum_\sigma P(\sigma)\left(\frac{Q(\sigma)}{P(\sigma)} - 1\right)\left(W(\sigma) - \frac{1}{9}\right)\right| = \left|\sum_\sigma P(\sigma)\delta(\sigma)\beta(\sigma)\right| \le \|\beta\|_\infty \sum_\sigma P(\sigma)|\delta(\sigma)|, \qquad (7.41)$$

which implies $|z_q - z_p| \le \|\beta\|_\infty\|\delta\|_\infty$. For the second term, direct algebraic manipulations give

$$\frac{1}{9}\left(\frac{Q^1(\sigma^1)}{z_q} - \frac{P^1(\sigma^1)}{z_p}\right) = \sum_{\sigma \in \{\sigma^1\} \times \{\mathbf{r}, \mathbf{y}, \mathbf{c}\}}\left\{\frac{1}{z_q}Q(\sigma)\left(W(\sigma) - \beta(\sigma)\right) - \frac{1}{z_p}P(\sigma)\left(W(\sigma) - \beta(\sigma)\right)\right\}$$

$$= \sum_{\sigma \in \{\sigma^1\} \times \{\mathbf{r}, \mathbf{y}, \mathbf{c}\}}\beta(\sigma)\left\{\frac{P(\sigma)}{z_p} - \frac{Q(\sigma)}{z_q}\right\} + \underbrace{(\pi_q)^1(\sigma^1) - (\pi_q)^1(\sigma^1)}_{\text{zero}}$$

$$= \sum_{\sigma \in \{\sigma^1\} \times \{\mathbf{r}, \mathbf{y}, \mathbf{c}\}}\beta(\sigma)P(\sigma)\left\{\frac{z_q - z_p}{z_p z_q} - \frac{1}{z_q}\left(\frac{Q(\sigma)}{P(\sigma)} - 1\right)\right\},$$

from which it follows (dividing through by $P^1(\sigma^1)$) that

$$\frac{1}{P^1(\sigma^1)}\left|\frac{Q^1(\sigma^1)}{z_q} - \frac{P^1(\sigma^1)}{z_p}\right| \le \|\beta\|_\infty \cdot O\left(|z_q - z_p| + \sum_\tau \frac{\mathbf{1}\{\tau^1 = \sigma^1\}}{P^1(\sigma^1)}P(\tau)\delta(\tau)\right),$$

having used that $z_p, z_q \in [1/18, 3/18]$. Combining with (7.41) proves the lemma. □

Proof of Corollary 7.14. Let P, Q be the reweightings of $\dot{p}, \dot{q}$ defined by (7.10).

Step 1. Non-marginal errors between P and Q. We claim that Proposition 7.13 implies, for each $e \in \delta v$,

$$\left|\frac{Q_e(\sigma)}{P_e(\sigma)} - 1\right| \le k^{O(1)}\left(\mathrm{err}_v + \mathbf{1}\{\mathbf{r}[\sigma] \ge 1\}\left\{\dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}}\right\}\right) \qquad (7.42)$$

We will only briefly sketch the proof of (7.42): let $g_e(\sigma) = \theta_e(\sigma)\hat{q}_e(\sigma)$ and $h_e(\sigma) = \lambda_e(\sigma)\hat{p}_e(\sigma)$, and let $Z_{Q,e}$ and $Z_{P,e}$ be the normalizing constants such that

$$\frac{Z_{Q,e}Q_e(\mathbf{rr})}{Z_{P,e}P_e(\mathbf{rr})} = \left\{\prod_{j=1,2}\frac{\lambda^j(\mathbf{L}_e) \cdot (\lambda_e)^j}{\theta^j(\mathbf{L}_e) \cdot (\theta_e)^j}\right\}\prod_{e' \in \delta v(+e)}\frac{h_{e'}(\mathbf{pp})}{g_{e'}(\mathbf{pp})}\prod_{e' \in \delta v(-e)}\frac{h_{e'}(\mathbf{yy})}{g_{e'}(\mathbf{yy})}.$$

By the assumptions of Proposition 7.13 together with the resulting bounds on the error between θ_e and λ_e, we have

$$\left| \frac{h_e(\mathrm{pp})}{g_e(\mathrm{pp})} - 1 \right| \leq k^{O(1)} \left(\epsilon_e + \frac{1}{2^k} \left(\dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} + \mathrm{err}_v \right) \right)$$

for all $e \in \delta v$. Substituting into the previous expression, and using Proposition 7.13 again, we conclude

$$\left| \frac{Z_{Q,e} Q_e(\mathrm{rr})}{Z_{P,e} P_e(\mathrm{rr})} - 1 \right| \leq k^{O(1)} \left(\mathrm{err}_v + \left\{ \dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} + \mathrm{err}_v \right\} + \sum_{e' \in \delta v} \left\{ \epsilon_{e'} + \frac{1}{2^k} \left(\dot{\epsilon}_{e'} + \frac{\min\{\ddot{\epsilon}_{e'}, 1\}}{(2^{k\zeta})_{e'}} + \mathrm{err}_v \right) \right\} \right)$$

$$\leq k^{O(1)} \left(\mathrm{err}_v + \dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} \right).$$

A similar calculation can be made for the other spins in $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$, and (7.42) straightforwardly follows.

Step 2. Marginal errors between P and Q. Fix an edge $e \in \delta v$. Let $\hat{s}_e = \hat{p}_e$, and $\hat{s}_{e'} = \hat{q}_{e'}$ for $e' \in \delta v \setminus e$. We first apply Proposition 7.13 with $\hat{s}$ in place of $\hat{q}$: let Γ be the resulting weights such that $\nu_{\delta v}[\Gamma; \hat{s}]$ is judicious, and let $\dot{s} \equiv \mathrm{BP}[\Gamma; \hat{s}]$ denote the corresponding outgoing messages from the variable, with reweighted versions S as defined by (7.10). Then we are exactly in the situation of Lemma 7.18. The reweighting W of $\hat{r}_e = \hat{p}_e$ satisfies

$$\left\| W - \frac{1}{9} \right\|_\infty \leq \frac{k^{O(1)}}{2^{k\zeta}}$$

by assumption (7.25). We also note that $P(\sigma) \leq O(1)/2^{k\mathtt{r}[\sigma]}$, and likewise for $S(\sigma)$. It follows from (7.42) that

$$\left\| \frac{S}{P} - 1 \right\|_\infty \leq k^{O(1)} \mathrm{err}_v.$$

Combining this with the result of Lemma 7.18 gives the marginal error bound between S and P,

$$\left\| \frac{S^j}{P^j} - 1 \right\|_\infty \leq O(1) \left\| W - \frac{1}{9} \right\|_\infty \left\| \frac{S}{P} - 1 \right\|_\infty \leq \frac{k^{O(1)} \mathrm{err}_v}{2^{k\zeta}} \tag{7.43}$$

for both $j = 1, 2$. It also follows from (7.42) that

$$\left| \frac{Q(\sigma)}{S(\sigma)} - 1 \right| \leq k^{O(1)} \left(\epsilon_e + \frac{1}{(2^k)^{\mathbf{1}\{\mathtt{r}[\sigma]=0\}}} \left\{ \dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} \right\} \right), \tag{7.44}$$

from which we obtain the marginal error bound

$$\left| \frac{Q^j(\tau)}{S^j(\tau)} - 1 \right| \leq k^{O(1)} \left(\epsilon_e + \frac{1}{(2^k)^{\mathbf{1}\{\tau \neq \mathtt{r}\}}} \left\{ \dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} \right\} \right). \tag{7.45}$$

The result follows by combining (7.42), (7.43), and (7.45). □

7.4. **Contraction in non-defective trees.** In this subsection we complete the proof of Proposition 7.3. As in the statement of the proposition, let U be a compound enclosure, and take a subtree $T \subseteq U$ of the form described in Definition 7.1, such that T contains no defective variables. Recall from (7.6) that for any clause a in T, and for any given set of boundary marginals $\omega_{\delta T}$, we defined the quantity

$$\xi_a(T; \omega_{\delta T}) \equiv \sum_{e \in \delta T} \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{\mathsf{b}_T(e,a)} \mathrm{disc}_e(\omega), \tag{7.46}$$

for $\mathrm{disc}_e(\omega)$ as given by (7.4). The main technical result of this subsection is the following:

Proposition 7.19. *In the setting of Proposition 7.3, for the iterative construction of Definition 7.9, on any edge (av) in T we have (using the notations from Definitions 7.5 and 7.10)*

$$\mathrm{E\ddot{R}R}(\dot{q}_{va,t-1}, \dot{q}_{va,t}) \equiv \boldsymbol{\delta}_{va}(t) \equiv \begin{pmatrix} \delta_{va,t} \\ \dot{\delta}_{va,t} \\ \ddot{\delta}_{va,t} \\ {}^{\mathrm{m}}\delta_{va,t} \\ {}^{\mathrm{m}}\dot{\delta}_{va,t} \end{pmatrix} \le k^{O(1)} \xi_a (k^{O(1)}(\vartheta_*)^{1/3})^{t-1} \underline{\delta}, \quad \underline{\delta} \equiv \begin{pmatrix} \delta \\ \dot{\delta} \\ \ddot{\delta} \\ {}^{\mathrm{m}}\delta \\ {}^{\mathrm{m}}\dot{\delta} \end{pmatrix} \equiv \begin{pmatrix} 1 \\ (\vartheta_*)^{-1} \\ (\vartheta_*)^{-1} \\ (\vartheta_*)^{2} \\ \vartheta_* \end{pmatrix}, \qquad (7.47)$$

$$\widehat{\mathrm{ERR}}(\hat{q}_{av,t-1}, \hat{q}_{av,t}) \equiv \boldsymbol{\epsilon}_{av}(t) \equiv \begin{pmatrix} \epsilon_{av,t} \\ \dot{\epsilon}_{av,t} \\ \ddot{\epsilon}_{av,t} \end{pmatrix} \le k^{O(1)} \xi_a (k^{O(1)}(\vartheta_*)^{1/3})^{t+1} \underline{\epsilon}, \quad \underline{\epsilon} \equiv \begin{pmatrix} \epsilon \\ \dot{\epsilon} \\ \ddot{\epsilon} \end{pmatrix} \equiv \begin{pmatrix} 2^{-k} \\ \vartheta_* \\ (\vartheta_*)^{-1} \end{pmatrix}, \qquad (7.48)$$

where $\xi_a \equiv \xi_a(T; \omega_{\delta T})$ is defined by (7.46).

We first provide some lemmas towards the proof of Proposition 7.19. The following is an elementary bound:

Lemma 7.20. *In the setting of Propositions 7.3 and 7.19, if $N(a)$ denotes the clauses in T at unit distance from a clause a in T, then we have the bound*

$$\sum_{b \in N(a)} \xi_b(T; \omega_{\delta T}) \le O\left(\frac{k^{O(1)} 2^k \xi_a(T; \omega_{\delta T})}{(\vartheta_*)^{1/4}} \right).$$

Proof. All the variables in T are assumed to be non-defective, hence nice. Let $d_{\max} = O(k2^k)$ denote the maximum degree of a nice variable (Definition 3.8). Then, for any edge $e \in \delta T$ and any clause a in T, we have

$$\left| \left\{ b \in N(a) : \mathfrak{b}_T(e,b) = \mathfrak{b}_T(e,a) - 1 \right\} \right| \le 1,$$

$$\left| \left\{ b \in N(a) : \mathfrak{b}_T(e,b) = \mathfrak{b}_T(e,a) \right\} \right| \le d_{\max},$$

$$\left| \left\{ b \in N(a) : \mathfrak{b}_T(e,b) = \mathfrak{b}_T(e,a) + 1 \right\} \right| \le k d_{\max}.$$

Combining these bounds with the definition (7.46) gives

$$\sum_{b \in N(a)} \xi_b(T; \omega_{\delta T}) \le \sum_{e \in \delta T} \mathrm{disc}_e(\omega) \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{\mathfrak{b}_T(e,a)} \left\{ \frac{2^k}{(\vartheta_*)^{1/4}} + d_{\max} + k d_{\max} \cdot \frac{(\vartheta_*)^{1/4}}{2^k} \right\} \le O\left(\frac{2^k \xi_a(T; \omega_{\delta T})}{(\vartheta_*)^{1/4}} \right),$$

as claimed. $\qquad \square$

Next, we give a bound on the errors after the first round of boundary updates:

Lemma 7.21. *Recall the notations of Definitions 7.5 and 7.10. In the setting of Propositions 7.3 and 7.19, for any boundary edge $e = (av) \in \delta T$, suppose that we have, for an absolute constant $0 < \zeta \le 1/20$, the bounds*

$$\widehat{\mathrm{ERR}}(_*\omega_e, \omega_e) \le \begin{pmatrix} \gamma_e \\ \dot{\gamma}_e \\ \ddot{\gamma}_e \end{pmatrix} \le k^{O(1)} \min\left\{ \mathrm{disc}_e(\omega) \begin{pmatrix} 1 \\ (\vartheta_*)^{-1} \\ (\vartheta_*)^{-1} \end{pmatrix}, \frac{1}{2^{2k\zeta}} \begin{pmatrix} 1 \\ 1 \\ 2^k \end{pmatrix} \right\}, \qquad (7.49)$$

where the last bound follows from the definition (7.4) together with the assumption (7.5) (taking the entrywise minimum of the two vectors). Then, after the first round of boundary updates (i.e., after applying Definition 7.9 Step I), we have

$$\mathrm{E\ddot{R}R}(\dot{q}_{e,0}, \dot{q}_{e,1}) \le k^{O(1)} \begin{pmatrix} 1 & 2^{-k} & 4^{-k} \\ 1 & 1 & 4^{-k} \\ 1 & 2^{-k} & 1 \\ 2^{-k/10} & 2^{-k/10} 2^{-k} & 2^{-k/10} 4^{-k} \\ 2^{-k/10} & 2^{-k/10} & 2^{-k/10} 2^{-k} \end{pmatrix} \begin{pmatrix} \gamma_e \\ \dot{\gamma}_e \\ \ddot{\gamma}_e \end{pmatrix} \le k^{O(1)} \min\left\{ \mathrm{disc}_e(\omega) \begin{pmatrix} 1 \\ (\vartheta_*)^{-1} \\ (\vartheta_*)^{-1} \\ 2^{-k/10}(\vartheta_*)^{-1} \\ 2^{-k/10}(\vartheta_*)^{-1} \end{pmatrix}, \frac{1}{2^{k\zeta}} \begin{pmatrix} 1 \\ 1 \\ 2^k \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Moreover, if $Q_{e,1}$ is the reweighted version of $\dot{q}_{e,1}$ defined by (7.10), then it satisfies the bounds (7.11).

Proof. Recall that $\dot{q}_{e,0} = {}_*\dot{q}_e = {}_\bullet\dot{q}_e \otimes {}_\bullet\dot{q}_e$. From Definition 7.9, the first updates (Step I) yield boundary messages

$$\dot{q}_{e,1}(\sigma) \cong \frac{\omega_e(\sigma)}{{}_*\hat{q}_e(\sigma)} = \frac{\omega_e(\sigma)}{{}_\bullet\hat{q}_e(\sigma^1){}_\bullet\hat{q}_e(\sigma^2)}.$$

We use (7.10) to define the reweighted versions

$$Q_{e,1}(\sigma) \cong \frac{1}{(2^{|\partial a|-1})^{\mathbf{r}[\sigma]}} \frac{\omega_e(\sigma)}{{}_*\hat{q}_e(\sigma)} \cong \frac{\omega_e(\sigma)}{{}_*\omega_e(\sigma)} \frac{{}_*\dot{q}_e(\sigma)}{(2^{|\partial a|-1})^{\mathbf{r}[\sigma]}} \cong \frac{\omega_e(\sigma)}{{}_*\omega_e(\sigma)}{}_*Q_e(\sigma),$$

where the second-to-last step uses that ${}_*\omega_e(\sigma) \cong {}_*\dot{q}_e(\sigma){}_*\hat{q}_e(\sigma)$, and ${}_*Q_e = Q_{e,0}$ is the reweighting of ${}_*\dot{q}$ by (7.10). Therefore, there exists a normalizing constant Z_e such that

$$Z_e Q_{e,1}(\sigma) = \frac{\omega_e(\sigma)}{{}_*\omega_e(\sigma)}{}_*Q_e(\sigma).$$

Summing the equation over $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$, and recalling the assumption (7.49), we find

$$Z_e = \sum_{\sigma:\mathbf{r}[\sigma]=0} {}_*Q_e(\sigma)\Big(1 + O(\gamma_e)\Big) + \sum_{\sigma:\mathbf{r}[\sigma]=1} {}_*Q_e(\sigma)\Big(1 + O(\dot{\gamma}_e)\Big) + {}_*Q_e(\mathtt{rr})\Big(1 + O(\ddot{\gamma}_e)\Big)$$

$$= 1 + O\Big(\gamma_e + \frac{\dot{\gamma}_e}{2^k} + \frac{\ddot{\gamma}_e}{4^k}\Big).$$

Recalling that $Q_{e,0} = {}_*Q_e$, it follows by combining the above estimates that

$$\left| \frac{Q_{e,1}(\sigma)}{Q_{e,0}(\sigma)} - 1 \right| \le k^{O(1)}\Big(\gamma_e + \frac{\dot{\gamma}_e}{2^k} + \frac{\ddot{\gamma}_e}{4^k} + \mathbf{1}\{\mathbf{r}[\sigma] = 1\}\dot{\gamma}_e + \mathbf{1}\{\mathbf{r}[\sigma] = 2\}\ddot{\gamma}_e\Big).$$

Next recall that from the initialization of Definition 7.9 we have ${}_*\omega_e \cong {}_*\hat{q}_e{}_*\dot{q}_e = \hat{q}_{e,0}\dot{q}_{e,0}$; and after the first round of boundary updates we have $\omega_e \cong \hat{q}_{e,0}\dot{q}_{e,1}$. Both ${}_*\omega_e$ and ω_e have single-copy marginals ${}_\star\pi_e$, so we are precisely in the situation of Lemma 7.18 with $h = {}_*\hat{q}_e$. Applying Lemma 7.18 gives, for both $j = 1, 2$,

$$\left| \frac{(Q_{e,1})^j(\sigma)}{(Q_{e,0})^j(\sigma)} - 1 \right| \le \frac{k^{O(1)}}{2^{k/10}} \begin{cases} \gamma_e + \dot{\gamma}_e/2^k + \ddot{\gamma}_e/4^k & \text{for } \sigma \ne \mathtt{r}, \\ \gamma_e + \dot{\gamma}_e + \ddot{\gamma}_e/2^k & \text{for } \sigma = \mathtt{r}. \end{cases}$$

This proves the first claimed bound on $\ddot{\mathrm{ERR}}(\dot{q}_{e,0}, \dot{q}_{e,1})$. The remaining bounds follow using (7.49). $\qquad\square$

Corollary 7.22. *In the setting of Propositions 7.3 and 7.19, after the first update of the clauses just above the boundary variables (i.e., after applying Definition 7.9 Step I at each boundary edge, followed by Step II(a) at each clause incident to a boundary edge), we have*

$$\widehat{\mathrm{ERR}}(\hat{q}_{e',1}, \hat{q}_{e',0}) \le k^{O(1)} \begin{pmatrix} 2^{-k(1+\zeta)} \\ 2^{-k\zeta} \\ 2^{-k\zeta} \end{pmatrix}$$

for every edge $e' \in \delta a \setminus \delta T$ for any clause a incident to δT.

Proof. For all $e \in \delta T$, the error $\ddot{\mathrm{ERR}}(\dot{q}_{e,1}, \dot{q}_{e,0})$ is bounded by Lemma 7.21. Applying Proposition 7.12 gives, after applying Step II(a) at the clauses above the boundary variables,

$$\widehat{\mathrm{ERR}}(\hat{q}_{e',1}, \hat{q}_{e',0}) \le k^{O(1)} \begin{pmatrix} 2^{-k(1+\zeta)} & 4^{-k} & 2^{-k(1+\zeta)} & 2^{-k} & 2^{-k} \\ 2^{-k} & 2^{-k} & 2^{-k(1+\zeta)} & 1 & 2^{-k} \\ 1 & 4^{-k} & 2^{-k(1+\zeta)} & 2^{-k} & 2^{-k} \end{pmatrix} \begin{pmatrix} 2^{-k\zeta} \\ 2^{-k\zeta} \\ \min\{2^{k(1-\zeta)}, 1\} \\ 2^{-k\zeta} \\ 2^{-k\zeta} \end{pmatrix} \le k^{O(1)} \begin{pmatrix} 2^{-k(1+\zeta)} \\ 2^{-k\zeta} \\ 2^{-k\zeta} \end{pmatrix}$$

for the edges e' just above δT. This proves the claim. $\qquad\square$

Proof of Proposition 7.19. We will prove the result by induction. We divide the proof into a few steps, indicated by Arabic numerals below. We use Roman numerals to refer to the steps of Definition 7.9.

Step 1. Verification of base case. By Lemma 7.21, for any boundary edge $e = (av) \in \delta T$, we have

$$\boldsymbol{\delta}_{va}(1) \equiv \begin{pmatrix} \delta_{va,1} \\ \dot{\delta}_{va,1} \\ \ddot{\delta}_{va,1} \\ {}^{\mathrm{m}}\delta_{va,1} \\ {}^{\mathrm{m}}\dot{\delta}_{va,1} \end{pmatrix} \leq k^{O(1)} \operatorname{disc}_e(\omega) \begin{pmatrix} 1 \\ (\vartheta_*)^{-1} \\ (\vartheta_*)^{-1} \\ 2^{-k/10}(\vartheta_*)^{-1} \\ 2^{-k/10}(\vartheta_*)^{-1} \end{pmatrix} \leq k^{O(1)} \xi_a \begin{pmatrix} 1 \\ (\vartheta_*)^{-1} \\ (\vartheta_*)^{-1} \\ (\vartheta_*)^2 \\ \vartheta_* \end{pmatrix} = k^{O(1)} \xi_a \underline{\delta}, \qquad (7.50)$$

which verifies the bound (7.47) for $t = 1$.

Step 2. Bounds for clause updates. For any clause $a \neq a_{\mathrm{RT}}$ in the tree T, for both steps II(a) and III(b), our earlier result Proposition 7.12 implies the bound

$$\boldsymbol{\epsilon}_{av}(t) \equiv \begin{pmatrix} \epsilon_{av,t} \\ \dot{\epsilon}_{av,t} \\ \ddot{\epsilon}_{av,t} \end{pmatrix} \leq k^{O(1)} \begin{pmatrix} 2^{-k(1+\zeta)} & 4^{-k} & 2^{-k(1+\zeta)} & 2^{-k} & 2^{-k} \\ 2^{-k} & 2^{-k} & 2^{-k(1+\zeta)} & 1 & 2^{-k} \\ 1 & 4^{-k} & 2^{-k(1+\zeta)} & 2^{-k} & 2^{-k} \end{pmatrix} \sum_{u \in \partial a \setminus v} \begin{pmatrix} \delta_{ua,t-1} \\ \dot{\delta}_{ua,t-1} \\ \min\{\ddot{\delta}_{ua,t-1}, 1\} \\ {}^{\mathrm{m}}\delta_{ua,t-1} \\ {}^{\mathrm{m}}\dot{\delta}_{ua,t-1} \end{pmatrix}.$$

(On the right-hand side of the above, for Step III(b) one can omit the contribution from the parent variable of a; but we will not use this fact.) Substituting the inductive hypothesis (7.47) gives the simplified bound

$$\boldsymbol{\epsilon}_{av}(t) \leq k^{O(1)} \xi_a (k^{O(1)}(\vartheta_*)^{1/3})^{t-2} \underbrace{\begin{pmatrix} 2^{-k(1+\zeta)} & 4^{-k} & 2^{-k(1+\zeta)} & 2^{-k} & 2^{-k} \\ 2^{-k} & 2^{-k} & 2^{-k(1+\zeta)} & 1 & 2^{-k} \\ 1 & 4^{-k} & 2^{-k(1+\zeta)} & 2^{-k} & 2^{-k} \end{pmatrix}}_{\text{denote this } \widehat{\mathsf{M}}} \underline{\delta}. \qquad (7.51)$$

At the root clause a_{RT} we make the trivial update (7.18), so $\widehat{\mathrm{ERR}}(\hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}, t-1}, \hat{q}_{a_{\mathrm{RT}} v_{\mathrm{RT}}, t}) = (0, 0, 0) \in \mathbb{R}^3$.

Step 3. Bounds for internal variable updates. For any internal variable v in T, for both steps II(b) and III(a), applying Corollary 7.14 gives the bound

$$\boldsymbol{\delta}_{va}(t) \equiv \begin{pmatrix} \delta_{va,t} \\ \dot{\delta}_{va,t} \\ \ddot{\delta}_{va,t} \\ {}^{\mathrm{m}}\delta_{va,t} \\ {}^{\mathrm{m}}\dot{\delta}_{va,t} \end{pmatrix} \leq k^{O(1)} \sum_{c \in \partial v \setminus a} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2^{-k\zeta} \\ 2^{-k\zeta} \end{pmatrix} \begin{pmatrix} 1 & 2^{-k} & 2^{-k(1+\zeta)} \end{pmatrix} \begin{pmatrix} \epsilon_{cv,t-1} \\ \dot{\epsilon}_{cv,t-1} \\ \min\{\ddot{\epsilon}_{cv,t-1}, 1\} \end{pmatrix}$$

$$+ k^{O(1)} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2^{-k\zeta} \\ 0 & 1 & 2^{-k\zeta} \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 1 & 2^{-k\zeta} \end{pmatrix} \begin{pmatrix} \epsilon_{av,t-1} \\ \dot{\epsilon}_{av,t-1} \\ \min\{\ddot{\epsilon}_{av,t-1}, 1\} \end{pmatrix}.$$

(On the right-hand side of the above, for Step II(b) one only needs the first term that sums over $c \in \partial v \setminus a$. For Step III(a) one can omit the contribution to the sum from the parent clause of v. We will not use this fact.) We will bound the above sum over $c \in \partial v \setminus a$ by the sum over all $c \in N(a)$. Then, substituting the inductive hypothesis

([7.48](#)) and combining with Lemma [7.20](#) gives the simplified bound

$$\delta_{va}(t) \leq k^{O(1)}\xi_a(k^{O(1)}(\vartheta_*)^{1/3})^t \left\{ \frac{2^k}{(\vartheta_*)^{1/4}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2^{-k\zeta} \\ 2^{-k\zeta} \end{pmatrix} \begin{pmatrix} 1 & 2^{-k} & 2^{-k(1+\zeta)} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2^{-k\zeta} \\ 0 & 1 & 2^{-k\zeta} \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 1 & 2^{-k\zeta} \end{pmatrix} \right\} \underline{\epsilon}$$

$$\leq k^{O(1)}\xi_a(k^{O(1)}(\vartheta_*)^{1/3})^t \frac{1}{(\vartheta_*)^{1/4}} \underbrace{\begin{pmatrix} 2^k & 1 & 2^{-k\zeta} \\ 2^k & 1 & 2^{-k\zeta} \\ 2^k & 1 & 2^{-k\zeta} \\ 2^{k(1-\zeta)} & 2^{-k\zeta} & 2^{-2k\zeta} \\ 2^{k(1-\zeta)} & (\vartheta_*)^{1/4} & (\vartheta_*)^{1/4}2^{-k\zeta} \end{pmatrix}}_{\text{denote this } \dot{\mathsf{M}}_{\text{int}}} \underline{\epsilon} \, . \tag{7.52}$$

Step 4. Bounds for boundary updates. For each boundary variable v in T with parent clause a, in Step I we have

$$\delta_{va}(t) \leq k^{O(1)} \begin{pmatrix} 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 1 & 2^{-k(1+\zeta)} \\ 1 & 2^{-k} & 1 \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 1 & 2^{-k\zeta} \end{pmatrix} \begin{pmatrix} \epsilon_{av,t-1} \\ \dot{\epsilon}_{av,t-1} \\ \ddot{\epsilon}_{av,t-1} \end{pmatrix} \leq k^{O(1)}\xi_a(k^{O(1)}(\vartheta_*)^{1/3})^t \underbrace{\begin{pmatrix} 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 1 & 2^{-k(1+\zeta)} \\ 1 & 2^{-k} & 1 \\ 1 & 2^{-k} & 2^{-k(1+\zeta)} \\ 1 & 1 & 2^{-k\zeta} \end{pmatrix}}_{\text{denote this } \dot{\mathsf{M}}_{\text{bd}}} \underline{\epsilon} \, , \tag{7.53}$$

where the last bound again uses the inductive hypothesis ([7.48](#)).

Step 5. Verification of induction. For the matrices defined in ([7.51](#)), ([7.52](#)), and ([7.53](#)), we have

$$\widehat{\mathsf{M}}\underline{\delta} \leq k^{O(1)}\vartheta_*\underline{\epsilon}, \quad \left(\dot{\mathsf{M}}_{\text{int}} + \dot{\mathsf{M}}_{\text{bd}}\right)\underline{\epsilon} \leq \frac{k^{O(1)}}{(\vartheta_*)^{1/4}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ k^c/2^{k\zeta} \\ (\vartheta_*)^{5/4} \end{pmatrix} + k^{O(1)} \begin{pmatrix} 2^{-k} \\ \vartheta_* \\ (\vartheta_*)^{-1} \\ 2^{-k} \\ \vartheta_* \end{pmatrix} \leq \frac{k^{O(1)}}{(\vartheta_*)^{1/4}}\underline{\delta} \, . \tag{7.54}$$

Using these inequalities, the bounds ([7.51](#)), ([7.52](#)), and ([7.53](#)) simplify to

$$\epsilon_{av}(t) \leq k^{O(1)}\xi_a(k^{O(1)}(\vartheta_*)^{1/3})^{t-2}\vartheta_*\underline{\epsilon} \leq k^{O(1)}\xi_a(k^{O(1)}(\vartheta_*)^{1/3})^{t+1}\underline{\epsilon} \, ,$$

$$\delta_{va}(t) \leq k^{O(1)}\xi_a(k^{O(1)}(\vartheta_*)^{1/3})^t(\vartheta_*)^{-1/4}\underline{\delta} \leq k^{O(1)}\xi_a(k^{O(1)}(\vartheta_*)^{1/3})^{t-1}\underline{\delta} \, .$$

This verifies the inductive hypotheses ([7.47](#)) and ([7.48](#)).

Step 6. Conclusion. We conclude the proof by briefly addressing some points which we neglected in the above. In Step 2, in order to apply Proposition [7.12](#) we needed to check that the messages incoming to the clauses satisfy the estimates ([7.11](#)). This can be seen by induction, with the base case given directly by Lemma [7.21](#). Similarly, in Step 3, in order to apply Corollary [7.14](#) we needed to check that the messages incoming to the internal variables satisfy the estimates ([7.25](#)). This can also be seen by induction, with the base case given directly by Corollary [7.22](#). This finishes the proof of the proposition. □

Remark 7.23 (explanation of choices in ([7.47](#)) and ([7.48](#))). In this remark we give some explanation for the choice of the vectors $\underline{\delta}$ and $\underline{\epsilon}$ in Proposition [7.19](#). If we take the product of the 3×5 matrix in ([7.51](#)) with the 5×3 matrix in ([7.52](#)), the result can be (entrywise) upper bounded as

$$\widehat{\mathsf{M}}\dot{\mathsf{M}}_{\text{int}} \leq k^{O(1)} \begin{pmatrix} 2^{-k\zeta} & (\vartheta_*)^{1/4}2^{-k} & (\vartheta_*)^{1/4}2^{-k(1+\zeta)} \\ 2^{k(1-\zeta)} & 2^{-k\zeta} & 2^{-k\zeta} \\ 2^k & 1 & 2^{-k\zeta} \end{pmatrix} \leq k^{O(1)} \begin{pmatrix} 2^{-k} \\ \vartheta_* \\ (\vartheta_*)^{-1} \end{pmatrix} \begin{pmatrix} 2^k\vartheta_* & (\vartheta_*)^{1/4} & (\vartheta_*)^5 \end{pmatrix} \, .$$

For the last bound on the right-hand side, the vector $\underline{\epsilon}$ from ([7.48](#)) from Proposition [7.19](#) is a right eigenvector with eigenvalue $O(\vartheta_*)$. This explains why $\underline{\epsilon}$ is a good choice for our purposes (although it is certainly not the unique

choice that would give a sufficiently good bound). On the other hand, given $\underline{\epsilon}$, we chose the vector $\underline{\delta}$ to satisfy the bounds (7.54) and (7.50). (Again, it is certainly not the unique one that suffices for our purposes.)

Proof of Proposition 7.3. It is clear from Proposition 7.19 that the construction of Definition 7.9 converges to the desired Lagrangian weights $\Lambda \equiv \Lambda(T; \omega_{\delta T})$. The corresponding Gibbs measure $\nu = \nu[T; \Lambda]$ has then edge marginals $\nu_{av} \cong \dot{q}_{va} \hat{q}_{av}$ where $\dot{q}_{va} \equiv \dot{q}_{va,\infty}$ and $\hat{q}_{av} \equiv \hat{q}_{av,\infty}$ are the limiting BP messages. It remains only to verify the estimate (7.6). By Proposition 7.19, on any edge (av) we have, abbreviating $\xi_a \equiv \xi_a(T; \omega_{\delta T})$,

$$\ddot{\mathrm{ERR}}(\ast\dot{q}_{va}, \dot{q}_{va}) \leq \sum_{t \geq 1} \ddot{\mathrm{ERR}}(\dot{q}_{va,t-1}, \dot{q}_{va,t}) \leq O(1)\xi_a \underline{\delta},$$

$$\widehat{\mathrm{ERR}}(\ast\hat{q}_{av}, \hat{q}_{av}) \leq \sum_{t \geq 1} \widehat{\mathrm{ERR}}(\hat{q}_{av,t-1}, \hat{q}_{av,t}) \leq O(1)\xi_a (k^{2c}(\vartheta_\ast)^{1/3})^2 \underline{\epsilon}.$$

This implies, for all edges (av) in T and all $\sigma \in \{\mathrm{r}, \mathrm{y}, \mathrm{g}, \mathrm{b}\}^2$, the bound

$$\left| \frac{\dot{q}_{va}(\sigma)}{\ast\dot{q}_{va}(\sigma)} - 1 \right| + \left| \frac{\hat{q}_{av}(\sigma)}{\ast\hat{q}_{av}(\sigma)} - 1 \right| \leq O\left(\frac{\xi_a}{(\vartheta_\ast)^{\mathbf{1}\{\mathrm{r} \in \{\sigma^1, \sigma^2\}\}}} \right).$$

Since $\nu_{av} \cong \dot{q}_{va} \hat{q}_{av}$, it follows straightforwardly that

$$\left| \frac{\nu_{av}(\sigma)}{\ast\omega_{av}(\sigma)} - 1 \right| \leq O\left(\frac{\xi_a}{(\vartheta_\ast)^{\mathbf{1}\{\mathrm{r} \in \{\sigma^1, \sigma^2\}\}}} \right).$$

Recalling (7.4) gives $\mathrm{disc}_{av}(\nu_{av}) \leq O(\xi_a)$, as claimed.   □

7.5. **Contraction with multiple clause types.** In this subsection we prove Proposition 7.4, which concerns the entropy maximization problem from Proposition 3.46. The following is an analogue of Definition 7.8:

Definition 7.24 (parametrization of Lagrangian weights Ψ on a depth-one neighborhood of a non-compound variable). Let U be the depth-one tree from Definition 3.45, rooted at a variable v of (non-compound) total type $\boldsymbol{T}$. For an augmented pair coloring $(\underline{\sigma}, \underline{L})$ of U, we will define weight functions $\Psi \equiv \Psi_U$ parametrized as

$$\Psi(\underline{\sigma}, \underline{L}) = \Psi_v(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \cdot \left\{ \prod_{e \in \delta U} \psi_e(\sigma_e, L_e) \right\},$$

where the root variable weight Ψ_v takes the form

$$\Psi_v(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \equiv \varphi_v(\underline{\sigma}_{\delta v}) \underbrace{\left\{ \prod_{j=1,2} \psi^j(x^j) \right\} \left\{ \prod_{e \in \delta v} \prod_{j=1,2} (\psi_e)^j((\sigma_e)^j) \right\}}_{\text{denote this } \psi_v(\underline{\sigma}_{\delta v})} \left\{ \prod_{e \in \delta v} \underbrace{\left(\psi_e(L_e) \prod_{j=1,2} (\psi_e)^j((\sigma_e)^j \mid L_e) \right)}_{\text{denote this } \psi_e(\sigma_e, L_e)} \right\}, \quad (7.55)$$

— in the above, $\varphi_v(\underline{\sigma}_{\delta v})$ is the indicator of a valid pair coloring $\underline{\sigma}_{\delta v}$ (i.e., the pair version of (2.20)), and $x \in \{\text{-}, \text{+}, \text{f}\}^2$ denotes the pair frozen spin corresponding to $\underline{\sigma}_{\delta v}$. If Ψ is any weight on U of the functional form just described, then Ψ is a Lagrangian weight for the optimization problem from Proposition 3.46, meaning that $\langle \ln \Lambda, \nu \rangle$ is constant over $\nu \in J_{\mathscr{D}}(U; \omega_{\delta U})$. Conversely, it is easy to see that any Lagrangian weight can be expressed in this form. Moreover, for our convenience we have chosen weights that are somewhat over-parametrized, since we have edge weights $(\psi_e)^j((\sigma_e)^j)$ as well as $(\psi_e)^j((\sigma_e)^j \mid L_e)$.

We next give the analogue of Definition 7.9:

Definition 7.25 (iterative construction of weights for a non-compound variable). In the setting of Proposition 7.4, we again let U be the depth-one tree from Definition 3.45. We now define a sequence of weights Ψ_t on U (parametrized as in Definition 7.24 for each $t \geq 0$), which will be shown to converge as $t \to \infty$ to the Lagrangian weights $\Psi_\infty \equiv \Psi_{\mathscr{D}}(U; \omega_{\delta U})$ of Proposition 7.4. At the same time we will define messages q_t which will converge as $t \to \infty$ to the BP solution $q \equiv q_{\mathrm{BP}}(U; \Psi_\infty)$ for the Ψ_∞-weighted model on U. (On each edge e of U, the BP messages $\dot{q}_{e,t}$ and $\hat{q}_{e,t}$ will now be probability measures on tuple (σ, L) where $\sigma \in \{\mathrm{r}, \mathrm{y}, \mathrm{g}, \mathrm{b}\}^2$ while L ranges over all possible types for the clause incident to e.)

To start the construction, let $_\star\Lambda_U$ denote the single-copy weights on U given by Corollary 3.56 (and explicitly constructed in Corollary 6.2) — similarly to (7.16), if $\underline{\tau} \equiv \underline{\tau}_U$ is a single-copy coloring of U, then

$$_\star\Lambda_U(\underline{\tau}) = \overbrace{\left\{ _\star\lambda_v(x_v) \prod_{e \in \delta v} {_\star\lambda_e}(\tau_e) \right\}}^{_\star\Lambda_v(\underline{\tau}_{\delta v})} \left\{ \prod_{e \in \delta U} {_\star\lambda_e}(\tau_e) \right\}$$

where $x_v \in \{\text{-}, \text{+}, \text{f}\}$ denotes the frozen spin corresponding to $\underline{\tau}_{\delta v}$, and for $e \in \delta U$ we take $_\star\lambda_e = {_\bullet\dot{q}_e}$ as defined by (3.92). Next, recalling (3.76), we define $_\star\psi_e(L) \equiv \pi_{\mathscr{D}}(L \mid t_e)$. We initialize our construction at $t = 0$ with

$$\Psi_0(\underline{\sigma}, \underline{L}) \equiv \left\{ \prod_{j=1,2} {_\star\Lambda_U}(\underline{\sigma}^j) \right\} \left\{ \prod_{e \in \delta v} {_\star\psi_e}(L_e) \right\},$$

where $(\underline{\sigma}, \underline{L}) \equiv (\underline{\sigma}_U, \underline{L}_U)$ denotes an augmented pair coloring of U. This fits the functional form prescribed in Definition 7.24 above, since we can rewrite

$$\Psi_0(\underline{\sigma}, \underline{L}) = \Psi_{v,0}(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \cdot \left\{ \prod_{e \in \delta U} \overbrace{\left(\prod_{j=1,2} {_\star\lambda_e}((\sigma_e)^j) \right)}^{\psi_{e,0}(\sigma_e, L_e)} \right\},$$

$$\Psi_{v,0}(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) = \underbrace{\left\{ \prod_{j=1,2} {_\star\Lambda_v}((\underline{\sigma}_{\delta v})^j) \right\}}_{\psi_{v,0}(\underline{\sigma}_{\delta v})} \left\{ \prod_{e \in \delta v} {_\star\psi_e}(L_e) \right\},$$

which is consistent with (7.55) if for all $e \in \delta v$ we take $\psi_{e,0}(\sigma, L) \equiv {_\star\psi_e}(\sigma, L)$, where

$$_\star\psi_e(\sigma, L) \equiv {_\star\psi_e}(L) \left\{ \prod_{j=1,2} ({_\star\psi_e})^j(\sigma^j \mid L) \right\} = {_\star\psi_e}(L) = \pi(L \mid t_e). \tag{7.56}$$

Having defined Ψ_0, we let $q_0 \equiv q_{\text{BP}}(U; \Psi_0)$. Recall that $\omega_{\delta U}$ is given, and abbreviate $\Omega \equiv \text{aug}(\omega)$, as defined by (3.76). For $t \geq 1$, given Ψ_{t-1} we define updated weights Ψ_t by making the following series of updates, started from the boundary δU and working up to the root v, then working back down to δU:

I. **Boundary updates.** For each leaf edge $e = (au) \in \delta U$, we update its weight by setting

$$\psi_{e,t}(\sigma, L) \equiv \dot{q}_{e,t}(\sigma, L) \equiv \frac{\Omega_e(\sigma, L)}{\hat{q}_{e,t}(\sigma, L)} \left/ \left(\sum_{\sigma', L'} \frac{\Omega_e(\sigma', L')}{\hat{q}_{e,t}(\sigma', L')} \right) \right. .$$

II. **Upward pass through clauses.** For each clause a in U, update $\hat{q}_{av,t} \equiv \text{BP}_{av}[\dot{q}_t]$.

III. **Root variable update.** Recall (3.75) that $\varphi_v(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v})$ denotes the indicator of a valid augmented pair coloring of δv. Suppose inductively that the measure

$$\left(\nu_{\delta v}[\Psi_{v,t-1}; \hat{q}_{t-1}] \right)(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \equiv \frac{1}{z_{v,t-1}} \varphi_v(\underline{\sigma}_{\delta v}; \underline{L}_{\delta v}) \Psi_{v,t-1}(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \prod_{e \in \delta v} \hat{q}_{e,t-1}(\sigma_e, L_e) \tag{7.57}$$

is fully judicious on δv. We will show below, in Proposition 7.28, how to define updated weights $\Psi_{v,t}$ such that $\nu_{\delta v}[\Psi_{v,t}; \hat{q}_t]$ is also fully judicious on δv. We then use this to update $\dot{q}_{va,t} \equiv \text{BP}_{va}[\hat{q}_t; \Psi_{v,t}]$ for all $a \in \partial v$.

IV. **Downward pass through clauses.** For each clause a in U, update

$$\hat{q}_{au,t} \equiv \text{BP}_{au}\left[\left(\dot{q}_{va,t-1}, (\dot{q}_{wa,t})_{w \in \partial a \setminus \{u,v\}} \right) \right]$$

for each child variable $u \in \partial a \setminus v$. This completes the definition of Ψ_t and q_t.

In summary, by iterating the above steps, we obtain (Ψ_t, q_t) for all $t \geq 0$ on T. The remainder of this subsection is devoted to analyzing this iteration under the assumptions of Proposition 7.4.

Definition 7.26. In the setting of Definition 3.45, let ν be a probability measure on augmented colorings $(\underline{\sigma}, \underline{L})$ of U. For any edge e of U we let ν_e be the marginal law of (σ_e, L_e) under ν. We say ν_e is **judicious on average** if

$$\bar{\nu}_e(\sigma_e) \equiv \sum_{L_e} \bar{\nu}_e(\sigma_e, L_e)$$

is a judicious measure on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$, in the sense of (3.65). We say that ν_e is **conditionally judicious** if for all clause types L_e that can appear on edge e, the conditional measure $\nu_e(\sigma_e \mid L_e)$ is a judicious measure on $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$, again in the sense of (3.65). Note that **fully judicious** (Definition 3.45) implies conditionally judicious which implies judicious on average.

The following extends the notation from Definitions 7.5 and 7.10:

Definition 7.27. Let $\hat{p}$ and $\hat{q}$ be two nonnegative measures on pairs (σ, L) where $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$ and L ranges over clause types. In what follows we shall write

$$\widehat{\mathbf{ERR}}(\hat{p}, \hat{q}) \leq \underline{\epsilon} = \begin{pmatrix} \epsilon \\ \dot{\epsilon} \\ \ddot{\epsilon} \end{pmatrix}$$

to indicate that $\widehat{\mathrm{ERR}}(\hat{p}(\cdot \mid L), \hat{q}(\cdot \mid L)) \leq \underline{\epsilon}$ for all L in the sense of Definition 7.5, and that

$$\max_L \left\{ \left| \left| \frac{\hat{q}(L)}{\hat{p}(L)} - 1 \right| \right| \right\} \leq \epsilon \,.$$

If $\dot{p}$ and $\dot{q}$ are two probability measures on pairs (σ, L), we define $\mathbf{\ddot{E}RR}(\dot{p}, \dot{q})$ in an analogous fashion, extending the notation from Definition 7.10.

In the following we use $_*\hat{Q}_e$ to be the probability measure on pairs (σ, L) such that

$$_*\hat{Q}_e(\sigma, L) = \frac{_*\hat{q}_e(\sigma)}{\hat{Z}_e} \tag{7.58}$$

where $\hat{Z}_e$ is the normalization. The following result will be applied to the analysis of Step III from Definition 7.25. It builds on the analysis of Proposition 7.13 and Corollary 7.14.

Proposition 7.28. *Let v be a nice variable with incoming messages $\hat{p}$ (in the augmented pair coloring model). Recalling the notation of Definition 7.27, assume that for all $e \in \delta v$ we have (cf. (7.25))*

$$\widehat{\mathbf{ERR}}(_*\hat{Q}_e, \hat{p}_e) \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix} . \tag{7.59}$$

Suppose v has weight Φ_v, of the functional form (7.55), such that the measure $\mu \equiv \nu_{\delta v}[\Phi_v; \hat{p}]$ (defined using the notation of (7.57)) is fully judicious. For the variable weights that do not depend on the clause types, assume that (cf. (7.26))

$$\sum_{j=1,2} \sum_{x \in \{-,+,\mathtt{f}\}} \left| \phi^j(x) - 1 \right| \leq \frac{k^{O(1)}}{2^{k\zeta}} \,, \quad \max_{e \in \delta v} \left\{ \sum_{j=1,2} \left| (\phi_e)^j(\mathtt{r}) - 1 \right| \right\} \leq \frac{k^{O(1)}}{2^{k\zeta}} \,.$$

For the variable weights that do depend on the clause types, assume that

$$\widehat{\mathbf{ERR}}(_\star\psi_e, \phi_e) \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix} \tag{7.60}$$

for $_\star\psi_e$ as defined by (7.56). Now suppose we have a new set of messages $\hat{q}_e$, such that (7.59) also holds with $\hat{q}$ in place of $\hat{p}$, such that

$$\widehat{\mathbf{ERR}}(\hat{p}_e, \hat{q}_e) \leq \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \ddot{\epsilon}_e \end{pmatrix} \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix} \tag{7.61}$$

for each $e \in \delta v$. Then there exists a new set of weights Ψ_v, also of the functional form (7.55), such that $\nu_{\delta v}[\Psi_v; \hat{q}]$ is fully judicious. In addition, the clause-independent weights in Ψ_v satisfy the bounds (7.27) from Proposition 7.13. The clause-dependent weights in Ψ_v satisfy

$$\max_L \left\{ \left| \frac{\psi_e(L)}{\phi_e(L)} - 1 \right| + \max_{j=1,2} \max_{\tau \neq r} \left| \frac{(\psi_e)^j(\tau \mid L)}{(\phi_e)^j(\tau \mid L)} - 1 \right| \right\} \leq \epsilon_e + \frac{\dot{\epsilon}_e}{2^k} + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k(1+\zeta)}} + \frac{\mathrm{err}_v}{2^{k(1+\zeta)}},$$

$$\max_L \max_{j=1,2} \left| \frac{(\psi_e)^j(r \mid L)}{(\phi_e)^j(r \mid L)} - 1 \right| \leq \epsilon_e + \dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} + \frac{\mathrm{err}_v}{2^{k\zeta}} \qquad (7.62)$$

If $\dot{p}_e$ are the original outgoing messages and $\dot{q}_e$ are the new ones, then (cf. Corollary 7.14)

$$\mathbf{\ddot{E}RR}(\dot{p}_e, \dot{q}_e) \leq k^{O(1)} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2^{-k\zeta} \\ 2^{-k\zeta} \end{pmatrix} \mathrm{err}_v + k^{O(1)} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \min\{\ddot{\epsilon}_e, 1\}/2^{k\zeta} \end{pmatrix}, \qquad (7.63)$$

again using the notation from Definition 7.27.

Proof. Since the proof is somewhat involved, we divide it into a few numbered parts. We again recall from (3.75) that $\varphi_v(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v})$ denotes the indicator of a valid augmented pair coloring of δv. It can be expressed as

$$\varphi_v(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) = \varphi_v(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \mathbf{1}\{L_e \ni t_e\},$$

where $\varphi_v(\underline{\sigma}_{\delta v})$ is the indicator of a valid pair coloring on δv (not augmented with clause types) as defined in (2.20); and we write $L \ni t$ to indicate that $L_j = t$ for $j = j(t)$.

Part 1. Iterative construction of Ψ_v. We first give an iterative definition for a sequence of weights $\Psi_{v,t}$ ($t \geq 0$), all of the functional form (7.55). We emphasize that this index t is **purely local to the proof of this proposition**, and is not the same as the t that indexes the up-and-down passes in Definition 7.25. We will show in the remainder of the proof that this sequence converges as $t \to \infty$ to the desired weights $\Psi_v = \Psi_{v,\infty}$. Initialize $\Psi_{v,0} \equiv \Phi_v$. For each $t \geq 0$, we will update from $\Psi_{v,t}$ to $\Psi_{v,t+1}$ in three stages, summarized by the following table:

	$\psi_v(\underline{\sigma}_{\delta v})$	$(\psi_e)^j(\sigma^j \mid L)$	$\psi_e(L)$
$\Psi_{v,t}$	t	t	t
$\Psi_{v,t+1/3}$	$t+1$	t	t
$\Psi_{v,t+2/3}$	$t+1$	$t+1$	t

(7.64)

— e.g., the last row of the table indicates that the weight $\Psi_{v,t+2/3}$ is defined by (7.55) with the $(t+1)$-versions of $\psi_v(\underline{\sigma}_{\delta v})$ and $(\psi_e)^j(\sigma^j \mid L)$ (for both $j = 1, 2$ and all $e \in \delta v$), but with the t-version of $\psi_e(L)$ (again for all $e \in \delta v$). For notational convenience, we also define $\Psi_t \equiv \Phi$ for all $t < 0$, and

$$\hat{q}_{e,t}(\sigma, L) \equiv \begin{cases} \hat{p}_e(\sigma, L) & \text{if } t < 0, \\ \hat{q}_e(\sigma, L) & \text{if } t \geq 0. \end{cases} \qquad (7.65)$$

We then abbreviate $\nu_t \equiv \nu_{\delta v}[\Psi_{v,t}; \hat{q}_t]$, using the notation of (7.57); in particular this means $\nu_t = \nu_{\delta v}[\Phi; \hat{p}]$ for all $t < 0$, while $\nu_0 = \nu_{\delta v}[\Phi; \hat{q}]$. For $t > 0$, the weights of (7.64) are defined by the following procedure (with t denoting an integer time from now on):

A. From t to $t + 1/3$: we will suppose inductively that the measure $\nu_{t-2/3}$ is judicious on average (in the terminology of Definition 7.26). (For $t = 0$, recall that we defined $\nu_{-2/3} \equiv \nu_{\delta v}[\Phi, \hat{p}]$, which by hypothesis is fully judicious and therefore also judicious on average.) At time t (for $t \geq 0$) we have

$$\nu_t(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) = \frac{1}{z_{v,t}} \varphi_v(\underline{\sigma}_{\delta v}) \psi_{v,t}(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \left\{ \mathbf{1}\{L_e \ni t_e\} \psi_{e,t}(\sigma_e, L_e) \hat{q}_e(\sigma_e, L_e) \right\}.$$

The marginal law of $\underline{\sigma}_{\delta v}$ under ν_t can then be expressed as

$$\bar{\nu}_t(\underline{\sigma}_{\delta v}) \cong \varphi_v(\underline{\sigma}_{\delta v})\psi_{v,t}(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \bar{q}_{e,t}(\sigma_e),$$

(7.66)

where $\bar{q}_{e,t}$ denotes the message at time t averaged over L:

$$\bar{q}_{e,t}(\sigma) \equiv \frac{1}{\bar{z}_{e,t}} \sum_L \mathbf{1}\{L \ni t_e\}\psi_{e,t}(\sigma, L)\hat{q}_{e,t}(\sigma, L),$$

(7.67)

where $\hat{q}_{e,t}$ is defined by (7.65). Since we also defined $\psi_{e,t} \equiv \phi_e$ for negative t, this means that $\bar{q}_{e,t}(\sigma_e)$ is also well-defined for negative t. For instance, the marginal law of $\underline{\sigma}_{\delta v}$ at time $t - 2/3$ is

$$\bar{\nu}_{t-2/3}(\underline{\sigma}_{\delta v}) \cong \varphi_v(\underline{\sigma}_{\delta v})\psi_{v,t}(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \bar{q}_{e,t-1}(\sigma_e),$$

(7.68)

including in the case $t = 0$. Now, from the inductive assumption, $\nu_{t-2/3}$ is judicious on average, which means that the measure $\bar{\nu}_{t-2/3}$ in (7.68) is judicious in the sense of Definition 3.43. Comparing (7.66) with (7.68), we see that we can apply Proposition 7.13 — with the variable weight function $\psi_{v,t}(\underline{\sigma}_{\delta v})$ and the two sets of incoming messages $\bar{q}_{e,t-1}$ and $\bar{q}_{e,t}$ — to define a new variable weight function $\psi_{v,t+1}(\underline{\sigma}_{\delta v})$ such that the measure

$$\bar{\nu}_{t+1/3}(\underline{\sigma}_{\delta v}) \cong \varphi_v(\underline{\sigma}_{\delta v})\psi_{v,t+1}(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \bar{q}_{e,t}(\sigma_e)$$

is judicious in the sense of Definition 3.43, i.e., such that $\nu_{t+1/3}$ is judicious on average. Note that, since $t + 1/3 = (t+1) - 2/3$, this verifies the inductive assumption that $\nu_{t-2/3}$ is judicious on average.

B. From $t + 1/3$ to $t + 2/3$: for all $e \in \delta v$ and both $j = 1, 2$, update

$$\frac{(\psi_{e,t+1})^j(\tau \mid L)}{(\psi_{e,t})^j(\tau \mid L)} = \frac{\star \pi_e(\tau)}{(\nu_{e,t+1/3})^j(\tau \mid L)}$$

(7.69)

for all $\tau \in \{r, y, g, b\}$.

C. From $t + 2/3$ to $t + 1$: for all $e \in \delta v$ and both $j = 1, 2$, update

$$\frac{\psi_{e,t+1}(L)}{\psi_{e,t}(L)} = \frac{\pi_{\mathscr{D}}(L \mid t_e)}{\nu_{e,t+2/3}(L_e)},$$

(7.70)

using the notation from (3.76).

The rest of the proof is devoted to the analysis of this iterative procedure. For the calculations that follow, it is useful to note that if $L \not\ni t_e$, then we must have $\psi_{e,t}(L) = 0$ for all t.

Part 2. Error notations. In the proof we will track the following error quantities for all times t. Let κ measure the deviation from being conditionally judicious: for $e \in \delta v$, $j = 1, 2$, and $\tau \in \{r, y, g, b\}$, let

$$(\kappa_{e,t})^j(\tau \mid L) \equiv \frac{(\nu_{e,t})^j(\tau \mid L)}{\star \pi_e(\tau)} - 1.$$

(7.71)

Taking the marginal over L gives

$$(\nu_{e,t})^j(\tau) = \sum_L \nu_{e,t}(L) \cdot (\nu_{e,t})^j(\tau \mid L) = \sum_L \nu_{e,t}(L)\star\pi_e(\tau)\left\{1 + (\kappa_{e,t})^j(\tau \mid L)\right\}$$

$$= \star\pi_e(\tau)\left\{1 + \underbrace{\left[\sum_L \nu_{e,t}(L) \cdot (\kappa_{e,t})^j(\tau \mid L)\right]}_{\text{denote this } (\chi_{e,t})^j(\tau)}\right\}.$$

Notice also that since $\star\pi_e$ and $(\nu_{e,t})^j(\cdot \mid L)$ are both probability measures over $\{r, y, g, b\}$, we must have

$$\sum_\tau \star\pi_e(\tau)(\kappa_{e,t})^j(\tau \mid L) = 0 = \sum_\tau \star\pi_e(\tau)(\chi_{e,t})^j(\tau).$$

(7.72)

Next let ρ measure the deviations in the clause proportions: for $e \in \delta v$,

$$\rho_{e,t}(L) \equiv \frac{\pi_{\mathscr{D}}(L \,|\, t_e)}{\nu_{e,t}(L)} - 1 \,. \tag{7.73}$$

Next, recall the definition (7.58) of $_*\hat{Q}_e(\sigma, L)$, and let

$$\alpha_{e,t}(\sigma, L) \equiv \frac{\psi_{e,t}(\sigma, L)\hat{q}_e(\sigma, L)}{\psi_{e,0}(\sigma, L)_*\hat{Q}_e(\sigma, L)} - 1 = \frac{\psi_{e,t}(\sigma, L)\hat{q}_e(\sigma, L)}{_\star\psi_e(L)_*\hat{q}_e(\sigma)/\hat{Z}_e} - 1 \,. \tag{7.74}$$

Lastly, let ξ measure the deviation from being a product measure: for $e \in \delta v$,

$$\xi_{e,t}(\sigma \,|\, L) \equiv \frac{\nu_{e,t}(\sigma \,|\, L)}{_*\omega_e(\sigma)} - 1$$

We will also make some shorthand notation for the maximum absolute values of the above quantities: let

$$\kappa_{e,t}(\tau) \equiv \max_{L,j} \left\{ \left|(\kappa_{e,t})^j(\tau \,|\, L)\right| \right\}, \quad \begin{pmatrix} \kappa_{e,t} \\ \dot{\kappa}_{e,t} \end{pmatrix} \equiv \begin{pmatrix} \max\{\kappa_{e,t}(\tau) : \tau \neq \mathtt{r}\} \\ \kappa_{e,t}(\mathtt{r}) \end{pmatrix}, \tag{7.75}$$

and let $\rho_{e,t} \equiv \max_L\{|\rho_{e,t}(L)|\}$. Finally let

$$\alpha_{e,t}(\sigma) \equiv \max_L \left\{ \left|\alpha_{e,t}(\sigma \,|\, L)\right| \right\}, \quad \begin{pmatrix} \alpha_{e,t} \\ \dot{\alpha}_{e,t} \\ \ddot{\alpha}_{e,t} \end{pmatrix} = \begin{pmatrix} \max\{\alpha_{e,t}(\sigma) : \mathtt{r}[\sigma] = 0\} \\ \max\{\alpha_{e,t}(\sigma) : \mathtt{r}[\sigma] = 1\} \\ \max\{\alpha_{e,t}(\sigma) : \mathtt{r}[\sigma] = 2\} \end{pmatrix}, \tag{7.76}$$

and make the analogous notation with ξ in place of α. We will show by induction that for all $e \in \delta v$,

$$\sum_{t \in \mathbb{Z}/3, t \geq 1/3} \begin{pmatrix} \rho_{e,t} + \kappa_{e,t} + \alpha_{e,t} \\ \dot{\kappa}_{e,t} + \dot{\alpha}_{e,t} + \xi_{e,t} + \dot{\xi}_{e,t} \\ \ddot{\alpha}_{e,t} + \ddot{\xi}_{e,t} \end{pmatrix} \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix} \tag{7.77}$$

for all $t \geq 1/3$. We also record a calculation that we will use repeatedly in what follows: for any integer t, at time $t + 1/3$, according to the table (7.64) we have

$$\nu_{t+1/3}(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \cong \varphi_v(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v})\psi_{v,t+1}(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \left\{ \psi_{e,t}(\sigma_e, L_e)\hat{q}_{e,t}(\sigma_e, L_e) \right\},$$

where we recall that $\hat{q}_{e,t}$ is defined by (7.65) for all t. From this it is straightforward to verify that the marginal on an edge $e \in \delta v$ is given by

$$\nu_{e,t+1/3}(\sigma, L) = \frac{1}{\bar{z}_{e,t+1/3}}\psi_{e,t}(\sigma, L)\hat{q}_{e,t}(\sigma, L) \overbrace{\left[\frac{1}{\dot{z}_{e,t+1/3}} \sum_{\underline{\sigma}_{\delta v \backslash e}} \varphi_v(\underline{\sigma}_{\delta v})\psi_{v,t+1}(\underline{\sigma}_{\delta v}) \prod_{e'} \bar{q}_{e,t}(\sigma_{e'})\right]}^{\text{denote this } u_{e,t+1/3}(\sigma)} \tag{7.78}$$

where $\dot{z}_{e,t+1/3}$ is the normalizing constant that makes $u_{e,t+1/3}$ a probability measure over $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$, and $\bar{z}_{e,t+1/3}$ is the normalizing constant that makes $\nu_{e,t+1/3}$ a probability measure over pairs (σ, L). For integers t we will compare the following measures:

$$\bar{z}_{e,t-2/3}\nu_{e,t-2/3}(\sigma, L) = \hat{q}_{e,t-1}(\sigma, L)\psi_{e,t-1}(L)\psi_{e,t-1}(\sigma \,|\, L)u_{e,t-2/3}(\sigma, L)\,,$$

$$\bar{z}_{e,t-1/3}\nu_{e,t-1/3}(\sigma, L) = \hat{q}_{e,t-1}(\sigma, L)\psi_{e,t-1}(L)\psi_{e,t}(\sigma \,|\, L)u_{e,t-1/3}(\sigma, L)\,,$$

$$\bar{z}_{e\bullet,t}\nu_{e\bullet,t}(\sigma, L) = \hat{q}_{e,t}(\sigma, L)\psi_{e,t}(L)\psi_{e,t}(\sigma \,|\, L)u_{e,t-2/3}(\sigma, L)\,,$$

$$\bar{z}_{e,t}\nu_{e,t}(\sigma, L) = \hat{q}_{e,t}(\sigma, L)\psi_{e,t}(L)\psi_{e,t}(\sigma \,|\, L)u_{e,t}(\sigma, L)\,,$$

$$\bar{z}_{e,t+1/3}\nu_{e,t+1/3}(\sigma, L) = \hat{q}_{e,t}(\sigma, L)\psi_{e,t}(L)\psi_{e,t}(\sigma \,|\, L)u_{e,t+1/3}(\sigma, L)\,. \tag{7.79}$$

The remaining parts of the proof are organized as follows:

– Part 3a proves some preliminary bounds.
– Part 3b controls $\nu_{e,t+1/3}$ in terms of $\nu_{e\bullet,t}$ for integers $t \geq 0$, in the bounds (7.97) and (7.98). The bounds also depend on the error $\underline{\epsilon}(t)$ (see (7.87)) between $\bar{q}_{e,t-2/3}$ and $\bar{q}_{e,t}$.

- Part 4a controls $\bar{q}_{e,t-1/3}$ in terms of $\nu_{e,t-2/3}$ and $\bar{q}_{e,t-2/3}$, in the bound (7.104).
- Part 4b controls $\bar{q}_{e,t}$ in terms of $\nu_{e,t-1/3}$ and $\bar{q}_{e,t-1/3}$, in the bound (7.107).
- Part 5a controls $\nu_{e,t-1/3}$ in terms of $\nu_{e,t-2/3}$ for integers $t \geq 1$ (in the bounds (7.111) and (7.112)).
- Part 5b compares $\nu_{e,t-1/3}$ with $\nu_{e\bullet,t}$ and $\nu_{e,t}$ for integers $t \geq 1$ (in the bounds (7.115) and (7.116)).
- Part 6 combines the results of the preceding parts to show that the iteration of Part 1 converges to the desired weights $\Psi_v \equiv \Psi_{v,\infty}$, and moreover that these weights satisfy the claimed bounds (7.27) and (7.62).
- Part 7 shows that the messages outgoing from the variable satisfy the claimed bound (7.63), thereby concluding the proof.

Part 3a. Preliminary bounds. Recall the definition (7.74) of $\alpha_{e,t}$, and observe that $\alpha_{e,1/3} = \alpha_{e,0}$. It follows from the assumptions (7.59) and (7.60) that

$$\left| \alpha_{e,1/3}(\sigma, L) \right| \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{cases} 2^{-k} & \text{if } r[\sigma] = 0, \\ 1 & \text{if } r[\sigma] = 1, \\ 2^k & \text{if } r[\sigma] = 2, \end{cases}$$

which verifies that $\alpha_{e,1/3}(\sigma)$ satisfies the bound from (7.77). The marginal on L at time $t = 1/3$ satisfies

$$\nu_{e,1/3}(L) \overset{(7.78)}{=} \frac{1}{\bar{z}_{e,1/3}} \sum_\sigma \phi_e(\sigma, L) \hat{q}_e(\sigma, L) u_{e,1/3}(\sigma)$$

$$\overset{(7.74)}{=} \frac{\star \psi_e(L)}{\bar{z}_{e,1/3} \hat{Z}_e} \sum_\sigma {}_\star \hat{q}_e(\sigma) \left(1 + \alpha_e(\sigma, L) \right) u_{e,1/3}(\sigma) = {}_\star \psi_e(L) \left\{ 1 + \frac{k^{O(1)}}{2^{k(1+\zeta)}} \right\}, \tag{7.80}$$

which proves that $\rho_{e,1/3}$ also satisfies the bound from (7.77). Next, the marginal on σ at time $t = 1/3$ is given by

$$\nu_{e,1/3}(\sigma) = \frac{\bar{z}_{e,1/3}}{\bar{z}_{e,1/3}} \overbrace{\left[\frac{1}{\bar{z}_{e,1/3}} \sum_L \mathbf{1}\{L \ni t_e\} \phi_e(\sigma, L) \hat{q}_e(\sigma, L) \right]}^{\bar{q}_{e,0}(\sigma) = \bar{q}_{e,1/3}(\sigma)} u_{e,1/3}(\sigma), \tag{7.81}$$

where we recall that $\bar{q}_{e,t}$ is defined by (7.67). It follows from the assumptions (7.59) and (7.60) that

$$\widehat{ERR}({}_\star \hat{q}_e, \bar{q}_{e,1/3}) \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix}. \tag{7.82}$$

Substituting (7.82) into (7.81) shows that $\nu_{e,1/3}(\sigma)$ is very close to the normalization of the measure ${}_\star \hat{q}_e(\sigma) u_{e,1/3}(\sigma)$. On the other hand, the same must be true for the conditional measure $\nu_{e,1/3}(\sigma \mid L)$, since

$$\nu_{e,1/3}(\sigma \mid L) \overset{(7.78)}{=} \frac{\phi_e(\sigma, L) \hat{q}_e(\sigma, L) u_{e,1/3}(\sigma)}{\bar{z}_{e,1/3} \nu_{e,1/3}(L)} \overset{(7.74)}{=} \frac{\star \psi_e(L) {}_\star \hat{q}_e(\sigma) u_{e,1/3}(\sigma)}{\bar{z}_{e,1/3} \hat{Z}_e \nu_{e,1/3}(L)} \left\{ 1 + \alpha_e(\sigma, L) \right\}$$

$$\overset{(7.80)}{=} {}_\star \psi_e(L) {}_\star \hat{q}_e(\sigma) u_{e,1/3}(\sigma) \left\{ 1 + \alpha_e(\sigma, L) \right\} \left\{ 1 + \frac{k^{O(1)}}{2^{k(1+\zeta)}} \right\}.$$

Thus, by comparing both $\nu_{e,1/3}(\sigma)$ and $\nu_{e,1/3}(\sigma \mid L)$ with the normalization of ${}_\star \hat{q}_e(\sigma) u_{e,1/3}(\sigma)$, we deduce that

$$\left| \frac{\nu_{e,1/3}(\sigma \mid L)}{\nu_{e,1/3}(\sigma)} - 1 \right| \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{cases} 2^{-k} & \text{if } r[\sigma] = 0, \\ 1 & \text{if } r[\sigma] = 1, \\ 2^k & \text{if } r[\sigma] = 2. \end{cases} \tag{7.83}$$

Since the measure $\nu_{e,1/3}$ is judicious by construction, this implies that $\kappa_{e,1/3}(\tau)$ satisfies the bounds from (7.77). Finally, we note that the variable BP recursion together with (7.82) implies

$$\left| \frac{u_{e,1/3}(\sigma)}{{}_\star \hat{q}_e(\sigma)} - 1 \right| \leq \frac{k^{O(1)}}{2^{k\zeta}}$$

for all $\sigma \in \{r, y, g, b\}^2$. Combining this with (7.59) and (7.60) gives

$$\left| \frac{v_{e,1/3}(\sigma \mid L)}{{}_*\omega_e(\sigma)} - 1 \right| \leq \frac{k^{O(1)}}{2^{k\zeta}} \times \begin{cases} 1 & \text{if } r[\sigma] \leq 1, \\ 2^k & \text{if } r[\sigma] = 2. \end{cases}$$

This implies that $\xi_{e,1/3}(\sigma)$ satisfy the bounds from (7.77). Thus we have proved (7.77) for the case $t = 1/3$. We note for later use that the inductive hypothesis (7.77) implies that (7.82) and (7.83) hold more generally:

$$\widehat{\mathrm{ERR}}({}_*\hat{q}_e, \bar{q}_{e,t}) \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix} \tag{7.84}$$

for all $t \geq 1/3$, and likewise

$$\left| \frac{v_{e,t}(\sigma \mid L)}{v_{e,t}(\sigma)} - 1 \right| \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{cases} 2^{-k} & \text{if } r[\sigma] = 0, \\ 1 & \text{if } r[\sigma] = 1, \\ 2^k & \text{if } r[\sigma] = 2 \end{cases} \tag{7.85}$$

for all $t \geq 1/3$.

Part 3b. Analysis of general applications of update A. We next analyze the update from $t - 2/3$ to $t + 1/3$ for integer times $t \geq 0$. At times $t - 2/3$ and $t + 1/3$ we have (cf. (7.78), and using the notation from (7.65))

$$\bar{z}_{e,t-2/3} v_{e,t-2/3}(\sigma, L) = \psi_{e,t-1}(\sigma, L) \hat{q}_{e,t-1}(\sigma, L) \overbrace{\left[\frac{1}{\dot{z}_{e,t-2/3}} \sum_{\underline{\sigma}_{\delta v \backslash e}} \varphi_v(\underline{\sigma}_{\delta v}) \psi_{v,t}(\underline{\sigma}_{\delta v}) \prod_{e' \in \delta v \backslash e} \bar{q}_{e',t-1}(\sigma_{e'}) \right]}^{u_{e,t-2/3}(\sigma)},$$

$$\bar{z}_{e,t+1/3} v_{e,t+1/3}(\sigma, L) = \psi_{e,t}(\sigma, L) \hat{q}_{e,t}(\sigma, L) \underbrace{\left[\frac{1}{\dot{z}_{e,t+1/3}} \sum_{\underline{\sigma}_{\delta v \backslash e}} \varphi_v(\underline{\sigma}_{\delta v}) \psi_{v,t+1}(\underline{\sigma}_{\delta v}) \prod_{e' \in \delta v \backslash e} \bar{q}_{e',t}(\sigma_{e'}) \right]}_{u_{e,t+1/3}(\sigma)}. \tag{7.86}$$

We then estimate the error between $u_{e,t-2/3}$ and $u_{e,t+1/3}$. Recall the notation of Definition 7.5, and let

$$\underline{\epsilon}(t) \equiv \begin{pmatrix} \epsilon_e(t) \\ \dot{\epsilon}_e(t) \\ \ddot{\epsilon}_e(t) \end{pmatrix} \equiv \widehat{\mathrm{ERR}}(\bar{q}_{t-2/3}, \bar{q}_{t-1/3}) + \widehat{\mathrm{ERR}}(\bar{q}_{t-1/3}, \bar{q}_t). \tag{7.87}$$

(Note that $\bar{q}_t = \bar{q}_{t+1/3}$.) Recall also the notation of Definition 7.10, and let

$$\underline{\delta}(t + 1/3) \equiv \begin{pmatrix} \delta_e(t) \\ \dot{\delta}_e(t + 1/3) \\ \ddot{\delta}_e(t + 1/3) \\ {}^m\delta_e(t + 1/3) \\ {}^m\dot{\delta}_e(t + 1/3) \end{pmatrix} \equiv \mathrm{E\ddot{R}R}(u_{e,t-2/3}, u_{e,t+1/3}). \tag{7.88}$$

The measure $v_{e,t-2/3}$ is judicious on average for all $t \geq 0$ (and fully judicious for $t = 0$). It then follows by Corollary 7.14 (whose conditions are satisfied, in view of (7.84)) that

$$\underline{\delta}(t + 1/3) \leq k^{O(1)} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2^{-k\zeta} \\ 2^{-k\zeta} \end{pmatrix} \mathrm{err}_v(t) + k^{O(1)} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e(t) \\ \dot{\epsilon}_e(t) \\ \min\{\ddot{\epsilon}_e(t), 1\}/2^{k\zeta} \end{pmatrix}, \tag{7.89}$$

with $\mathrm{err}_v(t)$ defined analogously to (7.27). Now, recalling (7.86), we will consider also an intermediate measure $v_{e\bullet,t}$, which we define by

$$v_{e\bullet,t}(\sigma, L) = \frac{1}{\bar{z}_{e\bullet,t}} \psi_{e,t}(\sigma, L) \hat{q}_{e,t}(\sigma, L) u_{e,t-2/3}(\sigma). \tag{7.90}$$

Let $\chi_{e\bullet,t}$, $\kappa_{e\bullet,t}$, etc. be defined analogously to the quantities in Part 2 with $\nu_{e\bullet,t}$ in place of $\nu_{e,t}$. In this part of the proof we shall abbreviate (recalling (7.75) and (7.76))

$$\begin{pmatrix} \kappa_{e\bullet}(\tau) \\ \alpha_{e\bullet}(\sigma) \end{pmatrix} \equiv \begin{pmatrix} \max_{L,j}\{|(\kappa_{e\bullet,t})^j(\tau\,|\,L)|\} \\ \max_L\{|\alpha_{e\bullet,t}(\sigma\,|\,L)|\} \end{pmatrix}, \tag{7.91}$$

Comparing with $\nu_{e,t+1/3}$ gives

$$\gamma_{e,t}(\sigma) \equiv \frac{\nu_{e,t+1/3}(\sigma,L)}{\nu_{e\bullet,t}(\sigma,L)} - 1 = \frac{\bar{z}_{e\bullet,t}}{\bar{z}_{e,t+1/3}} \cdot \frac{u_{e,t+1/3}(\sigma)}{u_{e,t-2/3}(\sigma)} - 1, \tag{7.92}$$

where we emphasize that this error does not depend on L. As a result, marginally on σ we also have

$$\gamma_{e,t}(\sigma) = \frac{\nu_{e,t+1/3}(\sigma)}{\nu_{e\bullet,t}(\sigma)} - 1. \tag{7.93}$$

Substituting (7.89) into the expression (7.92) for $\gamma_{e,t}$ gives

$$\left|\gamma_{e,t}(\sigma)\right| \le k^{O(1)}\left\{\mathrm{err}_v(t) + \mathbf{1}\{\mathfrak{r}[\sigma] \ge 1\}\left(\dot{\epsilon}_e(t) + \frac{\min\{\ddot{\epsilon}_e(t), 1\}}{2^{k\zeta}}\right)\right\}. \tag{7.94}$$

We then compare the marginal laws of (σ^1, L) under $\nu_{e\bullet,t}$ and $\nu_{e,t+1/3}$: by the definition of $\gamma_{e,t}$, we have

$$\nu_{e,t+1/3}(\sigma^1, L) - \nu_{e\bullet,t}(\sigma^1, L) = \sum_{\sigma^2} \nu_{e\bullet,t}(\sigma, L)\gamma_{e,t}(\sigma)$$

$$= \nu_{e\bullet,t}(L)\star\pi_e(\sigma^1)\left\{\underbrace{\left[\sum_{\sigma^2} \frac{\nu_{e\bullet,t}(\sigma)}{\star\pi_e(\sigma^1)}\gamma_{e,t}(\sigma)\right]}_{\text{denote this }(X_{e,t})^1(\sigma^1)} + \underbrace{\left[\sum_{\sigma^2} \frac{\nu_{e\bullet,t}(\sigma\,|\,L) - \nu_{e\bullet,t}(\sigma)}{\star\pi_e(\sigma^1)}\gamma_{e,t}(\sigma)\right]}_{\text{denote this }(Y_{e,t})^1(\sigma^1\,|\,L)}\right\}. \tag{7.95}$$

We note that the X term in (7.95) can be simplified using (7.93) and the fact that $\nu_{e,t+1/3}$ is judicious on average:

$$(X_{e,t})^1(\sigma^1) = \sum_{\sigma^2} \frac{\nu_{e,t+1/3}(\sigma) - \nu_{e\bullet,t}(\sigma)}{\star\pi_e(\sigma^1)} = 1 - \frac{(\nu_{e\bullet,t})^1(\sigma^1)}{\star\pi_e(\sigma^1)} = -(\chi_{e\bullet,t})^1(\sigma^1).$$

It follows from the definition of χ that for all $\tau \in \{\mathfrak{r}, \mathfrak{y}, \mathfrak{g}, \mathfrak{b}\}$ we have

$$\left|(X_{e,t})^1(\tau)\right| = \left|(\chi_{e\bullet,t})^1(\tau)\right| \le \max_L\left\{\left|(\kappa_{e\bullet,t})^1(\tau\,|\,L)\right|\right\} = \kappa_{e\bullet}(\tau).$$

Summing (7.95) over $\sigma^1 = \tau \in \{\mathfrak{r}, \mathfrak{y}, \mathfrak{g}, \mathfrak{b}\}$ and recalling (7.72) gives

$$\nu_{e,t+1/3}(L) - \nu_{e\bullet,t}(L) = \nu_{e\bullet,t}(L)\left\{\sum_\tau \star\pi_e(\tau)(Y_{e,t})^1(\tau\,|\,L)\right\}. \tag{7.96}$$

Now, the Y term in (7.95) can be bounded as

$$\left|(Y_{e,t})^1(\sigma^1\,|\,L)\right| \le \sum_{\sigma^2} \frac{\nu_{e\bullet,t}(\sigma)}{\star\pi_e(\sigma^1)}\left|\frac{\nu_{e\bullet,t}(\sigma\,|\,L)}{\nu_{e\bullet,t}(\sigma)} - 1\right|\gamma_{e,t}(\sigma).$$

It follows from (7.85) and (7.94) that if $\sigma^1 \ne \mathfrak{r}$,

$$\left|(Y_{e,t})^1(\sigma^1\,|\,L)\right| \le \frac{k^{O(1)}}{2^{k(1+\zeta)}}\left(\gamma_{e,t} + \dot{\gamma}_{e,t}\right) \le \frac{k^{O(1)}}{2^{k(1+\zeta)}}\left\{\mathrm{err}_v(t) + \dot{\epsilon}_e(t) + \frac{\min\{\dot{\epsilon}_e(t), 1\}}{2^{k\zeta}}\right\}.$$

If $\sigma^1 = \mathtt{r}$ then we recall from (7.79) that $\nu_{e\bullet,t} \cong \hat{q}_{e,t}\psi_{e,t}u_{e\bullet,t}$, from which it follows that $\nu_{e\bullet,t}(\mathtt{rr} \mid L) \le k^{O(1)}/2^{-(1+\zeta)}$. It then follows that

$$
\left| (Y_{e,t})^1(\sigma^1 \mid L) \right| \le \frac{k^{O(1)}}{2^{k\zeta}}\dot{\gamma}_{e,t} + O\left(\frac{\max_L\{\nu_{e\bullet,t}(\mathtt{rr} \mid L)\}}{\star\pi_e(\mathtt{r})}\ddot{\gamma}_{e,t} \right) \le \frac{k^{O(1)}}{2^{k\zeta}}\left(\dot{\gamma}_{e,t} + \ddot{\gamma}_{e,t} \right)
$$

$$
\le \frac{k^{O(1)}}{2^{k\zeta}}\left\{ \mathrm{err}_v(t) + \dot{\epsilon}_e(t) + \frac{\min\{\dot{\epsilon}_e(t), 1\}}{2^{k\zeta}} \right\}
$$

Substituting these bounds for Y back into (7.96) gives, for $t \ge 1$,

$$
\rho_{e,t+1/3} = \max_L\left\{ \left| \rho_{e,t+1/3}(L) \right| \right\} \le \rho_{e\bullet,t} + \frac{k^{O(1)}}{2^{k(1+\zeta)}}\left\{ \mathrm{err}_v(t) + \dot{\epsilon}_e(t) + \frac{\min\{\dot{\epsilon}_e(t), 1\}}{2^{k\zeta}} \right\}. \tag{7.97}
$$

Substituting the bounds for X and Y back into (7.95), and combining with (7.96), gives

$$
\kappa_{e,t+1/3} \le \rho_{e\bullet,t} + \kappa_{e\bullet,t} + \frac{k^{O(1)}}{2^{k(1+\zeta)}}\left\{ \mathrm{err}_v(t) + \dot{\epsilon}_e(t) + \frac{\min\{\ddot{\epsilon}_e(t), 1\}}{2^{k\zeta}} \right\},
$$

$$
\dot{\kappa}_{e,t+1/3} \le \rho_{e\bullet,t} + \dot{\kappa}_{e\bullet,t} + \frac{k^{O(1)}}{2^{k\zeta}}\left\{ \mathrm{err}_v(t) + \dot{\epsilon}_e(t) + \frac{\min\{\ddot{\epsilon}_e(t), 1\}}{2^{k\zeta}} \right\}. \tag{7.98}
$$

for all $t \ge 0$. We next turn to the question of bounding $\underline{\epsilon}(t)$ (as defined by (7.87)) for $t \ge 1$.

Part 4a. Errors in averaged messages incurred by update B. We now estimate the error between the averaged messages $\bar{q}_{e,t+1/3}$ and $\bar{q}_{e,t+2/3}$ for $t \ge 0$. Between times $t + 1/3$ and $t + 2/3$ we only make the update (7.69). In this part of the proof we will fix $e \in \delta v$ and abbreviate (recalling (7.75) and (7.76))

$$
\begin{pmatrix} \kappa_e(\tau) \\ \alpha_e(\sigma) \end{pmatrix} \equiv \begin{pmatrix} \kappa_{e,t+1/3}(\tau) \\ \alpha_{e,t+1/3}(\sigma) \end{pmatrix} = \begin{pmatrix} \kappa_{e,t+1/3}(\tau) \\ \alpha_{e,t}(\sigma) \end{pmatrix}.
$$

By the definition (7.71) of κ, it holds for all $\tau \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$ that

$$
0 = \sum_L \nu_{e,t+1/3}(L)\left\{ (\nu_{e,t+1/3})^j(\tau \mid L) - \star\pi_e(\tau) \right\} = \sum_L \nu_{e,t+1/3}(L)\star\pi_e(\tau)(\kappa_{e,t+1/3})^j(\tau \mid L), \tag{7.99}
$$

where the first equality holds because $\nu_{e,t+1/3}$ is judicious on average. We now turn to comparing the messages. At time $t + 1/3$ we have $\bar{q}_{e,t+1/3} = \bar{q}_{e,t}$, defined by (7.67). More explicitly, it follows using (7.78) that

$$
\bar{z}_{e,t}\bar{q}_{e,t}(\sigma) = \sum_L \psi_{e,t}(\sigma, L)\hat{q}_e(\sigma, L) = \frac{\bar{z}_{e,t+1/3}\nu_{e,t+1/3}(\sigma)}{u_{e,t+1/3}(\sigma)}, \tag{7.100}
$$

On the other hand, at time $t + 2/3$ we have (by Taylor expansion)

$$
\bar{z}_{e,t+2/3}\bar{q}_{e,t+2/3}(\sigma) = \sum_L \psi_{e,t}(\sigma, L)\hat{q}_e(\sigma, L) \prod_{j=1,2} \frac{1}{1 + (\kappa_{e,t+1/3})^j(\sigma^j \mid L)}
$$

$$
= \bar{z}_{e,t}\left\{ \bar{q}_{e,t}(\sigma)\left\{ 1 + O\left(\sum_{j=1,2} \kappa_e(\sigma^j)^2 \right) \right\} - \sum_{j=1,2} \underbrace{\left[\sum_L \frac{\psi_{e,t}(\sigma, L)\hat{q}_e(\sigma, L)}{\bar{z}_{e,t}}(\kappa_{e,t+1/3})^j(\sigma^j \mid L) \right]}_{\text{denote this } K^j(\sigma)} \right\} \tag{7.101}
$$

From the definition of $K^j(\sigma)$ we can easily deduce the bound

$$
\left| K^j(\sigma) \right| \le O\left(\frac{\kappa_e(\sigma^j)}{\bar{z}_{e,t}\hat{Z}_e}\bar{q}_{e,t}(\sigma) \sum_L \star\psi_e(L) \right) = O\left(\bar{q}_{e,t}(\sigma) \cdot \kappa_e(\sigma^j) \right), \tag{7.102}
$$

where the last step uses that $\bar{z}_{e,t}\hat{Z}_e = \Theta(1)$ by a similar calculation as in (7.80). On the other hand, we also have

$$\left|K^j(\sigma)\right| = \left|\frac{{}_{\star}\hat{q}_e(\sigma)}{\hat{Z}_e} \sum_L {}_{\star}\psi_e(L)\left(1 + \alpha_{e,t}(\sigma, L)\right)(\kappa_{e,t+1/3})^j(\sigma^j \mid L)\right|$$

$$\overset{(7.99)}{\leq} O\left({}_{\star}\hat{q}_e(\sigma) \cdot \kappa_e(\sigma^j) \cdot \alpha_e(\sigma)\right). \tag{7.103}$$

Combining (7.102) with (7.103) gives

$$K^j(\sigma) = O\left(\bar{q}_{e,t}(\sigma) \cdot \kappa_e(\sigma^j) \cdot \min\left\{\frac{\alpha_e(\sigma)_{\star}\hat{q}_e(\sigma)}{\bar{q}_{e,t}(\sigma)}, 1\right\}\right).$$

Substituting this bound into (7.101) gives

$$\bar{z}_{e,t+2/3}\bar{q}_{e,t+2/3}(\sigma) = \bar{z}_{e,t}\bar{q}_{e,t}(\sigma)\left\{1 + \sum_{j=1,2} \kappa_e(\sigma^j) \cdot \min\left\{\frac{\alpha_e(\sigma)_{\star}\hat{q}_e(\sigma)}{\bar{q}_{e,t}(\sigma)}, 1\right\}\right\}.$$

It then follows using the hypothesis (7.77) that

$$\left|\frac{\bar{z}_{e,t+2/3}\bar{q}_{e,t+2/3}(\sigma)}{\bar{z}_{e,t}\bar{q}_{e,t}(\sigma)} - 1\right| \leq k^{O(1)}\begin{cases} \dfrac{\kappa}{2^{k(1+\zeta)}} & \text{if } \mathfrak{r}[\sigma] = 0, \\[2mm] \dfrac{\kappa + \dot{\kappa}}{2^{k\zeta}} & \text{if } \mathfrak{r}[\sigma] = 1, \\[2mm] \dot{\kappa} & \text{if } \mathfrak{r}[\sigma] = 2. \end{cases}$$

Altogether, the error in averaged messages between times $t + 1/3$ and $t + 2/3$ satisfies

$$\widehat{\mathrm{ERR}}(\bar{q}_{e,t+1/3}, \bar{q}_{e,t+2/3}) = \widehat{\mathrm{ERR}}(\bar{q}_{e,t}, \bar{q}_{e,t+2/3}) \leq \frac{k^{O(1)}}{2^{k\zeta}}\begin{pmatrix} 2^{-k} & 2^{-k} \\ 1 & 1 \\ 2^{-k} & 2^{k\zeta} \end{pmatrix}\begin{pmatrix} \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix}. \tag{7.104}$$

Substituting (7.104) into the definition of u from (7.78) gives

$$\left|\frac{u_{e,t+2/3}(\sigma)}{u_{e,t+1/3}(\sigma)} - 1\right| \leq \frac{k^{O(1)}}{2^{k(1+\zeta)}}\left\{\underbrace{\sum_{e\in\delta v}\kappa_{e,t+1/3}}_{\text{denote this }\kappa_{t+1/3}} + \underbrace{\sum_{e\in\delta v}\dot{\kappa}_{e,t+1/3}}_{\text{denote this }\dot{\kappa}_{t+1/3}}\right\} \equiv \frac{k^{O(1)}(\kappa_{t+1/3} + \dot{\kappa}_{t+1/3})}{2^{k(1+\zeta)}} \tag{7.105}$$

for all $\sigma \in \{\mathbf{r}, \mathbf{y}, \mathbf{g}, \mathbf{b}\}^2$.

Part 4b. Errors in averaged messages incurred by update C. By a similar (but somewhat simpler) calculation, we can also estimate the error between the averaged messages $\bar{q}_{e,t+2/3}$ and $\bar{q}_{e,t+1}$ for $t \geq 0$. Between times $t + 2/3$ and $t + 1$ we only make the update (7.70). In this part of the proof we will fix $e \in \delta v$ and abbreviate

$$\begin{pmatrix} \kappa_e(\tau) \\ \alpha_e(\sigma) \\ \rho_e \end{pmatrix} \equiv \begin{pmatrix} \kappa_{e,t+2/3}(\tau) \\ \alpha_{e,t+2/3}(\sigma) \\ \rho_{e,t+2/3} \end{pmatrix}.$$

By the definition (7.73) of ρ, we have

$$0 = \sum_L\left\{\pi_{\mathscr{D}}(L \mid t_e) - \nu_{e,t+2/3}(L)\right\} = \sum_L \pi_{\mathscr{D}}(L \mid t_e)\left\{\rho_{e,t+2/3}(L) + O\left((\rho_e)^2\right)\right\}. \tag{7.106}$$

We now turn to the comparison of the messages: by the update rule (7.70),

$$\bar{z}_{e,t+1}\bar{q}_{e,t+1}(\sigma) = \sum_L \psi_{e,t+2/3}(\sigma, L)\hat{q}_e(\sigma, L)\Big\{1 + \rho_{e,t+2/3}(L)\Big\}$$

$$= \bar{z}_{e,t+2/3}\Bigg\{\bar{q}_{e,t+2/3}(\sigma) + \underbrace{\Bigg[\sum_L \frac{\psi_{e,t+2/3}(\sigma, L)\hat{q}_e(\sigma, L)}{\bar{z}_{e,t+2/3}}\rho_{e,t+2/3}(L)\Bigg]}_{\text{denote this } R(\sigma)}\Bigg\}.$$

From the definition of $R(\sigma)$ we can easily deduce the bound

$$\Big|R(\sigma)\Big| \leq O\Big(\bar{q}_{e,t+2/3}(\sigma) \cdot \rho_e\Big).$$

On the other hand we also have

$$R(\sigma) = \frac{{}_*\hat{q}_e(\sigma)}{\bar{z}_{e,t+2/3}\hat{Z}_e}\sum_L {}_*\psi_e(L)\Big(1 + \alpha_{e,t+1/3}(\sigma, L)\Big)\rho_{e,t+2/3}(L)$$

$$= O\Bigg({}_*\hat{q}_e(\sigma) \cdot \rho_e \cdot \Big(\rho_e + \alpha_e(\sigma)\Big)\Bigg),$$

where the last step uses (7.106). Substituting back into the preceding calculation gives

$$\bar{z}_{e,t+1}\bar{q}_{e,t+1}(\sigma) = \bar{z}_{e,t+2/3}\bar{q}_{e,t+2/3}(\sigma)\Bigg\{1 + O\Bigg(\rho_e \cdot \min\Bigg\{\frac{(\rho_e + \alpha_e(\sigma)){}_*\hat{q}_e(\sigma)}{\bar{q}_{e,t+2/3}(\sigma)}, 1\Bigg\}\Bigg)\Bigg\}.$$

It follows using the inductive bounds (7.77) that

$$\widehat{\mathrm{ERR}}(\bar{q}_{e,t+2/3}, \bar{q}_{e,t+1}) \leq \frac{k^{O(1)}}{2^{k\zeta}}\rho_{e,t+2/3}\begin{pmatrix}2^{-k}\\1\\2^{k\zeta}\end{pmatrix}. \tag{7.107}$$

Substituting this again into the definition (7.78) of u gives

$$\left|\frac{u_{e,t+1}(\sigma)}{u_{e,t+2/3}(\sigma)} - 1\right| \leq \frac{k^{O(1)}}{2^{k(1+\zeta)}}\sum_{e\in\delta v}\rho_{e,t+2/3} \equiv \frac{k^{O(1)}}{2^{k(1+\zeta)}}\boldsymbol{\rho}_{t+2/3} \tag{7.108}$$

for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$.

Part 5a. Errors in marginals incurred by update B. We now estimate the change in the edge marginals between times $t + 1/3$ and $t + 2/3$, resulting from update (7.69). To this end we introduce an intermediate measure $\nu_{e,t+1/2}$, in which we use the new weight $\psi_{e,t+1}(\sigma \mid L)$ but the old message $u_{e,t+1/3}$. Thus, recalling (7.78), we have

$$\bar{z}_{e,t+1/3}\nu_{e,t+1/3}(\sigma, L) = \hat{q}_e(\sigma, L)\psi_{e,t}(L)\psi_{e,t}(\sigma \mid L)u_{e,t+1/3}(\sigma),$$

$$\bar{z}_{e,t+1/2}\nu_{e,t+1/2}(\sigma, L) = \hat{q}_e(\sigma, L)\psi_{e,t}(L)\psi_{e,t+1}(\sigma \mid L)u_{e,t+1/3}(\sigma),$$

$$\bar{z}_{e,t+2/3}\nu_{e,t+2/3}(\sigma, L) = \hat{q}_e(\sigma, L)\psi_{e,t}(L)\psi_{e,t+1}(\sigma \mid L)u_{e,t+2/3}(\sigma).$$

We first estimate the error between times $t + 1/3$ and $t + 2/3$. For this purpose, note that (7.71) implies (similarly to the relation (7.72) that was used earlier) that for all L,

$$0 = \sum_\tau \Big\{(\nu_{e,t+1/3})^j(\tau \mid L) - {}_*\pi_e(\tau)\Big\} = \sum_\tau \Bigg\{(\nu_{e,t+1/3})^j(\tau \mid L) - \frac{(\nu_{e,t+1/3})^j(\tau \mid L)}{1 + (\kappa_{e,t+1/3})^j(\tau \mid L)}\Bigg\}$$

$$= \sum_\tau (\nu_{e,t+1/3})^j(\tau \mid L)(\kappa_{e,t+1/3})^j(\tau \mid L) + O\Bigg(\sum_\tau {}_*\pi_e(\tau)\kappa_{e,t+1/3}(\tau)^2\Bigg). \tag{7.109}$$

It follows using (7.109) that the marginal error on L between times $t + 1/3$ and $t + 1/2$ is given by

$$\bar{z}_{e,t+1/2}v_{e,t+1/2}(L) = \sum_{\sigma}\bar{z}_{e,t+1/3}v_{e,t+1/3}(\sigma, L)\prod_{j=1,2}\frac{1}{1 + (\kappa_{e,t+1/3})^j(\sigma^j \mid L)}$$

$$= \bar{z}_{e,t+1/3}v_{e,t+1/3}(L)\left\{1 + O\left(\sum_{\tau}{}_{\star}\pi_e(\tau)\kappa_{e,t+1/3}(\tau)^2\right)\right\}. \qquad (7.110)$$

We note the error term can be simplified using (7.77) as

$$\sum_{\tau}{}_{\star}\pi_e(\tau)\kappa_{e,t+1/3}(\tau)^2 \le k^{O(1)}\left((\kappa_{e,t+1/3})^2 + \frac{(\dot{\kappa}_{e,t+1/3})^2}{2^k}\right) \le \frac{k^{O(1)}(\kappa_{e,t+1/3} + \dot{\kappa}_{e,t+1/3})}{2^{k(1+\zeta)}}.$$

The error between $u_{e,t+1/3}$ and $u_{e,t+2/3}$ is bounded by (7.105), and this gives the marginal error on L between times $t + 1/2$ and $t + 2/3$. Combining these bounds gives altogether

$$\rho_{e,t+2/3} - \rho_{e,t+1/3} \le \left|\frac{v_{e,t+2/3}(L)}{v_{e,t+1/3}(L)} - 1\right| \le \frac{k^{O(1)}(\kappa_{t+1/3} + \dot{\kappa}_{t+1/3})}{2^{k(1+\zeta)}}. \qquad (7.111)$$

Next we argue that the update (7.69) in fact brings the measures closer to being conditionally judicious. Indeed, focusing on the first copy $j = 1$, we can express

$$\frac{\bar{z}_{e,t+1/2}v_{e,t+1/2}(\sigma, L)}{\bar{z}_{e,t+1/3}v_{e,t+1/3}(L)} = \frac{{}_{\star}\pi_e(\sigma^1)}{(v_{e,t+1/3})^1(\sigma^1 \mid L)}\frac{v_{e,t+1/3}(\sigma \mid L)}{1 + (\kappa_{e,t+1/3})^2(\sigma^2 \mid L)},$$

and summing over σ^2 gives

$$\frac{\bar{z}_{e,t+1/2}v_{e,t+1/2}(\sigma^1, L)}{\bar{z}_{e,t+1/3}v_{e,t+1/3}(L)} = \frac{{}_{\star}\pi_e(\sigma^1)}{(v_{e,t+1/3})^1(\sigma^1 \mid L)}\sum_{\sigma^2}v_{e,t+1/3}(\sigma \mid L)\left\{1 + O(\kappa_{e,t+1/3}(\sigma^2)^2) - (\kappa_{e,t+1/3})^2(\sigma^2 \mid L)\right\}$$

$$= {}_{\star}\pi_e(\sigma^1)\left\{1 + O\left(\sum_{\tau}{}_{\star}\pi_e(\tau)\kappa_{e,t+1/3}(\tau)^2\right) - \sum_{\sigma^2}\frac{v_{e,t+1/3}(\sigma \mid L)}{(v_{e,t+1/3})^1(\sigma^1 \mid L)}(\kappa_{e,t+1/3})^2(\sigma^2 \mid L)\right\}$$

$$= {}_{\star}\pi_e(\sigma^1)\left\{1 + O\left(\sum_{\tau}{}_{\star}\pi_e(\tau)\kappa_{e,t+1/3}(\tau)^2\right) - \underbrace{\sum_{\sigma^2}\frac{v_{e,t+1/3}(\sigma \mid L)}{{}_{\star}\pi_e(\sigma^1)}(\kappa_{e,t+1/3})^2(\sigma^2 \mid L)}_{\text{denote this } S^1(\sigma^1 \mid L)}\right\}.$$

We then use (7.77) and (7.99) to estimate (with $\kappa \equiv \kappa_{e,t+1/3}$ and $\xi \equiv \xi_{e,t+1/3}$)

$$S(\sigma^1 \mid L) = \sum_{\sigma^2}\left\{1 + \xi_{e,t+1/3}(\sigma \mid L)\right\}{}_{\star}\pi_e(\sigma^2)(\kappa_{e,t+1/3})^2(\sigma^2 \mid L)$$

$$\le \begin{cases} k^{O(1)}\left(\kappa\xi + \dfrac{\dot{\kappa}\dot{\xi}}{2^k}\right) \le \dfrac{k^{O(1)}}{2^{k\zeta}}\left(\kappa + \dfrac{\dot{\kappa}}{2^k}\right) & \text{if } \sigma^1 \ne \mathtt{r}\,, \\[3mm] k^{O(1)}\left(\kappa\dot{\xi} + \dfrac{\dot{\kappa}\ddot{\xi}}{2^k}\right) \le \dfrac{k^{O(1)}}{2^{k\zeta}}(\kappa + \dot{\kappa}) & \text{if } \sigma^1 = \mathtt{r}\,. \end{cases}$$

By substituting this into the preceding calculation and combining with (7.110), we conclude that

$$\begin{pmatrix}\kappa_{e,t+1/2} \\ \dot{\kappa}_{e,t+1/2}\end{pmatrix} \le \frac{k^{O(1)}}{2^{k\zeta}}\begin{pmatrix}1 & 2^{-k} \\ 1 & 1\end{pmatrix}\begin{pmatrix}\kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3}\end{pmatrix}.$$

We combine this with the error incurred by going from time $t + 1/2$ to $t + 2/3$ — this results from the change in the message u, which again is bounded by (7.105). Altogether we conclude

$$\begin{pmatrix}\kappa_{e,t+2/3} \\ \dot{\kappa}_{e,t+2/3}\end{pmatrix} \le \frac{k^{O(1)}}{2^{k\zeta}}\begin{pmatrix}1 & 2^{-k} \\ 1 & 1\end{pmatrix}\begin{pmatrix}\kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3}\end{pmatrix} + \frac{k^{O(1)}(\kappa_{t+1/3} + \dot{\kappa}_{t+1/3})}{2^{k(1+\zeta)}}\begin{pmatrix}1 \\ 1\end{pmatrix} \qquad (7.112)$$

for all $e \in \delta v$.

Part 5b. Errors in marginals incurred by update C. We now estimate the change in edge marginals between times $t + 2/3$ and $t + 1$, resulting from update (7.70). To this end, we can express

$$\bar{z}_{e,t+1}\nu_{e,t+1}(\sigma, L) = \bar{z}_{e,t+2/3}\nu_{e,t+2/3}(\sigma, L)\frac{\pi_{\mathscr{D}}(L \mid t_e)}{\nu_{e,t+2/3}(L)}\frac{u_{e,t+1}(\sigma)}{u_{e,t+2/3}(\sigma)}$$

$$= \bar{z}_{e,t+2/3}\nu_{e,t+2/3}(\sigma, L)\frac{\pi_{\mathscr{D}}(L \mid t_e)}{\nu_{e,t+2/3}(L)}\left\{1 + O\left(\frac{k^{O(1)}\rho_{t+2/3}}{2^{k(1+\zeta)}}\right)\right\},$$

where the last estimate comes from (7.108). Summing over σ gives

$$\nu_{e,t+1}(L) = \frac{\bar{z}_{e,t+2/3} \cdot \pi_{\mathscr{D}}(L \mid t_e)}{\bar{z}_{e,t+1}}\left\{1 + O\left(\frac{k^{O(1)}\rho_{t+2/3}}{2^{k(1+\zeta)}}\right)\right\} = \pi_{\mathscr{D}}(L \mid t_e)\left\{1 + O\left(\frac{k^{O(1)}\rho_{t+2/3}}{2^{k(1+\zeta)}}\right)\right\}, \tag{7.113}$$

where the last equality is because both $\pi_{\mathscr{D}}(L \mid t_e)$ and $\nu_{e,t+1}(L)$ are probability measures over L. On the other hand, the conditional measures given L change very little as a result of update (7.70): it follows from the above that

$$\nu_{e,t+1}(\sigma \mid L) = \nu_{e,t+2/3}(\sigma \mid L)\left\{1 + O\left(\frac{k^{O(1)}\rho_{t+2/3}}{2^{k(1+\zeta)}}\right)\right\} \tag{7.114}$$

for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2$. Combining (7.113) and (7.114) gives

$$\begin{pmatrix} \rho_{e,t+1} \\ \kappa_{e,t+1} \\ \dot{\kappa}_{e,t+1} \end{pmatrix} \leq \begin{pmatrix} 0 \\ \kappa_{e,t+2/3} \\ \dot{\kappa}_{e,t+2/3} \end{pmatrix} + \frac{k^{O(1)}}{2^{k(1+\zeta)}}\rho_{t+2/3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{7.115}$$

Finally, recalling (7.90), we compare the measures

$$\bar{z}_{e,t+1}\nu_{e,t+1}(\sigma, L) = \psi_{e,t+1}(\sigma, L)\hat{q}_e(\sigma, L)u_{e,t+1}(\sigma),$$

$$\bar{z}_{e\bullet,t+1}\nu_{e\bullet,t+1}(\sigma, L) = \psi_{e,t+1}(\sigma, L)\hat{q}_e(\sigma, L)u_{e,t+1/3}(\sigma).$$

It follows by combining (7.105) and (7.108) that

$$\begin{pmatrix} \rho_{e\bullet,t+1} \\ \kappa_{e\bullet,t+1} \\ \dot{\kappa}_{e\bullet,t+1} \end{pmatrix} \leq \begin{pmatrix} \rho_{e,t+1} \\ \kappa_{e,t+1} \\ \dot{\kappa}_{e,t+1} \end{pmatrix} + \frac{k^{O(1)}}{2^{k(1+\zeta)}}\left(\kappa_{t+1/3} + \dot{\kappa}_{t+1/3} + \rho_{t+2/3}\right)\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{7.116}$$

This concludes our analysis of update C.

Part 6. Convergence of iterative procedure. We now collect the bounds obtained above to prove that the iterative procedure described in Part 1 converges. Recall from (7.87) the definition of $\underline{\epsilon}(t)$. For $t \geq 1$, it follows by combining (7.104) (from Part 4a) and (7.107) (from Part 4b) that we have

$$\underline{\epsilon}(t) \leq \frac{k^{O(1)}}{2^{k\zeta}}\begin{pmatrix} 2^{-k} & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \\ 2^{k\zeta} & 2^{-k} & 2^{k\zeta} \end{pmatrix}\begin{pmatrix} \rho_{e,t-1/3} \\ \kappa_{e,t-2/3} \\ \dot{\kappa}_{e,t-2/3} \end{pmatrix}. \tag{7.117}$$

It will be useful to abbreviate $K \equiv \kappa + \dot{\kappa}$. Recall that $\mathrm{err}_v(t)$ is defined analogously to (7.27), so

$$\mathrm{err}_v(t) \leq \sum_{e \in \delta v} \frac{k^{O(1)}}{2^{k\zeta}}\begin{pmatrix} 1 & 2^{-k} & 2^{-k(1+\zeta)} \end{pmatrix}\begin{pmatrix} 2^{-k} & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \\ 2^{k\zeta} & 2^{-k} & 2^{k\zeta} \end{pmatrix}\begin{pmatrix} \rho_{e,t-1/3} \\ \kappa_{e,t-2/3} \\ \dot{\kappa}_{e,t-2/3} \end{pmatrix} \leq \frac{k^{O(1)}(K_{t-2/3} + \rho_{t-1/3})}{2^{k(1+\zeta)}}.$$

We then substitute these bounds into the analysis of update A, from Part 3b: from (7.97) and (7.98) we obtain

$$\begin{pmatrix} \rho_{e,t+1/3} \\ \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix} \leq \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} \rho_{e\bullet,t} \\ \kappa_{e\bullet,t} \\ \dot{\kappa}_{e\bullet,t} \end{pmatrix} + \frac{k^{O(1)}}{2^{k(1+\zeta)}}\left\{\frac{\rho_{t-1/3} + K_{t-2/3}}{2^{k(1+\zeta)}} + \frac{\kappa_{e,t-2/3} + \dot{\kappa}_{e,t-2/3} + \rho_{e,t-1/3}}{2^{k\zeta}}\right\}\begin{pmatrix} 1 \\ 1 \\ 2^k \end{pmatrix}. \tag{7.118}$$

Next, from the analysis of the marginal errors resulting from update B (Part 5a), we have the bounds (7.111) and (7.112), which we recall give

$$\begin{pmatrix} \rho_{e,t+2/3} \\ \kappa_{e,t+2/3} \\ \dot{\kappa}_{e,t+2/3} \end{pmatrix} \leq \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{k\zeta} & 0 & 0 \\ 0 & 1 & 2^{-k} \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \rho_{e,t+1/3} \\ \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix} + \frac{k^{O(1)} K_{t+1/3}}{2^{k(1+\zeta)}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{7.119}$$

From the analysis of marginal errors resulting from update C (Part 5b), we have the bounds (7.115) and (7.116), which combine to give

$$\begin{pmatrix} \rho_{e\bullet,t+1} \\ \kappa_{e\bullet,t+1} \\ \dot{\kappa}_{e\bullet,t+1} \end{pmatrix} \leq \begin{pmatrix} 0 \\ \kappa_{e,t+2/3} \\ \dot{\kappa}_{e,t+2/3} \end{pmatrix} + \frac{k^{O(1)}}{2^{k(1+\zeta)}} \Big(\rho_{t+2/3} + K_{t+1/3} \Big) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{7.120}$$

Combining the above bounds gives, for all $t \geq 1$,

$$\begin{pmatrix} \rho_{e\bullet,t} \\ \kappa_{e\bullet,t} \\ \dot{\kappa}_{e\bullet,t} \end{pmatrix} \overset{(7.120)}{\leq} \begin{pmatrix} 0 \\ \kappa_{e,t-1/3} \\ \dot{\kappa}_{e,t-1/3} \end{pmatrix} + \frac{k^{O(1)}}{2^{k(1+\zeta)}} \Big(\rho_{t-1/3} + K_{t-2/3} \Big) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\overset{(7.119)}{\leq} \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2^{-k} \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \rho_{e,t-2/3} \\ \kappa_{e,t-2/3} \\ \dot{\kappa}_{e,t-2/3} \end{pmatrix} + \frac{k^{O(1)}}{2^{k(1+\zeta)}} \Big(\rho_{t-1/3} + K_{t-2/3} \Big) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Substituting these bounds (along with (7.119)) into (7.118) gives

$$\begin{pmatrix} \rho_{e,t+1/3} \\ \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix} \leq \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \rho_{e,t-2/3} \\ \kappa_{e,t-2/3} \\ \dot{\kappa}_{e,t-2/3} \end{pmatrix} + \frac{\Upsilon_{t-2/3}}{2^k} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}, \tag{7.121}$$

where $\Upsilon \equiv \rho + K$. (In the 3×3 matrix on the right-hand side of (7.121), some entries are zero because we absorbed the errors into the $\Upsilon_{t-2/3}$ term.) Aggregating the last bound over $e \in \delta v$ gives

$$\Upsilon_{t+1/3} \equiv \rho_{t+1/3} + K_{t+1/3} \leq \frac{k^{O(1)}}{2^{k\zeta}} \Big\{ \rho_{t+1/3} + K_{t+1/3} \Big\} = \frac{k^{O(1)}}{2^{k\zeta}} \Upsilon_{t-2/3},$$

for all $t \geq 1$. This implies that the iterative procedure of Part 1 converges to the desired weights $\Psi_v \equiv \Psi_{v,\infty}$. We now turn to proving that the weights satisfy the bounds (7.27) and (7.62). Summing (7.121) over $t \geq 1$ gives

$$S_{\geq 1} \equiv \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \sum_{t \geq 1} \begin{pmatrix} \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix} \leq \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \sum_{t \geq 1} \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \kappa_{e,t-2/3} \\ \dot{\kappa}_{e,t-2/3} \end{pmatrix} + \frac{\Upsilon_{t-2/3}}{2^k} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

$$\leq \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \sum_{t \geq 1} \begin{pmatrix} \kappa_{e,1/3} \\ \dot{\kappa}_{e,1/3} \end{pmatrix} + \frac{\Upsilon_{1/3}}{2^k} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + S_{\geq 1} \right\},$$

and rearranging this inequality gives an upper bound on $S_{\geq 1}$. It follows using (7.121) again that

$$\sum_{t \geq 1} \begin{pmatrix} \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix} \leq \begin{pmatrix} \kappa_{e,1/3} \\ \dot{\kappa}_{e,1/3} \end{pmatrix} + \frac{k^{O(1)}}{2^{k\zeta}} \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \kappa_{e,1/3} \\ \dot{\kappa}_{e,1/3} \end{pmatrix} + \frac{\Upsilon_{1/3}}{2^k} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + S_{\geq 1} \right\}$$

$$\leq k^{O(1)} \left\{ \begin{pmatrix} 1 & 0 \\ 2^{-k\zeta} & 1 \end{pmatrix} \begin{pmatrix} \kappa_{e,1/3} \\ \dot{\kappa}_{e,1/3} \end{pmatrix} + \frac{\Upsilon_{1/3}}{2^{k(1+\zeta)}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\},$$

where the last step makes use of the upper bound on $S_{\geq 1}$. We can also use (7.119) and (7.121) to bound the sum of $\rho_{e,t+2/3}$ over all $t \geq 1$, so altogether we have

$$\sum_{t \geq 1} \begin{pmatrix} \rho_{e,t+2/3} \\ \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix} \leq k^{O(1)} \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2^{-k\zeta} & 1 \end{pmatrix} \begin{pmatrix} \rho_{e,1/3} \\ \kappa_{e,1/3} \\ \dot{\kappa}_{e,1/3} \end{pmatrix} + \frac{\Upsilon_{1/3}}{2^{k(1+\zeta)}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \tag{7.122}$$

We can then apply (7.97) and (7.98) with $t = 0$ to obtain

$$\begin{pmatrix} \rho_{e,1/3} \\ \kappa_{e,1/3} \\ \dot{\kappa}_{e,1/3} \end{pmatrix} \le \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \rho_{e\bullet,0} \\ \kappa_{e\bullet,0} \\ \dot{\kappa}_{e\bullet,0} \end{pmatrix} + \frac{k^{O(1)}}{2^{k(1+\zeta)}} \left\{ \mathrm{err}_v + \dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k(1+\zeta)}} \right\} \begin{pmatrix} 1 \\ 1 \\ 2^k \end{pmatrix}.$$

Finally, since we see from (7.79) that $\nu_{e\bullet,-2/3} \cong \hat{p}_e \phi_e u_{e,-2/3}$ while $\nu_{e\bullet,0} \cong \hat{q}_e \phi_e u_{e,-2/3}$, we conclude that

$$\begin{pmatrix} \rho_{e\bullet,0} \\ \kappa_{e\bullet,0} \\ \dot{\kappa}_{e\bullet,0} \end{pmatrix} \le \underbrace{\begin{pmatrix} \rho_{e,-2/3} \\ \kappa_{e,-2/3} \\ \dot{\kappa}_{e,-2/3} \end{pmatrix}}_{\text{zero}} + \begin{pmatrix} 1 & 2^{-k} & 2^{-k} \\ 1 & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \min\{\ddot{\epsilon}_e, 1\}/2^{k\zeta} \end{pmatrix},$$

where the first term on the right-hand side vanishes because $\nu_{e,-2/3}$ is fully judicious by the assumption. Combining the last two bounds gives

$$\begin{pmatrix} \rho_{e,1/3} \\ \kappa_{e,1/3} \\ \dot{\kappa}_{e,1/3} \end{pmatrix} \le \begin{pmatrix} 1 & 2^{-k} & 2^{-k} \\ 1 & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \min\{\ddot{\epsilon}_e, 1\}/2^{k\zeta} \end{pmatrix} + \frac{k^{O(1)}\mathrm{err}_v}{2^{k(1+\zeta)}} \begin{pmatrix} 1 \\ 1 \\ 2^k \end{pmatrix}.$$

This implies $\Upsilon_{1/3} \le O(\mathrm{err}_v)$. It follows by combining with (7.122) that

$$\sum_{t\in\mathbb{Z}, t\ge 0} \begin{pmatrix} \rho_{e,t+2/3} \\ \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix} \le \begin{pmatrix} 1 & 2^{-k} & 2^{-k} \\ 1 & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \min\{\ddot{\epsilon}_e, 1\}/2^{k\zeta} \end{pmatrix} + \frac{k^{O(1)}\mathrm{err}_v}{2^{k(1+\zeta)}} \begin{pmatrix} 1 \\ 1 \\ 2^k \end{pmatrix}. \tag{7.123}$$

We remark that (7.123) implies that the quantities $\rho, \kappa, \dot{\kappa}$ satisfy the claimed bound (7.77). The quantities α, ξ can be bounded in terms of $\rho, \kappa, \dot{\kappa}$, so it is straightforward to deduce that (7.77) indeed holds. To conclude, it follows by recalling (7.69) and (7.71) that

$$\left| \frac{(\psi_{e,\infty})^j(\tau \mid L)}{(\phi_e)^j(\tau \mid L)} - 1 \right| \le O\left(\sum_{t\in\mathbb{Z}, t\ge 0} \left| \frac{(\psi_{e,t+1})^j(\tau \mid L)}{(\psi_{e,t})^j(\tau \mid L)} - 1 \right| \right) \le O\left(\sum_{t\ge 0} \kappa_{e,t+1/3}(\tau) \right).$$

It follows by recalling (7.70) and (7.73) that

$$\sum_{t\in\mathbb{Z}, t\ge 0} \left| \frac{\psi_{e,t+1}(L)}{\psi_{e,t}(L)} - 1 \right| \le O(1) \sum_{t\ge 0} \rho_{e,t+2/3}.$$

These quantities are bounded by (7.123), and this implies the claimed bound (7.62) for the clause-dependent weights. The errors in the clause-independent weights $\psi_v(\underline{\sigma}_{\delta v})$ can be bounded by applying Proposition 7.13 with input errors $\underline{\epsilon}(t)$ for integers $t \ge 0$, where (recalling (7.87)) we have

$$\underline{E} \equiv \sum_{t\in\mathbb{Z}, t\ge 0} \underline{\epsilon}(t) = \sum_{t\in\mathbb{Z}, t\ge 0} \left\{ \widehat{\mathrm{ERR}}(\bar{q}_{t-2/3}, \bar{q}_{t-1/3}) + \widehat{\mathrm{ERR}}(\bar{q}_{t-1/3}, \bar{q}_t) \right\}$$

$$\overset{(7.117)}{\le} \underline{\epsilon}(0) + \frac{k^{O(1)}}{2^{k\zeta}} \begin{pmatrix} 2^{-k} & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \\ 2^{k\zeta} & 2^{-k} & 2^{k\zeta} \end{pmatrix} \sum_{t\in\mathbb{Z}, t\ge 0} \begin{pmatrix} \rho_{e,t+2/3} \\ \kappa_{e,t+1/3} \\ \dot{\kappa}_{e,t+1/3} \end{pmatrix}$$

$$\overset{(7.123)}{\le} k^{O(1)} \begin{pmatrix} 1 & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \min\{\ddot{\epsilon}_e, 1\}/2^{k\zeta} \end{pmatrix} + \frac{k^{O(1)}\mathrm{err}_v}{2^{k\zeta}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^{k\zeta} \end{pmatrix}. \tag{7.124}$$

It follows from this that

$$E_e + \frac{\dot{E}_e}{2^k} + \frac{\min\{\ddot{E}_e, 1\}}{2^{k(1+\zeta)}} \le k^{O(1)} \left\{ \epsilon_e + \frac{\dot{\epsilon}_e}{2^k} + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k(1+\zeta)}} + \frac{\mathrm{err}_v}{2^{k(1+\zeta)}} \right\},$$

$$\dot{E}_e + \frac{\min\{\ddot{E}_e, 1\}}{2^{k\zeta}} \le k^{O(1)} \left\{ \epsilon_e + \dot{\epsilon}_e + \min\{\ddot{\epsilon}_e, 1\} + \frac{\mathrm{err}_v}{2^{k\zeta}} \right\}.$$

Thus, if we apply the bound Proposition 7.13 with input errors $\underline{\epsilon}(t)$ summed over all $t \geq 0$, the final bound will still be of the form (7.27), as claimed.

Part 7. Error bound for outgoing messages. It remains to verify (7.63). This follows by a similar argument as in Step 2 of the proof of Corollary 7.14. Fix an edge $e \in \delta v$, and define $\hat{s}_e = \hat{p}_e$ and $\hat{s}_e = \hat{q}_{e'}$ for $e' \in \delta v \setminus e$. We can apply the above result to obtain weights Θ_v such that $\nu_{\delta v}[\Theta_v; \hat{s}]$ is fully judicious. On the same edge e, the outgoing BP messages for the measures $\nu_{\delta v}[\Phi_v; \hat{p}]$, $\nu_{\delta v}[\Theta_v; \hat{s}]$, and $\nu_{\delta v}[\Psi_v; \hat{q}]$ are given by

$$\dot{p}_e(\sigma, L) \cong \phi_e(\sigma, L) u_p(\sigma)$$
$$\dot{s}_e(\sigma, L) \cong \theta_e(\sigma, L) u_s(\sigma)$$
$$\dot{q}_e(\sigma, L) \cong \psi_e(\sigma, L) u_q(\sigma)$$

where $u_p \equiv u_{e,-2/3}$, $u_q \equiv u_{e,\infty}$, and analogously

$$u_s(\sigma_e) \cong \sum_{\underline{\sigma}_{\delta a \setminus e}} \varphi_v(\underline{\sigma}_{\delta v}) \theta_v(\underline{\sigma}_{\delta v}) \prod_{e' \in \delta a \setminus e} \left\{ \sum_L \theta_{e'}(\sigma_{e'}, L_{e'}) \hat{s}_{e'}(\sigma_{e'}, L_{e'}) \right\}.$$

Conditional on L, we use (7.10) to define the reweighted messages $P_e(\sigma \mid L)$, $S_e(\sigma \mid L)$, $Q_e(\sigma \mid L)$. Recall (7.88), and note that

$$\underline{D} \equiv \text{E\"{R}R}(u_p, u_q) \leq O\left(\sum_{t \in \mathbb{Z}, t \geq 0} \text{E\"{R}R}(u_{e,t-2/3}, u_{e,t+1/3}) \right) = O\left(\sum_{t \in \mathbb{Z}, t \geq 0} \underline{\delta}(t + 1/3) \right).$$

Let $\underline{D}'$ denote the first three entries of $\underline{D}$. Then

$$\underline{D}' \overset{(7.89)}{\leq} k^{O(1)} \sum_{t \in \mathbb{Z}, t \geq 0} \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{err}_v(t) + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e(t) \\ \dot{\epsilon}_e(t) \\ \min\{\ddot{\epsilon}_e(t), 1\}/2^{k\zeta} \end{pmatrix} \right\}$$

$$\overset{(7.124)}{\leq} k^{O(1)} \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{err}_v + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_e \\ \dot{\epsilon}_e \\ \min\{\ddot{\epsilon}_e, 1\}/2^{k\zeta} \end{pmatrix} \right\}$$

Meanwhile, the errors between the edge weights $\phi_e(\sigma, L)$, $\theta_e(\sigma, L)$, and $\psi_e(\sigma, L)$ are bounded by (7.62). From this it is easy to see that

$$\max_L \left\{ \left| \frac{\dot{s}_e(L)}{\dot{p}_e(L)} - 1 \right| + \left| \frac{\dot{q}_e(L)}{\dot{p}_e(L)} - 1 \right| + \left\| \frac{S_e(\cdot \mid L)}{P_e(\cdot \mid L)} - 1 \right\|_\infty \right\} \leq k^{O(1)} \text{err}_v.$$

Combining with the result of Lemma 7.18 gives (cf. (7.43))

$$\max_L \left\{ \left| \frac{(S_e)^j(\tau \mid L)}{(P_e)^j(\tau \mid L)} - 1 \right| \right\} \leq \frac{k^{O(1)} \text{err}_v}{2^{k\zeta}}.$$

The error between S_e and Q_e is bounded by (cf. (7.44) and (7.45))

$$\max_L \left\{ \left| \frac{Q_e(\sigma \mid L)}{S_e(\sigma \mid L)} - 1 \right| \right\} \leq k^{O(1)} \left(\text{err}_v + \mathbf{1}\{\mathfrak{r}[\sigma] \geq 1\} \left(\dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} \right) \right)$$

$$\max_L \left\{ \left| \frac{(Q_e)^j(\tau \mid L)}{(S_e)^j(\tau \mid L)} - 1 \right| \right\} \leq k^{O(1)} \left(\epsilon_e + \frac{1}{(2^k)^{\mathbf{1}\{\tau \neq \mathfrak{r}\}}} \left\{ \dot{\epsilon}_e + \frac{\min\{\ddot{\epsilon}_e, 1\}}{2^{k\zeta}} \right\} \right).$$

Combining these bounds gives the claim (7.63). $\qquad \square$

Proof of Proposition 7.4. We must argue that the iterative construction of Definition 7.25 converges to the desired weights $\Psi \equiv \Psi_{\mathscr{D}}(U; \omega_{\delta U})$ asserted by the proposition. The argument closely follows the outline of the proofs of Propositions 7.3 and 7.19. For each leaf edge $e = (au) \in \delta U$, it is easy to check that at the initial stage $t = 0$ we have

$$\hat{q}_{e,0}(\sigma, L) \cong \pi_{\mathscr{D}}(L \mid t_{p(e)})_* \hat{q}_e(\sigma),$$

where $p(e) = (av)$ (as in (3.76)). On the other hand, for each $e \in \delta U$ we are given the edge marginal

$$\Omega_e(\sigma, L) = \Omega_e(L)\Omega_e(\sigma \mid L) = \pi_{\mathscr{D}}(L \mid t_{p(e)})\omega_{L,j}(\sigma),$$

with $j = j(t_e)$. Applying Step I in Definition 7.25 gives

$$\dot{q}_{e,1}(\sigma, L) \cong \frac{\Omega_e(\sigma, L)}{\hat{q}_{e,0}(\sigma, L)} \cong \frac{\omega_{L,j}(\sigma)}{{}_*\hat{q}_e(\sigma)} \,.$$

It follows that $\mathbf{E\ddot{R}R}(\dot{q}_{e,0}, \dot{q}_{e,1})$ (in the notation of Definition 7.27) can be bounded in the same way as $\text{E\ddot{R}R}(\dot{q}_{e,0}, \dot{q}_{e,1})$ is bounded by Lemma 7.21, using assumption (7.7) in place of assumption (7.5). We can then repeat the argument from the proof of Proposition 7.19, with the following modifications:

– The base case is the bound on $\mathbf{E\ddot{R}R}(\dot{q}_{e,0}, \dot{q}_{e,1})$, (as opposed to $\text{E\ddot{R}R}(\dot{q}_{e,0}, \dot{q}_{e,1})$ in the original argument);
– We apply Proposition 7.12 conditional on the clause type L; and
– We apply Proposition 7.28 in place of Proposition 7.13 and Corollary 7.14.

This modification gives that for any edge e in U, the variable-to-clause message error $\mathbf{E\ddot{R}R}(\dot{q}_{e,t-1}, \dot{q}_{e,t})$ satisfies the bound (7.47), while the clause-to-variable message error $\widehat{\mathbf{ERR}}(\hat{q}_{e,t-1}, \hat{q}_{e,t})$ satisfies the bound (7.48). Substituting these bounds into the proof of Proposition 7.3 gives the result. □

8. Solution of second moment optimization

In this section we complete the proof of the key second moment estimate Proposition 3.32, assuming an *a priori* estimate (Proposition 8.4 below) which will be proved in Section 9. For an edge $e = (av)$ in the processed graph $\mathcal{G}$ with $L_a = L$ and $j(e) = j \in [k]$, we will write $\omega_e \equiv \omega_{L,j}$. Recall that in Proposition 7.4, we let U represent the depth-one neighborhood of a non-compound variable v (see Definition 3.45), and considered the optimization problem $\nu = \nu^{\text{op}}(U; \omega_{\delta U})$ for ω satisfying (7.7). The main result (7.9) from Proposition 7.4 shows that for the optimizer ν, the discrepancy $\text{disc}_{av}(\nu)$ for $a \in \partial v$ is small relative to the maximum discrepancy $\text{disc}_e(\nu)$ over all edges $e \in \delta U$: indeed we can summarize (7.9) more simply as

$$\text{disc}_{av}(\nu) \le k^{O(1)}(\vartheta_*)^{1/4}\left\{\max_{e \in \delta U} \text{disc}_e(\omega)\right\}. \tag{8.1}$$

(Recall from Definition 7.2 that $\vartheta_* \equiv 2^{-k\zeta/6}$ for an absolute constant $0 < \zeta \le 1/20$. Recall also that the discrepancy measure disc is defined by (7.4) and (7.8) for the compound and non-compound settings, respectively.) The main technical result of this section is the analogue of Proposition 7.4 for compound regions:

Proposition 8.1 (contraction result for compound regions). *Assume that $R \ge k$, where R is the neighborhood radius in (3.1). In the same setting as Proposition 3.44, consider the constrained optimization problem*

$$\nu^{\text{op}}(U; \omega_{\delta U}) = \arg\max_\nu\left\{\mathcal{H}(\nu) : \nu \in J(U, \omega_{\delta U})\right\},$$

where U is a compound enclosure (Definition 3.16), and we assume that (cf. (7.7))

$$\max_{\sigma \in \{\mathbf{r,y,g,b}\}^2}\left\{\frac{1}{(2^k)\mathbf{1}\{\sigma=\mathbf{rr}\}}\left|\frac{\omega_e(\sigma)}{{}_*\omega_e(\sigma)} - 1\right|\right\} \le \frac{1}{2^{2k\zeta}} \tag{8.2}$$

for all $e \in \delta U$. For $\nu \equiv \nu^{\text{op}}(U; \omega_{\delta U})$, let $\text{disc}_e(\nu)$ be as in (7.4). For any edge e, define

$$C_e \equiv \begin{cases} 1 & \text{if } e \text{ is nice,} \\ C(k,R) \equiv \dfrac{1}{\min_t \min\{{}_*\pi_t(\sigma)^6 : \sigma \in \text{supp }{}_*\pi_t\}} & \text{if } e \text{ is not nice.} \end{cases} \tag{8.3}$$

(The choice of $C \equiv C(k,R)$ is very crude; but we emphasize that it is a finite constant independent of n.) Then for every edge e in $U \setminus \delta U$ we have

$$\frac{\text{disc}_e(\nu)}{C_e} \le \frac{1}{4}\left\{\max_{e \in \delta U} \text{disc}_e(\omega)\right\}, \tag{8.4}$$

provided that the constants κ_* and δ_* from Definitions 3.10 and 3.11 satisfy

$$\kappa_* \ge \left(\frac{240}{\zeta}\right)^4, \quad \delta_* \le \min\left\{\frac{\zeta}{30}, \frac{\ln 2}{2(\kappa_*)^{1/2}}\right\}.$$

We let $\Psi \equiv \Psi(U; \omega_{\delta U})$ denote the Lagrangian weights such that the solution $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ of the above coincides with $\nu[U; \Psi]$, i.e., the Ψ-weighted Gibbs measure on U for the pair coloring model. If ω is judicious and close to $_*\omega$, then ω defines a nonempty subspace $J(U; \omega_{\delta U})$ by the considerations of Lemma 3.54, so Ψ exists by the general theory of Lagrange multipliers.

The proof of Proposition 8.1 occupies most of this section, and uses the result of Proposition 7.3 as the main inductive building block. Before turning to this, we first explain in §8.1 how the results of Propositions 7.4 and 8.1 can be combined with the *a priori* estimate (Proposition 8.4) to prove Proposition 3.32.

8.1. **From contraction to optimization.** The *a priori* estimate requires some conditions on the processed neighborhood profile $\mathscr{D}$ (Definition 3.20), which we now formally introduce.

Definition 8.2 (expansion condition). In the processed graph $\mathscr{G} = (V, F, E)$, for any subset of variables $S \subseteq V$, let $F_\bullet(S)$ denote the subset of clauses $a \in F$ with $|S \cap \partial a| \geq 9k/10$. We say that a subset of variables $S \subseteq V$, is a **type-subset** if membership in S can be determined by the variable type alone — i.e., for all $v \in S$ and $w \notin S$, we have $T_v \neq T_w$. We say that $\mathscr{G}$ **expands on type-subsets** if every type-subset $S \subseteq V$ satisfies the bound

$$|F_\bullet(S)| \leq \begin{cases} |V|2^{29k/30} & \text{if } |S|/|V| \leq 4/5, \\ |S| & \text{if } |S|/|V| \leq 1/16 \end{cases} \tag{8.5}$$

Whether $\mathscr{G}$ expands on type-subsets can be determined from the neighborhood profile $\mathscr{D}$, so we say equivalently that $\mathscr{D}$ **expands on type-subsets**.

Lemma 8.3 (expansion result; proved in Section 9). *Let $\mathscr{G}'$ be the random k-SAT instance, and $\mathscr{G} = \mathrm{pr}\,\mathscr{G}'$ the processed graph, with neighborhood profile $\mathscr{D}$. It holds with high probability that $\mathscr{D}$ expands on type-subsets in the sense of Definition 8.2.*

For the next proposition, we recall that I_0 appears in Lemma 3.42, and denotes the subset of edge marginals ω that are judicious and consistent with $z \in I_0$, where I_0 is defined by (3.31).

Proposition 8.4 (*a priori* estimate; proved in Section 9). *Let $\mathscr{G}'$ be the random k-SAT instance, and $\mathscr{G} = \mathrm{pr}\,\mathscr{G}'$ the processed graph, with neighborhood profile $\mathscr{D}$. Recall Lemma 3.42, and let*

$$\omega(\mathscr{D}) \in \arg\max_\omega \left\{ \boldsymbol{\Psi}_{\mathscr{D},2}(\omega) : \omega \in I_0 \right\}. \tag{8.6}$$

(We do not yet argue that $\omega(\mathscr{D})$ is unique; that will be shown in the proof of Proposition 3.32.) There exists an absolute constant $\zeta_* > 0$ such that if $\mathscr{D}$ expands on type-subsets in the sense of Definition 8.2, then any maximizer $\omega(\mathscr{D})$ must satisfy the "a priori" bounds

$$\left| \frac{\omega_{L,j}(\sigma)}{_*\omega_{L,j}(\sigma)} - 1 \right| \leq \begin{cases} 2^{-k\zeta_*} & \text{if } \mathrm{r}[\sigma] \leq 1, \\ 2^{k(1-\zeta_*)} & \text{if } \mathrm{r}[\sigma] = 2. \end{cases} \tag{8.7}$$

whenever $L(j)$ is a strongly non-defective edge type in the sense of Remark 6.5.

We conclude this subsection by explaining how Propositions 7.4, 8.1, and 8.4 combine to give the key second moment estimate, Proposition 3.32. We begin with an elementary lemma:

Lemma 8.5 (relative entropy identity). *Let $\mathcal{A}$ be any finite set, and suppose $\nu(a) = \Lambda(a)/z$ is a probability measure over $a \in \mathcal{A}$. We then have*

$$\mathcal{H}(\mu \mid \nu) = \left\{ \mathcal{H}(\nu) - \mathcal{H}(\mu) \right\} + \left\langle \nu - \mu, \ln \Lambda \right\rangle$$

for any probability measure μ over $\mathcal{A}$.

Proof. The relative entropy of μ with respect to ν is

$$\mathcal{H}(\mu \mid \nu) = \left\langle \mu, \ln \frac{\mu}{\nu} \right\rangle = -\mathcal{H}(\mu) - \left\langle \mu, \ln \nu \right\rangle = -\mathcal{H}(\mu) - \left\langle \mu, \ln \frac{\Lambda}{z} \right\rangle = -\mathcal{H}(\mu) - \left\langle \mu, \ln \Lambda \right\rangle + \ln z \,.$$

Applying this identity with $\mu = \nu$ gives

$$0 = \mathcal{H}(\nu \mid \nu) = -\mathcal{H}(\nu) - \left\langle \nu, \ln \Lambda \right\rangle + \ln z \,.$$

Subtracting the two expressions gives the conclusion. □

Remark 8.6 (applications of Lemma 8.5). We now informally describe the two main ways that we will apply the identity obtained in Lemma 8.5. Let LIN be an affine subspace of the simplex of probability measures over $\mathcal{A}$.

A. Suppose that $\nu(a) = \Lambda(a)/z$ where $\ln \Lambda$ are the Lagrange multipliers for the subspace LIN, such that ν is the solution of the optimization problem

$$\arg\max_{\nu'} \left\{ \mathcal{H}(\nu') : \nu' \in \mathsf{LIN} \right\} . \tag{8.8}$$

If μ is any element of LIN, then Lemma 8.5 gives

$$\mathcal{H}(\mu \mid \nu) = \left\{ \underbrace{\mathcal{H}(\nu) - \mathcal{H}(\mu)}_{\text{nonnegative}} \right\} + \underbrace{\left\langle \nu - \mu, \ln \Lambda \right\rangle}_{\text{zero}} = \mathcal{H}(\nu) - \mathcal{H}(\mu) \,, \tag{8.9}$$

where the last equality uses that $\langle \mu, \ln \Lambda \rangle$ must be constant over $\mu \in \mathsf{LIN}$, by the nature of Lagrange multipliers. In the proof of Proposition 3.32, we will lower bound $\mathcal{H}(\mu \mid \nu)$, and use the above identity to deduce a lower bound on $\mathcal{H}(\nu) - \mathcal{H}(\mu)$.

B. Now suppose instead that μ solves (8.8), while $\nu(a) = \Lambda(a)/z$ also belongs to LIN, but no longer with the assumption that $\ln \Lambda$ are the Lagrange multipliers for LIN. Then Lemma 8.5 gives

$$\mathcal{H}(\mu \mid \nu) = \left\{ \underbrace{\mathcal{H}(\nu) - \mathcal{H}(\mu)}_{\text{nonpositive}} \right\} + \left\langle \nu - \mu, \ln \Lambda \right\rangle \leq \left\langle \nu - \mu, \ln \Lambda \right\rangle \,, \tag{8.10}$$

where the inequality holds since $\mathcal{H}(\mu) \geq \mathcal{H}(\nu)$ by assumption. If Λ are the Lagrange multipliers for LIN, then $\langle \mu, \ln \Lambda \rangle$ must be constant over $\mu \in \mathsf{LIN}$, so we would have $\langle \nu - \mu, \ln \Lambda \rangle = 0$ and thus $\mathcal{H}(\mu \mid \nu) = 0$, as already noted. However, in some of the applications that follow, we will only have Λ "close to" the Lagrange multipliers for LIN, such that $\langle \mu, \ln \Lambda \rangle$ is roughly constant over $\mu \in \mathsf{LIN}$. This will imply that $\langle \nu - \mu, \ln \Lambda \rangle$ is small, so that μ must be close to ν. For the formal details, see the proof of Proposition 8.10.

We emphasize that although Lemma 8.5 is a trivial identity, it plays an essential part in the proof of the key results of this paper.

Proof of Proposition 3.32. By the result of Lemma 3.42, it suffices to show that when we restrict to $\omega \in I_0$, the function $\Psi = \Psi_{\mathscr{D},2}$ is uniquely maximized at $\omega(\mathscr{D}) = {}_*\omega$ (in the notation of (8.6)), with negative-definite Hessian at the maximizer. For the proof, we assume that $\mathscr{D}$ is bounded away from zero in the sense of (3.45); this holds with high probability by Proposition 3.23. We also assume that the profile $\mathscr{D}$ expands on type-subsets, which by Lemma 8.3 occurs with high probability. Now take any maximizer $\omega = \omega(\mathscr{D})$ from (8.6) — we have not yet argued that it is unique. By the expansion condition and the *a priori* estimate Proposition 8.4 (and recalling from Definition 7.2 that $\zeta = \zeta_*/4$), we know that ω satisfies the estimate (8.7) on all (L, j) such that $L(j)$ is a strongly non-defective edge type in the sense of Remark 6.5.

Proof that $\omega(\mathscr{D}) = {}_*\omega$ uniquely maximizes $\Psi = \Psi_{\mathscr{D},2}$ on I_0. We start by letting $\omega = \omega(\mathscr{D})$ be any maximizer from (8.6). Consequently, in Proposition 3.44, for any choice of compound enclosure U, on the left-hand side of (3.70) the maximum over $\acute{\omega}$ must be $\Psi(\acute{\omega}) - \Psi(\omega) = 0$. Therefore the right-hand side of (3.70) also vanishes:

$$\max_{\nu'} \left\{ \mathcal{H}(\nu') : \nu' \in J(U, \omega_{\delta U}) \right\} = \max_{\mu'} \left\{ \mathcal{H}(\mu') : \mu' \in \Delta(U; \omega_U) \right\}$$

Since the entropy function $\mathcal{H}$ is strictly concave on $J(U, \omega_{\delta U}) \supseteq \Delta(U; \omega_U)$, this implies $\mu = \nu$ where

$$\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U}) = \arg\max_{\nu'} \left\{ \mathcal{H}(\nu') : \nu' \in J(U, \omega_{\delta U}) \right\},$$

$$\mu = \mu(U; \omega) = \arg\max_{\mu'} \left\{ \mathcal{H}(\mu') : \mu' \in \Delta(U; \omega_U) \right\}. \tag{8.11}$$

(Note that μ, ν are uniquely defined by the strict concavity of $\mathcal{H}$.) Likewise, by Proposition 3.46, if U is the depth-one neighborhood of any non-compound variable, then we must have $\mu = \nu$ where $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ and $\mu = \mu(U; \omega_U)$ (the only difference being that these are now in the augmented pair coloring model). We now argue that $\omega_{L,j} = {}_*\omega_{L,j}$ in all cases where L is nice. Indeed, let

$$D(\omega) = \max_e \left\{ \mathrm{disc}_e(\omega) : e \text{ is nice} \right\}, \tag{8.12}$$

and let $e_\star$ be any nice edge that attains the maximum. We have two cases:

1. Suppose $e_\star$ is an interior edge in a compound enclosure U. Recall from Definition 3.16 that the boundary of U consists of variables which are perfect (Definition 3.13), hence also orderly. We saw in Definition 3.12 that an orderly variable cannot lie within distance $(\delta_*)^3$ of any defect, so every $e \in \delta U$ is strongly non-defective in the sense of Remark 6.5. Thus Proposition 8.4 can be applied, and guarantees that the condition (8.2) of Proposition 8.1 is satisfied by every $e \in \delta U$ (again, recall from Definition 7.2 that we take $\zeta = \zeta_*/4$). Then the result (8.4) of Proposition 8.1 (with $C_{e_\star} = 1$, since $e_\star$ is nice) gives that $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ must satisfy

$$\mathrm{disc}_{e_\star}(\nu) \leq \frac{1}{4} \left\{ \max_{e \in \delta U} \mathrm{disc}_e(\omega) \right\} \leq \frac{D(\omega)}{4},$$

where the last bound uses that all edges of δU must be nice. On the other hand, as we noted in (8.11) above, ν must coincide with $\mu = \mu(U; \omega)$ by Proposition 3.44, so in fact we must have

$$D(\omega) = \mathrm{disc}_{e_\star}(\omega) = \mathrm{disc}_{e_\star}(\nu) \leq \frac{D(\omega)}{4}.$$

This proves $D(\omega) = 0$ for the case that e lies in the interior of a compound enclosure.

2. Suppose instead that $e_\star = (av)$ where v is a non-compound variable. It follows from Definition 3.16 that v must be perfect, so by the same reasoning as above it cannot lie within distance $(\delta_*)^3$ of any defect. Thus, if U is the depth-one neighborhood of v, every $e \in \delta U$ must be strongly non-defective in the sense of Remark 6.5. It follows by Proposition 8.4 that the condition (7.7) of Proposition 7.7 is satisfied for every $e \in \delta U$ (again using that $\zeta = \zeta_*/4$). Then, the result (7.9) of Proposition 7.4 (or its consequence (8.1)) implies that $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ must satisfy

$$\mathrm{disc}_{e_\star}(\nu) \leq k^{O(1)} (\vartheta_*)^{1/4} \left\{ \max_{e \in \delta U} \mathrm{disc}_e(\omega) \right\} \leq k^{O(1)} (\vartheta_*)^{1/4} D(\omega),$$

where the last bound again uses that all edges of δU must be nice. On the other hand, we again have that ν must coincide with $\mu = \mu(U; \omega)$ by Proposition 3.46, so in fact we obtain the bound

$$D(\omega) = \mathrm{disc}_{e_\star}(\omega) = \mathrm{disc}_{e_\star}(\nu) \leq \frac{D(\omega)}{4}.$$

This proves $D(\omega) = 0$ for the case that e does not lie in the interior of a compound enclosure.

The above proves $D(\omega) = 0$, i.e., the optimizer $\omega = \omega(\mathscr{D})$ coincides with the canonical product measure ${}_*\omega$ on all nice edges. If e is a non-nice edge, then it must lie in the interior of a compound enclosure U. In this case, the result (8.4) of Proposition 8.1 gives, with C as in (8.3),

$$\mathrm{disc}_e(\omega) \leq \frac{C}{4} \left\{ \max_{e \in \delta U} \mathrm{disc}_e(\omega) \right\} \leq \frac{CD(\omega)}{4} = 0, \tag{8.13}$$

again using that all edges of δU must be nice. This proves that $\omega = \omega(\mathscr{D})$ coincides with $_*\omega$ on all edges, so that $_*\omega$ is the unique maximizer of $\mathbf{\Psi}$ on I_0 as claimed.

Proof that $\mathbf{\Psi} = \mathbf{\Psi}_{\mathscr{D},2}$ has negative-definite Hessian at $_*\omega$. Now suppose that $\omega \in I_0$ is close in euclidean norm to the optimizer $\omega(\mathscr{D}) = _*\omega$. As in (8.12), let $D(\omega)$ denote the maximal discrepancy under ω over all nice edges, and let $e_\star$ be any nice edge with $\mathrm{disc}_{e_\star}(\omega) = D(\omega)$. If $e_\star$ lies in the interior of a compound enclosure then let U be that enclosure. If instead $e_\star = (av)$ for a non-compound variable v, then let U be the depth-one neighborhood of v. In either case we let $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ and $\mu = \mu(U; \omega_U)$, as in (8.11). Note that since $\mathcal{H}$ is strictly concave, if ω is very close to $_*\omega$ then $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ must be very close to $_*\nu \equiv \nu^{\mathrm{op}}(U; _*\omega_{\delta U})$, and likewise $\mu = \mu(U; \omega_U)$ must be very close to $_*\mu \equiv \mu(U; _*\omega_U)$. But since we proved above that $_*\omega$ is the unique optimizer for $\mathbf{\Psi}$ on I_0, we must have $_*\nu = _*\mu$, and therefore μ must be very close to ν. Next, by Lemma 8.5 and the first calculation (8.9) in Remark 8.6, we have

$$\mathcal{H}(\nu) - \mathcal{H}(\mu) = \left\{\mathcal{H}(\nu) - \mathcal{H}(\mu)\right\} + \underbrace{\left\langle \ln \Lambda, \nu - \mu \right\rangle}_{\text{zero}} = \mathcal{H}(\mu \,|\, \nu) \geq \max_e \mathcal{H}(\mu_e \,|\, \nu_e), \qquad (8.14)$$

where $\Lambda = \Psi(U; \omega_{\delta U})$ from Proposition 8.1 in the compound case, and $\Lambda = \Psi_{\mathscr{D}}(U; \omega_{\delta U})$ from Proposition 7.4 in the non-compound case. In the compound case, for μ sufficiently close to ν, we have

$$\mathcal{H}(\mu_{e_\star} \,|\, \nu_{e_\star}) = \sum_\sigma \mu_{e_\star}(\sigma) \ln \frac{\mu_{e_\star}(\sigma)}{\nu_{e_\star}(\sigma)} \geq \frac{1}{3} \sum_\sigma \frac{(\mu_{e_\star}(\sigma) - \nu_{e_\star}(\sigma))^2}{\nu_{e_\star}(\sigma)} \geq \frac{1}{3}\left(\|\mu_{e_\star} - \nu_{e_\star}\|_\infty\right)^2.$$

In the non-compound case we can instead bound

$$\mathcal{H}(\mu_{e_\star} \,|\, \nu_{e_\star}) = \sum_{\sigma, L} \mu_{e_\star}(\sigma, L) \ln \frac{\mu_{e_\star}(\sigma, L)}{\nu_{e_\star}(\sigma, L)} \geq \frac{1}{3}\left(\|\mu_{e_\star} - \nu_{e_\star}\|_\infty\right)^2. \qquad (8.15)$$

We remark that $\mu_e = \omega_e$ for all edges e in U, since $\mu = \mu(U; \omega)$. Next, in the compound case we have from (7.4) and the triangle inequality that

$$\mathrm{disc}_{e_\star}(\omega) = \sum_\sigma (\vartheta_*)^{\mathbf{1}\{\mathbf{r}[\sigma] \geq 1\}} \left| \frac{\omega_{e_\star}(\sigma)}{_*\omega_{e_\star}(\sigma)} - 1 \right| = \sum_\sigma (\vartheta_*)^{\mathbf{1}\{\mathbf{r}[\sigma] \geq 1\}} \left| \frac{\mu_{e_\star}(\sigma)}{_*\omega_{e_\star}(\sigma)} - 1 \right|$$

$$\leq \underbrace{\sum_\sigma (\vartheta_*)^{\mathbf{1}\{\mathbf{r}[\sigma] \geq 1\}} \left| \frac{\nu_{e_\star}(\sigma)}{_*\omega_{e_\star}(\sigma)} - 1 \right|}_{\mathrm{disc}_{e_\star}(\nu)} + \underbrace{\sum_\sigma (\vartheta_*)^{\mathbf{1}\{\mathbf{r}[\sigma] \geq 1\}} \left| \frac{\mu_{e_\star}(\sigma)}{_*\omega_{e_\star}(\sigma)} - \frac{\nu_{e_\star}(\sigma)}{_*\omega_{e_\star}(\sigma)} \right|}_{\leq C_1(k)\|\mu_{e_\star} - \nu_{e_\star}\|_\infty},$$

where $C_1(k)$ is a constant depending only on k (this uses the $e_\star$ is a nice edge). Rearranging gives

$$\frac{1}{C_1(k)}\left\{\mathrm{disc}_{e_\star}(\omega) - \mathrm{disc}_{e_\star}(\nu)\right\} \leq \|\mu_{e_\star} - \nu_{e_\star}\|_\infty.$$

Similarly, in the non-compound case we have from (7.8) and the triangle inequality that

$$\frac{1}{C_1}\left\{\mathrm{disc}_{e_\star}(\omega) - \mathrm{disc}_{e_\star}(\nu)\right\} \leq \|\mu_{e_\star} - \nu_{e_\star}\|_\infty \qquad (8.16)$$

where C_1 depends only on k and on the constant c_1 from Proposition 3.23. By adjusting C_1 we can assume it is larger than the $C_1(k)$ from the preceding bound, so that (8.16) holds for $e_\star$ in both the compound and non-compound cases. Next, by the contraction results from Propositions 7.4 and 8.1, we have

$$\mathrm{disc}_{e_\star}(\nu) \leq \frac{1}{4} \max_e \left\{\mathrm{disc}_e(\omega) : e \text{ is nice}\right\} = \frac{1}{4}\,\mathrm{disc}_{e_\star}(\omega), \qquad (8.17)$$

where the last equality holds because $e_\star$ was chosen to maximize $\mathrm{disc}_e(\omega)$ over all nice edges. (Observe that (8.17) crucially uses the restriction to nice edges; we recall from (8.4) that we have an extra factor C (from (8.3)) for non-nice

edges.) It follows that

$$\|\mu_{e_\star} - \nu_{e_\star}\|_\infty \overset{(8.16)}{\geq} \frac{1}{C_1} \left\{ \mathrm{disc}_{e_\star}(\omega) - \mathrm{disc}_{e_\star}(\nu) \right\} \overset{(8.17)}{\geq} \frac{3}{4C_1} \mathrm{disc}_{e_\star}(\omega) = \frac{3D(\omega)}{4C_1} . \tag{8.18}$$

Next, let $e_{\max}$ be the edge e with maximal discrepancy $\mathrm{disc}_e(\omega) = D_{\max}(\omega)$ (this is over all edges, both nice and non-nice). If $e_{\max}$ is nice, then $D_{\max}(\omega) = D(\omega)$. More generally, if $D_{\max}(\omega) \leq C^3 D(\omega)$ for C as in (8.3), then the preceding bound (8.18) readily implies

$$\|\mu_{e_\star} - \nu_{e_\star}\|_\infty \geq \frac{3D(\omega)}{4C_1} \geq \frac{3D_{\max}(\omega)}{4C_1 C^3} . \tag{8.19}$$

Now suppose instead that $D_{\max}(\omega) \geq C^3 D(\omega)$. Then

$$\|\nu_{e_{\max}} - {}_\star\omega_{e_{\max}}\|_\infty \leq \frac{\mathrm{disc}_{e_{\max}}(\nu)}{(\vartheta_\star)^2} \overset{(8.4)}{\leq} \frac{CD(\omega)}{4(\vartheta_\star)^2} \overset{\odot}{\leq} \frac{\mathrm{disc}_{e_{\max}}(\omega)}{4(\vartheta_\star)^2 C^2} \leq \frac{\|\omega_{e_{\max}} - {}_\star\omega_{e_{\max}}\|_\infty}{4(\vartheta_\star)^2 C} .$$

(In the above, the bound marked $\odot$ uses the assumed lower bound on $D_{\max}(\omega)$. The last bound uses the definition (8.3) of C.) Since $(\vartheta_\star)^2 C$ is a large constant, we can conclude for this case that

$$\|\mu_{e_{\max}} - \nu_{e_{\max}}\|_\infty \geq \|\omega_{e_{\max}} - {}_\star\omega_{e_{\max}}\|_\infty - \|\nu_{e_{\max}} - {}_\star\omega_{e_{\max}}\|_\infty \geq \frac{\|\omega_{e_{\max}} - {}_\star\omega_{e_{\max}}\|_\infty}{2} \geq \frac{D_{\max}(\omega)}{C} . \tag{8.20}$$

Combining (8.19) and (8.20) gives, for some constant $C_2 = C_2(k, R)$, the lower bound

$$\max_e \|\mu_e - \nu_e\|_\infty \geq \frac{D_{\max}(\omega)}{C_2} . \tag{8.21}$$

Since we assumed that each clause type L occurs at least c_1 fraction of times, the total number of indices (L, j) is at most k/c_1, so the euclidean distance between ω and ${}_\star\omega$ can be upper bounded as

$$\left(\|\omega - {}_\star\omega\|_2 \right)^2 = \sum_{L,j} \left(\|\omega_{L,j} - {}_\star\omega_{L,j}\|_2 \right)^2 \leq \frac{C_3 k}{c_1} \left(\max_e \mathrm{disc}_e(\omega) \right)^2 \leq \frac{C_3 k}{c_1} D_{\max}(\omega)^2 . \tag{8.22}$$

Combining the bounds obtained so far gives

$$\mathcal{H}(\nu) - \mathcal{H}(\mu) \overset{(8.14)}{\geq} \max_e \mathcal{H}(\mu_e \mid \nu_e) \overset{(8.15)}{\geq} \max_e \frac{(\|\mu_e - \nu_e\|_\infty)^2}{3} \overset{(8.21)}{\geq} \frac{1}{3} \left(\frac{D_{\max}(\omega)}{C_3} \right)^2$$

$$\overset{(8.22)}{\geq} \frac{1}{3} \frac{c_1 (\|\omega - {}_\star\omega\|_2)^2}{C_3 k (C_2)^2} \equiv \frac{1}{C_4} \left(\|\omega - {}_\star\omega\|_2 \right)^2 .$$

To conclude, let $\omega_\nu \in I_0$ be the edge empirical marginals of $\nu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$. Since we already showed that ${}_\star\omega$ optimizes Ψ over I_0, we must have $\Psi({}_\star\omega) \geq \Psi(\omega_\nu)$. Combining with the results of Propositions 3.44 and 3.46 gives

$$\Psi({}_\star\omega) - \Psi(\omega) \geq \Psi(\omega_\nu) - \Psi(\omega) \geq c_1 \left\{ \mathcal{H}(\nu) - \mathcal{H}(\mu) \right\} \geq \frac{c_1}{C_4} \left(\|\omega - {}_\star\omega\|_2 \right)^2 .$$

This proves that Ψ has negative-definite Hessian at ${}_\star\omega$, as claimed. As noted above, the result follows by applying Lemma 3.42. $\square$

8.2. Decomposition of compound enclosures.

We now turn to the proof of Proposition 8.1.

Definition 8.7 (decomposition of a compound enclosure). Let U be a compound enclosure: recall that this means $U = U^\circ \cup \partial_\circ U$ where $\partial_\circ U$ is the boundary of perfect variables. Suppose that an internal variable $v_\star \in U^\circ$ is designated as the "root." Let $\kappa_\star$ be the defect buffer depth from Definition 3.10, and let $\mathcal{M} \equiv (\kappa_\star)^{1/4}$. If a variable $u \in U^\circ$ lies at distance $\ell \mathcal{M}^2$ from $v_\star$ for some positive integer ℓ, and moreover does not lie within distance $5\mathcal{M}^2$ of any non-nice variable, then we call u a **terminal**. Let X denote the set of all terminals in U. Let $\mathcal{A}(X)$ denote the set of all clauses a in U such that a is the parent of a terminal variable (again, regarding U as being rooted at $v_\star$).

– If $J = J^\circ \cup \partial J^\circ$ where J° is the maximal connected component of $U \setminus \mathcal{A}(X)$ containing $v_\star$, and $\partial J^\circ \equiv \partial_\circ J$ is the set of variables in $U \setminus J^\circ$ at unit distance from J°, then we call J the **root join**.

– If instead $J = J^\circ \cup (av) \cup (\partial(J^\circ \setminus v))$ where J° is a maximal connected component of $U \setminus \mathcal{A}(X)$ not containing $v_\star$, v is the terminal variable in J° that lies closest to $v_\star$, and $a \in \mathcal{A}(X)$ is the parent clause of v, then we call J a **non-root join**. We call (av) the root edge of J.

With a mild abuse of notation, we use J to denote both the subset of variables and the induced subgraph of U. The join J is termed **non-defective** if it does not intersect any defect. Otherwise we say that the join J is **defective**. For any choice of $v_\star \in U^\circ$, this definition decomposes U into joins. The joins naturally have a tree structure, where the root of the tree is the root join.

Definition 8.8 (recursive merging of joins). To prove Proposition 8.1, we will merge the joins recursively upwards. At each merge step, we fix an **uppermost join** J which will connect the subtrees below. The root of J is either $v_{\text{RT}} = v_\star$, or it is an edge $(a_{\text{RT}} v_{\text{RT}})$ where $a_{\text{RT}} \in \mathcal{A}(X)$ and $v_{\text{RT}} \in X$. A leaf variable $u \in \mathcal{L}J$ can lie arbitrarily close to v_{RT}, but the set $X \cap \mathcal{L}J$ of leaves that are terminal variables lie at distance at least $\mathcal{M}^2$ from v_{RT}. We let $\delta'J$ denote the edges in J that are incident to $X \cap \mathcal{L}J$. For each $e \in \delta'J$, we let T_e be the subtree of U descended from e. We then let T be the merged tree formed from the union of J with all the T_e ($e \in \delta'J$). In the merge step, our goal will be to bound edge discrepancies (defined by (7.4)) for the measure $\mu = \nu^{\text{op}}(T; \omega_{\delta T})$ in terms of edge discrepancies for the measures $\nu(e) = \nu^{\text{op}}(T_e; \omega_{\delta T_e})$.

Remark 8.9. Let $\mathcal{L}U$ denote the set of all leaf variables of the compound enclosure U. The subset $\partial_\circ U \subseteq \mathcal{L}U$ is the boundary of perfect variables, so these variables are not leaves in the full graph $\mathcal{G}$. The variables in the complement $\mathcal{L}U \setminus U$ are leaves in the full graph $\mathcal{G}$ — in particular, this means they are not nice (Definition 3.8), hence they are defective. We thus have two cases:

– If J is a non-defective join, then all its leaf variables must belong either to $\partial_\circ U$ or to the terminal set X. In this case we let δJ denote the set of edges incident to all leaf variables of J.
– If J is a defective join, then it cannot intersect $\partial_\circ U$: this is because any defect has at its boundary a buffer of nice variables of depth at least $\kappa_\ast$ (Definition 3.10), so the $\kappa_\ast$-neighborhood of any variable in $\partial_\circ U$ is nice. In this case we let $\delta J \equiv \delta'J$.

These two cases will be treated separately in the analysis below.

Proposition 8.10 (merge through a defective join). *In the setting of Proposition 8.1 and Definition 8.8, consider a merge step where the uppermost join J is defective. As in Definition 8.8, let T_e denote the subtrees descended from $e \in \delta J$, and $\nu(e) \equiv \nu^{\text{op}}(T_e; \omega_{\delta T_e})$. Let T be the merged tree, and $\mu \equiv \nu^{\text{op}}(T; \omega_{\delta T})$. Then for every edge e' in $J \setminus \delta J$ we have*

$$\text{disc}_{e'}(\mu) \leq C_{e'} \cdot 2^{5k} \sum_{e \in \delta J} \text{disc}_e(\nu(e))$$

for $C_{e'}$ as defined by (8.3).

Proposition 8.11 (merge through a non-defective join). *In the setting of Proposition 8.1 and Definition 8.8, consider a merge step where the uppermost join J is non-defective. As in Definition 8.8, let T_e denote the subtrees descended from $e \in \delta'J$, and $\nu(e) \equiv \nu^{\text{op}}(T_e; \omega_{\delta T_e})$. Let T be the merged tree, and $\mu \equiv \nu^{\text{op}}(T; \omega_{\delta T})$. Then*

$$\sum_{e \in \delta'J} \text{disc}_e(\mu) \leq 4^{k\mathcal{M}} \sum_{e \in \delta'J} \text{disc}_e(\nu(e)),$$

where we recall that $\mathcal{M} = (\kappa_\ast)^{1/4}$, and the result holds provided $\kappa_\ast \geq (240/\zeta)^4$ for ζ as in Definition 7.2.[20]

Proposition 8.10 is proved in §8.3, while Proposition 8.11 is proved in §8.4. We now explain how they combine to give Proposition 8.1:

Proof of Proposition 8.1. Fix any edge $e_\star = (av)$ in $U \setminus \delta U$, and decompose U as in Definition 8.7 around $v_\star = v$. As in Proposition 8.11, consider a merge step where the uppermost join J is non-defective. Let T be the merged tree, and $\mu = \nu^{\text{op}}(T; \omega_{\delta T})$. If J contains $v_\star$, then $T = U$ and $\mu = \nu^{\text{op}}(U; \omega_{\delta U})$. If J does not contain $v_\star$, then T is rooted at an edge $e_{\text{RT}} = (a_{\text{RT}} v_{\text{RT}})$, and indeed

$$\mu = \nu^{\text{op}}(T; \omega_{\delta T}) = \nu^{\text{op}}(T_{e_{\text{RT}}}; \omega_{\delta T_{e_{\text{RT}}}}) = \nu(e_{\text{RT}}),$$

---

[20]In fact we will see from the proof that for Proposition 8.11 it suffices here to have $\kappa_\ast > (108/\zeta)^4$, see (8.41). We put the stronger condition $\kappa_\ast \geq (240/\zeta)^4$ since this was already required for Proposition 8.1.

where $\nu(e_{\mathrm{RT}})$ is defined analogously to the measure $\nu(e)$ from Definition 8.8. In either case, the marginal of μ on J must be given by $\mu = \nu^{\mathrm{op}}(J; \mu_{\delta J})$. It follows by Proposition 7.3 that for any edge (au) in J we have

$$\mathrm{disc}_{au}(\mu) \leq k^{O(1)} \sum_{e \in \delta J} \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{\mathsf{b}_J(e,a)} \mathrm{disc}_e(\mu) . \tag{8.23}$$

(To be pedantic, (8.23) follows directly from Proposition 7.3 in the case that J is rooted at an edge e_{RT}. In the final merge step where J is rooted at the variable $v_\star$, (8.23) follows from the variable-rooted analogue of Proposition 7.3 — this can be obtained by an extremely similar proof, so we will not elaborate on it here.) Recall that δJ denotes the set of edges in J incident to all the leaf variables of J, while $\delta' J$ is the subset that are incident to terminal variables. If $e \in \delta J \setminus \delta' J$, then $\mu_e = \omega_e$. On the other hand, the contribution to the above sum from $e \in \delta' J$ is bounded by Proposition 8.11. It follows that

$$\mathrm{disc}_{au}(\mu) \leq k^{O(1)} \sum_{e \in \delta J \setminus \delta' J} \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{\mathsf{b}_J(e,a)} \mathrm{disc}_e(\omega) + k^{O(1)} 4^{kM} \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{M^2} \sum_{e \in \delta' J} \mathrm{disc}_e(\nu(e))$$

$$\leq k^{O(1)} \sum_{e \in \delta J \setminus \delta' J} \left(\frac{(\vartheta_*)^{1/5}}{2^k} \right)^{\mathsf{b}_J(e,a)} \mathrm{disc}_e(\omega) + k^{O(1)} \left(\frac{(\vartheta_*)^{1/5}}{2^k} \right)^{M^2} \sum_{e \in \delta' J} \mathrm{disc}_e(\nu(e)) , \tag{8.24}$$

where the last bound holds for M large enough: to be precise, we require

$$4^{kM} \leq \left(\frac{(\vartheta_*)^{1/5}}{(\vartheta_*)^{1/4}} \right)^{M^2} = \left(\frac{1}{(\vartheta_*)^{1/20}} \right)^{M^2} = 2^{k\zeta M^2/120} ,$$

where the last equality uses that $\vartheta_* \equiv 2^{-k\zeta/6}$ from Definition 7.2. Thus it suffices here to have

$$\kappa_* \equiv M^4 \geq \left(\frac{240}{\zeta} \right)^4 .$$

Next note for $e \in \delta J$, the discrepancy $\mathrm{disc}_e(\nu(e))$ can be bounded either by a recursive application of (8.24) if e is the root of another non-defective join, or by Proposition 8.10 if e is the root of a defective join. Altogether it gives that for the edge $e_\star = (av)$ where $v = v_\star$ (and with $C_{e_\star}$ as in (8.3)), the maximal-entropy measure $\mu = \nu^{\mathrm{op}}(U; \omega_{\delta U})$ for the full enclosure satisfies

$$\frac{\mathrm{disc}_{e_\star}(\mu)}{C_{e_\star}} \leq \sum_{e \in \delta U} \left(\frac{(\vartheta_*)^{1/5}}{2^k} \right)^{\mathsf{b}_U(e,a)} \left\{ 2^{5k} \left(\frac{2^k}{(\vartheta_*)^{1/5}} \right)^{M^2} \right\}^{\mathfrak{B}(v_\star, e)} \mathrm{disc}_e(\omega) ,$$

where $\mathfrak{B}(v_\star, e)$ is the number of defective variables between $v_\star$ and e, and is a crude upper bound on the number of defective joins intersecting the path between $v_\star$ and e. The above can be upper bounded as

$$\frac{\mathrm{disc}_{e_\star}(\mu)}{C_{e_\star}} \leq \left(\max_{e \in \delta U} \mathrm{disc}_e(\omega) \right) \sum_{u \in \partial_\circ U} \frac{(2^{2kM^2})^{\mathfrak{B}(v_\star, u)}}{(2^k (\vartheta_*)^{-1/5})^{d(v_\star, u)}}$$

$$\leq \left(\max_{e \in \delta U} \mathrm{disc}_e(\omega) \right) \sum_{u \in \partial_\circ U} \frac{\exp\{k(\delta_*)^{-1} \mathfrak{B}(v_\star, v)\}}{\exp\{(k \ln 2)(1 + \delta_*) d(v_\star, v)\}} , \tag{8.25}$$

where for the last inequality in (8.25) to hold we make the following choices. First, in the numerator we can ensure that 2^{2M^2} is upper bounded by $\exp(1/\delta_*)$ by requiring

$$\frac{\ln 2}{2} \geq M^2 \delta_* = (\kappa_*)^{1/2} \delta_* .$$

In the denominator we can ensure that $(\vartheta_*)^{-1/5} = 2^{k\zeta/30}$ upper bounds $2^{k\delta_*}$ by requiring

$$\delta_* \leq \frac{\zeta}{30} .$$

These choices together guarantee the last bound in (8.25). It follows by combining with the definition of a compound enclosure (cf. (3.19)) that

$$\frac{\text{disc}_{e_\star}(\mu)}{C_{e_\star}} \le \frac{1}{4}\left(\max_{e\in\delta U}\text{disc}_e(\omega)\right).$$

This proves the result. □

8.3. **Merging through defective joins.** In this subsection we prove Proposition 8.10. The idea of the proof is to reweight the boundary of a defective join in such a way as to circumvent difficulty of analyzing tree recursions within a defect.

Proof of Proposition 8.10. Recall from Definition 8.8 the root of J is either $v_{\text{RT}} = v_\star$, or it is an edge $(a_{\text{RT}}v_{\text{RT}})$. In the first case we let $_\star\Lambda_J$ be as given by Corollary 3.56. In the second case we let $_\star\Lambda_J$ be given by (7.13) from Remark 7.7, i.e., we give weight $_\star\hat{q}_{a_{\text{RT}}v_{\text{RT}}}$ to the root clause a_{RT}. Note that $_\star\Lambda_J$ includes a weight $_\bullet\dot{q}_e$ for each $e \in \delta J$. Therefore, for a single-copy coloring $\underline{\tau}_{J\setminus\delta J}$ on $J \setminus \delta J$, we can let

$$_\star\Theta_{J\setminus\delta J}(\underline{\tau}_{J\setminus\delta J}) \equiv {}_\star\Lambda_J(\underline{\tau}_J)\prod_{e\in\delta J}\frac{1}{_\bullet\dot{q}_e(\tau_e)},$$

where this is well-defined because the right-hand side does not depend on $\underline{\tau}_{\delta J}$. We then let $_*\Theta \equiv {}_\star\Theta \otimes {}_\star\Theta$.

Next, recall from Remark 8.9 that $\delta J = \delta' J$ is the set of boundary edges of J. Let A be the subset of clauses in J that are incident to δJ, so A is a subset of the set $\mathcal{A}(X)$ from Definition 8.7. Let $e = (au) \in \delta J$, so $a \in A$ and u is a terminal variable. Recall that T_e is the subtree of U descended from e. Let $\Lambda_{T_e} = \Lambda(T_e; \omega_{\delta T_e})$ be the Lagrangian weights on T_e such that $v(e) = v^{\text{op}}(T_e, \omega_{\delta T_e})$ agrees with $v[T_e; \Lambda_{T_e}]$. We assume moreover that they are parametrized as in Definition 7.8, meaning that the root clause $a(e)$ of T_e receives weight $_*\hat{q}_e$ under Λ_{T_e}. Write $q \equiv q_{\text{BP}}(T_e; \Lambda_{T_e})$ for the associated BP messages, so in particular the downward message on e is $\hat{q}_e = {}_*\hat{q}_e$. It will follow by induction that the upward message $\dot{q}_e$ satisfies the bounds (7.11) for each edge $e \in \delta J$.

Step 1. Weights on merged tree. We now apply Proposition 7.6 to each clause $a \in A$ where we take incoming messages $\dot{q}_{ua}$ from below as just discussed, but the canonical product message $\dot{q}_{va} = {}_*\dot{q}_{va}$ from the parent variable v. Let h denote the resulting outgoing messages from the unweighted clause a, that is, $h = \text{BP}_a[\dot{q}]$. Proposition 7.6 gives a clause weighting $\Gamma_a \equiv (\gamma_e)_{e\in\delta a}$ such that for these input messages to a, the Γ_a-weighted clause BP recursion **outputs the canonical product message** $_*\hat{q}$ **on every edge in** δa. For a pair configuration $\underline{\sigma} \equiv \underline{\sigma}_T$ on the merged tree T, we define the weight

$$\Theta(\underline{\sigma}) \equiv {}_*\Theta_{J\setminus\delta J}(\underline{\sigma}_{J\setminus\delta J})\left\{\prod_{a\in A}\Gamma_a(\underline{\sigma}_{\delta a})\right\}\left\{\prod_{e\in\delta J}\frac{\Lambda_{T_e}(\underline{\sigma}_{T_e})}{_*\hat{q}_e(\sigma_e)}\right\}.$$

Recall from above that the clause $a(e)$ receives weight $_*\hat{q}_e$ under Λ_{T_e}, and in the above expression we divided by $\hat{q}_e(\sigma_e)$ to remove this weight. Likewise, $_*\Lambda_J$ includes a weight $_*\dot{q}_e$ for each $e \in \delta J$, and we also divided by $_*\dot{q}_e(\sigma_e)$ to remove this weight. Let $\nu \equiv \nu[T; \Theta]$. Let $e(a)$ denote the parent edge of a, i.e., the unique element of $\delta a \setminus \delta J$. The marginal of ν on $J \setminus \delta J$ is

$$\nu_{J\setminus\delta J}(\underline{\sigma}_{J\setminus\delta J}) \cong {}_*\Theta_{J\setminus\delta J}(\underline{\sigma}_{J\setminus\delta J})\left\{\sum_{\underline{\sigma}_{\delta J}}\prod_{a\in A}\left(\Gamma_a(\underline{\sigma}_{\delta a})\prod_{e\in\delta a\setminus e(a)}\left[\sum_{\underline{\sigma}_{T_e\setminus J}}\frac{\Lambda_{T_e}(\underline{\sigma}_{T_e})}{_*\hat{q}_e(\sigma_e)}\right]\right)\right\}$$

$$\overset{(a)}{\cong} {}_*\Theta_{J\setminus\delta J}(\underline{\sigma}_{J\setminus\delta J})\left\{\prod_{a\in A}\left(\sum_{\underline{\sigma}_{\delta a\setminus e(a)}}\Gamma_a(\underline{\sigma}_{\delta a})\prod_{e\in\delta a\setminus e(a)}\dot{q}_e(\sigma_e)\right)\right\}$$

$$\overset{(b)}{\cong} {}_*\Theta_{J\setminus\delta J}(\underline{\sigma}_{J\setminus\delta J})\left\{\prod_{e'\in\delta A\setminus\delta J}{}_*\hat{q}_{e'}(\sigma_{e'})\right\} \cong {}_*\nu_{J\setminus\delta J}(\underline{\sigma}_{J\setminus\delta J}), \tag{8.26}$$

where $_*\nu = {}_\star\nu \otimes {}_\star\nu$ is the canonical product measure. (In the above calculation, the step marked *(a)* uses that under the Λ_{T_e}-weighted measure on T_e, the marginal on edge e is proportional to $_*\hat{q}_e\dot{q}_e$. The step marked *(b)* uses

that the Γ_a-weighted BP recursion outputs the canonical product message $_*\hat{q}$ on the edge $e(a)$.) Similarly, for any $e = (bu) \in \delta J$, the marginal of ν on T_e is given by

$$\nu_{T_e}(\underline{\sigma}_{T_e}) \cong \frac{\Lambda_{T_e}(\underline{\sigma}_{T_e})}{_*\hat{q}_e(\sigma_e)} \sum_{\underline{\sigma}_{J\backslash\delta J}} {_*\Theta}_{J\backslash\delta J}(\underline{\sigma}_{J\backslash\delta J}) \left\{ \sum_{\underline{\sigma}_{\delta b\backslash\{e,e(b)\}}} \Gamma_b(\underline{\sigma}_{\delta b}) \prod_{e'\in\delta b\backslash\{e,e(b)\}} \left[\sum_{\underline{\sigma}_{T_{e'}\backslash J}} \frac{\Lambda_{T_{e'}}(\underline{\sigma}_{T_{e'}})}{_*\hat{q}_{e'}(\sigma_{e'})} \right] \right\}$$

$$\times \left\{ \prod_{a\in A\backslash b} \left(\sum_{\underline{\sigma}_{\delta a\backslash e(a)}} \Gamma_a(\underline{\sigma}_{\delta a}) \prod_{e''\in\delta a\backslash e(a)} \left[\sum_{\underline{\sigma}_{T_{e''}\backslash J}} \frac{\Lambda_{T_{e''}}(\underline{\sigma}_{T_{e''}})}{_*\hat{q}_{e''}(\sigma_{e''})} \right] \right) \right\}$$

$$\cong \frac{\Lambda_{T_e}(\underline{\sigma}_{T_e})}{_*\hat{q}_e(\sigma_e)} \sum_{\underline{\sigma}_{J\backslash\delta J}} {_*\Theta}_{J\backslash\delta J}(\underline{\sigma}_{J\backslash\delta J}) \left\{ \sum_{\underline{\sigma}_{\delta b\backslash\{e,e(b)\}}} \Gamma_b(\underline{\sigma}_{\delta b}) \prod_{e'\in\delta b\backslash\{e,e(b)\}} \dot{q}_{e'}(\sigma_{e'}) \right\} \prod_{e''\in\delta A\backslash(\delta J\cup\delta b)} {_*\hat{q}_{e''}(\sigma_{e''})}$$

$$\cong \frac{\Lambda_{T_e}(\underline{\sigma}_{T_e})}{_*\hat{q}_e(\sigma_e)} \sum_{\underline{\sigma}_{\delta b\backslash e}} \Gamma_b(\underline{\sigma}_{\delta b}) \left\{ {_*\dot{q}_{e(b)}}(\sigma_{e'}) \prod_{e'\in\delta b\backslash\{e,e(b)\}} \dot{q}_{e'}(\sigma_{e'}) \right\} \cong \Lambda_{T_e}(\underline{\sigma}_{T_e}) \,, \qquad (8.27)$$

where the last step uses that the Γ_a-weighted BP recursion outputs the canonical product message $_*\hat{q}$ on the edge e. This shows that the marginal of ν on T_e is simply $\nu(e) = \nu^{\mathrm{op}}(T_e; \omega_{\delta T_e})$. It follows that $\nu \in J(T; \omega_{\delta T})$. Consequently, $\mu = \nu^{\mathrm{op}}(T; \omega_{\delta T})$ is the entropy maximizer over $J(T; \omega_{\delta T})$, then $\mathcal{H}(\mu) \geq \mathcal{H}(\nu)$. It follows by recalling Lemma 8.5 and Remark 8.6 (in particular, the calculation (8.10)) that

$$\mathcal{H}(\mu \,|\, \nu) = \left\{ \mathcal{H}(\nu) - \mathcal{H}(\mu) \right\} + \left\langle \nu - \mu, \ln \Theta \right\rangle \leq \left\langle \nu - \mu, \ln \Theta \right\rangle. \qquad (8.28)$$

Recall that the weights in Θ are of product form, except on the boundary edges δT and on the edges δA. Since μ and ν agree on the subtrees T_e, and have the same single-copy marginals (given by $_\star\pi$) on all edges of the merged tree T, the only contribution to the right-hand side of (8.28) comes from the edges $e' \in \delta A \setminus \delta J$:

$$\mathcal{H}(\mu \,|\, \nu) \leq \sum_{e'\in\delta A\backslash\delta J} \sum_\sigma \left| \left\{ \nu_{e'}(\sigma) - \mu_{e'}(\sigma) \right\} \ln \gamma_{e'}(\sigma) \right|$$

$$\leq \sum_{e'\in\delta A\backslash\delta J} \|\ln \gamma_{e'}\|_\infty \|\mu_{e'} - {_*\omega_{e'}}\|_\infty \,, \qquad (8.29)$$

where the last step uses that $\nu_{e'} = {_*\omega_{e'}}$ for all $e' \in J \setminus \delta J$.

Step 2. Bound on clause weights. We first bound the $\ln \gamma_{e'}$ term in (8.29). For each $a \in A$, there is a single edge $e' = e(a)$ in $\delta a \setminus \delta J$. A crude application of Proposition 7.6 gives

$$\|\ln \gamma_{e(a)}\|_\infty \leq k^{O(1)} \sum_{e''\in\delta a\backslash e(a)} \left(\epsilon_{e''} + \dot{\epsilon}_{e''} + \ddot{\epsilon}_{e''} \right), \qquad (8.30)$$

where $\underline{\epsilon} = (\epsilon_{e''}, \dot{\epsilon}_{e''}, \ddot{\epsilon}_{e''}) = \widehat{\mathrm{ERR}}(_*\hat{q}_{e''}, h_{e''})$ is the error between the unweighted messages $h = \mathrm{BP}_a[\dot{q}]$ and the weighted messages $_*\hat{q} = \mathrm{BP}_a[\dot{q}; \Gamma_a] = \mathrm{BP}_a[_*\dot{q}]$. A crude application of Proposition 7.12 gives, for $e'' \in \delta a \setminus e(a)$,

$$\epsilon_{e''} + \dot{\epsilon}_{e''} + \ddot{\epsilon}_{e''} \leq k^{O(1)} \sum_{e\in\delta a\backslash e(a)} \left(\delta_e + \dot{\delta}_e + \ddot{\delta}_e + {^{\mathrm{m}}\delta_e} + {^{\mathrm{m}}\dot{\delta}_e} \right), \qquad (8.31)$$

where $\underline{\delta} = (\delta_e, \dot{\delta}_e, \ddot{\delta}_e, {^{\mathrm{m}}\delta_e}, {^{\mathrm{m}}\dot{\delta}_e}) = \mathrm{E\ddot{R}R}(_*\dot{q}_e, \dot{q}_e)$. To bound the error between $\dot{q}_e$ and $_*\dot{q}_e$ for $e \in \delta J$, note

$$\dot{q}_e(\sigma) \cong \frac{\nu_e(\sigma)}{_*\hat{q}_e(\sigma)} \overset{(7.4)}{=} \frac{_*\omega_e(\sigma)}{_*\hat{q}_e(\sigma)} \left\{ 1 + O\left(\frac{\mathrm{disc}_e(\nu)}{\vartheta_*} \right) \right\} \cong {_*\dot{q}_e(\sigma)} \left\{ 1 + O\left(\frac{\mathrm{disc}_e(\nu)}{\vartheta_*} \right) \right\},$$

so each entry of $\underline{\delta}$ is $O((\vartheta_*)^{-1} \mathrm{disc}_e(\nu))$. Combining the last few bounds gives (again very crudely)

$$\|\ln \gamma_{e(a)}\|_\infty \leq 2^k \sum_{e\in\delta a\backslash e(a)} \mathrm{disc}_e(\nu) = 2^k \sum_{e\in\delta a\backslash e(a)} \mathrm{disc}_e(\nu(e)), \qquad (8.32)$$

for each edge $e' \in \delta A \setminus \delta J$.

Step 3. Conclusion. To conclude we note that for each $e \in \delta A$, we have

$$\|\mu_e - {}_\star\omega_e\|_\infty \leq \|{}_\star\omega_e\|_\infty \left\|\frac{\mu_e}{{}_\star\omega_e} - 1\right\|_\infty \leq \left\|\frac{\mu_e}{{}_\star\omega_e} - 1\right\|_\infty \overset{(7.4)}{\leq} 2^k \operatorname{disc}_e(\mu).$$

Combining with (8.32) and substituting into (8.29) gives

$$\mathcal{H}(\mu \mid \nu) \leq 2^{2k} \sum_{a \in A} \left(\sum_{e \in \delta a \setminus e(a)} \operatorname{disc}_e(\nu(e))\right)\left(\sum_{e' \in \delta a \setminus \delta J} \operatorname{disc}_{e'}(\mu)\right)$$

$$\leq 2^{2k} \left(\max_{e' \in J \setminus \delta J} \frac{\operatorname{disc}_{e'}(\mu)}{C_{e'}}\right)\left(\sum_{e \in \delta J} \operatorname{disc}_e(\nu(e))\right) \tag{8.33}$$

where $C_{e'}$ is as in the statement of the proposition (and the last bound holds simply because $C_{e'} = 1$ for all $e' \in \delta A$, since those edges must be nice). On the other hand, for any $e' \in J \setminus \delta J$, we have

$$\mathcal{H}(\mu \mid \nu) \geq \mathcal{H}(\mu_{e'} \mid \nu_{e'}) = \mathcal{H}(\mu_{e'} \mid {}_\star\omega_{e'}) \geq \frac{1}{2^{3k}} \left(\frac{\operatorname{disc}_{e'}(\mu)}{C_{e'}}\right)^2,$$

where we take $C_{e'}$ as in (8.3) to account for non-nice edges where the minimum of ${}_\star\pi_{e'}$ on its support may be small. Combining the last two displays gives

$$\max_{e' \in J \setminus \delta J} \frac{\operatorname{disc}_{e'}(\mu)}{C_{e'}} \leq 2^{5k} \sum_{e \in \delta J} \operatorname{disc}_e(\nu(e)),$$

as claimed. $\qquad\square$

8.4. **Merging through non-defective joins.** We conclude the section with the proof of Proposition 8.11, showing that for a non-defective join, the total boundary discrepancy under $\mu = \nu^{\mathrm{op}}(T; \omega_{\delta T})$ (the optimizer on the merged tree) is not too much larger than the total boundary discrepancy under the measures $\nu(e) = \nu^{\mathrm{op}}(T_e; \omega_{\delta T_e})$ (the optimizer on the subtrees T_e, for $e \in \delta'J$).

Proof of Proposition 8.11. Recall that $\mu \equiv \nu^{\mathrm{op}}(T; \omega_{\delta T})$, while $\nu(e) = \nu^{\mathrm{op}}(T_e; \omega_{\delta T_e})$ for $e \in \delta'J$.

Step 1. Identification of small subtree with large discrepancy. For the sake of contradiction, let us suppose the desired bound false, i.e., that we have

$$\sum_{e \in \delta'J} \operatorname{disc}_e(\mu) > 4^{k\mathcal{M}} \sum_{e \in \delta'J} \operatorname{disc}_e(\nu(e)). \tag{8.34}$$

Recall from Definition 8.8 that each variable in $\mathcal{X} \cap \mathcal{L}J$ lies at distance exactly $\mathcal{M}^2$ from the root of J, and $\delta'J$ denotes the edges in J that are incident to $\mathcal{X} \cap \mathcal{L}J$. Similarly to (7.46), for any clause a in J we define

$$\xi_a(J; \mu_{\delta'J}) \equiv \sum_{e \in \delta'J} \left(\frac{(\vartheta_\star)^{1/4}}{2^k}\right)^{\mathsf{b}_J(e,a)} \operatorname{disc}_e(\mu).$$

Now let A be the subset of clauses in J at distance exactly $\mathcal{M}$ from $\delta'J$. For any $e \in \delta'J$ and any $j \geq 0$, we have

$$\left|\left\{a \in A : \mathsf{b}_J(e,a) = \mathcal{M} + j\right\}\right| \leq (dk)^{\lceil j/2 \rceil}.$$

It follows from this that

$$\sum_{a \in A} \xi_a(J; \mu_{\delta'J}) = \sum_{e \in \delta'J} \mathrm{disc}_e(\mu) \sum_{a \in A} \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{\mathrm{b}_T(e,a)} \leq \sum_{e \in \delta'J} \mathrm{disc}_e(\mu) \sum_{j=0}^{\infty} \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{\mathcal{M}+j} (dk)^{\lceil j/2 \rceil}$$

$$\leq \left\{ \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{\mathcal{M}} \sum_{e \in \delta'J} \mathrm{disc}_e(\mu) \right\} \left(1 + 2 \sum_{\ell=0}^{\infty} (dk)^{\ell} \left(\frac{(\vartheta_*)^{1/4}}{2^k} \right)^{2\ell-1} \right)$$

$$\leq \left(\frac{k(\vartheta_*)^{1/4}}{2^k} \right)^{\mathcal{M}} \sum_{e \in \delta'J} \mathrm{disc}_e(\mu). \tag{8.35}$$

For $a \in A$, write J_a for the subtree of J descended from a. Write $\delta J_a \equiv J_a \cap \delta'J$. Combining (8.34) and (8.35) gives

$$\sum_{a \in A} \left\{ 2 \sum_{e \in \delta J_a} \mathrm{disc}_e(\mu) - \left(\frac{2^k}{k(\vartheta_*)^{1/4}} \right)^{\mathcal{M}} \xi_a(J; \mu_{\delta'J}) - 4^{k\mathcal{M}} \sum_{e \in \delta J_a} \mathrm{disc}_e(\nu(e)) \right\} \geq 0,$$

so we can find some $a \in A$ for which the expression in braces is non-negative. This implies two bounds: first,

$$\xi_a(J; \mu_{\delta'J}) \leq 2 \cdot \left(\frac{k(\vartheta_*)^{1/4}}{2^k} \right)^{\mathcal{M}} \sum_{e \in \delta J_a} \mathrm{disc}_e(\mu) \leq (k^4(\vartheta_*)^{1/4})^{\mathcal{M}} \left\{ \max_{e \in \delta'J_a} \mathrm{disc}_e(\mu) \right\}, \tag{8.36}$$

where the last bound uses that $|\delta J_a| \leq (k^3 2^k)^{\mathcal{M}}$. Secondly, we must also have

$$\max_{e \in \delta J_a} \mathrm{disc}_e(\nu(e)) \leq \sum_{e \in \delta J_a} \mathrm{disc}_e(\nu(e)) \leq \frac{2}{4^{k\mathcal{M}}} \sum_{e \in \delta J_a} \mathrm{disc}_e(\mu) \leq \left(\frac{k^4}{2^k} \right)^{\mathcal{M}} \left\{ \max_{e \in \delta J_a} \mathrm{disc}_e(\mu) \right\}. \tag{8.37}$$

In the remainder of the proof we derive a contradiction.

Step 2. Weights on merged tree. First, let $\Lambda = \Lambda_T = \Lambda(T; \omega_{\delta T})$ denote the Lagrangian weights such that $\nu[T; \Lambda]$ coincides with $\mu = \nu^{\mathrm{op}}(T; \omega_{\delta T})$. We write q for the corresponding BP messages. For the distinguished clause a, apply Proposition 7.6 to obtain a clause weighting $\Gamma_a \equiv (\gamma_e)_{e \in \delta a}$ such that

$$\hat{q}_{av} = \mathrm{BP}_{av} \left[\left({}_*\dot{q}_{ua} : u \in \partial a \setminus v \right); \Gamma_a \right] \quad \text{for the parent } v \text{ of } a,$$

$${}_*\hat{q}_{au} = \mathrm{BP}_{au} \left[\left(\dot{q}_{av}, ({}_*\dot{q}_{u'a} : u' \in \partial a \setminus \{u, v\}) \right); \Gamma_a \right] \quad \text{for each child } u \text{ of } a.$$

Next, for each $e \in \delta J$, recall that $\Lambda_{T_e} \equiv \Lambda(T_e; \omega_{\delta T_e})$ denotes the Lagrangian weights such that $\nu[T_e; \Lambda_{T_e}]$ coincides with $\nu(e) = \nu^{\mathrm{op}}(T_e; \omega_{\delta T_e})$. We write $\tilde{q}$ for the corresponding BP messages. Let A_a denote the clauses in J_a incident to δJ_a. For each clause $b \in A_a$, apply Proposition 7.6 again to obtain a clause weighting $\Gamma_b = (\gamma_e)_{e \in \delta b}$ such that

$${}_*\hat{q}_{bw} = \mathrm{BP}_{bw} \left[\left(\tilde{q}_{ua} : u \in \partial b \setminus w \right) \right] \quad \text{for the parent variable } w \text{ of } b,$$

$${}_*\hat{q}_{bu} = \mathrm{BP}_{bu} \left[\left({}_*\dot{q}_{wb}, (\tilde{q}_{u'b} : u' \in \partial b \setminus \{u, w\}) \right) \right] \quad \text{for each child } u \text{ of } b.$$

Let T_a denote the subtree of T descended from a. Recall that $\Lambda_T = \Lambda(T; \omega_{\delta T})$ is a product of variable factors. Let $\Lambda_{T,T_a}(\underline{\sigma}_{T_a})$ denote the product of those factors over the variables in T_a only, and define

$$\Lambda_{T,T \setminus T_a}(\underline{\sigma}_{T \setminus T_a}) \equiv \frac{\Lambda_T(\underline{\sigma})}{\Lambda_{T,T_a}(\underline{\sigma}_{T_a})}.$$

For a pair configuration $\underline{\sigma} \equiv \underline{\sigma}_T$ on the merged tree T, we define the weight

$$\Theta(\underline{\sigma}) \equiv \left\{ \Lambda_{T,T \setminus T_a}(\underline{\sigma}_{T \setminus T_a}) \Gamma_a(\underline{\sigma}_{\delta a}) \frac{{}_*\Theta_{J_a \setminus \delta J_a}(\underline{\sigma}_{J_a \setminus \delta J_a})}{{}_*\hat{q}_{av}(\sigma_{av})} \right\} \left\{ \prod_{b \in A_a} \Gamma_b(\underline{\sigma}_{\delta b}) \right\} \left\{ \prod_{e \in \delta J_a} \frac{\Lambda_{T_e}(\underline{\sigma}_{T_e})}{{}_*\hat{q}_e(\sigma_e)} \right\}.$$

Let $\tilde{\mu} = \nu[T; \Theta]$. It follows by similar calculations as (8.26) and (8.27) that $\tilde{\mu}$ satisfies the following:

(i) Its marginal on $T \setminus T_a$ agrees with that of $\mu = \nu^{\mathrm{op}}(T; \omega_{\delta T})$;

(ii) Its marginal on J_a agrees with that of $_*\nu = \nu^{\mathrm{op}}(T; _*\omega_{\delta T})$;

(iii) Its marginal on T_e (for each $e \in \delta J_a$) agrees with that of $\nu(e) = \nu^{\mathrm{op}}(T_e; \omega_{\delta T_e})$.

In particular, $\tilde{\mu} \in J(T; \omega_{\delta T})$. Recalling Lemma 8.5 and Remark 8.6 (in particular the calculation (8.10)), we have

$$\mathcal{H}(\mu \mid \tilde{\mu}) = \left\{ \underbrace{\mathcal{H}(\tilde{\mu}) - \mathcal{H}(\mu)}_{\text{nonpositive}} \right\} + \left\langle \tilde{\mu} - \mu, \ln \Theta \right\rangle \leq \left\langle \tilde{\mu} - \mu, \ln \Theta \right\rangle.$$

Since μ and $\tilde{\mu}$ are both judicious, and agree on $T \setminus T_a$, we have

$$\mathcal{H}(\mu \mid \tilde{\mu}) \leq \left\langle \tilde{\mu} - \mu, \ln \Theta \right\rangle \leq \left\{ \underbrace{\sum_{e \in \delta a \cap J_a} \left\langle \tilde{\mu}_e - \mu_e, \ln \gamma_e \right\rangle}_{\text{denote this } \mathcal{A}} \right\} + \left\{ \underbrace{\sum_{b \in A_a} \sum_{e \in \delta b} \left\langle \tilde{\mu}_e - \mu_e, \ln \gamma_e \right\rangle}_{\text{denote this } \mathcal{B}} \right\}. \tag{8.38}$$

We now turn to bounding these quantities in terms of $\mathrm{disc}(\mu)$ and $\mathrm{disc}(\nu)$.

Step 3. Bounds on entropy and clause weights. Let e be the edge in J_a with maximal $\mathrm{disc}_e(\mu)$. Note

$$\left(\mathcal{H}(\mu \mid \tilde{\mu}) \right)^{1/2} \geq \left(\mathcal{H}(\mu_e \mid \tilde{\mu}_e) \right)^{1/2} \geq \frac{\|\mu_e - \tilde{\mu}_e\|_\infty}{O(2^k)} \geq \frac{\|\mu_e - _*\omega_e\|_\infty - \|\tilde{\mu}_e - _*\omega_e\|_\infty}{O(2^k)}.$$

If e belongs to $J_a \setminus \delta J_a$, then $\tilde{\mu}_e = _*\omega_e$. If $e \in \delta J_a$, then $\tilde{\mu}_e = \nu(e)_e$, and it follows using (7.4) and (8.37) that

$$\|\tilde{\mu}_e - _*\omega_e\|_\infty \overset{(7.4)}{\leq} \frac{\mathrm{disc}_e(\tilde{\mu})}{\vartheta_*} = \frac{\mathrm{disc}_e(\nu(e))}{\vartheta_*} \overset{(8.37)}{\leq} \left(\frac{k^4}{2^k} \right)^{\mathcal{M}} \frac{\mathrm{disc}_e(\mu)}{\vartheta_*}.$$

This is negligible in comparison to $\|\mu_e - _*\omega_e\|_\infty$, which is lower bounded by $\Omega(4^{-k}) \, \mathrm{disc}_e(\mu)$. Therefore

$$\mathcal{H}(\mu \mid \tilde{\mu}) \geq \frac{1}{2^{6k}} \left(\max_{e \in J_a} \mathrm{disc}_e(\mu) \right)^2. \tag{8.39}$$

We next turn to bounding the clause weights. Recall that we use q to denote the BP messages for the Λ-weighted model. Similarly as in (8.30) and (8.31), a crude application of Proposition 7.6 gives

$$\sum_{e \in \delta a} \| \ln \gamma_e \|_\infty \leq k^{O(1)} \sum_{e \in \delta a} \left(\delta_e + \dot{\delta}_e + \ddot{\delta}_e + {}^{\mathrm{m}}\delta_e + {}^{\mathrm{m}}\dot{\delta}_e \right)$$

where $\underline{\delta} = (\delta_e, \dot{\delta}_e, \ddot{\delta}_e, {}^{\mathrm{m}}\delta_e, {}^{\mathrm{m}}\dot{\delta}_e) = \ddot{\mathrm{ERR}}(_*\dot{q}_e, \dot{q}_e)$. In order to bound $\underline{\delta}$ we can argue as follows. Let $\hat{p}_e \cong _*\omega_e / \dot{q}_e$ for all $e \in \delta a$, and let $Z_a = (\zeta_e)_{e \in \delta a}$ be the reweighting of a such that $\hat{p}_{au} = \mathrm{BP}_{au}[(\dot{q}_{u'a} : u' \in \partial a \setminus u); Z_a]$ for all $u \in \partial a$. It follows by another crude application of Proposition 7.6 that, similarly as in (8.30),

$$\sum_{e \in \delta a} \| \ln \zeta_e \|_\infty \leq k^{O(1)} \sum_{e \in \delta a} (\varepsilon_e + \dot{\varepsilon}_e + \ddot{\varepsilon}_e)$$

where $\underline{\varepsilon} = (\varepsilon_e, \dot{\varepsilon}_e, \ddot{\varepsilon}_e) = \widehat{\mathrm{ERR}}(\hat{p}_e, \hat{q}_e)$. We can bound $\underline{\varepsilon}$ by noting that

$$\frac{\hat{p}_e}{\hat{q}_e} \cong \frac{_*\omega_e / \dot{q}_e}{\mu_e / \dot{q}_e} \cong \frac{_*\omega_e}{\mu_e} = 1 + O\left(\frac{\mathrm{disc}_e(\mu)}{\vartheta_*} \right),$$

and substituting into the previous expression gives the bound

$$\sum_{e \in \delta a} \| \ln \zeta_e \|_\infty \leq k^{O(1)} \frac{\mathrm{disc}_e(\mu)}{\vartheta_*}.$$

From the definition of Z_a, for each $u \in \partial a$ we have

$$_*\omega_{au} \cong \dot{q}_{ua} \hat{p}_{au} = \dot{q}_{ua} \mathrm{BP}_{au}\left[\left(\dot{q}_{u'a} : u' \in \partial a \setminus u \right); Z_a \right] \cong g_{ua} \mathrm{BP}_{au}\left[\left(g_{u'a} : u' \in \partial a \setminus u \right) \right],$$

where g_e is the probability measure on $\{\mathrm{r}, \mathrm{y}, \mathrm{g}, \mathrm{b}\}^2$ such that $g_e \cong \zeta_e \dot{q}_e$. It follows that the measure

$$\hat{\nu}_a(\underline{\sigma}_{\delta a}) \cong \hat{\varphi}_a(\underline{\sigma}_{\delta a}) \prod_{e \in \delta a} g_e(\sigma_e) \tag{8.40}$$

has marginal $_*\omega_e$ on each $e \in \delta a$. This can only occur if $g_e = {}_*\dot{q}_e$ for all $e \in \delta a$ — this is because $(\ln g_e)_{e \in \delta a}$ must be the Lagrange multipliers for the constrained optimization problem

$$\max \left\{ \mathcal{H}(\hat{v}) : \hat{v} \text{ has marginals } {}_*\omega_e \right\},$$

and the Lagrange multipliers are unique so we must have $g_e = {}_*\dot{q}_e$ for all $e \in \delta a$. It follows that

$$\zeta_e \cong \frac{g_e}{\dot{q}_e} \cong \frac{{}_*\dot{q}_e}{\dot{q}_e},$$

so the above bound on $\| \ln \zeta \|_\infty$ implies a bound on $\underline{\delta}$. Altogether we obtain

$$\sum_{e \in \delta a} \| \ln \zeta_e \|_\infty \le \frac{k^{O(1)}}{\vartheta_*} \sum_{e \in \delta a} \mathrm{disc}_e(\mu).$$

Substituting into the definition of $\mathcal{A}$ from (8.38) gives

$$\mathcal{A} \le \frac{k^{O(1)}}{(\vartheta_*)^2} \left(\sum_{e \in \delta a} \mathrm{disc}_e(\mu) \right)^2 \le \frac{k^{O(1)}}{(\vartheta_*)^2} \xi_a(J; \mu_{\delta' J})^2,$$

where the last step follows by Proposition 7.3. Combining with (8.36) gives

$$\mathcal{A} \le \frac{k^{O(1)}}{(\vartheta_*)^2} \xi_a(J; \mu_{\delta' J})^2 \le \frac{k^{O(1)}(k^4(\vartheta_*)^{1/4})^{2\mathcal{M}}}{(\vartheta_*)^2} \left(\max_{e \in \delta J_a} \mathrm{disc}_e(\mu) \right)^2.$$

Next, very similarly to the derivation of (8.32), for $b \in A_a$ and $e \in \delta b$ we have

$$\| \ln \gamma_e \|_\infty \le 2^k \sum_{e \in \delta b \cap \delta J_a} \mathrm{disc}_e(\nu(e)).$$

(Indeed, a clause $b \in A_a$ takes in the canonical product message $_*\dot{q}$ from above, and takes from below the messages $\dot{q}_e$ from the Λ_{T_e}-weighted measures (which are precisely the measures $\nu(e)$. Thus $\| \ln \gamma_e \|_\infty$ can be bounded in terms of the errors between $\dot{q}_e$ and $_*\dot{q}_e$ for $e \in \delta b \cap \delta J_a$. These in turn can be bounded in terms of the marginal discrepancies $\mathrm{disc}_e(\nu(e))$, similarly as in (8.32).) Substituting into the definition of $\mathcal{B}$ from (8.38) gives

$$\mathcal{B} \le 2^{2k} \sum_{b \in A_a} \left(\sum_{e \in \delta b \cap \delta J_a} \mathrm{disc}_e(\nu(e)) \right) \left(\sum_{e \in \delta b} \mathrm{disc}_e(\mu) + \sum_{e \in \delta b \cap \delta J_a} \mathrm{disc}_e(\nu(e)) \right)$$

$$\overset{(8.37)}{\le} 2^{2k} 2k \left(\max_{e' \in J_a} \mathrm{disc}_{e'}(\mu) \right) \sum_{e \in \delta J_a} \mathrm{disc}_e(\nu(e)) \overset{(8.37)}{\le} 2^{2k} 2k \left(\frac{k^4}{2^k} \right)^{\mathcal{M}} \left(\max_{e \in J_a} \mathrm{disc}_e(\mu) \right)^2.$$

Substituting these bounds back into (8.38) and combining with (8.39) gives

$$\frac{1}{2^{6k}} \left(\max_{e \in J_a} \mathrm{disc}_e(\mu) \right)^2 \le \mathcal{H}(\mu \mid \tilde{\mu}) \le \mathcal{A} + \mathcal{B} \le (\vartheta_*)^{\mathcal{M}/3} \left(\max_{e \in J_a} \mathrm{disc}_e(\mu) \right)^2.$$

Again recall from Definition 3.11 that $\vartheta_* \equiv 2^{-k\zeta/6}$. Thus, as long as we have

$$\kappa_* = \mathcal{M}^4 > \left(\frac{108}{\zeta} \right)^4, \tag{8.41}$$

we obtain the required contradiction. $\square$

9. A priori estimates for edge marginals

In this section we prove Lemma 8.3 and Proposition 8.4, which were used in the proof of Proposition 3.32. Lemma 8.3 is a fairly easy expansion result, whose proof appears at the start of §9.7 below. Proposition 8.4 is an *a priori* estimate whose proof occupies the majority of this section. In §9.1 we make some preliminary definitions and give an overview of the proof of Proposition 8.4.

9.1. **Preliminaries.** In preparation for the proof of Proposition 8.4, we introduce a richer set of colors, as follows:

Definition 9.1 (expanded alphabet of colors). Write $\mathscr{X} \equiv \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$. Define the expanded alphabets

$$\mathscr{A} \equiv \left\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}, \mathtt{w} \equiv \texttt{white}, \mathtt{v} \equiv \texttt{violet}\right\}, \quad \mathscr{S} \equiv \left\{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\right\}, \quad \mathscr{T} \equiv \left\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}, \mathtt{v}\right\}.$$

Let $\mathscr{G} = (V, F, E)$ be any (processed) k-sat instance, and let $\underline{\sigma} \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^E$ denote a valid (single-copy) coloring on $\mathscr{G}$. Given $\underline{\sigma}$, let $\underline{\varsigma} \in \mathscr{S}^E$ be defined by setting

$$\varsigma_e \equiv \begin{cases} \mathtt{w} \equiv \texttt{white} & \text{if } \sigma_e \in \{\mathtt{y}, \mathtt{g}, \mathtt{b}\} \text{ and } \underline{\sigma}_{\delta a \backslash e} \text{ contains at least two } \texttt{cyan} \text{ edges,} \\ \mathtt{y} & \text{if } \sigma_e = \mathtt{y}, \\ \mathtt{c} & \text{if } \sigma_e \in \{\mathtt{g}, \mathtt{b}\}. \end{cases}$$

Let $\underline{v} \in \mathscr{T}^E$ be defined by setting

$$v_e = \begin{cases} \mathtt{v} \equiv \texttt{violet} & \text{if } \sigma_e = \mathtt{r} \text{ and } \underline{\sigma}_{\delta v \backslash e} \text{ contains at least one } \texttt{red} \text{ edge,} \\ \sigma_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\} & \text{otherwise.} \end{cases}$$

Finally, let $\underline{\Sigma} \in \mathscr{A}^E$ be defined by setting

$$\Sigma_e \equiv \begin{cases} \mathtt{w} \equiv \texttt{white} & \text{if } \sigma_e \in \{\mathtt{y}, \mathtt{g}, \mathtt{b}\} \text{ and } \underline{\sigma}_{\delta a \backslash e} \text{ contains at least two } \texttt{cyan} \text{ edges,} \\ \mathtt{v} \equiv \texttt{violet} & \text{if } \sigma_e = \mathtt{r} \text{ and } \underline{\sigma}_{\delta v \backslash e} \text{ contains at least one } \texttt{red} \text{ edge,} \\ \sigma_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\} & \text{otherwise.} \end{cases}$$

The new colors $\{\mathtt{w}, \mathtt{v}\}$ indicate edges that are "flexible" in some sense: from a clause's perspective, $\texttt{white}$ indicates an edge that can be $\texttt{yellow}$, $\texttt{green}$, or $\texttt{blue}$ without violating the clause constraint. Similarly, from a variable's perspective, $\texttt{violet}$ indicates an edge that is $\texttt{red}$ but can also be $\texttt{blue}$ without violating the variable constraint.

Definition 9.2 (pair empirical measure in expanded alphabet). Throughout the remainder of this section, on a (processed) k-sat instance $\mathscr{G} = (V, F, E)$, we denote pair colorings in the expanded alphabets as

$$\begin{aligned} \underline{\sigma} &\equiv (\underline{\sigma}^1, \underline{\sigma}^2) \in \mathscr{X}^E \times \mathscr{X}^E, \\ \underline{\varsigma} &\equiv (\underline{\varsigma}^1, \underline{\varsigma}^2) \in \mathscr{S}^E \times \mathscr{S}^E, \\ \underline{v} &\equiv (\underline{v}^1, \underline{v}^2) \in \mathscr{T}^E \times \mathscr{T}^E, \\ \underline{\Sigma} &\equiv (\underline{\Sigma}^1, \underline{\Sigma}^2) \in \mathscr{A}^E \times \mathscr{A}^E. \end{aligned}$$

As before, we use $\omega \equiv (\omega_{L,j})_{L,j}$ to denote the pair empirical measure, which we assume to be judicious in the sense of Definition 3.41. We let $\varpi \equiv (\varpi_{L,j})_{L,j}$ denote the pair empirical measure of $(\underline{\sigma}, \underline{\Sigma})$, so that each entry $\varpi_{L,j}$ is a probability measure over elements $(\sigma, \Sigma) \in \mathscr{X}^2 \times \mathscr{A}^2$. The marginal of ϖ on the $\mathscr{X}^2$-coordinate is given by ω. Clearly, by the mappings of Definition 9.1, each $\varpi_{L,j}$ also induces a probability measure over $(\sigma, \varsigma) \in \mathscr{X}^2 \times \mathscr{S}^2$ and $(\sigma, v) \in \mathscr{X}^2 \times \mathscr{T}^2$. We let $\zeta_{L,j}$ denote the marginal law of ς under $\varpi_{L,j}$.

Recall that Proposition 8.4 concerns the optimization of $\boldsymbol{\Psi}_{\mathscr{D},2}(\omega)$ over $\omega \in \boldsymbol{I}_0$. The function $\boldsymbol{\Psi}_{\mathscr{D},2}$ is defined in the discussion leading up to Lemma 3.42, and we review it briefly here. If $v = (\dot{v}, \hat{v})$ is a (judicious) vertex empirical measure in the pair coloring model, then analogously to (3.48), its contribution to the second moment is given by

$$\mathbb{E}_{\mathscr{D}} Z^2[v] = \underbrace{\left\{\prod_T \binom{n_T}{n_T \dot{v}_T} \prod_L \binom{m_L}{m_L \hat{v}_L}\right\}}_{\substack{\text{number of colorings} \\ \text{prior to matching}}} \underbrace{\left\{\prod_t \binom{n_t}{n_t \pi_t}\right\}^{-1}}_{\substack{\text{probability of matching} \\ \text{to respect colorings}}} = \frac{\exp\{n\boldsymbol{\Phi}_{\mathscr{D},2}(v)\}}{n^{O(1)}}, \tag{9.1}$$

where $\boldsymbol{\Phi}_{\mathscr{D},2}$ is the analogue of (3.42) for the pair model:

$$\boldsymbol{\Phi}_{\mathscr{D},2}(v) = \frac{1}{n} \left\{ |V| \, \mathbb{E}_{\dot{\mathscr{D}}}[\mathcal{H}(\dot{v}_T)] + |F| \, \mathbb{E}_{\hat{\mathscr{D}}}[\mathcal{H}(\hat{v}_L)] - |E| \, \mathbb{E}_{\bar{\mathscr{D}}}[\mathcal{H}(\pi_t)] \right\}. \tag{9.2}$$

(Note that ν denotes a vertex empirical measure for the single-copy model in (3.42), but for the pair model in (9.2).) Write $\nu \sim \omega$ if ν has marginals ω. Then the contribution to the second moment from any (judicious) ω is

$$\mathbb{E}_{\mathscr{D}}\mathbf{Z}^2(\omega) = \sum_{\nu:\nu\sim\omega} \mathbb{E}_{\mathscr{D}}\mathbf{Z}^2[\nu] = \frac{\exp\{n\mathbf{\Psi}_{\mathscr{D},2}(\omega)\}}{n^{O(1)}},$$

where $\mathbf{\Psi}_{\mathscr{D},2}$ is the analogue of (3.49) for the pair model:

$$\mathbf{\Psi}_{\mathscr{D},2}(\omega) \equiv \mathbf{\Phi}_{\mathscr{D},2}(\nu^{\mathrm{op}}[\omega]), \quad \nu^{\mathrm{op}}[\omega] \equiv \arg\max_{\nu}\left\{\mathbf{\Phi}_{\mathscr{D},2}(\nu) : \nu \text{ is consistent with } \omega\right\}. \tag{9.3}$$

We will prove Proposition 8.4 by analyzing the constrained entropy maximization problems involved in the above definition of $\nu^{\mathrm{op}}[\omega]$. We now make two definitions which will be used throughout the section:

Definition 9.3. For a variable v of type T, let $\pi_v \equiv \pi_T$ denote the marginal on the (pair) frozen configuration spin x_v. Thus π_v is a probability measure on $\{+,-,\mathtt{f}\}^2$, which can be computed from π_e for any $e \in \delta v$: for example, $\pi_v(\mathtt{++}) = \pi_e(\mathtt{yy})$ for any $e \in \delta v(\mathtt{-})$. We define a **diverse variable** to be a variable v of type T such that

$$\pi_v\left(\{\mathtt{+-},\mathtt{-+}\}\right) \equiv \pi_T\left(\{\mathtt{+-},\mathtt{-+}\}\right) = 2\pi_T(\mathtt{+-}) \geq \frac{1}{4}.$$

We then define a **diverse clause** to be a clause that neighbors at least $k/10$ diverse variables — equivalently, it is a clause of type L such that

$$\sum_{j=1}^{k(L)} \mathbf{1}\left\{\pi_{L(j)}\left(\{\mathtt{yp},\mathtt{py}\}\right) \geq \frac{1}{4}\right\} = \sum_{j=1}^{k(L)} \mathbf{1}\left\{\pi_{L(j)}(\mathtt{yp}) \geq \frac{1}{8}\right\} \geq \frac{k}{10}.$$

The definition of diversity depends only on type, so we can also speak of **diverse variable types** and **diverse clause types**. Let $\mathbb{D}$ denote the collection of all diverse clause types. For a variable v of type T, we write

$$\not{\mathbb{D}}(v) \equiv \not{\mathbb{D}}(T) \equiv \sum_{t\in T}\sum_{L} \mathbf{1}\{L \notin \mathbb{D}\}\pi_{\mathscr{D}}(L\,|\,t) \tag{9.4}$$

for the expected number of non-diverse clauses incident to v under $\mathbb{P}_{\mathscr{D}}$.

Definition 9.4. Let $\epsilon_\circ > 0$ be a small absolute constant. A clause type L is termed **light** if

$$\max_{1\leq j\leq k(L)} \omega_{L,j}(\mathtt{rr}) \leq \frac{1}{2^{k(1+\epsilon_\circ)}}.$$

Otherwise we say that L is **heavy**. Let $\mathbb{L}$ denote the collection of light clause types. Recall from (9.4) in Definition 9.3 that $\not{\mathbb{D}}(v)$ denotes the expected number of non-diverse clauses next to v. Analogously define $\not{\mathbb{L}}(v)$ to be the expected number of non-light (i.e., heavy) clauses next to v.

We conclude this preliminary subsection with an overview of the proof of Proposition 8.4. Recall that the goal is to show that if the neighborhood profile $\mathscr{D}$ satisfies the expansion condition (8.5) from Definition 8.2, and $\omega = \omega(\mathscr{D})$ is any maximizer of $\mathbf{\Psi}_{\mathscr{D},2}$ over $\omega \in \mathbf{I}_0$ (see (8.6)), then $\omega_{L,j}$ satisfies the "*a priori*" estimates (8.7) whenever $L(j)$ is a strongly non-defective edge type in the sense of Remark 6.5. In what follows, if $e = (av)$ with $t_e \equiv t$, $j(t) = j$, and $L_a = L$, we will denote $\pi_e \equiv \pi_t$, $\omega_e \equiv \omega_{L,j}$, $\varpi_e \equiv \varpi_{L,j}$, and $\zeta_e \equiv \zeta_{L,j}$.

In Proposition 9.13 we will show that if v is a non-defective variable with $\not{\mathbb{D}}(v) = \not{\mathbb{L}}(v) = 0$, then ω_e satisfies the required estimates (8.7) for all $e \in \delta v$. The proof of Proposition 9.13 uses a reweighting argument that builds on the result of Proposition 7.28 from Section 7. However, unlike in Section 7, the analysis of Proposition 9.13 uses some additional information concerning white edges, which are only weakly dependent on the other edges sharing the same clause. In particular, in a nice clause a, Lemma 9.5 gives that the expected number of edges $e \in \delta a$ with $\varsigma_e \neq \mathtt{ww}$ is small. If the clause is furthermore diverse and light, then Lemma 9.7 gives that the expected number of edges $e \in \delta a$ with $(\varsigma_e)^i \neq \mathtt{w}$ for both $i = 1, 2$ is very small. Both these estimates are used in the proof of Proposition 9.13.

Of course, the main challenge in applying Proposition 9.13 is that it relies on the assumption $\not{\mathbb{D}}(v) = \not{\mathbb{L}}(v) = 0$ — that is, *all* clauses neighboring any variable of this type must be diverse and light. Thus a large part of this section is devoted to proving estimates leading to Proposition 9.17, which uses an expansion argument (under the

condition (8.5)) to show that *all* strongly non-defective clauses must be diverse and light. Proposition 9.13 can then be applied to give the conclusion of Proposition 8.4.

The proof of Proposition 9.17 proceeds roughly as follows. For the purposes of this overview we will ignore the presence of defective variables. For any $\omega \in I_0$, the empirical distribution of frozen spins $x_v \in \{\text{-},\text{+},\text{f}\}^2$ among all $v \in V$ must give weight roughly $1/4$ to each $x \in \{\text{-},\text{+}\}^2$. Of course, these may not be evenly distributed among the different variable types, so this does not imply that all variable types are diverse. However, an easy application (9.186) of Markov's inequality shows that at least a quarter of all variables must be diverse. The expansion condition (8.5) then implies that most clauses are diverse, since most clauses will be incident to more than $k/10$ diverse variables. We then show in Lemma 9.11 and Corollary 9.12 that if a variable v has a small value of $\mathbb{D}(v)$, then it will also have a small value of $\mathbb{L}(v)$. We show in Proposition 9.16 that if $\mathbb{D}(v)$ and $\mathbb{L}(v)$ are both small, then v will be diverse. To summarize, write S for the set of non-diverse variables, $\mathbb{D}$ for the set of non-diverse clauses, and $\mathbb{L}$ for the set of heavy clauses. Then

- $|S|$ can be bounded in terms of $|\mathbb{D}| + |\mathbb{L}|$ by Proposition 9.16;
- $|\mathbb{L}|$ can be bounded in terms of $|\mathbb{D}|$ by Lemma 9.11 and Corollary 9.12; and
- $|\mathbb{D}|$ can be bounded in terms of $|S|$ by the expansion condition (8.5).

Combining these gives a bound for $|S|$ in terms of $|S|$ itself, which we find is satisfied only if $|S| = 0$. This yields the conclusion of Proposition 9.17.

This concludes our overview for the proof of Proposition 8.4, and we now turn to the details of the proof. For the reader's reference, the dependency diagram of results in this section is given in Figure 8.

9.2. Entropy maximization around non-forcing clauses.

The goal of this subsection is to prove that, for clause types L satisfying certain conditions, $\omega_{L,j}(\varsigma \neq \text{ww})$ must be small, and $\omega_{L,j}(\varsigma^1 \neq \text{w}, \varsigma^2 \neq \text{w})$ must be even smaller, for all $1 \leq j \leq k(L)$. In later subsections we will prove estimates restricted to white edges — these results will transfer easily to estimates concerning all edges, since the results of this subsection show that most edges are white. The current subsection is organized as follows:

- In Lemma 9.5 we show that if L is a nice clause, then $\omega_{L,j}(\varsigma \neq \text{ww})$ must be small (at most $O(k^2/2^k)$) for all $1 \leq j \leq k(L)$.
- In Lemma 9.6 we show that if t is a non-compound edge type, and $\sigma \in \{\text{y},\text{b}\}^2$, then a lower bound on $\pi_t(\sigma)$ implies a comparable lower bound on $\omega_{L,j}(\sigma)$ for all $L(j) = t$.
- In Lemma 9.7 we prove that if a clause type L is nice, diverse, and light, then $\omega_{L,j}(\varsigma^1 \neq \text{w}, \varsigma^2 \neq \text{w}) \leq 2^{-k(1+\epsilon_\circ)}$ for all $1 \leq j \leq k(L)$. The proof of this result makes use of Lemma 9.6, which gives a lower bound on $\omega_{L,j}(\{\text{yb},\text{by}\})$ for at least $k/10$ indices $1 \leq j \leq k(L)$.

We now turn to the precise statements and proofs.

Lemma 9.5. *Let a be a nice clause of type L (meaning all its incident edges are nice in the sense of Definition 3.8). Given $\omega \in I_0$, let $\hat{v}_L$ be the probability measure on valid (pair) colorings $\underline{\sigma}_{\delta a}$ which maximizes entropy subject to edge marginals $(\omega_{L,j})_j$. By Definition 9.1, each $\underline{\sigma}_{\delta a}$ maps to a configuration $\underline{\varsigma}_{\delta a}$, so $\hat{v}_L$ also induces a probability measure on $(\mathcal{S}^2)^{\delta a}$. Under this measure we have*

$$\omega_{L,j}\left(\varsigma = \text{ww}\right) = \hat{v}_L\left(\varsigma_j = \text{ww}\right) \geq 1 - O\left(\frac{k^2}{2^k}\right)$$

for all $1 \leq j \leq k(L)$.

Proof. For the proof we will mostly suppress L from the notation, and write $\hat{v} \equiv \hat{v}_L$. We claim that for this proof it suffices to take σ in the reduced alphabet $\{\text{r},\text{y},\text{c}\}$. Indeed, each $\omega_{L,j}$ is a measure over $\{\text{r},\text{y},\text{g},\text{b}\}^2$, and naturally induces a measure $\{\text{r},\text{y},\text{c}\}^2$, which we temporarily denote $\tilde{\omega}_{L,j}$. Let $\tilde{v} \equiv \tilde{v}_L$ be the probability measure on valid (pair) colorings $\underline{\tilde{\sigma}}_{\delta a} \in (\{\text{r},\text{y},\text{c}\}^{\delta a})^2$ which maximizes entropy subject to edge marginals $(\tilde{\omega}_{L,j})_j$. Since the clause factor (2.19) does not distinguish between green and blue, the optimal $\hat{v} \equiv \hat{v}_L$ must take the form

$$\hat{v}_L(\underline{\sigma}_{\delta a}) = \tilde{v}(\underline{\tilde{\sigma}}_{\delta a}) \prod_{j=1}^{k(L)} \frac{\omega_{L,j}(\sigma_j)}{\tilde{\omega}_{L,j}(\tilde{\sigma}_j)}, \tag{9.5}$$
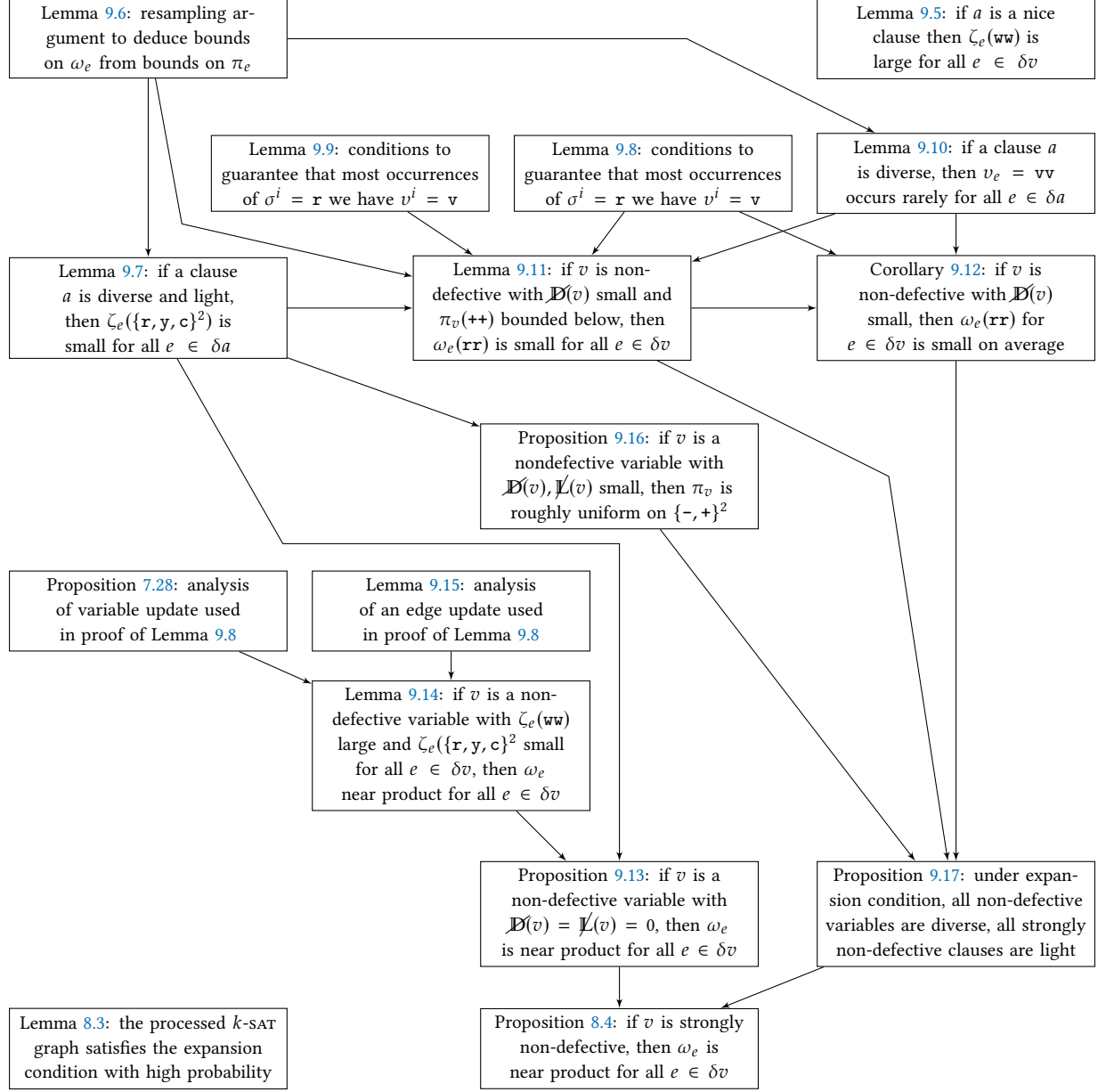
FIGURE 8. Dependency diagram of results of Section 9. Lemma 9.5 is a basic estimate used in the proofs of several subsequent claims, so the arrows leaving Lemma 9.5 are not shown to avoid cluttering the diagram further. Lemma 8.3 and Proposition 8.4 were both used in the proof of Proposition 3.32.

where $\underline{\sigma}_{\delta a} \in (\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^{\delta a})^2$ and $\underline{\tilde{\sigma}}_{\delta a}$ is its representative in $(\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^{\delta a})^2$. Thus it suffices to study $\tilde{v}$. For the remainder of the proof we will abuse notation slightly and write $\hat{v}, \underline{\sigma}_{\delta a}$ when we technically mean $\tilde{v}, \underline{\tilde{\sigma}}_{\delta a}$. The remainder of the proof is divided into a few numbered steps.

Step 1. Reduction to consideration of non-forcing clause colorings. The single-copy marginals of ω are given by the canonical measure $\star\pi$. Since L is assumed to be nice, it follows from Definition 3.8 that

$$\star\pi_{L(j)}(\mathtt{r}) \le \frac{O(1)}{2^k}, \quad \left|\star\pi_{L(j)}(\mathtt{c}) - \frac{1}{2}\right| \le O\left(\frac{1}{2^{k/10}}\right). \tag{9.6}$$

Let $\mathtt{u}$ denote the set of all **valid** pair colorings $\underline{\sigma}_{\delta a} \in \{\mathtt{y}, \mathtt{c}\}^{2K}$, where we abbreviate $K \equiv k(L)$. Note that if $\underline{\sigma}_{\delta a}$ is a valid coloring belonging to the complement $\mathtt{u}^c$, then $\underline{\sigma}_{\delta a}$ must have at least one edge colored red in at least one of

the two copies. It follows by the judicious condition together with (9.6) that

$$\hat{v}(\mathsf{u}^c) \leq \sum_{i=1,2} \sum_{j=1}^{K} (\omega_{L,j})^i(\mathsf{r}) = 2 \sum_{j=1}^{K} {}_\star\pi_{L(j)}(\mathsf{r}) \overset{(9.6)}{\leq} O\left(\frac{k}{2^k}\right). \tag{9.7}$$

Define the conditional measure $\mu(\underline{\sigma}_{\delta a}) \equiv \hat{v}(\underline{\sigma}_{\delta a} \mid \mathsf{u})$. Let μ_j be the marginal of μ on the j-th edge in δa, so μ_j is a probability measure on $\{\mathsf{y}, \mathsf{c}\}^2$. Since $\hat{v}(\mathsf{u})$ is close to one by (9.7), the single-copy marginals of μ_j must be close to ${}_\star\pi_{L(j)}$; in particular, we must have

$$(\mu_j)^i(\mathsf{c}) = \frac{{}_\star\pi_{L(j)}(\mathsf{c}) - \hat{v}(\mathsf{u}^c)\hat{v}((\sigma_j)^i = \mathsf{c} \mid \mathsf{u}^c)}{1 - \hat{v}(\mathsf{u}^c)} \overset{(9.7)}{=} {}_\star\pi_{L(j)}(\mathsf{c}) + O\left(\frac{1}{2^k}\right) \overset{(9.6)}{=} \frac{1}{2} + O\left(\frac{1}{2^{k/10}}\right). \tag{9.8}$$

Since $\hat{v} = \hat{v}_L$ maximizes entropy subject to marginals $(\omega_{L,j})_j$, it must be that μ maximizes entropy subject to marginals $(\mu_j)_j$. Consequently, by the method of Lagrange multipliers, there must exist probability measures q_j over $\{\mathsf{y}, \mathsf{c}\}^2$ (for $1 \leq j \leq K$) such that

$$\mu(\underline{\sigma}_{\delta a}) = \frac{\mathbf{1}\{\underline{\sigma}_{\delta a} \in \mathsf{u}\} Q(\underline{\sigma}_{\delta a})}{Q(\mathsf{u})}, \quad Q(\underline{\sigma}_{\delta a}) \equiv \prod_{j=1}^{K} q_j(\sigma_e). \tag{9.9}$$

In the next step we will estimate the q_j to show that under μ, each edge is ww with probability close to one. Since μ takes up most of the mass of $\hat{v}$ by (9.7), the result will follow.

Step 2. Construction of Lagrangian weights. We will iteratively construct a sequence $q_{j,t}$ that converges to the desired q_j of (9.9) in the limit $t \to \infty$. We initialize $q_{j,0} = \mu_j$ for all $1 \leq j \leq K$. Analogously to (9.9) let

$$\mu_t(\underline{\sigma}_{\delta a}) = \frac{\mathbf{1}\{\underline{\sigma}_{\delta a} \in \mathsf{u}\} Q_t(\underline{\sigma}_{\delta a})}{Q_t(\mathsf{u})}, \quad Q_t(\underline{\sigma}_{\delta a}) \equiv \prod_{j=1}^{K} q_{j,t}(\sigma_e).$$

Let $\mu_{j,t}$ denote the marginal of μ_t on the j-th edge. Writing $\underline{\sigma}_{-j} \equiv (\sigma_\ell)_{\ell \in [K] \setminus j}$, we have

$$\mu_{j,t}(\sigma) = \frac{1}{\bar{z}_{j,t}} q_{j,t}(\sigma) \overbrace{\sum_{\underline{\sigma}_{-j}} \mathbf{1}\{\underline{\sigma} \in \mathsf{u}\} \prod_{\ell \in [K] \setminus j} q_{\ell,t}(\sigma_\ell)}^{\text{denote this } \xi_{j,t}(\sigma)}. \tag{9.10}$$

Note that $\xi_{j,t}(\sigma)$ has a simple expression for each $\sigma \in \{\mathsf{y}, \mathsf{c}\}^2$: for instance, for $\sigma = \mathsf{cc}$ we have

$$\xi_{j,t}(\mathsf{cc}) = 1 - \sum_{i=1,2} \prod_{\ell \in [K] \setminus j} (q_{\ell,t})^i(\mathsf{y}) + \prod_{\ell \in [K] \setminus j} q_{\ell,t}(\mathsf{yy}),$$

and we have similar expressions for $\sigma \in \{\mathsf{yy}, \mathsf{yc}, \mathsf{cy}\}$. It is easily verified that $1 - k^{O(1)}/2^k \leq \xi_{j,0}(\sigma) \leq 1$ for all $\sigma \in \{\mathsf{y}, \mathsf{c}\}^2$, and substituting this estimate into (9.10) gives (crudely)

$$\left\| \frac{\mu_{j,0}}{\mu_j} - 1 \right\|_\infty \overset{(9.10)}{=} \left\| \frac{\xi_{j,0}}{\bar{z}_{j,0}} - 1 \right\|_\infty \leq \frac{1}{2^{2k/3}}. \tag{9.11}$$

Moreover, it follows by a straightforward calculation that

$$\left| \frac{\partial \xi_{j,t}(\sigma)}{\partial q_{\ell,t}(\sigma')} \right| \leq \frac{k^{O(1)}}{2^k} \tag{9.12}$$

for all $j \neq \ell$ and all $\sigma, \sigma' \in \{\mathsf{y}, \mathsf{c}\}^2$. We now define $q_{j,t+1}$ to be the probability measure such that

$$q_{j,t+1}(\sigma) \cong \frac{\mu_j(\sigma)}{\xi_{j,t}(\sigma)} \overset{(9.10)}{\cong} \frac{\mu_j(\sigma)}{\mu_{j,t}(\sigma)} q_{j,t}(\sigma). \tag{9.13}$$

Let $z_{j,t+1}$ denote the normalizing constant such that

$$q_{j,t+1}(\sigma) = \frac{\mu_j(\sigma)}{\mu_{j,t}(\sigma)} \frac{q_{j,t}(\sigma)}{z_{j,t+1}}. \tag{9.14}$$

Now suppose inductively that for all $t \geq 0$ we have

$$\left\|\frac{\mu_{j,t}}{\mu_j} - 1\right\|_\infty \leq \frac{1}{(2^{2k/3})^{t+1}}, \tag{9.15}$$

where the base case $t = 0$ is given by (9.11). Note that substituting (9.15) into (9.14) implies that $q_{j,t}$ is close to $z_{j,t+1}q_{j,t+1}$; since all the q's are probability measures, it further implies that $z_{j,t+1}$ is close to one. Thus, for $t \geq 0$, the inductive hypothesis (9.15) leads to

$$\left\|\xi_{j,t+1} - \xi_{j,t}\right\|_\infty \overset{(9.12)}{\leq} \frac{k^{O(1)}}{2^k}\left\|q_{j,t+1} - q_{j,t}\right\|_\infty \overset{(9.14)}{=} \frac{k^{O(1)}}{2^k}\left\|q_{j,t}\left(\frac{\mu_j}{\mu_{j,t}}\frac{1}{z_{j,t+1}} - 1\right)\right\|_\infty \overset{(9.15)}{\leq} \frac{k^{O(1)}}{2^k}\frac{1}{(2^{2k/3})^{t+1}}. \tag{9.16}$$

Recall that $\mu_{j,t+1} = q_{j,t+1}\xi_{j,t+1}/\bar{z}_{j,t+1}$ by (9.10), while (9.13) implies that there exists a normalizing constant $\acute{z}_{j,t+1}$ such that $\mu_j = q_{j,t+1}\xi_{j,t}/\acute{z}_{j,t+1}$. It follows that

$$\left\|\frac{\mu_{j,t+1}}{\mu_j} - 1\right\|_\infty = \left\|\frac{\xi_{j,t+1}/\bar{z}_{j,t+1}}{\xi_{j,t}/\acute{z}_{j,t+1}} - 1\right\|_\infty \overset{(9.16)}{\leq} \frac{1}{(2^{2k/3})^{t+2}},$$

which verifies the inductive hypothesis (9.15).

Step 3. Conclusion. Let $q_j = \lim_{t\to\infty} q_{j,t}$, where it is clear from (9.15) that the limit is well-defined. Moreover, summing (9.15) over $t \geq 0$ implies

$$\left\|\frac{q_j}{\mu_j} - 1\right\|_\infty = \left\|\frac{q_j}{q_{j,0}} - 1\right\|_\infty \leq \sum_{t\geq 0}\left\|\frac{q_{j,t+1}}{q_{j,t}} - 1\right\|_\infty$$

$$\overset{(9.14)}{\leq} O\left(\sum_{t\geq 0}\left\|\frac{\mu_{j,t}}{\mu_j} - 1\right\|_\infty\right) \overset{(9.15)}{\leq} O\left(\sum_{t\geq 0}\frac{1}{(2^{2k/3})^{t+1}}\right) \leq \frac{1}{2^{k/2}}.$$

We can then straightforwardly derive from (9.8), (9.9), and the last bound that

$$\mu\left(\varsigma_j = \mathtt{ww}\right) \geq 1 - O\left(\frac{k^2}{2^k}\right)$$

for all $1 \leq j \leq K$. The result follows since μ is the measure $\hat{\nu}$ conditioned on event $\mathtt{U}$, and we know that $\hat{\nu}(\mathtt{U}^c)$ is negligible by (9.7). □

Lemma 9.6. *Suppose $\omega \in I_0$, and let $\pi \equiv (\pi_t)_t$ be the marginal of ω — each π_t is obtained by averaging over $\omega_{L,j}$ such that $L \ni t$ (see Definition 3.25). Let*

$$\nu^{\mathrm{op}}[\pi] \equiv \arg\max_\nu\left\{\Phi_{\mathscr{D},2}(\nu) : \nu \text{ is consistent with } \pi\right\}.$$

This is a relaxation of (9.3), since if ν is consistent with ω then it is also consistent with π. The relaxation is also a strictly convex problem, so the maximizer $\nu^{\mathrm{op}}[\pi]$ is uniquely defined. Let $\acute{\omega}$ denote the marginals of $\nu^{\mathrm{op}}[\pi]$. Then, for every $\sigma \in \{\mathtt{y}, \mathtt{b}\}^2$ and every non-compound edge t with $\pi_t(\sigma) \geq 2^{-k/15}$, we have

$$\min\left\{\frac{\acute{\omega}_{L,j}(\sigma)}{\pi_t(\sigma)} : L(j) = t\right\} \geq \frac{1}{8},$$

where $j \equiv j(t)$. (If t is a compound edge type then there is only one clause type L with $L(j) = t$, so in this case $\omega_{L,j} = \pi_t$ and there is nothing to prove.)

Proof. The proof follows a familiar outline: we first reduce to a simplified constrained entropy maximization problem, then estimate the Lagrangian weights solving that problem to derive the conclusion.

Step 1. Simplified entropy maximization problem. From the expression (9.2) for $\Phi_{\mathscr{D},2}$, we see that to optimize $\Phi_{\mathscr{D},2}(\nu)$ given fixed π, we can optimize separately over the variable and clause empirical measures, $\dot{\nu}$ and $\hat{\nu}$. We can therefore consider $\hat{\nu}$ alone, since it determines $\acute{\omega}$. The optimal $\hat{\nu}$ is given by

$$\hat{\nu} = \arg\max_{\hat{\nu}}\left\{\mathbb{E}_{\hat{\mathscr{D}}}[\mathcal{H}(\hat{\nu}_L)] \equiv \sum_L \hat{\mathscr{D}}(L)\mathcal{H}(\hat{\nu}_L) : \hat{\nu} \text{ is judicious and consistent with } \pi\right\} \tag{9.17}$$

— this is because $\mathbb{E}_{\hat{\mathscr{D}}}[\mathcal{H}(\hat{v}_L)]$ is the only term of (9.2) that varies with $\hat{v}$ when π is fixed. Equivalently, recalling the combinatorial calculation (9.1), $\hat{v}$ must satisfy

$$\hat{v} = \arg\max_{\hat{v}} \left\{ \prod_L \binom{m_L}{m_L \hat{v}_L} \right\}, \tag{9.18}$$

that is to say, $\hat{v}$ is the empirical measure of clause colorings $\underline{\sigma}_{\delta a}$ that maximizes entropy subject to marginals π.

Now, as in the statement of the lemma, let us fix a non-compound edge type t such that $\pi_t(\sigma) \geq 2^{-k/20}$. Denote $j = j(t)$, and recall that we write $L \ni t$ if and only if $L(j) = t$. Given $\mathscr{G} = (V, F, E)$, define the subset of clauses

$$F(t) = \left\{ a \in F : L_a \ni t \right\} \subseteq F.$$

Note for all $L \ni t$ the clause width $k(L)$ equals the same value $K \in \{k-1, k\}$, since L must be compatible with t. Recall from Definition 9.1 that a clause coloring $\underline{\sigma} \in (\mathscr{X}^K)^2$ defines an element $\underline{\varsigma} \in (\mathscr{S}^K)^2$. Let

$$\mathbb{N} \equiv \mathbb{N}_j \equiv \left\{ \underline{\sigma} \in (\mathscr{X}^K)^2 : \sigma_j \in \{\mathtt{y}, \mathtt{b}\}^2 \text{ and } \varsigma_j = \mathtt{ww} \right\}.$$

Then, for any (pair) coloring $\underline{\sigma}$ on $\mathscr{G}$, let

$$F(t, \underline{\sigma}) \equiv \left\{ a \in F(t) : \underline{\sigma}_{\delta a} \in \mathbb{N} \right\} \subseteq F(t).$$

Within $F(t, \underline{\sigma})$ only, because each clause has $\varsigma_j = \mathtt{ww}$, we are free to reassign the value of σ_j to any other color in $\{\mathtt{y}, \mathtt{b}\}^2$, provided we continue to respect π and the judicious constraints. Let $\mathbf{P} \equiv \mathbf{P}_t$ be the probability measure

$$\mathbf{P}\left(L, \underline{\sigma}_{\delta a} = \underline{\sigma}\right) = \pi_{\mathscr{D}}(L \,|\, t) \cdot \hat{v}_L\left(\underline{\sigma}_{\delta a} = \underline{\sigma}\right),$$

so $\mathbf{P}$ represents the empirical measure of clause types and colorings within $F(t)$. Note that marginal of $\mathbf{P}$ on L is $\mathbf{P}(L) = \pi_{\mathscr{D}}(L \,|\, t)$, while the marginal on σ_j (for $j = j(t)$) is $L(\sigma_j = \sigma) = \pi_t(\sigma)$. Let ζ be the empirical measure for the j-th edges of the clauses in $F(t, \underline{\sigma})$, i.e.,

$$\zeta(\sigma, L) = \mathbf{P}\left(L, \sigma_j = \sigma \,\middle|\, \mathbb{N}\right),$$

where $\mathbb{N}$ is shorthand for the event that $\underline{\sigma} \in \mathbb{N}$. If $\hat{v}$ is the maximizer as in (9.17) or (9.18), then ζ must satisfy

$$\zeta = \arg\max_{\zeta} \left\{ \begin{array}{c} \mathcal{H}(\zeta) : \zeta(L) = \mathbf{P}(L \,|\, \mathbb{N}) \; \forall L, \\ \zeta(\sigma) = \mathbf{P}(\sigma_j = \sigma \,|\, \mathbb{N}) \; \forall \sigma \in \{\mathtt{y}, \mathtt{b}\}^2, \\ \zeta(\tau \,|\, L) = \mathbf{P}((\sigma_j)^i = \tau \,|\, \mathbb{N}, L) \; \forall \tau \in \{\mathtt{y}, \mathtt{b}\} \end{array} \right\}. \tag{9.19}$$

By the method of Lagrange multipliers, there exist (real-valued) weights $\gamma, \beta(L), (\beta_L)^1, (\beta_L)^2$ such that

$$\zeta(L, \sigma) \cong \exp\left\{ \gamma \mathbf{1}\{\sigma^1 \neq \sigma^2\} + \beta(L) + 2 \sum_{i=1,2} (\beta_L)^i \mathbf{1}\{\sigma^i = \mathtt{b}\} \right\}, \tag{9.20}$$

where γ can be chosen independently of L since its purpose is to enforce the constraint on $\zeta(\sigma)$. (The weights $(\beta_L)^i$ are multiplied by a factor of two for convenience in subsequent calculations.)

Step 2. Estimation of marginals for (9.20). Note that since t was assumed to be a non-compound edge type, any clause type $L \ni t$ must be nice. It then follows from the judicious condition, Definition 3.8, and Lemma 9.5 that for all $L \ni t$ we have

$$1 \geq \hat{v}_L(\mathbb{N}) \geq \hat{v}_L\left(\varsigma_j = \mathtt{ww}\right) - \sum_{i=1,2} \overbrace{\hat{v}_L\left((\sigma_j)^i = \mathtt{g}\right)}^{\star \pi_t(\mathtt{g})} \geq 1 - O\left(\frac{k^2}{2^k}\right). \tag{9.21}$$

As a result, the quantities appearing in (9.19) can be written more explicitly and estimated as follows:

$$\zeta(L) = \mathbf{P}(L \,|\, \mathbb{N}) = \frac{\pi_{\mathscr{D}}(L \,|\, t)\hat{v}_L(\mathbb{N})}{\sum_{L'} \pi_{\mathscr{D}}(L' \,|\, t)\hat{v}_{L'}(\mathbb{N})} \overset{(9.21)}{=} \pi_{\mathscr{D}}(L \,|\, t)\left\{ 1 - O\left(\frac{k^2}{2^k}\right) \right\},$$

i.e., the clause type proportions within $F(t, \underline{\sigma})$ are close to those within $F(t)$. Next, for each $\sigma \in \{y, b\}^2$, we have

$$\zeta(\sigma) = \mathbf{P}\left(\sigma_j = \sigma \,\Big|\, \mathbb{N}\right) = \frac{\pi_t(\sigma) - \mathbf{P}(\sigma_j = \sigma, \mathbb{N}^c)}{\mathbf{P}(\mathbb{N})} \overset{(9.21)}{=} \pi_t(\sigma) - O\left(\frac{k^2}{2^k}\right),$$

where we note that the right-hand side must be positive by the assumption that $\pi_t \geq 2^{-k/20}$. Finally, for each $\sigma \in \{y, b\}^2$ and each L, we have

$$\zeta(\sigma \mid L) = \mathbf{P}\left(\sigma_j = \sigma \,\Big|\, \mathbb{N}, L\right) = \frac{\omega_{L,j}(\sigma) - \mathbf{P}(\sigma_j = \sigma, \mathbb{N}^c \mid L)}{\mathbf{P}(\mathbb{N} \mid L)} \overset{(9.21)}{=} \omega_{L,j}(\sigma) - O\left(\frac{k^2}{2^k}\right), \tag{9.22}$$

where we have not yet shown the right-hand side to be positive. Taking the marginal on the i-th copy gives

$$\zeta(\sigma^i = \tau \mid L) = \mathbf{P}\left((\sigma_j)^i = \tau \,\Big|\, \mathbb{N}, L\right) = \frac{_\star\pi_t(\tau) - \mathbf{P}((\sigma_j)^i = \tau, \mathbb{N}^c \mid L)}{\mathbf{P}(\mathbb{N} \mid L)} \overset{(9.21)}{=} {}_\star\pi_t(\tau) - O\left(\frac{k^2}{2^k}\right)$$

for each $\tau \in \{y, b\}$. This concludes our estimates for the quantities appearing in (9.19). Next, we note that by the judicious condition, for all L, j we have

$$\omega_{L,j}(yb) - \omega_{L,j}(by) = \left\{\omega_{L,j}(yb) + \omega_{L,j}(yy)\right\} - \left\{\omega_{L,j}(by) + \omega_{L,j}(yy)\right\} = {}_\star\pi_t(y) - {}_\star\pi(y) = 0.$$

By the judicious condition together with the assumption that $t = L(j)$ is nice, we also have

$$\omega_{L,j}(yy) - \omega_{L,j}(bb) = \left\{\omega_{L,j}(yy) + \omega_{L,j}(by)\right\} - \left\{\omega_{L,j}(bb) + \omega_{L,j}(by)\right\} = {}_\star\pi(y) - {}_\star\pi(b) = O\left(\frac{1}{2^{k/10}}\right).$$

Combining with (9.22) gives, for all L, j with $L(j) = t$, the bound

$$\left|\zeta(yy \mid L) - \zeta(bb \mid L)\right| + \left|\zeta(yb \mid L) - \zeta(by \mid L)\right| \leq O\left(\frac{1}{2^{k/10}}\right). \tag{9.23}$$

With these estimates in hand, we now turn to estimating the weights in (9.20).

Step 3. Estimation of Lagrangian weights. If L satisfies $\zeta(\{yb, by\} \mid L) \geq 2^{-k/15}$, then it follows by combining with (9.23) that

$$\left|\frac{\zeta(yb \mid L)}{\zeta(by \mid L)} - 1\right| \leq \frac{|\zeta(yb \mid L) - \zeta(by \mid L)|}{\min\{\zeta(yb \mid L), \zeta(by \mid L)\}} \overset{(9.23)}{\leq} O\left(\frac{2^{-k/10}}{2^{-k/15}}\right) = O\left(\frac{1}{2^{k/30}}\right).$$

On the other hand, from the Lagrangian solution (9.20), we have

$$\frac{\zeta(yb \mid L)}{\zeta(by \mid L)} = \frac{\exp\{\beta(L) + \gamma + 2(\beta_L)^2\}}{\exp\{\beta(L) + \gamma + 2(\beta_L)^1\}} = \exp\left\{2\left[(\beta_L)^2 - (\beta_L)^1\right]\right\}.$$

Comparing the last two displays, we see that

$$\max\left\{\left|(\beta_L)^2 - (\beta_L)^1\right| : L \ni t \text{ and } \zeta(\{yb, by\} \mid L) \geq \frac{1}{2^{k/15}}\right\} \leq O\left(\frac{1}{2^{k/30}}\right). \tag{9.24}$$

An entirely similar argument gives

$$\max\left\{\left|(\beta_L)^2 + (\beta_L)^1\right| : L \ni t \text{ and } \zeta(\{yy, bb\} \mid L) \geq \frac{1}{2^{k/15}}\right\} \leq O\left(\frac{1}{2^{k/30}}\right). \tag{9.25}$$

Now suppose t is such that $\pi_t(\{yb, by\}) \geq \chi$ for some $2^{-k/15} \leq \chi \leq 1/2$. It implies that for **some** clause type $L \ni t$, we have $\zeta(\{yb, by\} \mid L) \geq \chi$. For this particular L, writing ch for the hyperbolic cosine function, we have

$$\chi \leq \frac{\chi}{1 - \chi} \leq \frac{\zeta(\{yb, by\} \mid L)}{\zeta(\{yy, bb\} \mid L)} = e^\gamma \frac{\exp\{2(\beta_L)^2\} + \exp\{2(\beta_L)^1\}}{1 + \exp\{2(\beta_L)^1 + 2(\beta_L)^2\}} = e^\gamma \frac{\text{ch}((\beta_L)^1 - (\beta_L)^2)}{\text{ch}((\beta_L)^1 + (\beta_L)^2)}$$

$$\overset{(9.24)}{\leq} \frac{e^\gamma[1 + O(2^{-k/30})]}{\text{ch}((\beta_L)^1 + (\beta_L)^2)} \leq e^\gamma\left(1 + O\left(\frac{1}{2^{k/30}}\right)\right). \tag{9.26}$$

Next, for **every** clause type L such that $L \ni t$, since we assumed $\chi \le 1/2$, at least one of the two quantities $\zeta(\{\mathtt{yb},\mathtt{by}\} \mid L)$ and $\zeta(\{\mathtt{yy},\mathtt{bb}\} \mid L)$ must be $\ge \chi$. If $\zeta(\{\mathtt{yy},\mathtt{bb}\} \mid L) \ge \chi$, then

$$\frac{\zeta(\{\mathtt{yb},\mathtt{by}\} \mid L)}{\zeta(\{\mathtt{yy},\mathtt{bb}\} \mid L)} = e^\gamma \frac{\mathrm{ch}((\beta_L)^1 - (\beta_L)^2)}{\mathrm{ch}((\beta_L)^1 + (\beta_L)^2)} \ge \frac{e^\gamma}{\mathrm{ch}((\beta_L)^1 + (\beta_L)^2)}$$

$$\overset{(9.25)}{\ge} \frac{e^\gamma}{1 + O(2^{-k/30})} \overset{(9.26)}{\ge} \chi\left(1 - O\left(\frac{1}{2^{k/30}}\right)\right).$$

Combining the two cases $\zeta(\{\mathtt{yb},\mathtt{by}\} \mid L) \ge \chi$ and $\zeta(\{\mathtt{yy},\mathtt{bb}\} \mid L) \ge \chi$ gives

$$\zeta(\{\mathtt{yb},\mathtt{by}\} \mid L) \ge \min\left\{\chi, \frac{\chi[1 - O(2^{-k/30})]}{1 + \chi[1 - O(2^{-k/30})]}\right\} \ge \frac{2\chi}{3}\left(1 - O\left(\frac{1}{2^{k/30}}\right)\right).$$

Recalling (9.22) again, this proves that if $\pi_t(\sigma) \ge 2^{-k/15}$ for $\sigma \in \{\mathtt{yb},\mathtt{by}\}$, then

$$\omega_{L,j}(\sigma) \ge \frac{\chi}{4} \ge \frac{\pi_t(\sigma)}{8},$$

where the last bound follows by taking $\chi = \min\{\pi_t(\sigma), 1/2\}$. The analogous result for $\sigma \in \{\mathtt{yy},\mathtt{bb}\}$ by a symmetric argument, and this concludes the proof. $\qquad\square$

Lemma 9.7. *Let a be a clause of type L which is both nice (meaning t is nice for all $t \in L$) and diverse (Definition 9.3). Given $\omega \in I_0$, let $\hat{v}_L$ be the probability measure on valid (pair) colorings $\underline{\sigma}_{\delta a}$ which maximizes entropy subject to edge marginals $(\omega_{L,j})_j$. Recall from Definition 9.1 that each $\underline{\sigma}_{\delta a}$ maps to a configuration $\underline{\varsigma}_{\delta a} \in (\mathcal{S}^2)^{\delta a}$. Let*

$$\mathtt{A} \equiv \{\underline{\sigma}_{\delta a} : \sigma_e \ne \mathtt{rr} \text{ for all } e \in \delta a\},$$
$$\mathtt{W}^i \equiv \{\underline{\sigma}_{\delta a} : (\varsigma_e)^i = \mathtt{w} \text{ for all } e \in \delta a\}.$$

There exists an absolute constant $\epsilon_1 > 0$ such that

$$1 - \hat{v}_L\left(\mathtt{W}^1 \cup \mathtt{W}^2 \,\middle|\, \mathtt{A}\right) \le \frac{1}{2^{k(1+\epsilon_1)}}. \tag{9.27}$$

If in addition L is light (Definition 9.4), then it follows immediately that

$$1 - \hat{v}_L\left(\mathtt{W}^1 \cup \mathtt{W}^2\right) \le \frac{1}{2^{k(1+\epsilon_1)}} + \frac{k}{2^{k(1+\epsilon_\circ)}} \le O\left(\frac{k}{2^{k(1+\epsilon_\circ)}}\right), \tag{9.28}$$

where we can assume that $\epsilon_\circ \le \epsilon_1$.

Proof. It is an immediate consequence of Lemma 9.6 that if L is both nice and diverse, then

$$\sum_{j=1}^{k(L)} \mathbf{1}\left\{\omega_{L,j}\left(\{\mathtt{by},\mathtt{yb}\}\right) \ge \frac{1}{33}\right\} \ge \sum_{j=1}^{k(L)} \mathbf{1}\left\{\omega_{L,j}(\mathtt{yb}) \ge \frac{1}{65}\right\} \ge \frac{k}{10}. \tag{9.29}$$

In this proof, by the same reasoning as in the proof of Lemma 9.5 (see (9.5)), it suffices to take σ in the reduced alphabet $\{\mathtt{r},\mathtt{y},\mathtt{c}\}$. We also abbreviate $\mathtt{u} \equiv \{\mathtt{y},\mathtt{c}\}$. Throughout the following, ϵ_1 denotes a small positive number, whose value may change from one occurrence to the next, but ultimately is taken as an absolute constant. We denote the clause width by $K \equiv k(L) \in \{k-1, k\}$.

Step 1. Reduction to entropy maximization for a conditional measure. As in the statement of the lemma, let $\hat{v}_L$ be the optimizer given marginals ω. Note that since ω is judicious and L is nice,

$$\hat{v}_L(\mathtt{A}^c) \le \sum_{j=1}^{K} \omega_{L,j}(\mathtt{rr}) \le \sum_{i=1,2} \sum_{j=1}^{K} (\omega_{L,j})^i(\mathtt{r}) = 2 \sum_{j=1}^{K} {}_\star \pi_{L(j)}(\mathtt{r}) \le O\left(\frac{k}{2^k}\right).$$

Let $\mu(\underline{\sigma}) \equiv \hat{v}_L(\underline{\sigma} \mid \mathtt{A})$ for $\underline{\sigma} \in \{\mathtt{r},\mathtt{y},\mathtt{c}\}^{2K}$. For each edge $1 \le j \le K$, let μ_j be the marginal of μ on the j-th edge in δa. Thus μ_j is a probability measure over $\{\mathtt{r},\mathtt{y},\mathtt{c}\}^2 \setminus \{\mathtt{rr}\}$, and the preceding estimate implies

$$\mu_j(\sigma) = \frac{\hat{v}_L(\sigma_j = \sigma) - \hat{v}_L(\sigma_j = \sigma; \mathtt{A}^c)}{1 - \hat{v}_L(\mathtt{A}^c)} = \begin{cases} \omega_{L,j}(\sigma) + O(k/2^k) & \text{if } \mathtt{r}[\sigma] = 0, \\ \omega_{L,j}(\sigma) + O(1/2^k) & \text{if } \mathtt{r}[\sigma] = 1, \end{cases} \tag{9.30}$$

where the last bound uses that $\hat{v}_L(\sigma_j = \sigma) \le O(1/2^k)$ if $\mathbf{r}[\sigma] = 1$. The measure μ maximizes entropy subject to the edge marginals μ_j. Similarly as in the preceding proofs of this section, we shall construct a sequence of Lagrangian weights $b_{j,t}$ such that the measure

$$\mu_t(\underline{\sigma}) = \frac{\mathbf{1}\{\underline{\sigma} \in \mathtt{A}\}}{Z_t} \prod_{j=1}^{K} b_{j,t}(\sigma_j) \tag{9.31}$$

converges to the desired μ as $t \to \infty$, where Z_t is the normalizing constant. Let $\mu_{j,t}$ denote the marginals of the measure μ_t.

Step 2. Initalization of Lagrangian weights. We initialize the construction with

$$b_{j,0}(\sigma) \equiv \frac{\mu_j(\sigma)}{1/(2^{K-1})^{\mathbf{r}[\sigma]}}, \quad \sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2 \setminus \{\mathtt{rr}\}. \tag{9.32}$$

We first estimate the marginals $\mu_{j,0}$ of the resulting measure μ_0. Recall that we denote $\mathtt{u} \equiv \{\mathtt{y}, \mathtt{c}\}$. Note that (9.30), together with the judicious condition and the assumption that L is nice, gives

$$\mu_j(\mathtt{uu}) = 1 - O\left(\frac{k}{2^k}\right), \quad \sum_{\tau \in \{\mathtt{y},\mathtt{c}\}} \left\{ \left| \mu_j(\tau\mathtt{u}) - \frac{1}{2} \right| + \left| \mu_j(\mathtt{u}\tau) - \frac{1}{2} \right| \right\} \le \frac{1}{2^{k\epsilon_1}}. \tag{9.33}$$

Moreover, it follows from the diverse clause condition (9.32) together with (9.30) that

$$\sum_{j=1}^{k(L)} \mathbf{1}\left\{ \min\left\{ \mu_j(\mathtt{yb}), \mu_j(\mathtt{by}) \right\} \ge \frac{1}{65} \right\} \ge \frac{k}{10}. \tag{9.34}$$

Without loss of generality we now focus on the marginal on the edge indexed $j = 1$: at time $t = 0$,

$$Z_0 \mu_{1,0}(\mathtt{cc}) = \mu_1(\mathtt{cc})\left\{ \prod_{\ell=2}^{K} \mu_\ell(\mathtt{uu}) - \prod_{\ell=2}^{K} \mu_\ell(\mathtt{uy}) - \prod_{\ell=2}^{K} \mu_\ell(\mathtt{yu}) + \prod_{\ell=2}^{K} \mu_\ell(\mathtt{yy}) \right\} \overset{(9.33)}{=} \mu_1(\mathtt{cc})\left\{ 1 + O\left(\frac{k^2}{2^k}\right) \right\},$$

Similar calculations, again using (9.33), give the analogous estimate for $\sigma \in \{\mathtt{y}, \mathtt{c}\}^2$. Next, for $\sigma = \mathtt{rc}$, it follows from (9.33) together with the diversity bound (9.34) gives

$$Z_0 \mu_{1,0}(\mathtt{rc}) = \left(\mu_1(\mathtt{rc}) 2^{K-1} \right)\left\{ \prod_{\ell=2}^{K} \mu_\ell(\mathtt{yu}) - \prod_{\ell=2}^{K} \mu_\ell(\mathtt{yy}) \right\}$$

$$= \left(\mu_1(\mathtt{rc}) 2^{K-1} \right) \frac{1 + O(k/2^{k\epsilon_1})}{2^{K-1}} \left\{ 1 - \left(1 - \frac{2}{65} \right)^{k/10} \right\} = \mu_1(\mathtt{rc})\left\{ 1 + O\left(\frac{1}{2^{k\epsilon_1}}\right) \right\}$$

(recalling our convention that ϵ_1 can change from one expression to the next, but remains bounded below by a positive absolute constant). A similar calculation (again using (9.33) and (9.34)) gives an analogous estimate for $\sigma = \mathtt{ry}$. Altogether we conclude

$$\mu_{1,0}(\sigma) = \frac{\mu_1(\sigma)}{Z_0}\left\{ 1 + O\left(\frac{1}{2^{k\epsilon_1}}\right) \right\} = \mu_1(\sigma)\left\{ 1 + O\left(\frac{1}{2^{k\epsilon_1}}\right) \right\}. \tag{9.35}$$

for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2 \setminus \{\mathtt{rr}\}$.

Step 3. Iterative analysis of Lagrangian weights. Suppose at time t that we have weights $b_{j,t}$, which define a measure μ_t as in (9.31) with edge marginals $\mu_{j,t}$. We define the updated weights at time $t + 1$ by

$$\frac{b_{j,t+1}(\sigma)}{b_{j,t}(\sigma)} = \frac{\mu_j(\sigma)}{\mu_{j,t}(\sigma)}. \tag{9.36}$$

We will prove by induction that for all $t \ge 0$, all $1 \le j \le k(L)$, and all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2 \setminus \{\mathtt{rr}\}$,

$$\left| \frac{b_{j,t+1}(\sigma)}{b_{j,t}(\sigma)} - 1 \right| = \left| \frac{\mu_j(\sigma)}{\mu_{j,t}(\sigma)} - 1 \right| \le \frac{y_{\mathbf{r}[\sigma]}}{2^{k(t+1)\epsilon_1/3}}, \quad \binom{y_0}{y_1} = \frac{1}{2^{3k\epsilon_1/2}} \binom{1}{2^{k\epsilon_1/2}}, \tag{9.37}$$

where the base case $t = 0$ is implied by (9.35) (adjusting ϵ_1 appropriately). Suppose then that (9.37) holds up to time $t - 1 \geq 0$, and note it implies that for all j, σ we have

$$\left| \ln \frac{b_{j,t}(\sigma)}{\mu_j(\sigma)(2^{K-1})^{\mathbf{r}[\sigma]}} \right| = \left| \ln \frac{b_{j,t}(\sigma)}{b_{j,0}(\sigma)} \right| \leq \sum_{s=0}^{t-1} \left| \ln \frac{b_{j,s+1}(\sigma)}{b_{j,s}(\sigma)} \right| \overset{(9.37)}{\leq} O\left(\frac{1}{2^{k\epsilon_1}} \right). \tag{9.38}$$

To analyze the update (9.36), let us focus on the first two edges in the clause, and note that

$$\frac{\partial \mu_{1,t}(\sigma)}{\partial \ln b_{2,t}(\sigma')} = \mathrm{Cov}_{\mu_t}\left(\mathbf{1}\{\sigma_1 = \sigma\}, \mathbf{1}\{\sigma_2 = \sigma'\} \right) \tag{9.39}$$

for all $\sigma, \sigma' \in \{\mathbf{r}, \mathbf{y}, \mathbf{c}\}^2 \setminus \{\mathbf{rr}\}$ (by direct calculation). We therefore define

$$C_t(\sigma, \sigma') \equiv \frac{Z_t \mu_t(\sigma_1 = \sigma, \sigma_2 = \sigma')}{\mu_{j=1}(\sigma)\mu_{j'=2}(\sigma')},$$

and proceed to estimate this quantity. In the simplest case $\sigma = \sigma' = \mathbf{cc}$, we have

$$C_t(\mathbf{cc}, \mathbf{cc}) = \frac{b_{1,t}(\mathbf{cc})b_{2,t}(\mathbf{cc})}{\mu_1(\mathbf{cc})\mu_2(\mathbf{cc})} \prod_{j=3}^{K} b_{j,t}(\mathbf{uu}) \overset{(9.38)}{=} \left\{ \prod_{j=3}^{K} \mu_j(\mathbf{uu}) \right\} \left\{ 1 + O\left(\frac{1}{2^{k\epsilon_1}} \right) \right\} \overset{(9.33)}{=} 1 + O\left(\frac{1}{2^{k\epsilon_1}} \right).$$

A similar estimate holds for all cases where $\mathbf{r}[\sigma] + \mathbf{r}[\sigma'] = 0$. Next,

$$C_t(\mathbf{rc}, \mathbf{yc}) = \frac{b_{1,t}(\mathbf{rc})b_{2,t}(\mathbf{yc})}{\mu_1(\mathbf{rc})\mu_2(\mathbf{yc})} \prod_{j=3}^{K} b_{j,t}(\mathbf{yu}) \overset{(9.38)}{=} \left\{ 2^{K-1} \prod_{j=3}^{K} \mu_j(\mathbf{yu}) \right\} \left\{ 1 + O\left(\frac{1}{2^{k\epsilon_1}} \right) \right\} \overset{(9.33)}{=} 2 \left\{ 1 + O\left(\frac{1}{2^{k\epsilon_1}} \right) \right\},$$

and a similar estimate holds for $C_t(\sigma, \sigma')$ in all cases where $\mathbf{r}[\sigma] + \mathbf{r}[\sigma'] = 1$. Finally,

$$C_t(\mathbf{ry}, \mathbf{yr}) = \frac{b_{1,t}(\mathbf{ry})b_{2,t}(\mathbf{yr})}{\mu_1(\mathbf{ry})\mu_2(\mathbf{yr})} \prod_{j=3}^{K} b_{j,t}(\mathbf{yy}) \overset{(9.38)}{=} (2^{K-1})^2 \prod_{j=3}^{K} b_{j,t}(\mathbf{yy}) \left\{ 1 + O\left(\frac{1}{2^{k\epsilon_1}} \right) \right\} \leq \frac{2^k}{2^{k\epsilon_1}},$$

where the last bound uses the diversity bound (9.34). Combining with (9.39) gives

$$\left| \frac{\partial \mu_{1,t}(\sigma)/\partial \ln b_{2,t}(\sigma')}{\mu_{1,t}(\sigma)\mu_{2,t}(\sigma')} \right| = \left| \overbrace{\left\{ \frac{Z_t \mu_t(\sigma_1 = \sigma, \sigma_2 = \sigma')}{\mu_1(\sigma)\mu_2(\sigma')} \right\}}^{C_t(\sigma,\sigma')} \overbrace{\left\{ \frac{\mu_1(\sigma)\mu_2(\sigma')}{Z_t \mu_{1,t}(\sigma)\mu_{2,t}(\sigma')} \right\}}^{1+O(2^{-k\epsilon_1})} - 1 \right|$$

$$\leq O(1) \begin{cases} 2^{-k\epsilon_1} & \text{if } \mathbf{r}[\sigma] + \mathbf{r}[\sigma'] = 0, \\ 1 & \text{if } \mathbf{r}[\sigma] + \mathbf{r}[\sigma'] = 1, \\ 2^{k(1-\epsilon_1)} & \text{if } \mathbf{r}[\sigma] = \mathbf{r}[\sigma'] = 1. \end{cases}$$

Rewriting the above in a more convenient form gives

$$\frac{1}{\mu_1(\sigma)} \left| \frac{\partial \mu_{1,t}(\sigma)}{\partial \ln b_{2,t}(\sigma')} \right| \leq O\left(A_{\mathbf{r}[\sigma],\mathbf{r}[\sigma']} \right), \quad A = \begin{pmatrix} A_{0,0} & A_{0,1} \\ A_{1,0} & A_{1,1} \end{pmatrix} = \begin{pmatrix} 2^{-k\epsilon_1} & 2^{-k} \\ 1 & 2^{-k\epsilon_1} \end{pmatrix}. \tag{9.40}$$

Now recall the definition (9.31) of μ_t, and define

$$\tilde{\mu}_t(\underline{\sigma}) = \frac{\mathbf{1}\{\underline{\sigma} \in A\}}{Z_t} b_{1,t}(\sigma_1) \prod_{j=2}^{K} b_{j,t-1}(\sigma_j).$$

It follows from the update rule (9.36) that the marginal of $\tilde{\mu}_t$ on the first edge is precisely μ_1. Consequently,

$$
\left| \frac{b_{1,t+1}(\sigma)}{b_{1,t}(\sigma)} - 1 \right| \overset{(9.36)}{=} \left| \frac{\tilde{\mu}_t(\sigma_1 = \sigma)}{\mu_t(\sigma_1 = \sigma)} - 1 \right| \leq \frac{O(1)}{\mu_1(\sigma)} \sum_{j=2}^{K} \sum_{\sigma'} \left| \frac{\partial \mu_{1,t}(\sigma)}{\partial \ln b_{j,t}(\sigma')} \right| \cdot \left| \ln \frac{b_{j,t}(\sigma')}{b_{j,t-1}(\sigma')} \right|
$$

$$
\overset{(9.40)}{\leq} \sum_{\sigma'} A_{\mathbf{r}[\sigma],\mathbf{r}[\sigma']} \sum_{j=2}^{K} \sum_{\sigma'} \left| \ln \frac{b_{j,t}(\sigma')}{b_{j,t-1}(\sigma')} \right| \overset{(9.37)}{\leq} \frac{k}{2^{kt\epsilon_1/3}} \frac{1}{2^{3k\epsilon_1/2}} \sum_{\sigma'} A_{\mathbf{r}[\sigma],\mathbf{r}[\sigma']} (2^{k\epsilon_1/2})^{\mathbf{r}[\sigma]}
$$

$$
\overset{(9.40)}{\leq} \frac{1}{2^{k(t+1)\epsilon_1/3}} \frac{(2^{k\epsilon_1/2})^{\mathbf{r}[\sigma]}}{2^{3k\epsilon_1/2}} .
$$

This verifies the induction and proves (9.37) for all $t \geq 0$. Taking $t \to \infty$ gives the weights $b = b_\infty$ for the optimal measure μ. We can then use this to estimate

$$
1 - \mu\left(\mathbf{w}^1 \cup \mathbf{w}^2 \right) \leq \mu\left(\sum_{j=1}^{K} \mathbf{1}\{\sigma_j \in \{\mathtt{cc}, \mathtt{cy}, \mathtt{yc}\}\} \leq 4 \right) \leq \mu\left(\sum_{j=1}^{K} \mathbf{1}\{\sigma_j = \mathtt{yy}\} \geq K - 6 \right)
$$

$$
\leq \sum_{|S| \leq 6} \prod_{j \in [K] \setminus S} b_j(\mathtt{yy}) \leq \frac{1}{2^{k(1+\epsilon_1)}} ,
$$

where the last bound follows from (9.37) together with the diversity bound (9.34). This implies the first assertion (9.27) of the lemma. The second assertion (9.28) (removing the conditioning on $\mathtt{A}$) follows trivially, since the assumption that L is light means (by Definition 9.4) that $\omega_{L,j}(\mathtt{rr}) \leq 2^{-k(1+\epsilon_\circ)}$ for all j, and so a union bound gives

$$
\hat{v}_L(\mathtt{A}^c) \leq \sum_{j=1}^{k(L)} \omega_{L,j}(\mathtt{rr}) \leq \frac{k}{2^{k(1+\epsilon_\circ)}} .
$$

It follows by combining with (9.27) that

$$
\hat{v}_L\left((\mathbf{w}^1 \cup \mathbf{w}^2)^c \right) \leq \hat{v}_L(\mathtt{A}^c) + \hat{v}_L\left((\mathbf{w}^1 \cup \mathbf{w}^2)^c \,\middle|\, \mathtt{A}^c \right) \leq \frac{1}{2^{k(1+\epsilon_1)}} + \frac{k}{2^{k(1+\epsilon_\circ)}} ,
$$

as claimed. $\qquad\qquad\square$

9.3. Entropy maximization around forced variables.

In this subsection we show that under certain conditions, edges with $\sigma^i = \mathtt{red}$ are likely to have $v^i = \mathtt{violet}$. The precise statements are given in Lemmas 9.8 and 9.9 below. In later subsections we will prove bounds on $v = \mathtt{violet}$ edges, and use the results from this subsection to deduce bounds on $\sigma = \mathtt{red}$ edges.

In §9.2 we worked with configurations $\underline{\varsigma} \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^{2E}$ — we noted that $\underline{\sigma}_{\delta a}$ determines $\underline{\varsigma}_{\delta a}$ for each clause a, so to estimate the joint distribution of (σ, ς) under $\varpi_{L,j}$ it suffices to consider only the clause measures $\hat{v}_L$. In this subsection, however, the aim is to prove estimates concerning the configurations $\underline{v} \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}, \mathtt{v}\}^{2E}$, which can no longer be determined from the clause measures $\hat{v}_L$. Instead, $\underline{\sigma}_{\delta v}$ determines $\underline{v}_{\delta v}$ for each variable v. However, the joint distribution of (σ, v) under $\varpi_{L,j}$ cannot be inferred from only the variables measures $\dot{v}_T$, since those do not account for the distribution of clause types.

To resolve these issues, we again use the device of augmenting the alphabet with the clause type, similarly as in Definition 3.45. On a graph $\mathscr{G} = (V, F, E)$ let $(\underline{\sigma}, \underline{L})$ denote an augmented pair coloring. Let Ω denote the vertex empirical measure for the augmented coloring: for each variable type T we let Ω_T be the empirical measure of augmented colorings $(\underline{\sigma}, \underline{L})_{\delta v}$ on variables v of type T. Likewise we define Ω_t for edge types t, and Ω_L for clause types L. The combinatorial calculation (9.1) also implies that the contribution of Ω to the second moment is

$$
\mathbb{E}_{\mathscr{D}} Z^2[\Omega] = \underbrace{\left\{ \prod_T \binom{n_T}{n_T \Omega_T} \prod_L \binom{m_L}{m_L \Omega_L} \right\}}_{\substack{\text{number of colorings} \\ \text{prior to matching}}} \underbrace{\left\{ \prod_t \binom{n_t}{n_t \Omega_t} \right\}^{-1}}_{\substack{\text{probability of matching} \\ \text{to respect colorings}}} = \frac{\exp\{n \mathbf{\Phi}_{\mathscr{D},2}(\Omega)\}}{n^{O(1)}} , \tag{9.41}
$$

We write $\Omega \sim \nu$ if Ω is consistent with $\nu = (\dot{\nu}, \hat{\nu})$: in this case,

$$\Omega_L(\underline{\sigma}, \underline{L}) = \left\{ \prod_{j=1}^{k(L)} \mathbf{1}\{L_j = L\} \right\} \hat{\nu}_L(\underline{\sigma}) \tag{9.42}$$

If $\Omega \sim \nu$ and $\nu \sim \omega$, then we must have $\Omega_t = (\mathrm{aug}_{\mathscr{D}}(\omega))_t$ in the notation of (3.76). Thus we see that only the variable measures Ω_T carry more information than ν.

Lemma 9.8. *Let ϵ_2 be a small absolute constant. Fix $x \in \{+, -\}$, and let v be a variable of type T. Suppose we are given $\omega \in I_0$ with marginal π such that for both $i = 1, 2$ we have*

$$\gamma^i \equiv \sum_{e \in \delta v(x)} \pi_e \left(\sigma^i = \mathbf{r} \,\middle|\, \sigma \in \{\mathbf{r}, \mathbf{b}\}^2 \right) \geq k\epsilon_2 \,, \tag{9.43}$$

where $\pi_e \equiv \pi_t$ for an edge e of type t. Let Ω be the maximizer of (9.41) that is consistent with ω, and let $\varpi_{L,j}$ be the resulting joint distribution of (σ, v). Then there exists an absolute constant ϵ_3 (depending only on ϵ_2) such that for every edge $e \in \delta v(x)$ we have

$$\left. \begin{array}{l} \varpi_{L,j}(v = \mathtt{vv} \mid \sigma = \mathtt{rr}) \\ \varpi_{L,j}(v^1 = \mathtt{v} \mid \sigma = \mathtt{rb}) \\ \varpi_{L,j}(v^2 = \mathtt{v} \mid \sigma = \mathtt{br}) \end{array} \right\} \geq 1 - \frac{1}{2^{k\epsilon_3}}$$

for $j = j(t)$ and all clause types L such that $L(j) = t$.

Proof. Throughout the proof, ϵ_3 denotes a small positive number, whose value may change from one occurrence to the next, but ultimately is taken as an absolute constant that depends only on ϵ_2.

Step 1. Reduction to entropy maximization for a conditional measure. Note from (9.41) that if we fix Ω_t, then we obtain separate entropy maximization problems over Ω_T and Ω_L. Let us therefore fix $\Omega_t = (\mathrm{aug}_{\mathscr{D}}(\omega))_t$, and consider the optimization over Ω_T. Let v denote a variable of type T, and write $\Omega_v \equiv \Omega_T$. By the method of Lagrange multipliers, the optimal Ω_v must take the form

$$\Omega_v \left((\underline{\sigma}, \underline{L})_{\delta v} \right) = \frac{\varphi_v(\underline{\sigma}_{\delta v})}{\dot{z}} \prod_{e \in \delta v} \hat{q}_e(\sigma_e, L_e) \,,$$

where each $\hat{q}_e$ is a probability measure over elements (σ, L), chosen such that Ω_v has marginals $\Omega_e \equiv \Omega_t$ for $e \in \delta v$ and $t \equiv t_e$. The marginal of Ω_v on $(\underline{\sigma}_{\delta v}, L_e)$ is

$$\Omega_v(\underline{\sigma}_{\delta v}, L_e) = \frac{\varphi_v(\underline{\sigma}_{\delta v})}{\dot{z}} \hat{q}_e(\sigma_e, L_e) \prod_{e' \in \delta v \backslash e} \bar{q}_{e'}(\sigma_{e'}) \,,$$

where $\bar{q}_{e'}$ denotes the marginal of $\hat{q}_{e'}$ on $\sigma_{e'}$ alone. It follows that

$$\Omega_T \left(\underline{\sigma}_{\delta v \backslash e} \,\middle|\, (\sigma_e, L_e) \right) = \dot{\nu}_T \left(\underline{\sigma}_{\delta v \backslash e} \,\middle|\, \sigma_e \right) \cong \varphi_v(\underline{\sigma}_{\delta v}) \prod_{e' \in \delta v \backslash e} \bar{q}_{e'}(\sigma_{e'}) . \tag{9.44}$$

Thus, for the purposes of this lemma, it suffices to estimate only $\dot{\nu}_T(\cdot \mid \sigma_e)$. To this end, assume without loss $x = +$, and consider the event $x_v = ++$. In this case $\sigma_e \in \{\mathbf{r}, \mathbf{b}\}^2$ for all $e \in \delta v(+)$, and $\sigma_e = \mathtt{yy}$ for all $e \in \delta v(-)$. Let

$$\mu(\underline{\sigma}) \equiv \dot{\nu}_T \left(\underline{\sigma}_{\delta v(+)} = \underline{\sigma} \,\middle|\, x_v = ++ \right). \tag{9.45}$$

Denote $D \equiv |\delta v(+)|$. Then μ is a probability measure over colorings $\underline{\sigma} \in \{\mathbf{r}, \mathbf{b}\}^{2D}$ with at least one red spin in each copy $i = 1, 2$. Moreover, μ must maximize entropy subject to its marginals μ_ℓ. It follows by the method of Lagrange multipliers that there exist probability measures h_ℓ over $\{\mathbf{r}, \mathbf{b}\}^2$ such that

$$\mu(\underline{\sigma}) \cong \left\{ \prod_{i=1,2} \mathbf{1}\{\mathbf{r}[\underline{\sigma}^i] \geq 1\} \right\} \left\{ \prod_{\ell=1}^{D} h_\ell(\sigma_\ell) \right\}. \tag{9.46}$$

We now turn to the construction and estimation of the h_ℓ.

Step 2. Construction of Lagrangian weights. As usual, we will construct $h_{\ell,t} \to h_\ell$ in the limit $t \to \infty$. We initialize the construction with $h_{\ell,0} \equiv \mu_\ell$. Let μ_t be defined by (9.46) with $h_{\ell,t}$ in place of h_ℓ, and let $\mu_{\ell,t}$ denote the marginal of μ_t on the ℓ-th edge in $\delta v(+)$. Similarly to (9.10) we have

$$
\mu_{\ell,t}(\sigma) \cong h_{\ell,t}(\sigma) \overbrace{\sum_{\underline{\sigma}_{-\ell}} \left\{ \prod_{i=1,2} \mathbf{1}\{\mathbf{r}[\underline{\sigma}^i] \geq 1\} \right\} \prod_{\ell' \neq \ell} h_{\ell',t}(\sigma_{\ell'})}^{\text{denote this } \xi_{\ell,t}(\sigma)} .
\tag{9.47}
$$

For the remainder of the proof, we assume without loss that $\ell = 1$. Then at $t = 0$ we have

$$
0 \leq 1 - \xi_{1,0}(\mathbf{rb}) = \prod_{\ell=2}^{D} \left\{ 1 - \mu_\ell(\mathbf{pr}) \right\} \leq \exp\left\{ -\sum_{\ell=2}^{D} \mu_\ell(\mathbf{pr}) \right\} \overset{(9.45)}{\leq} \frac{e}{\exp(\gamma^2)} \overset{(9.43)}{\leq} \frac{1}{2^{k\epsilon_3}} .
\tag{9.48}
$$

By similar calculations, $0 \leq 1 - \xi_{1,0}(\sigma) \leq 2^{-k\epsilon_3}$ for all $\sigma \in \{\mathbf{r}, \mathbf{b}\}^2$. (In fact $\xi_{1,0}(\mathbf{rr}) = 1$.) Given $h_{\ell,t}$ we define $h_{\ell,t+1}$ to be the probability measure such that

$$
h_{\ell,t+1}(\sigma) \cong \frac{\mu_\ell(\sigma)}{\xi_{\ell,t}(\sigma)} \overset{(9.47)}{\cong} h_{\ell,t}(\sigma) \frac{\mu_\ell(\sigma)}{\mu_{\ell,t}(\sigma)}
\tag{9.49}
$$

for all $\sigma \in \{\mathbf{r}, \mathbf{b}\}^2$.

Step 3. Estimation of Lagrangian weights. Suppose inductively that

$$
\left\| \frac{\mu_{\ell,t}}{\mu_\ell} - 1 \right\|_\infty \leq \frac{1}{2^{k\epsilon_3(t+1)/3}}
\tag{9.50}
$$

for all $1 \leq \ell \leq D$, where the base case $t = 0$ follows from the above bounds (see (9.48)) on $\xi_{\ell,0}$. If we assume (9.50) holds up to t, then taking a telescoping sum gives

$$
\left| \xi_{1,t+1}(\mathbf{rb}) - \xi_{1,t}(\mathbf{rb}) \right| = \left| \prod_{j=2}^{D} \left\{ 1 - h_{\ell,t+1}(\mathbf{pr}) \right\} - \prod_{j=2}^{D} \left\{ 1 - h_{\ell,t}(\mathbf{pr}) \right\} \right|
$$

$$
\leq \sum_{\ell=2}^{D} h_{\ell,t}(\mathbf{pr}) \left| \frac{h_{\ell,t+1}(\mathbf{pr})}{h_{\ell,t}(\mathbf{pr})} - 1 \right| \prod_{j=2}^{\ell-1} \left\{ 1 - h_{j,t+1}(\mathbf{pr}) \right\} \prod_{j=\ell+1}^{D} \left\{ 1 - h_{j,t}(\mathbf{pr}) \right\}
$$

$$
\leq O(1) \left\{ \max_{1 \leq \ell \leq D} \left| \frac{h_{\ell,t+1}(\mathbf{pr})}{h_{\ell,t}(\mathbf{pr})} - 1 \right| \right\} \left\{ \sum_{\ell=1}^{D} h_{\ell,t}(\mathbf{pr}) \right\} \Big/ \exp\left\{ \sum_{\ell=1}^{D} h_{\ell,t}(\mathbf{pr}) \right\}
$$

$$
\overset{(9.50)}{\leq} O\left(\frac{\gamma^2}{\exp(\gamma^2)} \right) \left\{ \max_{1 \leq \ell \leq D} \left| \frac{h_{\ell,t+1}(\mathbf{pr})}{h_{\ell,t}(\mathbf{pr})} - 1 \right| \right\} \overset{(9.49)}{=} O\left(\frac{\gamma^2}{\exp(\gamma^2)} \right) \max_{1 \leq \ell \leq D} \left| \frac{\mu_\ell(\mathbf{pr})}{\mu_{\ell,t}(\mathbf{pr})} - 1 \right| \overset{(9.50)}{\leq} \frac{2^{-k\epsilon_3}}{2^{k\epsilon_3(t+1)/3}} .
\tag{9.51}
$$

A similar estimate holds for $\xi_{1,t+1}(\sigma) - \xi_{1,t}(\sigma)$ for all $\sigma \in \{\mathbf{r}, \mathbf{b}\}^2$. Next note that if we let $\tilde{\mu}_{t+1}$ be defined by (9.46) with weights $h_{1,t+1}$ and $h_{\ell,t}$ for $\ell \geq 2$, then the marginal on the first edge is, similarly to (9.47),

$$
\tilde{\mu}_{1,t+1}(\sigma) \cong h_{1,t+1}(\sigma) \xi_{1,t}(\sigma) \overset{(9.49)}{\cong} \frac{\mu_1(\sigma)}{\xi_{1,t}(\sigma)} \xi_{1,t}(\sigma) = \mu_1(\sigma) .
$$

It follows from this that

$$
\left\| \frac{\mu_{1,t+1}}{\mu_1} - 1 \right\|_\infty = \left\| \frac{\mu_{1,t+1}}{\tilde{\mu}_{1,t+1}} - 1 \right\|_\infty \overset{(9.49)}{=} \left\| \frac{\xi_{1,t+1}}{\xi_{1,t}} - 1 \right\|_\infty \overset{(9.51)}{\leq} O\left(\frac{2^{-k\epsilon_3}}{2^{k\epsilon_3(t+1)/3}} \right) \leq \frac{1}{2^{k\epsilon_3(t+2)/3}} ,
$$

which verifies the induction. We then take $t \to \infty$ to obtain the limiting weights $h_\ell \equiv h_{\ell,\infty}$ that define the optimal measure μ by (9.46). Suppose e corresponds to the first edge in $\delta v(+)$. Then, for all $\sigma_e \in \{\mathbf{r}, \mathbf{b}\}^2$,

$$
\dot{v}_T\left(\sum_{e' \in \delta v(+) \backslash e} \mathbf{1}\{(\sigma_{e'})^2 = \mathbf{r}\} = 0 \,\Big|\, \sigma_e \right) \overset{(9.45)}{=} \mu\left(\underline{\sigma}_{-1} \in \{\mathbf{pb}\}^D \,\Big|\, \sigma_1 \right) \overset{(9.50)}{\leq} O\left(\prod_{\ell=2}^{D} \left\{ 1 - h_\ell(\mathbf{pr}) \right\} \right) \leq \frac{1}{2^{k\epsilon_3}} .
$$

The lemma follows by recalling (9.44). □

Lemma 9.9 (used only in proof of Lemma 9.11). *Let ϵ_2 be a small absolute constant. Fix $x \in \{+, -\}$ and $\tau \in \{y, g\}$, and let v be a variable of type T. Suppose we are given $\omega \in I_0$ with marginal π such that*

$$\sum_{e \in \delta v(x)} \pi_e \left(\sigma^1 = r \, \middle| \, \sigma \in \{r\tau, b\tau\}\right) \geq k\epsilon_2 , \tag{9.52}$$

where $\pi_e \equiv \pi_t$ for an edge e of type t. Let Ω be the maximizer of (9.41) that is consistent with ω, and let $\varpi_{L,j}$ be the resulting joint distribution of (σ, v). Then there exists a positive absolute constant ϵ_3 (depending only on ϵ_2) such that for every edge $e \in \delta v(x)$ we have

$$\varpi_{L,j} \left(v^1 = v \, \middle| \, \sigma = r\tau\right) \geq 1 - \frac{1}{2^{k\epsilon_3}}$$

for $j = j(t)$ and all clause types L such that $L(j) = t$. The same holds if we exchange the two copies $i = 1, 2$.

Proof. The proof is very similar to (but simpler than) that of Lemma 9.8, and we omit the details. □

9.4. **Bounds on doubly forced edges.** In this subsection we bound the incidence of edges e which are forced in both coordinates, meaning that $\sigma_e = rr$ and $v_e \in \{r, v\}^2$.

Lemma 9.10. *If the clause type L is nice and diverse (Definition 9.3), then*

$$\varpi_{L,j}(v = vv) \leq \frac{1}{2^{k(1+\epsilon_4)}}$$

for all $1 \leq j \leq k(L)$, where ϵ_4 is an absolute constant.

Proof. Throughout the proof, ϵ_4 denotes a small positive number, whose value may change from one occurrence to the next, but ultimately is taken as an absolute constant. Fix L, j, and let Ξ denote the empirical measure of configurations $(\Sigma_e, \underline{\sigma}_{(\delta v \cup \delta a) \backslash e})$ over all $e = (av)$ such that $L_a = L$ and $j(v; a) = j$. In the proof below we assume without loss of generality that the edge e has label $L_e = +$.

Step 1. Reduction to entropy maximization for a conditional measure. Let v denote the subset of all valid configurations $(\Sigma_e, \underline{\sigma}_{(\delta v \cup \delta a) \backslash e})$ that have $\Sigma_e \in \{w, v\}^2$. Define the conditional measure

$$\mu \left(\Sigma_e, \underline{\sigma}_{(\delta v \cup \delta a) \backslash e}\right) \equiv \Xi \left(\Sigma_e, \underline{\sigma}_{(\delta v \cup \delta a) \backslash e} \, \middle| \, v\right).$$

Note the marginals of μ must be close to those of the original measure Ξ, since Lemma 9.5 gives

$$\Xi(v) \geq \varpi_{L,j}(\varsigma = ww) \geq 1 - O\left(\frac{k^2}{2^k}\right). \tag{9.53}$$

The measure μ maximizes entropy subject to the marginal distributions of $\sigma_{e'}$ for $e' \neq e$, as well as of $(\Sigma_e)^i$ for $i = 1, 2$. By the method of Lagrange multipliers, we can express

$$\mu \left(\Sigma_e, \underline{\sigma}_{(\delta v \cup \delta a) \backslash e}\right) \cong \left\{ \prod_{i=1,2} (b_e)^i ((\Sigma_e)^i) \right\} \left\{ \prod_{e' \in (\delta v \cup \delta a) \backslash e} b_{e'}(\sigma_{e'}) \right\} \mathbf{1}_V , \tag{9.54}$$

where we fix $(b_e)^i(w) \equiv 1$ for both $i = 1, 2$. For $x \in \{-, +, f\}^2$, let us write $(\Sigma_e, \underline{\sigma}_{\delta v \backslash e}) \sim x$ if $(\Sigma_e, \underline{\sigma}_{\delta v \backslash e})$ is consistent with frozen spin $x_v = x$. Note that for $\Sigma_e \in \{w, v\}^2$, we have

$$\left\{ \underline{\sigma}_{\delta v \backslash e} : (\Sigma_e, \underline{\sigma}_{\delta v \backslash e}) \sim x \right\} = S(x) \equiv \prod_{i=1,2} S^i(x^i) ,$$

where $S^i(x^i)$ **does not depend on** Σ_e. For instance, for all $\Sigma_e \in \{w, v\}^2$,

$$\left\{ \underline{\sigma}_{\delta v \backslash e} : (\Sigma_e, \underline{\sigma}_{\delta v \backslash e}) \sim ++ \right\} = \prod_{i=1,2} \overbrace{\left\{ (\underline{\sigma}_{\delta v \backslash e})^i : (\underline{\sigma}_{\delta v(-)})^i \equiv y, (\underline{\sigma}_{\delta v(+) \backslash e})^i \equiv p, \sum_{e' \in \delta v(+) \backslash e} \mathbf{1}\{(\sigma_{e'})^i = r\} \geq 1 \right\}}^{S^i(+)}.$$

Consequently, the marginal of (9.54) on $(x_v, \Sigma_e, \underline{\sigma}_{\delta a \setminus e})$ can be written as

$$\mu(x_v, \Sigma_e, \underline{\sigma}_{\delta a \setminus e}) = \frac{\mathbf{1}_{\text{V}}}{Z} \left\{ \prod_{i=1,2} (b_e)^i ((\Sigma_e)^i) \right\} \overbrace{\left\{ \sum_{\underline{\sigma}_{\delta v \setminus e} \in S(x_v)} \prod_{e' \in \delta v \setminus e} b_{e'}(\sigma_{e'}) \right\}}^{\text{denote this } b_v(x_v)} \left\{ \prod_{e' \in \delta a \setminus e} b_{e'}(\sigma_{e'}) \right\}, \tag{9.55}$$

where Z is the normalizing constant. We next turn to the construction and estimation of the weights in (9.55).

Step 2. Initialization of Lagrangian weights. We will construct weights $b_t \to b$ in the limit $t \to \infty$. At $t = 0$, on the variable v and on the edges $\delta a \setminus e$, we put

$$\begin{aligned}
b_{v,0}(x) &\equiv \mu_v(x) \equiv \mu(x_v = x) &&\text{for } x \in \{\texttt{-}, \texttt{+}, \texttt{f}\}^2, \\
b_{e',0}(\sigma) &\equiv \mu_{e'}(\sigma) \equiv \mu(\sigma_{e'} = \sigma) &&\text{for } \sigma \in \{\texttt{y}, \texttt{c}\}^2 \text{ and } e' \in \delta a \setminus e.
\end{aligned} \tag{9.56}$$

Abbreviate $K \equiv k(L)$. On the central edge $e = (av)$, for $i = 1, 2$ we put

$$(b_{e,0})^i ((\Sigma_e)^i) \equiv \frac{(\mu_e)^i ((\Sigma_e)^i)}{1/(2^K)^{\mathbf{1}\{(\Sigma_e)^i = \texttt{v}\}}}.$$

Let μ_t be defined as μ in (9.55), but with b_t in place of b, and normalizing constant Z_t. We begin by estimating the marginals at time $t = 0$. Recalling the notation $\texttt{u} \equiv \{\texttt{y}, \texttt{c}\}$, we have

$$Z_0 \mu_{e,0}(\texttt{ww}) = (\mu_e)^1(\texttt{w})(\mu_e)^2(\texttt{w}) \left\{ \prod_{e' \in \delta a \setminus e} \mu_{e'}(\texttt{uu}) \right\} \left\{ 1 - O\left(\frac{k^2}{2^k}\right) \right\} = (\mu_e)^1(\texttt{w})(\mu_e)^2(\texttt{w}) \left\{ 1 - O\left(\frac{k^2}{2^k}\right) \right\},$$

where the estimate uses (9.53) and the assumption that L is nice, which implies (together with the judicious condition) $\Xi_{e'}(\texttt{uu}) \geq 1 - O(2^{-k})$ for all $e' \in \delta a$. (In the above calculation, the $O(k^2/2^k)$ error comes from the fact that among the configurations $\underline{\sigma}_{\delta a \setminus e}$ with $\sigma_{e'} \in \texttt{uu}$ for all e', some will not be compatible with $\sigma_e = \texttt{ww}$ because they will have too few $\texttt{c}$ spins.) Next recall that (as noted in the proof of Lemma 9.7) if L is both nice and diverse, then Lemma 9.6 implies (9.29), which says that $\omega_{L,j}(\texttt{yb}) \geq 1/65$ for at least $k/10$ indices $1 \leq j \leq k(L)$. Recall also that (9.53) implies that the marginals of μ are close to those of the original measure Ξ, so (9.29) implies

$$\min\left\{ \sum_{j=1}^{k(L)} \mathbf{1}\left\{ \mu_j(\texttt{yb}) \geq \frac{1}{66} \right\}, \sum_{j=1}^{k(L)} \mathbf{1}\left\{ \mu_j(\texttt{by}) \geq \frac{1}{66} \right\} \right\} \geq \frac{k}{10}. \tag{9.57}$$

It follows using (9.57) that

$$\begin{aligned}
Z_0 \mu_{e,0}(\texttt{vw}) &= (\mu_e)^1(\texttt{v})(\mu_e)^2(\texttt{w}) \frac{\mu_v(\texttt{+})}{1/2} \left\{ \prod_{e' \in \delta a \setminus e} \frac{\mu_{e'}(\texttt{yu})}{1/2} \right\} \left\{ 1 - O\left(\frac{1}{2^{k\epsilon_4}}\right) \right\} \\
&= (\mu_e)^1(\texttt{v})(\mu_e)^2(\texttt{w}) \left\{ 1 - O\left(\frac{1}{2^{k\epsilon_4}}\right) \right\},
\end{aligned}$$

where the factors of $1/2$ come from the 2^K term in the definition of $(b_{e,0})^1(\texttt{v})$. Lastly,

$$Z_0 \mu_{e,0}(\texttt{vv}) = \overbrace{\frac{(\mu_e)^1(\texttt{v})}{1/2^K} \frac{(\mu_e)^2(\texttt{v})}{1/2^K}}^{O(1)} \mu_v(\texttt{++}) \overbrace{\left\{ \prod_{e' \in \delta a \setminus e} \mu_{e'}(\texttt{yy}) \right\}}^{\text{at most } 2^{-k(1+\epsilon_4)} \text{ by (9.57)}} \leq O\left(\frac{\min\{(\mu_e)^1(\texttt{v}), (\mu_e)^2(\texttt{v})\}}{2^{k\epsilon_4}}\right). \tag{9.58}$$

The above estimates imply that the single-copy marginals of $\mu_{e,0}$ are close to those of μ_e: more precisely, we have

$$\left| \frac{(\mu_{e,0})^i(\Sigma^i)}{(\mu_e)^i(\Sigma^i)} - 1 \right| \leq \begin{cases} k^{O(1)}/2^k & \text{if } \Sigma^i = \texttt{w}, \\ O(2^{-k\epsilon_4}) & \text{if } \Sigma^i = \texttt{v}, \end{cases} \tag{9.59}$$

for both $i = 1, 2$. We then turn to estimating the marginals at $t = 0$ on the edges of $\delta a \setminus e$. For $e' \in \delta a \setminus e$ and $\sigma \in \{\mathtt{y}, \mathtt{c}\}^2$, we can represent the marginal probability of σ on e' as

$$\mu_{e',0}(\sigma) = \sum_{\Sigma \in \{\mathtt{w}, \mathtt{r}\}^2} \mu_{e,0}(\Sigma)\mu_0\Big(\sigma_{e'} = \sigma \,\Big|\, \Sigma_e = \Sigma\Big). \tag{9.60}$$

The term $\mu_{e,0}(\Sigma)$ was just estimated above, while the conditional probability $\mu_0(\sigma_{e'} = \sigma \,|\, \Sigma_e = \Sigma)$ can be obtained from the representation (9.55). For instance, for each $e' \in \delta a \setminus e$, we have

$$\mu_0\Big(\sigma_{e'} = \mathtt{yy} \,\Big|\, \Sigma_e = \mathtt{vw}\Big) \stackrel{(9.55)}{=} \frac{\mu_{e'}(\mathtt{yy}) \displaystyle\prod_{e'' \in \delta a \setminus \{e, e'\}} \mu_{e''}(\mathtt{yu})}{\mu_{e'}(\mathtt{yu}) \displaystyle\prod_{e'' \in \delta a \setminus \{e, e'\}} \mu_{e''}(\mathtt{yu})} \left\{1 - O\left(\frac{1}{2^{k\epsilon_4}}\right)\right\} = \frac{\mu_{e'}(\mathtt{yy})}{1/2}\left\{1 - O\left(\frac{1}{2^{k\epsilon_4}}\right)\right\},$$

where the last estimate uses that e' is a nice edge, so $\mu_{e''}(\mathtt{yu})$ is close to $1/2$. Similarly,

$$\mu_0\Big(\sigma_{e'} = \mathtt{yy} \,\Big|\, \Sigma_e = \mathtt{ww}\Big) \stackrel{(9.55)}{=} \frac{\mu_{e'}(\mathtt{yy})}{\mu_{e'}(\mathtt{uu})}\left\{1 - O\left(\frac{k^2}{2^k}\right)\right\} = \mu_{e'}(\mathtt{yy})\left\{1 - O\left(\frac{k^2}{2^k}\right)\right\}.$$

Lastly we have $\mu_0(\sigma_{e'} = \mathtt{yy} \,|\, \Sigma_e = \mathtt{vv}) = 1$. Substituting the last few estimates into (9.60) gives

$$\mu_{e',0}(\mathtt{yy}) \stackrel{(9.60)}{=} \mu_{e,0}(\mathtt{vv}) + \mu_{e,0}\Big(\{\mathtt{vw}, \mathtt{wv}\}\Big)\frac{\mu_{e'}(\mathtt{yy})}{1/2}\left\{1 - O\left(\frac{1}{2^{k\epsilon_4}}\right)\right\} + \mu_{e,0}(\mathtt{ww})\mu_{e'}(\mathtt{yy})\left\{1 - O\left(\frac{k^2}{2^k}\right)\right\}$$

$$\stackrel{(9.58)}{=} \mu_{e'}(\mathtt{yy})\left\{O\left(\frac{1}{2^{k(1+\epsilon_4)}}\right) + O\left(\frac{1}{2^k}\right) + 1 - O\left(\frac{k^2}{2^k}\right)\right\} = \mu_{e'}(\mathtt{yy})\left\{1 - O\left(\frac{k^2}{2^k}\right)\right\}.$$

By similar (but simpler) calculations, the same estimate holds for the other elements $\sigma \in \{\mathtt{y}, \mathtt{c}\}^2$. Lastly, we estimate the marginal law at $t = 0$ for the frozen spin x_v, using a similar conditioning as in (9.60). To this end, note that

$$\mu_0\Big(x_v = x \,\Big|\, \Sigma_e = \mathtt{ww}\Big) \stackrel{(9.55)}{=} \frac{b_{v,0}(x)\displaystyle\prod_{i=1,2}(b_{e,0})^i(\mathtt{w})}{\displaystyle\sum_{x' \in \{-,+,\mathtt{f}\}^2} b_{v,0}(x')\prod_{i=1,2}(b_{e,0})^i(\mathtt{w})} \stackrel{(9.56)}{=} \mu_v(x)$$

for any $x \in \{-, +, \mathtt{f}\}^2$. Likewise we have

$$\mu_0\Big(x_v = x \,\Big|\, \Sigma_e = \mathtt{vw}\Big) \stackrel{(9.55)}{=} \frac{\mathbf{1}\{x^1 = +\}b_{v,0}(x)(b_{e,0})^1(\mathtt{v})(b_{e,0})^2(\mathtt{w})}{\displaystyle\sum_{x' \in \{+\} \times \{-,+,\mathtt{f}\}} b_{v,0}(x')(b_{e,0})^1(\mathtt{v})(b_{e,0})^2(\mathtt{w})} \stackrel{(9.56)}{=} \frac{\mu_v(x)}{(\mu_v)^1(+)}\mathbf{1}\{x^1 = +\},$$

and a similar expression holds for the case $\Sigma_e = \mathtt{wr}$. Lastly, note that $\mu_0(x_v = x \,|\, \Sigma_e = \mathtt{vv}) = \mathbf{1}\{x = ++\}$. Substituting these into (9.60) gives

$$\mu_{v,0}(x) \stackrel{(9.60)}{=} \mu_{e,0}(\mathtt{vv})\mathbf{1}\{x = ++\} + \mu_{e,0}(\mathtt{vw})\frac{\mu_v(x)}{(\mu_v)^1(+)}\mathbf{1}\{x^1 = +\} + \mu_{e,0}(\mathtt{wv})\frac{\mu_v(x)}{(\mu_v)^2(+)}\mathbf{1}\{x^2 = +\} + \mu_{e,0}(\mathtt{ww})\mu_v(x)$$

$$\stackrel{(9.58)}{=} \mu_v(x)\left\{O\left(\frac{\mathbf{1}\{x = ++\}}{2^{k(1+\epsilon_4)}}\right) + O\left(\frac{\mathbf{1}\{x^1 = +\} + \mathbf{1}\{x^2 = +\}}{2^k}\right) + 1 - O\left(\frac{k^2}{2^k}\right)\right\} = \mu_v(x)\left\{1 - O\left(\frac{k^2}{2^k}\right)\right\}.$$

To summarize the last few estimates, we have shown that at $t = 0$ we have

$$\left(\begin{array}{c}\left\|\dfrac{\mu_{v,0}}{\mu_v} - 1\right\|_\infty + \displaystyle\sum_{i=1,2}\left|\dfrac{(\mu_{e,0})^i(\mathtt{w})}{(\mu_e)^i(\mathtt{w})} - 1\right| + \displaystyle\sum_{e' \in \delta a \setminus v}\left\|\dfrac{\mu_{e',0}}{\mu_{e'}} - 1\right\|_\infty \\ \displaystyle\sum_{i=1,2}\left|\dfrac{(\mu_{e,0})^i(\mathtt{v})}{(\mu_e)^i(\mathtt{v})} - 1\right|\end{array}\right) \leq k^{O(1)}\left(\begin{array}{c}2^{-k} \\ 2^{-k\epsilon_4}\end{array}\right). \tag{9.61}$$

We now turn to the construction of the sequence $b_t \to b_\infty \equiv b$, where we recall that b denotes the limiting weights in (9.55) that give the solution to the constrained entropy maximization problem.

Step 3. Iterative construction and covariance estimates. For each $t \geq 0$, suppose we have the weights b_t, which define the measure μ_t. We then define the weights at time $t + 1$ by setting

$$\frac{b_{v,t+1}(x)}{b_{v,t}(x)} = \frac{\mu_v(x)}{\mu_{v,t}(x)}, \quad \frac{b_{e',t+1}(\sigma)}{b_{e',t}(\sigma)} = \frac{\mu_{e'}(\sigma)}{\mu_{e',t}(\sigma)}, \quad \frac{(b_{e,t+1})^i(\Sigma^i)}{(b_{e,t})^i(\Sigma^i)} = \frac{\mu_e(\Sigma^i)}{\mu_{e,t}(\Sigma^i)} \tag{9.62}$$

for all $x \in \{\text{-},\text{+},\text{f}\}^2$, $e' \in \delta a \setminus e$, $\sigma \in \{\text{y},\text{c}\}^2$, $i = 1, 2$, and $\Sigma^i \in \{\text{w},\text{v}\}$. We will show by induction that for all $t \geq 0$ we have the bound

$$\left(\begin{array}{c} \left\| \dfrac{\mu_{v,t}}{\mu_v} - 1 \right\|_\infty + \displaystyle\sum_{i=1,2} \left| \dfrac{(\mu_{e,t})^i(\text{w})}{(\mu_e)^i(\text{w})} - 1 \right| + \displaystyle\sum_{e' \in \delta a \setminus e} \left\| \dfrac{\mu_{e',t}}{\mu_{e'}} - 1 \right\|_\infty \\ \displaystyle\sum_{i=1,2} \left| \dfrac{(\mu_{e,t})^i(\text{v})}{(\mu_e)^i(\text{v})} - 1 \right| \end{array} \right) \leq \frac{k^{O(1)}}{(2^{k\epsilon_4})^{t/2}} \binom{2^{-k}}{2^{-k\epsilon_4}}, \tag{9.63}$$

where the base case $t = 0$ follows from (9.61). Note that (9.62) and (9.63) together imply

$$\left(\begin{array}{c} \left\{ \left\| \dfrac{b_{v,t+1}}{b_{v,t}} - 1 \right\|_\infty + \displaystyle\sum_{i=1,2} \left| \dfrac{(b_{e',t+1})^i(\text{w})}{(b_{e',t})^i(\text{w})} - 1 \right| + \displaystyle\sum_{e' \in \delta a \setminus e} \left\| \dfrac{b_{e',t+1}}{b_{e',t}} - 1 \right\|_\infty \right\} \\ \displaystyle\sum_{i=1,2} \left| \dfrac{(b_{e',t+1})^i(\text{v})}{(b_{e',t})^i(\text{v})} - 1 \right| \end{array} \right) \leq \frac{k^{O(1)}}{(2^{k\epsilon_4})^{t/2}} \binom{2^{-k}}{2^{-k\epsilon_4}}, \tag{9.64}$$

Towards the proof of (9.63), we estimate covariances under the measure μ_t. First, for all $x \in \{\text{-},\text{+},\text{f}\}^2$ we have

$$\frac{Z_t \mu_t(x_v = x, (\Sigma_e)^1 = \text{w})}{\mu_v(x)(\mu_e)^1(\text{w})} = \frac{b_{v,t}(x)(b_{e,t})^1(\text{w})}{\mu_v(x)(\mu_e)^1(\text{w})} \left\{ (b_{e,t})^2(\text{w}) \prod_{e' \in \delta a \setminus e} b_{e',t}(\text{uu}) \left(1 - \frac{k^{O(1)}}{2^k} \right) \right.$$

$$\left. + \mathbf{1}\{x^2 = \text{+}\}(b_{e,t})^2(\text{v}) \prod_{e'} b_{e',t}(\text{uy}) \left(1 - \frac{O(1)}{2^{k\epsilon_4}} \right) \right\}$$

$$= \frac{b_{v,t}(x)(b_{e,t})^1(\text{w})}{\mu_v(x)(\mu_e)^1(\text{w})} \left\{ 1 - \frac{k^{O(1)}}{2^k} \right\} \overset{(9.64)}{=} 1 - \frac{k^{O(1)}}{2^k}.$$

In combination with (9.63), this implies that for all $x \in \{\text{-},\text{+},\text{f}\}^2$ and $i = 1, 2$ we have

$$\left| \frac{\text{Cov}_{\mu_t}(\mathbf{1}\{x_v = x\}, \mathbf{1}\{(\Sigma_e)^i = \text{w}\})}{\mu_v(x)(\mu_e)^i(\text{w})} \right| = \left| \frac{\mu_t(x_v = x, (\Sigma_e)^1 = \text{w})}{\mu_v(x)(\mu_e)^1(\text{w})} - \frac{\mu_{v,t}(x)(\mu_{e,t})^1(\text{w})}{\mu_v(x)(\mu_e)^1(\text{w})} \right| \leq \frac{k^{O(1)}}{2^k}.$$

Next, for $e' \in \delta a \setminus e$ and $\sigma \in \{\text{y},\text{c}\}^2$, we calculate

$$\frac{Z_t \mu_t((\Sigma_e)^1 = \text{w}, \sigma_{e'} = \sigma)}{(\mu_e)^1(\text{w})\mu_{e'}(\sigma)} = \frac{(b_{e,t})^1(\text{w})b_{e',t}(\sigma)}{(\mu_e)^1(\text{w})\mu_{e'}(\sigma)} \left\{ (b_{e,t})^2(\text{w}) \prod_{e'' \in \delta a \setminus \{e,e'\}} b_{e'',t}(\text{uu}) \left(1 - \frac{k^{O(1)}}{2^k} \right) \right.$$

$$\left. + \mathbf{1}\{\sigma^2 = \text{y}\}(b_{e,t})^2(\text{v})(b_{v,t})^2(\text{+}) \prod_{e'' \in \delta a \setminus \{e,e'\}} b_{e'',t}(\text{uy}) \left(1 - \frac{O(1)}{2^{k\epsilon_4}} \right) \right\}$$

$$= \frac{(b_{e,t})^1(\text{w})b_{e',t}(\sigma)}{(\mu_e)^1(\text{w})\mu_{e'}(\sigma)} \left\{ 1 - \frac{k^{O(1)}}{2^k} \right\} \overset{(9.64)}{=} 1 - \frac{k^{O(1)}}{2^k}.$$

This implies that for all $e' \in \delta a \setminus e$ and $\sigma \in \{\text{y},\text{c}\}^2$, we have

$$\frac{\text{Cov}_{\mu_t}(\mathbf{1}\{(\Sigma_e)^i = \text{w}\}, \mathbf{1}\{\sigma_{e'} = \sigma\})}{(\mu_e)^i(\text{w})\mu_{e'}(\sigma)} \leq \frac{k^{O(1)}}{2^k}.$$

Similar calculations for the case $(\Sigma_e)^i = \mathtt{v}$ give

$$\frac{\mathrm{Cov}_{\mu_t}(\mathbf{1}\{x_v = x\}, \mathbf{1}\{(\Sigma_e)^i = \mathtt{v}\})}{\mu_v(x)(\mu_e)^i(\mathtt{v})} \leq O(1),$$

$$\frac{\mathrm{Cov}_{\mu_t}(\mathbf{1}\{(\Sigma_e)^i = \mathtt{v}\}, \mathbf{1}\{\sigma_{e'} = \sigma\})}{(\mu_e)^i(\mathtt{v})\mu_{e'}(\sigma)} \leq O(1),$$

for all $x \in \{\text{-}, \text{+}, \mathtt{f}\}^2$, $e' \in \delta a \setminus e$, and $\sigma \in \{\mathtt{y}, \mathtt{c}\}^2$. Next, on the edge $e = (av)$, we have

$$\frac{Z_t \mu_{e,t}(\mathtt{ww})}{(\mu_e)^1(\mathtt{w})(\mu_e)^2(\mathtt{w})} = \frac{(b_{e,t})^1(\mathtt{w})(b_{e,t})^2(\mathtt{w})}{(\mu_e)^1(\mathtt{w})(\mu_e)^2(\mathtt{w})} \prod_{e' \in \delta a \setminus e} b_{e',t}(\mathtt{uu}) \left\{ 1 - \frac{k^{O(1)}}{2^k} \right\} \overset{(9.64)}{=} 1 - \frac{k^{O(1)}}{2^k}.$$

Next, using the diversity bound (9.57), we also have

$$\frac{Z_t \mu_{e,t}(\mathtt{vw})}{(\mu_e)^1(\mathtt{v})(\mu_e)^2(\mathtt{w})} \overset{(9.57)}{=} \frac{(b_{e,t})^1(\mathtt{v})(b_{e,t})^2(\mathtt{w})}{(\mu_e)^1(\mathtt{v})(\mu_e)^2(\mathtt{w})}(b_{v,t})^1(\text{+}) \prod_{e' \in \delta a \setminus e} b_{e',t}(\mathtt{yu}) \left\{ 1 - \frac{O(1)}{2^{k\epsilon_4}} \right\}$$

$$\overset{(9.64)}{=} \left\{ 1 - \frac{O(1)}{2^{k\epsilon_4}} \right\} 2^K (b_{v,t})^1(\text{+}) \prod_{e' \in \delta a \setminus e} b_{e',t}(\mathtt{yu}) \overset{(9.64)}{=} 1 - \frac{O(1)}{2^{k\epsilon_4}},$$

$$\frac{Z_t \mu_{e,t}(\mathtt{vv})}{(\mu_e)^1(\mathtt{v})(\mu_e)^2(\mathtt{v})} = \frac{(b_{e,t})^1(\mathtt{v})(b_{e,t})^2(\mathtt{v})}{(\mu_e)^1(\mathtt{v})(\mu_e)^2(\mathtt{v})}\mu_v(\text{++}) \prod_{e' \in \delta v \setminus e} b_{e',t}(\mathtt{yy})$$

$$\overset{(9.64)}{=} 2^K \cdot 2^K \cdot \mu_v(\text{++}) \prod_{e' \in \delta v \setminus e} b_{e',t}(\mathtt{yy}) \overset{(9.57)}{\leq} O\left(\frac{2^k}{2^{k\epsilon_4}} \right).$$

The last few estimates combined imply

$$\frac{\mathrm{Cov}_{\mu_t}(\mathbf{1}\{(\Sigma_e)^1 = \Sigma^1\}, \mathbf{1}\{(\Sigma_e)^2 = \Sigma^2\})}{(\mu_e)^1(\Sigma^1)(\mu_e)^2(\Sigma^2)} \leq k^{O(1)} \begin{cases} 2^{-k} & \text{if } \Sigma = \mathtt{ww}, \\ 2^{-k\epsilon_4} & \text{if } \Sigma \in \{\mathtt{vw}, \mathtt{wv}\}, \\ 2^{k(1-\epsilon_4)} & \text{if } \Sigma = \mathtt{vv}. \end{cases}$$

Lastly, it is straightforward to verify (details omitted) that

$$\frac{\mathrm{Cov}_{\mu_t}(\mathbf{1}\{x_v = x\}, \mathbf{1}\{\sigma_{e'} = \sigma\})}{\mu_v(x)\mu_{e'}(\sigma)} \leq \frac{k^{O(1)}}{2^k}$$

for all $x \in \{\text{-}, \text{+}, \mathtt{f}\}^2$, $e' \in \delta a \setminus e$, and $\sigma \in \{\mathtt{y}, \mathtt{c}\}^2$.

Step 4. Estimates on Lagrangian weights. Now recall from (9.39) that derivatives of μ_t-marginals with respect to b_t can be expressed as covariance. Let $\tilde{\mu}_{t+1}$ be defined by

$$\tilde{\mu}_{t+1}\left(x_v, \Sigma_e, \underline{\sigma}_{\delta a \setminus e}\right) = \frac{1}{\tilde{Z}_{t+1}} b_{v,t}(x)(b_{e,t+1})^1((\Sigma_e)^1)(b_{e,t})^2((\Sigma_e)^2) \prod_{e' \in \delta a \setminus e} b_{e',t}(\sigma_{e'}),$$

where $\tilde{Z}_{t+1}$ is the normalizing constant. It follows from the update rule (9.62) that the marginal of $\tilde{\mu}_{t+1}$ on $(\Sigma_e)^1$ is exactly the desired marginal $(\mu_e)^1$. Consequently,

$$\delta_{t+1}(\Sigma^1) \equiv \left| \frac{(\mu_{e,t+1})^1(\Sigma^1)}{(\mu_e)^1(\Sigma^1)} - 1 \right| = \left| \frac{(\mu_{e,t+1})^1(\Sigma^1)}{(\tilde{\mu}_{e,t+1})^1(\Sigma^1)} - 1 \right|$$

$$\leq \frac{1}{(\mu_e)^1(\Sigma^1)} O\left(\sum_x \frac{\partial \mu_{e,t+1}(\Sigma^1)}{\partial \ln b_{v,t+1}(x)} \left| \frac{b_{v,t+1}(x)}{b_{v,t}(x)} - 1 \right| + \sum_{\Sigma^2} \frac{\partial \mu_{e,t+1}(\Sigma^1)}{\partial \ln(b_{e,t+1})^2(\Sigma^2)} \left| \frac{(b_{e,t+1})^2(\Sigma^2)}{(b_{e,t})^2(\Sigma^2)} - 1 \right| \right.$$

$$\left. + \sum_{e',\sigma} \frac{\partial \mu_{e,t+1}(\Sigma^1)}{\partial \ln b_{e',t+1}(\sigma)} \left| \frac{b_{e',t+1}(\sigma)}{b_{e',t}(\sigma)} - 1 \right| \right).$$

Substituting the preceding covariance estimates into the last bound and applying (9.64) gives

$$\delta_{t+1}(\mathtt{w}) \le k^{O(1)} \left\{ \frac{1}{2^k} \cdot \frac{1}{2^k(2^{k\epsilon_4/2})^t} + \frac{1}{2^{k\epsilon_4}} \cdot (\mu_e)^2(\mathtt{v}) \cdot \left| \frac{(b_{e,t+1})^2(\mathtt{v})}{(b_{e,t})^2(\mathtt{v})} - 1 \right| \right\}$$

$$\le k^{O(1)} \left\{ \frac{1}{2^k} \cdot \frac{1}{2^k(2^{k\epsilon_4/2})^t} + \frac{1}{2^{k\epsilon_4}} \cdot \frac{1}{2^k} \cdot \frac{1}{2^{k\epsilon_4}((2^{k\epsilon_4/2})^t} \right\} \le \frac{k^{O(1)}}{2^k(2^{k\epsilon_4/2})^{t+2}},$$

$$\delta_{t+1}(\mathtt{v}) \le k^{O(1)} \left\{ \frac{1}{2^{k\epsilon_4}((2^{k\epsilon_4/2})^t} + 2^{k(1-\epsilon_4)} \cdot (\mu_e)^2(\mathtt{v}) \cdot \frac{1}{2^{k\epsilon_4}((2^{k\epsilon_4/2})^t} \right\} \le \frac{k^{O(1)}}{2^{k\epsilon_4}(2^{k\epsilon_4/2})^{t+2}}.$$

This verifies the inductive hypothesis (9.63) for the quantities

$$\delta_{t+1}(\Sigma^i) = \left| \frac{(\mu_{e,t+1})^i(\Sigma^i)}{(\mu_e)^i(\Sigma^i)} - 1 \right|.$$

The remaining estimates in (9.63) follow by similar calculations (details omitted). It follows that as $t \to \infty$ the weights b_t converge to the desired limiting weights $b_\infty \equiv b$ that define the optimal measure μ in (9.55). It then follows from the bound (9.64) that

$$\mu_e(\mathtt{vv}) \overset{(9.55)}{=} O\left(b_v(\mathtt{++}) \left\{ \prod_{i=1,2} (b_e)^i(\mathtt{v}) \right\} \left\{ \prod_{e' \in \delta a \backslash e} b_{e'}(\mathtt{yy}) \right\} \right)$$

$$\overset{(9.64)}{=} O\left(\mu_v(\mathtt{++}) \prod_{i=1,2} \left((\mu_e)^i(\mathtt{v}) 2^K \right) \prod_{e' \in \delta a \backslash e} \mu_{e'}(\mathtt{yy}) \right) \overset{(9.57)}{\le} \frac{1}{2^{k(1+\epsilon_4)}},$$

as claimed.   □

Lemma 9.11. *Suppose v is a non-defective variable of type T, satisfying the bounds $\pi_v(\mathtt{++}) = \pi_T(\mathtt{++}) \ge 2^{-k/16}$ and*

$$\mathcal{D}(v) \overset{(9.4)}{=} \sum_{t \in T} \sum_L \mathbf{1}\{L \notin \mathbb{D}\} \pi_{\mathscr{D}}(L \,|\, t) \le 2^k. \tag{9.65}$$

Then, for every edge $e \in \delta v(\mathtt{+})$, we have (with ϵ_4 as in Lemma 9.10) the bound

$$\max_{L,j} \left\{ \omega_{L,j}(\mathtt{rr}) : L \in \mathbb{D} \text{ with } L(j) = t_e \right\} \le \frac{1}{2^{k(1+\epsilon_4)}}.$$

The same statement holds if we replace + with – throughout.

Proof. We claim that

$$\min \left\{ \sum_{e \in \delta v(\mathtt{+})} \pi_e\left(\{\mathtt{rr}, \mathtt{rb}\} \right), \sum_{e \in \delta v(\mathtt{+})} \pi_e\left(\{\mathtt{rr}, \mathtt{br}\} \right) \right\} \ge \frac{k\pi_v(\mathtt{++})}{300}. \tag{9.66}$$

Let us first note that (9.66) implies the result of the lemma: indeed, if (9.66) holds, then condition (9.43) of Lemma 9.8 is satisfied with $\epsilon_2 = 1/300$, and applying that lemma gives

$$\omega_{L,j}(\mathtt{rr}) \le \frac{\varpi_{L,j}(v = \mathtt{vv})}{1 - 2^{-k\epsilon_3}} \le O\left(\varpi_{L,j}(v = \mathtt{vv}) \right)$$

for $j = j(t)$ and all L such that $L(j) = t$. If in addition $L \in \mathbb{D}$, then combining with Lemma 9.10 gives

$$\omega_{L,j}(\mathtt{rr}) \le O\left(\varpi_{L,j}(v = \mathtt{vv}) \right) \le O\left(\frac{1}{2^{k(1+\epsilon_4)}} \right).$$

Thus it suffices to prove (9.66).

Step 1. Preliminary bounds. For $e \in \delta v(\mathtt{+})$ we have $\pi_e(\mathtt{bb}) = \pi_v(\mathtt{++}) - O(2^{-k}) \ge 2^{-k/15}$, so Lemma 9.6 gives

$$\omega_{L,j}(\mathtt{bb}) \ge \frac{\pi_e(\mathtt{bb})}{8} \ge \frac{1}{8} \left\{ \pi_v(\mathtt{++}) - O\left(\frac{1}{2^k} \right) \right\}$$

for all L, j with $L(j) = t_e$. Combining with Lemma 9.5 gives

$$\varpi_{L,j}\begin{pmatrix} \sigma = \mathtt{bb}, \\ \Sigma = \mathtt{ww} \end{pmatrix} \geq \omega_{L,j}(\mathtt{bb}) - O\left(\frac{k^2}{2^k}\right) \geq \omega_{L,j}(\mathtt{bb})\{1 - o_k(1)\}. \tag{9.67}$$

Suppose for contradiction that (9.66) fails. Since v is non-defective, it must also be nice (Definition 3.8), so

$$\sum_{e \in \delta v(+)} \pi_e\Big(\{\mathtt{ry}, \mathtt{rg}\}\Big) \geq \sum_{e \in \delta v(+)} (\pi_e)^1(\mathtt{r}) - \frac{k\pi_v(\mathtt{++})}{300} \geq \frac{k2^{k-1}\ln 2}{2^k}\{1 - o_k(1)\} - \frac{k\pi_v(\mathtt{++})}{300} \geq \frac{k}{3}.$$

Therefore, it holds for some $\tau \in \{\mathtt{y}, \mathtt{g}\}$ that

$$\sum_{e \in \delta v(+)} \pi_e(\mathtt{r}\tau) \geq \frac{k}{6} \tag{9.68}$$

— suppose this is the case for $\tau = \mathtt{y}$. Then Lemma 9.9 (whose condition (9.52) is satisfied, due to (9.68)) gives

$$\varpi_{L,j}\begin{pmatrix} \sigma = \mathtt{ry}, \\ v^1 = \mathtt{v} \end{pmatrix} \geq \omega_{L,j}(\mathtt{ry})\left\{1 - \frac{1}{2^{k\epsilon_3}}\right\}$$

for all $e \in \delta v(+)$ and all $L(j) = t_e$. If L is also diverse, then combining with Lemma 9.7 gives

$$\varpi_{L,j}\begin{pmatrix} \sigma = \mathtt{ry}, \\ \Sigma = \mathtt{vw} \end{pmatrix} \geq \varpi_{L,j}\begin{pmatrix} \sigma = \mathtt{ry}, \\ v^1 = \mathtt{v} \end{pmatrix} - \varpi_{L,j}\begin{pmatrix} \sigma = \mathtt{ry}, \\ \varsigma^1 \neq \mathtt{w}, \varsigma^2 \neq \mathtt{w} \end{pmatrix} \geq \varpi_{L,j}(\mathtt{ry})\left\{1 - \frac{1}{2^{k\epsilon_3}}\right\} - \frac{1}{2^{k(1+\epsilon_1)}}. \tag{9.69}$$

Note that (9.69) used the assumption that v is non-defective, which implies that L is nice (i.e., neighbors only nice variables) whenever $L(j) = t_e$ for $e \in \delta v$. We now define a probability measure π on the space of pairs (e, L), where $e \in \delta v(+)$ and $t_e \in L$, such that

$$\pi(e, L) \equiv \frac{\mathbf{1}\{e \in \delta v(+)\}}{|\delta v(+)|}\pi_{\mathscr{D}}(L \,|\, t_e). \tag{9.70}$$

We can equivalently regard π as a probability measure on pairs (t, L) where $t = t_e$. Now, on this probability space, define the random variable

$$X \equiv X(e, L) \equiv \mathbf{1}\{L \in \mathbb{D}\}\frac{\omega_{L,j}(\mathtt{ry})}{\star\pi_e(\mathtt{r})} \equiv X(t, L),$$

and note that $0 \leq X \leq 1$ with probability one. The expectation of X with respect to π is

$$\mathbf{E}_\pi X \overset{(9.70)}{=} \frac{1}{|\delta v(+)|}\left\{\sum_{e \in \delta v(+)} \sum_L \pi_{\mathscr{D}}(L \,|\, t_e)\frac{\omega_{L,j}(\mathtt{ry})}{\star\pi_e(\mathtt{r})} - O\left(\sum_{e \in \delta v(+)} \sum_L \mathbf{1}\{L \notin \mathbb{D}\}\pi_{\mathscr{D}}(L \,|\, t_e)\right)\right\}$$

$$\overset{(9.65)}{\geq} \frac{1}{|\delta v(+)|}\sum_{e \in \delta v(+)} \sum_L \pi_{\mathscr{D}}(L \,|\, t_e)\frac{\omega_{L,j}(\mathtt{ry})}{\star\pi_e(\mathtt{r})} - \frac{1}{k} = \frac{1}{|\delta v(+)|}\sum_{e \in \delta v(+)} \frac{\pi_e(\mathtt{ry})}{\star\pi_e(\mathtt{r})} - \frac{1}{k}$$

$$\overset{(9.68)}{\geq} \{1 - o_k(1)\}\frac{2^{k-1} \cdot k/6}{k2^{k-1}\ln 2} \geq \frac{1}{5}.$$

On the other hand, since we noted above that $0 \leq X \leq 1$, we have

$$\mathbf{E}_\pi X \leq \pi\left(X \geq \frac{1}{10}\right) + \frac{1}{10}\left\{1 - \pi\left(X \geq \frac{1}{10}\right)\right\},$$

and rearranging gives $\pi(X \geq 1/10) \geq 1/9$. On the event $X(e, L) \geq 1/10$, it follows from (9.69) that

$$\varpi_{L,j}\begin{pmatrix} \sigma = \mathtt{ry}, \\ \Sigma = \mathtt{vw} \end{pmatrix} \geq \omega_{L,j}(\mathtt{ry})\{1 - o_k(1)\}. \tag{9.71}$$

In the next step we will perform a switching argument between (9.67) and (9.71).

Step 2. Edge switching argument. Let $\mathscr{G} = (V, F, E)$ be any (processed) k-SAT graph. Fix a clause type L and index $1 \leq j \leq k(L)$, and let $E(L, j)$ denote the subset of all edges $e = (av) \in E$ such that $L_a = L$ and $j(t_e) = j$.

Suppose $\underline{\sigma}$ is a valid pair coloring on $\mathscr{G}$, and let $\underline{\Sigma}$ be the corresponding configuration from Definition 9.1. Suppose we have two edges $e = (av)$ and $e' = (a'v')$ in $E(L, j)$, such that

$$\mathbf{X}_{av} \equiv \begin{pmatrix} \sigma_{av} \\ \Sigma_{av} \end{pmatrix} = \begin{pmatrix} \mathtt{bb} \\ \mathtt{ww} \end{pmatrix} \equiv A_1, \quad \mathbf{X}_{a'v'} \equiv \begin{pmatrix} \sigma_{a'v'} \\ \Sigma_{a'v'} \end{pmatrix} = \begin{pmatrix} \mathtt{ry} \\ \mathtt{vw} \end{pmatrix} \equiv A_2.$$

If we cut the edges e, e' and form new edges $(a'v), (av')$, then a valid configuration $(\underline{\sigma}', \underline{\Sigma}')$ on the switched graph $\mathscr{G}'$ is given by setting

$$\mathbf{X}_{a'v} \equiv \begin{pmatrix} \sigma_{a'v} \\ \Sigma_{a'v} \end{pmatrix} = \begin{pmatrix} \mathtt{rb} \\ \mathtt{vw} \end{pmatrix} \equiv B_1, \quad \mathbf{X}_{av'} \equiv \begin{pmatrix} \sigma_{av'} \\ \Sigma_{av'} \end{pmatrix} = \begin{pmatrix} \mathtt{by} \\ \mathtt{ww} \end{pmatrix} \equiv B_2,$$

keeping all other colors unchanged. Moreover, the switching preserves all single-copy marginals, so $\underline{\sigma}'$ is a judicious configuration on $\mathscr{G}'$. If $X(t, L) \geq 1/10$ for $t = L(j)$, then

$$p(A_1) \equiv \frac{|\{e \in E(L, j) : \mathbf{X}_{av} = A_1\}|}{|E(L, j)|} \overset{(9.67)}{\geq} \left\{ 1 - o_k(1) \right\} \omega_{L,j}(\mathtt{bb}),$$

$$p(A_2) \equiv \frac{|\{e \in E(L, j) : \mathbf{X}_{av} = A_2\}|}{|E(L, j)|} \overset{(9.71)}{\geq} \left\{ 1 - o_k(1) \right\} \omega_{L,j}(\mathtt{ry}).$$

Suppose ϖ (the empirical measure on the augmented spins (σ, Σ)) gives the maximal second moment contribution $\mathbb{E}_{\mathscr{D}}Z^2(\varpi)$, subject to the restriction that its projection ω lies in I_0. Let $\mathbf{\Omega}$ denote the space of all (valid) pairs $(\mathscr{G}, \underline{\sigma})$ that are consistent with this ϖ, and consider the probability measure

$$\mathbf{P}(\mathscr{G}, \underline{\sigma}) = \frac{\mathbb{P}_{\mathscr{D}}(\mathscr{G})\mathbf{1}\{(\mathscr{G}, \underline{\sigma}) \in \mathbf{\Omega}\}}{\mathbb{E}_{\mathscr{D}}Z^2(\varpi)}.$$

Let $(\mathscr{G}, \underline{\sigma}) \in \mathbf{\Omega}$ be sampled according to measure $\mathbf{P}$, and define the subsets of edges

$$E_A \equiv E_A(\mathscr{G}, \underline{\sigma}) \equiv \left\{ e \in E(L, j) : \mathbf{X}_e \in \{A_1, A_2\} \right\},$$

$$E_B \equiv E_B(\mathscr{G}, \underline{\sigma}) \equiv \left\{ e \in E(L, j) : \mathbf{X}_e \in \{B_1, B_2\} \right\}.$$

If we rematch the edges within E_A, and also rematch the edges within E_B (uniformly at random), the resulting $(\mathscr{G}', \underline{\sigma}')$ will also be distributed roughly according to $\mathbf{P}$. The number of B_1-edges in $(\mathscr{G}', \underline{\sigma}')$ is lower bounded by the number of switched edges from $E_A(\mathscr{G}, \underline{\sigma})$, so we conclude

$$\frac{p(B_1)}{1 - o_n(1)} \geq \frac{|E_A(\mathscr{G}, \underline{\sigma})|}{|E(L, j)|} \frac{p(A_1)p(A_2)}{p(A_1) + p(A_2)} \geq \left\{ 1 - o_n(1) \right\} p(A_1)p(A_2).$$

Recalling the definition of the A_i, B_i, the above can be rewritten as

$$\omega_{L,j}(\mathtt{rb}) \geq \frac{\omega_{L,j}(\mathtt{bb})\omega_{L,j}(\mathtt{ry})}{1 + o_k(1)} \geq \frac{\pi_v(\texttt{++})_\star \pi_e(\mathtt{r})}{11},$$

where the last step uses the assumption that $X(t, L) \geq 1/10$. It follows that

$$\sum_{e \in \delta v(\texttt{+})} \pi_e(\mathtt{rb}) \geq \sum_{e \in \delta v(\texttt{+})} \sum_L \mathbf{1}\left\{ X(e, L) \geq \frac{1}{10} \right\} \pi_{\mathscr{D}}(L \mid t_e)\omega_{L,j}(\mathtt{rb})$$

$$= |\delta v(\texttt{+})|\pi\left(X \geq \frac{1}{10} \right) \frac{\pi_v(\texttt{++})_\star \pi_e(\mathtt{r})}{11} \geq \frac{k2^{k-1}\ln 2}{9[1 - o_k(1)]} \frac{\pi_v(\texttt{++})}{11 \cdot 2^k} \geq \frac{k\pi_v(\texttt{++})}{300},$$

proving (9.66). A very similar argument proves (9.66) in the case that (9.68) holds for $\tau = \mathtt{g}$ instead of $\tau = \mathtt{y}$. As explained above, (9.66) implies the lemma, so this concludes the proof. $\qquad\square$

Corollary 9.12 (used only in proof of Proposition 9.17). *Consider the setting of Lemma 9.11, but without assuming a lower bound on $\pi_v(\mathtt{xx})$. Then*

$$\sum_{e \in \delta v} \sum_{L \in \mathbb{D}} \pi_{\mathscr{D}}(L \mid t_e)\omega_{L,j(t_e)}(\mathtt{rr}) \leq \frac{k^2}{2^{k\epsilon_4}}$$

for ϵ_4 as in Lemmas 9.10 and 9.11.

Proof. In view of Lemma 9.11 it suffices to consider the case $\pi_v(\text{++}) \le 2^{-k/16}$. Suppose the conclusion of this corollary fails; more explicitly, suppose that

$$\sum_{e \in \delta v} \sum_{L \in \mathbb{D}} \pi_{\mathscr{D}}(L \mid t_e) \omega_{L,j(t_e)}(\text{rr}) \ge \frac{k^4}{2^{k/16}} \,.$$

This implies that the assumption (9.43) of Lemma 9.8 holds, since

$$\sum_{e \in \delta v} \pi_e(\text{rr} \mid \sigma \in \{\text{r}, \text{b}\}^2) \ge \frac{k^4/2^{k/16}}{\pi_v(\text{++})} \ge k^4 \,.$$

In this case, combining Lemma 9.8 and Lemma 9.10 gives that for all L, j with $L(j) = t_e$ for $e \in \delta v(+)$,

$$\omega_{L,j}(\text{rr}) \le \frac{\omega_{L,j}(\text{vv})}{1 - o_k(1)} \le \frac{1}{2^{k(1+\epsilon_4)}} \,,$$

from which the conclusion of this corollary follows. $\square$

9.5. **Non-defective variables neighboring only diverse light clauses.** The main result of this subsection is the following proposition:

Proposition 9.13 (used only in proof of Proposition 8.4). *In the notation of Lemma 9.11, suppose v is a non-defective variable with $\mathbb{D}(v) = \mathbb{L}(v) = 0$. Then for every $e \in \delta v$, and every $L \ni_j t_e$ we have*

$$\left| \frac{\omega_{L,j}(\sigma)}{\star \pi_{L(j)}(\sigma^1) \star \pi_{L(j)}(\sigma^2)} - 1 \right| \le \frac{k^{O(1)}}{2^{k\epsilon_\circ}} \,, \tag{9.72}$$

for all $\sigma \in \{\text{r}, \text{y}, \text{g}, \text{b}\}^2 \setminus \{\text{rr}\}$ and all L, with $\epsilon_\circ$ as in Lemma 9.7.

The proof of Proposition 9.13 appears at the end of this subsection. Its main ingredient is the next lemma:

Lemma 9.14 (used only in proof of Proposition 9.13). *Let T be a non-defective variable type. Consider the second moment of judicious configurations under $\mathbf{P}_{\mathscr{D}}$, restricted to the near-independent regime I_0 (as has been the case throughout this section). Let $\mathbf{P}$ denote the empirical measure of configurations $(\underline{\sigma}_{\delta v}, \underline{\varsigma}_{\delta v}, \underline{L}_{\delta v})$, taken over all variables v of type T, that gives the maximal contribution to this restricted second moment. For $e \in \delta v$ let ζ_e denote the marginal law of (ς_e, L_e) under $\mathbf{P}$, and suppose it satisfies*

$$1 - \zeta_e(\text{ww} \mid L) \le \frac{k^{O(1)}}{2^k} \,, \tag{9.73}$$

$$\hat{\iota}_{e,L} \equiv \zeta_e\left(\varsigma \in \{\text{r}, \text{y}, \text{c}\}^2 \,\middle|\, L \right) + \frac{k^4}{4^k} \le \frac{1}{2^{k(1+\epsilon)}} \tag{9.74}$$

for all $e \in \delta v$ and all L. Then it must satisfy the estimate

$$\left| \frac{\mathbf{P}(\sigma_e = \sigma \mid L_e = L)}{\star \pi_e(\sigma^1) \star \pi_e(\sigma^2)} - 1 \right| \le \frac{k^{O(1)}}{2^{k\epsilon}} \tag{9.75}$$

for all $\sigma \in \{\text{r}, \text{y}, \text{g}, \text{b}\}^2 \setminus \{\text{rr}\}$ and all L.

Proof. As in the statement of the lemma, let us fix a variable v of type T that is non-defective. Write

$$\underline{X}_{\delta v} \equiv (X_e)_{e \in \delta v} \equiv \left(\left(\sigma_e, \varsigma_e \right) \right)_{e \in \delta v} \,. \tag{9.76}$$

Note that, by the rules of Definition 9.1, the configuration $\underline{X}_{\delta v}$ also implicitly encodes $\underline{\Sigma}_{\delta v}$. In the augmented model where each edge $e = (av)$ is also labelled with the clause type $L_e \equiv L_a$, denote

$$X \equiv X_v \equiv (\underline{X}_{\delta v}, \underline{L}_{\delta v}) \equiv \left((X_e)_{e \in \delta v}, (L_e)_{e \in \delta v} \right). \tag{9.77}$$

Recall that ζ_e denotes the marginal law of (L_e, ς_e) under $\mathbf{P}$. We can write it as

$$\zeta_e(L, \varsigma) \equiv \mathbf{P}\left(L_e = L, \varsigma_e = \varsigma \right) = \pi_{\mathscr{D}}(L \mid t_e) \zeta_e(\varsigma \mid L) \,.$$

The measure $\mathbf{P}$ must maximize entropy subject to the marginal constraints

$$\mathbf{P}(L_e = L) = \pi_{\mathscr{D}}(L \mid t_e) \quad \text{for all } L, \tag{9.78}$$

$$\mathbf{P}((\sigma_e)^i = \sigma^i \mid L_e = L) = {}_\star\pi_e(\sigma^i) \quad \text{for all } L \text{ and all } \sigma^i \in \{\text{r}, \text{y}, \text{g}, \text{b}\}, \tag{9.79}$$

$$\mathbf{P}(\varsigma_e = \varsigma \mid L_e = L) = \boldsymbol{\zeta}_e(\varsigma \mid L) \quad \text{for all } L \text{ and all } \varsigma \in \{\text{r}, \text{y}, \text{c}, \text{w}\}^2, \tag{9.80}$$

for all $e \in \delta v$. By the method of Lagrange multipliers, $\mathbf{P}$ must take the form

$$\mathbf{P}(\underline{\sigma}_{\delta v}, \underline{\varsigma}_{\delta v}, \underline{L}_{\delta v}) \cong \psi_v(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \left\{ \mathbf{1}\{\sigma_e \sim \varsigma_e\} \psi_e(L_e) \left(\prod_{i=1,2} (\beta_e)^i((\sigma_e)^i \mid L_e) \right) \chi_e(\varsigma_e \mid L_e) \right\} \tag{9.81}$$

where $\psi_v(\underline{\sigma}_{\delta v})$ is as in (7.55) from Definition 7.24, and we write $\sigma \sim \varsigma$ to indicate compatibility: formally, it means for both $i = 1, 2$ that we have $\sigma^i \sim \varsigma^i$ in the sense that the following holds:

$$\varsigma^i = \text{r} \quad \text{whenever } \sigma^i = \text{r},$$
$$\varsigma^i \in \{\text{y}, \text{w}\} \quad \text{whenever } \sigma^i = \text{y},$$
$$\varsigma^i \in \{\text{c}, \text{w}\} \quad \text{whenever } \sigma^i \in \{\text{g}, \text{b}\}.$$

We now turn to the construction of the weights in (9.81). We divide the remainder of the argument into a few parts.

Part 1. Single-copy estimates. Let $v \equiv (\dot{v}, \hat{v})$ be the vertex empirical measure that gives the maximal contribution to the first moment (of judicious configurations). The edge marginal of $\hat{v}_L$ is the canonical marginal $\omega_{L,j} = {}_\star\pi_{L(j)}$. Given the mapping from $\underline{\sigma}_{\delta a}$ to $\underline{\varsigma}_{\delta a}$ within each clause, the measure $\hat{v}_L$ induces a measure on configurations $\underline{\varsigma}_{\delta a}$. Let ${}_\star\zeta_{L,j} \equiv {}_\star\zeta_{L(j)}$ denote the marginal law of ς_j under this measure; we call this the canonical marginal on ς^i for an edge of type $t = L(j)$. If e is an edge of type t then we also write ${}_\star\zeta_e \equiv {}_\star\zeta_t$. Explicitly,

$$\star\zeta_e(\text{r}) \cong {}_\star\dot{q}_e(\text{r}) \left\{ \prod_{e' \in \delta v \setminus e} {}_\star\dot{q}_{e'}(\text{y}) \right\},$$

$$\star\zeta_e(\text{y}) \cong \dot{q}_e(\text{y}) \left\{ \sum_{e' \in \delta v \setminus e} \dot{q}_{e'}(\text{r}) \prod_{e'' \in \delta v \setminus \{e, e'\}} \dot{q}_{e''}(\text{y}) \right\},$$

$$\star\zeta_e(\text{c}) \cong \dot{q}_e(\text{c}) \left\{ \sum_{e' \in \delta v \setminus e} \dot{q}_{e'}(\text{c}) \prod_{e'' \in \delta v \setminus \{e, e'\}} \dot{q}_{e''}(\text{y}) \right\},$$

$$\star\zeta_e(\text{w}) \cong {}_\star\dot{q}_e(\{\text{y}, \text{c}\}) \prod_{e' \in \delta v \setminus e} {}_\star\dot{q}_{e'}(\{\text{y}, \text{c}\}) \left(1 - \frac{k^{O(1)}}{2^k} \right),$$

where the last estimate uses the assumption that v is non-defective, hence nice, so that ${}_\star\dot{q}_e$ satisfies the estimates of Definition 3.8 for all $e \in \delta v$. It follows that

$$\begin{pmatrix} \star\zeta_e(\text{r}) \\ \star\zeta_e(\text{y}) \\ \star\zeta_e(\text{c}) \\ \star\zeta_e(\text{w}) \end{pmatrix} = \begin{pmatrix} \Theta(1/2^k) \\ \Theta(k/2^k) \\ \Theta(k/2^k) \\ 1 - \Theta(k/2^k) \end{pmatrix}. \tag{9.82}$$

(We also remark that ${}_\star\zeta_e(\text{y}) = {}_\star\zeta_e(\text{c})$, although we will not use this fact in what follows.)

Part 2. Single-copy weights. We first consider the simpler problem of setting weights in the single-copy model. For this discussion, let $\sigma \in \{\text{r}, \text{y}, \text{g}, \text{b}\}$ and $\varsigma \in \{\text{r}, \text{y}, \text{c}, \text{w}\}$, and again write $\sigma \sim \varsigma$ to indicate compatibility. Given a probability measure ${}_\star\pi$ over $\{\text{r}, \text{y}, \text{g}, \text{b}\}$, along with a probability measure ${}_\star\zeta$ over spins $\{\text{r}, \text{y}, \text{c}, \text{w}\}$ such that ${}_\star\zeta(\text{r}) = {}_\star\pi(\text{r})$, we look for weights ${}_\star b$ and ${}_\star x$ such that

$$\mu(\sigma, \varsigma) \equiv \mathbf{1}\{\sigma \sim \varsigma\} {}_\star\dot{q}(\sigma) {}_\star b(\sigma) {}_\star x(\varsigma) \tag{9.83}$$

defines a probability measure over (σ, ς) whose marginal on σ is $_\star\pi$, and whose marginal on ς is $_\star\zeta$. Explicitly, for the σ-marginal to be $_\star\pi$ we must have the equations

$$\begin{aligned}
{}_\star\pi(\mathtt{r}) &= {}_\star\dot{q}(\mathtt{r})_\star b(\mathtt{r})_\star x(\mathtt{r}) = {}_\star\zeta(\mathtt{r})\,, \\
{}_\star\pi(\mathtt{y}) &= {}_\star\dot{q}(\mathtt{y})_\star b(\mathtt{y})[{}_\star x(\mathtt{y}) + {}_\star x(\mathtt{w})]\,, \\
{}_\star\pi(\mathtt{g}) &= {}_\star\dot{q}(\mathtt{g})_\star b(\mathtt{g})[{}_\star x(\mathtt{c}) + {}_\star x(\mathtt{w})]\,, \\
{}_\star\pi(\mathtt{b}) &= {}_\star\dot{q}(\mathtt{b})_\star b(\mathtt{b})[{}_\star x(\mathtt{c}) + {}_\star x(\mathtt{w})]\,.
\end{aligned}$$

For the ς-marginal to be $_\star\zeta$ we must have the equations

$$\begin{aligned}
{}_\star\zeta(\mathtt{y}) &= {}_\star\dot{q}(\mathtt{y})_\star b(\mathtt{y})_\star x(\mathtt{y})\,, \\
{}_\star\zeta(\mathtt{c}) &= [{}_\star\dot{q}(\mathtt{g})_\star b(\mathtt{g}) + {}_\star\dot{q}(\mathtt{b})_\star b(\mathtt{b})]_\star x(\mathtt{c})\,, \\
{}_\star\zeta(\mathtt{w}) &= [{}_\star\dot{q}(\mathtt{y})_\star b(\mathtt{y}) + {}_\star\dot{q}(\mathtt{g})_\star b(\mathtt{g}) + {}_\star\dot{q}(\mathtt{b})_\star b(\mathtt{b})]_\star x(\mathtt{w})\,.
\end{aligned}$$

Note also that if we multiply all the $_\star b$-weights by a scaling factor and divide all the $_\star x$-weights by the same factor, it has no effect on the right-hand side of (9.83), so without loss we can pin down the weights by requiring $_\star b(\mathtt{r}) = 1$ and $_\star x(\mathtt{w}) = 1$. Then, combining the equations for $_\star\pi(\mathtt{y})$ and $_\star\zeta(\mathtt{y})$ gives

$$_\star b(\mathtt{y}) = \frac{{}_\star\pi(\mathtt{y}) - {}_\star\zeta(\mathtt{y})}{{}_\star\dot{q}(\mathtt{y})}\,, \qquad _\star x(\mathtt{y}) = \frac{{}_\star\zeta(\mathtt{y})}{{}_\star\pi(\mathtt{y}) - {}_\star\zeta(\mathtt{y})}\,.$$

Next, comparing the equations for $_\star\pi(\mathtt{g})$ and $_\star\pi(\mathtt{b})$ gives

$$\frac{{}_\star b(\mathtt{g})}{{}_\star b(\mathtt{b})} = \frac{{}_\star\pi(\mathtt{g})/{}_\star\dot{q}(\mathtt{g})}{{}_\star\pi(\mathtt{b})/{}_\star\dot{q}(\mathtt{b})} = \frac{{}_\star\hat{q}(\mathtt{g})}{{}_\star\hat{q}(\mathtt{b})} = 1\,,$$

and we hereafter denote $_\star b(\mathtt{c}) \equiv {}_\star b(\mathtt{g}) \equiv {}_\star b(\mathtt{b})$. Combining with the equations for $_\star\pi(\mathtt{g})$, $_\star\pi(\mathtt{b})$, and $_\star\zeta(\mathtt{c})$ gives

$$_\star b(\mathtt{c}) = \frac{{}_\star\pi(\mathtt{g}) + {}_\star\pi(\mathtt{b}) - {}_\star\zeta(\mathtt{c})}{{}_\star\dot{q}(\mathtt{g}) + {}_\star\dot{q}(\mathtt{b})} = \frac{{}_\star\pi(\mathtt{c}) - {}_\star\zeta(\mathtt{c})}{{}_\star\dot{q}(\mathtt{c})}\,, \qquad _\star x(\mathtt{c}) = \frac{{}_\star\zeta(\mathtt{c})}{{}_\star\pi(\mathtt{c}) - {}_\star\zeta(\mathtt{c})}\,.$$

In summary, a valid solution is given by taking

$$\begin{pmatrix} {}_\star b(\mathtt{r}) \\ {}_\star b(\mathtt{y}) \\ {}_\star b(\mathtt{g}) = {}_\star b(\mathtt{b}) \equiv {}_\star b(\mathtt{c}) \end{pmatrix} = \begin{pmatrix} 1 \\ [{}_\star\pi(\mathtt{y}) - {}_\star\zeta(\mathtt{y})]/{}_\star\dot{q}(\mathtt{y}) \\ [{}_\star\pi(\mathtt{c}) - {}_\star\zeta(\mathtt{c})]/{}_\star\dot{q}(\mathtt{c}) \end{pmatrix} \overset{(9.82)}{=} \begin{pmatrix} \Theta(1) \\ \Theta(1) \\ \Theta(1) \end{pmatrix}\,, \tag{9.84}$$

$$\begin{pmatrix} {}_\star x(\mathtt{r}) \\ {}_\star x(\mathtt{y}) \\ {}_\star x(\mathtt{c}) \\ {}_\star x(\mathtt{w}) \end{pmatrix} = \begin{pmatrix} {}_\star\pi(\mathtt{r})/{}_\star\dot{q}(\mathtt{r}) \\ {}_\star\zeta(\mathtt{y})/[{}_\star\pi(\mathtt{y}) - {}_\star\zeta(\mathtt{y})] \\ {}_\star\zeta(\mathtt{c})/[{}_\star\pi(\mathtt{c}) - {}_\star\zeta(\mathtt{c})] \\ 1 \end{pmatrix} \overset{(9.82)}{=} \begin{pmatrix} \Theta(1/2^k) \\ \Theta(k/2^k) \\ \Theta(k/2^k) \\ 1 \end{pmatrix}\,. \tag{9.85}$$

Substituting these into (9.83) gives a measure with the desired marginals $_\star\pi$ and $_\star\zeta$.

Part 3. Initialization in pair model. Returning to (9.81), our goal is to construct a sequence of measures

$$\mathbf{P}_t(\underline{\sigma}_{\delta v}, \varsigma_{\delta v}, \underline{L}_{\delta v}) \cong \psi_{v,t}(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \left\{ \psi_{e,t}(L_e) \left(\prod_{i=1,2} (\beta_{e,t})^i((\sigma_e)^i \mid L_e) \right) \chi_{e,t}(\varsigma_e \mid L_e) \right\}$$

which converges as $t \to \infty$ to the desired solution $\mathbf{P}_\infty \equiv \mathbf{P}$. We will further decompose the β weights as

$$(\beta_{e,t})^i(\sigma^i \mid L_e) \equiv \left\{ (\dot{\beta}_{e,t})^i(\sigma^i \mid L_e) \right\} \cdot \left\{ (\hat{\beta}_{e,t})^i(\sigma^i \mid L_e) \right\}$$

such that $(\hat{\beta}_{e,t})^i$ does not distinguish between $\mathtt{g}$ and $\mathtt{b}$, that is, such that

$$(\hat{\beta}_{e,t})^i(\mathtt{g} \mid L_e) = (\hat{\beta}_{e,t})^i(\mathtt{b} \mid L_e) \equiv (\hat{\beta}_{e,t})^i(\mathtt{c} \mid L_e)\,.$$

This is clearly an over-parametrization, so the $\dot{\beta}$ and $\hat{\beta}$ weights will not be uniquely determined; we need only find one choice of weights such that the resulting measure (9.81) satisfies the constraints (9.78)–(9.80). The overparametrization will be useful below because it allows for some separation between the analysis of the ς-marginals and the analysis of the (σ, L)-marginals.

We initialize the construction at $t = 0$ as follows. We first set $\psi_{v,0}(\underline{\sigma}_{\delta v})$ to be equal to $\varphi_v(\underline{\sigma}_{\delta v})$, which we recall from (2.18) is simply the indicator of a valid pair coloring $\underline{\sigma}_{\delta v}$. For an edge e of type t, let $_\star\pi_e \equiv {}_\star\pi_t$ be the canonical

marginal on σ^i, and let $_\star\zeta_e \equiv {}_\star\zeta_t$ be the canonical marginal on ς^i, as discussed above. Let $_\star b_e$ and $_\star x_e$ be the corresponding weights defined by (9.84) and (9.85); in particular, we recall from (9.84) that $_\star b_e$ does not distinguish between $\mathtt{g}$ and $\mathtt{b}$. We then set

$$
\begin{aligned}
\psi_{e,0}(L) &\equiv \pi_{\mathscr{D}}(L \mid t_e) &&\text{for all } L, \\
(\dot\beta_{e,0})^i(\sigma^i \mid L) &\equiv 1 &&\text{for } i = 1, 2, \text{ all } L, \text{ and all } \sigma^i \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}, \\
(\hat\beta_{e,0})^i(\sigma^i \mid L) &\equiv {}_\star b_e(\sigma^i) &&\text{for } i = 1, 2, \text{ all } L, \text{ and all } \sigma^i \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}, \\
\chi_{e,0}(\varsigma \mid L) &\equiv {}_\star x_e(\varsigma^1){}_\star x_e(\varsigma^2) &&\text{for all } L \text{ and all } \varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2.
\end{aligned}
$$

Thus at $t = 0$ the measure $\mathbf{P}_0$ will satisfy constraints (9.78) and (9.79), but not (9.80), since we will have

$$
\zeta_{e,0}(\varsigma \mid L) \equiv \mathbf{P}_0(\varsigma_e = \varsigma \mid L_e = L) = \prod_{i=1,2} {}_\star\zeta_e(\varsigma^i),
$$

which in general is not the same as $\zeta_e(\varsigma \mid L)$.

Part 4. Update procedure. In this step we will make use of Lemma 9.15, which is stated and proved below. For $t \geq 0$ let $\delta_{e,t}$ and $\ddot\delta_{e,t}$ denote parameters such that the estimates (9.95)–(9.97) below hold for the error between $\zeta_{e,t}(\cdot \mid L)$ versus $\zeta_e(\cdot \mid L)$ — that is to say, for all clause types L that can appear incident to edge e, we assume that

$$
\max\left\{\left|\zeta_e(\varsigma \mid L) - \zeta_{e,t}(\varsigma \mid L)\right| : \varsigma \in \{\mathtt{rw}, \mathtt{yw}, \mathtt{cw}, \mathtt{wr}, \mathtt{wy}, \mathtt{wc}\}\right\} \leq \frac{k^{O(1)}\delta_{e,t}}{2^k}, \tag{9.86}
$$

$$
\max\left\{\left|\zeta_e(\varsigma \mid L) - \zeta_{e,t}(\varsigma \mid L)\right| : \varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2\right\} \leq \hat\imath\delta_{e,t}, \tag{9.87}
$$

$$
\left|\frac{\zeta_e(\mathtt{rr} \mid L)}{\zeta_{e,t}(\mathtt{rr} \mid L)} - 1\right| \leq \ddot\delta_{e,t}. \tag{9.88}
$$

At the initialization $t = 0$ it follows from (9.73) and (9.74) that $\delta_{e,t} \leq O(1)$ and $\ddot\delta_{e,t} \leq O(2^{k(1-\epsilon)})$. Denote

$$
\delta_t \equiv \sum_{e \in \delta v} \delta_{e,t}, \quad \ddot\delta_t \equiv \sum_{e \in \delta v} \min\{\ddot\delta_{e,t}, 1\}. \tag{9.89}
$$

Let $t \geq 0$ be an integer time, and suppose inductively that we have constructed the weights at time t. The marginal law on an edge $e \in \delta v$ is given by

$$
\mathbf{P}_t(\sigma_e = \sigma, \varsigma_e = \varsigma, L_e = L) \cong \mathbf{1}\{\sigma \sim \varsigma\}\dot q_{e,t}(\sigma)\psi_{e,t}(L)\left(\prod_{i=1,2}(\beta_{e,t})^i(\sigma^i \mid L)\right)\chi_{e,t}(\varsigma \mid L),
$$

where $\dot q_{e,t}$ is the probability measure over $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$ defined by

$$
\dot q_{e,t}(\sigma) \cong \sum_{\underline\sigma_{\delta v}:\sigma_e=\sigma}\psi_{v,t}(\underline\sigma)\prod_{e' \in \delta v \backslash e}\left\{\sum_{L_{e'}}\psi_{e',t}(L_{e'})\left(\prod_{i=1,2}\beta_{e',t}((\sigma_{e'})^i \mid L_{e'})\right)\sum_{\varsigma_{e'}:\varsigma_{e'}\sim\sigma_{e'}}\chi_{e',t}(\varsigma_{e'} \mid L_{e'})\right\}.
$$

The conditional law of (σ_e, ς_e) given L_e is then given by

$$
\mathbf{P}_t(\sigma_e = \sigma, \varsigma_e = \varsigma \mid L_e = L) \cong \mathbf{1}\{\sigma \sim \varsigma\}\dot q_{e,t}(\sigma)\left(\prod_{i=1,2}(\beta_{e,t})^i(\sigma^i \mid L)\right)\chi_{e,t}(\varsigma \mid L).
$$

For each L, we apply Lemma 9.15 (below) to find updated weights $(\hat\beta_{e,t+3/4})^i(\sigma^i \mid L)$ and $\chi_{e,t+3/4}(\varsigma \mid L)$ such that, if we define $(\beta_{e,t+3/4})^i \equiv (\dot\beta_{e,t})^i \cdot (\hat\beta_{e,t+3/4})^i$, then the probability measure

$$
\tilde\mu_{e,t+3/4}(\sigma, \varsigma \mid L) \cong \mathbf{1}\{\sigma \sim \varsigma\}\dot q_{e,t}(\sigma)\left(\prod_{i=1,2}(\beta_{e,t+3/4})^i(\sigma^i \mid L)\right)\chi_{e,t+3/4}(\varsigma \mid L)
$$

has ς-marginal exactly $\zeta_e(\,\cdot\mid L)$. The marginal law of $(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v})$ at time $t+3/4$ is then given by

$$\mathbf{P}_{t+3/4}(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}) \cong \psi_{v,t}(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \underbrace{\left\{ \psi_{e,t}(L_e) \sum_{\varsigma_e : \varsigma_e \sim \sigma_e} \left(\prod_{i=1,2} (\beta_{e,t+3/4})^i((\sigma_e)^i \mid L_e) \right) \chi_{e,t+3/4}(\varsigma_e \mid L_e) \right\}}_{\cong \hat{q}_{e,t+3/4}(\sigma_e, L_e)},$$

where $\hat{q}_{e,t}$ is a probability measure over pairs (σ, L). It follows from Lemma 9.15 that

$$\widehat{\mathbf{ERR}}(\hat{q}_{e,t}, \hat{q}_{e,t+3/4}) \leq \max_L \left\{ k^{O(1)} \begin{pmatrix} \hat{\iota}_{e,L} & 0 \\ 2^k \hat{\iota}_{e,L} & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \delta_{e,t} \\ \ddot{\delta}_{e,t} \end{pmatrix} \right\} \overset{(9.74)}{\leq} k^{O(1)} \begin{pmatrix} 2^{-k(1+\epsilon)} & 0 \\ 2^{-k\epsilon} & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \delta_{e,t} \\ \ddot{\delta}_{e,t} \end{pmatrix}. \tag{9.90}$$

Then apply Proposition 7.28 (where the $\hat{p}_e$ and $\hat{q}_e$ of Proposition 7.28 are given by $\hat{q}_{e,t}$ and $\hat{q}_{e,t+3/4}$ respectively) to find new weights $\psi_{v,t+1}(\underline{\sigma})$, $\psi_{e,t+1}(L)$, and $(\dot{\beta}_{e,t+1})^i(\sigma^i \mid L)$ such that, if we define $(\beta_{e,t+1})^i \equiv (\dot{\beta}_{e,t+1})^i(\hat{\beta}_{e,t+3/4})^i$, then the probability measure

$$\mathbf{P}_{t+1}(\underline{\sigma}_{\delta v}, \underline{\varsigma}_{\delta v}, \underline{L}_{\delta v}) \cong \psi_{v,t+1}(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \left\{ \psi_{e,t+1}(L_e) \left(\prod_{i=1,2} (\beta_{e,t+1})^i((\sigma_e)^i \mid L_e) \right) \chi_{e,t+1}(\varsigma_e \mid L_e) \right\},$$

has a marginal on $(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v})$ that is fully judicious (in the sense of Definition 3.45). In the notation of Proposition 7.28, in going between $\hat{q}_{e,t}$ and $\hat{q}_{e,t+3/4}$ we have

$$\mathrm{err}_v = \mathrm{err}_{v,t} = \sum_{e \in \delta v} \frac{\delta_{e,t} + \min\{\ddot{\delta}_{e,t}, 1\}}{2^{k(1+\epsilon)}} = \frac{\delta_t + \ddot{\delta}_t}{2^{k(1+\epsilon)}},$$

using the notation of (9.89). It then follows from the bound (7.62) of Proposition 7.28 that

$$\max_L \left\{ \left| \frac{\psi_{e,t+1}(L)}{\psi_{e,t+3/4}(L)} - 1 \right| + \max_{\tau \neq \mathbf{r}} \left| \frac{(\beta_{e,t+1})^i(\tau \mid L)}{(\beta_{e,t+3/4})^i(\tau \mid L)} - 1 \right| \right\} \leq \frac{k^{O(1)}}{2^{k(1+\epsilon)}} \left(\delta_{e,t} + \min\{\ddot{\delta}_{e,t}, 1\} + \mathrm{err}_{v,t} \right),$$

$$\max_L \left\{ \left| \frac{(\beta_{e,t+1})^i(\mathbf{r} \mid L)}{(\beta_{e,t+3/4})^i(\mathbf{r} \mid L)} - 1 \right| \right\} \leq \frac{k^{O(1)}}{2^{k\epsilon}} \left(\delta_{e,t} + \min\{\ddot{\delta}_{e,t}, 1\} + \mathrm{err}_{v,t} \right). \tag{9.91}$$

It follows from the bound (7.63) of Proposition 7.28 that

$$\mathbf{ERR}(\dot{q}_{e,t}, \dot{q}_{e,t+1}) \leq k^{O(1)} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2^{-k\epsilon} \\ 2^{-k\epsilon} \end{pmatrix} \mathrm{err}_{v,t} + k^{O(1)} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 2^{-k} & 2^{-k} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2^{-k(1+\epsilon)} \delta_{e,t} \\ 2^{-k\epsilon} \delta_{e,t} \\ 2^{-k\epsilon} \min\{\delta_{e,t} + \ddot{\delta}_{e,t}, 1\} \end{pmatrix}$$

$$\leq k^{O(1)} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2^{-k\epsilon} \\ 2^{-k\epsilon} \end{pmatrix} \frac{\delta_t + \ddot{\delta}_t}{2^{k(1+\epsilon)}} + k^{O(1)} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \frac{\delta_{e,t} + \min\{\ddot{\delta}_{e,t}, 1\}}{2^{k\epsilon}}. \tag{9.92}$$

At time $t+1$ we have

$$\frac{\zeta_e(\varsigma \mid L)}{\zeta_{e,t+1}(\varsigma \mid L)} = \frac{z_{e,t+1} \displaystyle\sum_{\sigma:\sigma \sim \varsigma} \dot{q}_{e,t}(\sigma)(\beta_{e,t+3/4})^1(\sigma^1)(\beta_{e,t+3/4})^2(\sigma^2)\chi_{e,t+3/4}(\varsigma \mid L)}{\tilde{z}_{e,t+3/4} \displaystyle\sum_{\sigma:\sigma \sim \varsigma} \dot{q}_{e,t+1}(\sigma)(\beta_{e,t+1})^1(\sigma^1)(\beta_{e,t+1})^2(\sigma^2)\chi_{e,t+3/4}(\varsigma \mid L)}$$

for normalizing constants $\tilde{z}_{e,t+3/4}$ and $z_{e,t+1}$. Combining with (9.91) and (9.92) then gives

$$\delta_{e,t+1} + \ddot{\delta}_{e,t+1} \leq \frac{k^{O(1)}}{2^{k\epsilon}} \left(\frac{\delta_t + \ddot{\delta}_t}{2^k} + \delta_{e,t} + \min\{\ddot{\delta}_{e,t}, 1\} \right).$$

This concludes our analysis of the update procedure.

Part 5. Conclusion. Summing the preceding bound over all $e \in \delta v$ gives

$$\sum_{t \geq 0} (\delta_t + \ddot{\delta}_t) \leq \delta_0 + \ddot{\delta}_0 + \frac{k^{O(1)}}{2^{k\epsilon}} \sum_{t \geq 0} (\delta_t + \ddot{\delta}_t) \leq O(\delta_0 + \ddot{\delta}_0) \leq O(k 2^k),$$

$$\sum_{t \geq 0} (\delta_{e,t} + \min\{\ddot{\delta}_{e,t}, 1\}) \leq \delta_{e,0} + \ddot{\delta}_{e,0} + \frac{k^{O(1)}}{2^{k\epsilon}} \sum_{t \geq 0} \left(\frac{\delta_t + \ddot{\delta}_t}{2^k} + \delta_{e,t} + \min\{\ddot{\delta}_{e,t}, 1\} \right) \leq O(1).$$

We can use the above bounds with (9.91) to obtain

$$\sum_{t \geq 0} \widehat{\mathsf{ERR}}(\hat{q}_{e,t+3/4}, \hat{q}_{e,t+1}) \leq \begin{pmatrix} 2^{-k} \\ 1 \\ 1 \end{pmatrix} \frac{k^{O(1)}}{2^{k\epsilon}} \sum_{t \geq 0} \left(\delta_{e,t} + \min\{\ddot{\delta}_{e,t}, 1\} + \mathrm{err}_{v,t} \right) \leq \begin{pmatrix} 2^{-k} \\ 1 \\ 1 \end{pmatrix} \frac{k^{O(1)}}{2^{k\epsilon}}.$$

Combining with (9.90) gives

$$\widehat{\mathsf{ERR}}(\hat{q}_{e,0}, \hat{q}_{e,\infty}) \leq O(1) \sum_{t \geq 0} \left\{ \widehat{\mathsf{ERR}}(\hat{q}_{e,t}, \hat{q}_{e,t+3/4}) + \widehat{\mathsf{ERR}}(\hat{q}_{e,t+3/4}, \hat{q}_{e,t+1}) \right\} \leq \frac{k^{O(1)}}{2^{k\epsilon}} \begin{pmatrix} 2^{-k} \\ 1 \\ 2^k \end{pmatrix}.$$

(We also see that condition (7.59) is satisfied by $\hat{q}_{e,t}$ for all $t \geq 0$, so that the above applications of Proposition 7.28 are justified.) We also obtain from (9.92) that

$$\ddot{\mathsf{ERR}}(\dot{q}_{e,0}, \dot{q}_{e,\infty}) \leq \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2^{-k\epsilon} \\ 1 \end{pmatrix} \frac{k^{O(1)}}{2^{k\epsilon}}.$$

The above estimates imply that for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2 \setminus \{\mathtt{rr}\}$ we have

$$\mathbf{P}(\sigma_e = \sigma \mid L_e = L) \cong \dot{q}_{e,\infty}(\sigma \mid L) \hat{q}_{e,\infty}(\sigma \mid L) = {}_\star \dot{q}_e(\sigma) {}_\star \hat{q}_e(\sigma) \left(1 + \frac{k^{O(1)}}{2^{k\epsilon}} \right) = {}_\star \omega_e(\sigma) \left(1 + \frac{k^{O(1)}}{2^{k\epsilon}} \right).$$

This implies the claimed bounds (9.75).   □

Lemma 9.15 (used only in proof of Lemma 9.14). *In this lemma we take σ in the reduced alphabet $\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$. We continue to take $\varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2$. We write $\sigma \sim \varsigma$ to indicate compatibility: if $\sigma^i = \mathtt{r}$ then $\varsigma^i = \mathtt{r}$; if $\sigma^i \in \{\mathtt{y}, \mathtt{c}\}$ then $\varsigma^i \in \{\sigma^i, \mathtt{w}\}$. On edge e let μ_0 be a probability measure on pairs (σ, ς) with $\sigma \sim \varsigma$, of the form*

$$\mu_0(\sigma, \varsigma) = \frac{\mathbf{1}\{\sigma \sim \varsigma\}}{z_0} \dot{q}(\sigma) (b_0)^1(\sigma^1) (b_0)^2(\sigma^2) x_0(\varsigma),$$

such that $\mu(\sigma^i = \tau) = {}_\star \pi_e(\tau)$ for both $i = 1, 2$ and all $\tau \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$. Assume that ${}_\star \pi_e$ is nice in the sense of Definition 3.8, that $\dot{q}(\sigma) = \Theta(1)$ for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$, and that $b^i(\tau) = \Theta(1)$ for $i = 1, 2$ and all $\tau \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$. Let ζ_0 denote the marginal law of ς under μ_0. Let ζ be another probability measure over $\varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2$ whose single-copy marginals satisfy $\zeta^1(\mathtt{r}) = {}_\star \pi_e(\mathtt{r}) = \zeta^2(\mathtt{r})$. Assume that

$$\max \left\{ 1 - \zeta_0(\mathtt{ww}), 1 - \zeta(\mathtt{ww}) \right\} \leq \frac{k^{O(1)}}{2^k}, \tag{9.93}$$

$$\hat{\imath} \equiv \frac{k^4}{4^k} + \max \left\{ \zeta_0(\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2), \zeta(\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2) \right\} \leq \frac{1}{2^{k(1+\epsilon)}}. \tag{9.94}$$

Assume moreover that we have parameters $\delta \leq O(1)$ and $\ddot{\delta} \leq 2^{k(1-\epsilon)}$ such that

$$\max\left\{\left|\zeta(\varsigma) - \zeta_0(\varsigma)\right| : \varsigma \in \{\mathtt{rw}, \mathtt{yw}, \mathtt{cw}, \mathtt{wr}, \mathtt{wy}, \mathtt{wc}\}\right\} \leq \frac{k^{O(1)}\delta}{2^k}, \tag{9.95}$$

$$\max\left\{\left|\zeta(\varsigma) - \zeta_0(\varsigma)\right| : \varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2\right\} \leq \hat{\imath}\delta, \tag{9.96}$$

$$\left|\frac{\zeta(\mathtt{rr})}{\zeta_0(\mathtt{rr})} - 1\right| \leq \ddot{\delta}. \tag{9.97}$$

Then there exist weights $(b_{3/4})^1$, $(b_{3/4})^2$, $x_{3/4}$ such that under the corresponding measure $\mu_{3/4}$ (see (9.99)) the spin ς has marginal $\zeta_{3/4} = \zeta$. Moreover, if we let $\hat{z}_t$ be the normalizing constant such that

$$\hat{q}_t(\sigma) = \frac{1}{\hat{z}_t}(b_t)^1(\sigma^1)(b_t)^2(\sigma^2)\sum_{\varsigma:\varsigma\sim\sigma}x_t(\varsigma)$$

is a probability measure, then the new weights can be chosen such that the error between $\hat{q}_0$ and $\hat{q}_{3/4}$ is very small:

$$\widehat{\mathrm{ERR}}(\hat{q}_0, \hat{q}_{3/4}) \leq k^{O(1)}\begin{pmatrix} \hat{\imath} & 0 \\ 2^k\hat{\imath} & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \delta \\ \ddot{\delta} \end{pmatrix},$$

using the notation of Definition 7.5.

Proof. We begin with some easy observations. By the assumption that $\dot{q} = \Theta(1)$ and $b^i = \Theta(1)$, we have

$$\mu_0(\sigma, \varsigma) = \frac{\mathbf{1}\{\sigma \sim \varsigma\}}{z_0}\Theta(1)x(\varsigma).$$

It then follows from (9.93) and (9.94) that $x(\mathtt{ww})/z_0 \geq 1 - k^{O(1)}/2^k$, while

$$\max\left\{\frac{x(\varsigma)}{z_0} : \varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2 \setminus \{\mathtt{ww}\}\right\} \leq \frac{k^{O(1)}}{2^k},$$

$$\max\left\{\frac{x(\varsigma)}{z_0} : \varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2\right\} \leq O(\hat{\imath}) \leq \frac{O(1)}{2^{k(1+\epsilon)}}.$$

This implies that $\mu_0(\sigma) = \Theta(1)$ for all $\sigma \in \{\mathtt{y}, \mathtt{c}\}^2$, since in this case

$$\mu_0(\sigma) = \frac{1}{z_0}\left\{\Theta(1)x(\mathtt{ww}) + \frac{k^{O(1)}}{2^k}\right\} = \Theta(1).$$

Next, by the assumption that $_{\star}\pi_e$ is nice and μ_0 is judicious, we have

$$\frac{\Theta(1)}{2^k} = {}_{\star}\pi_e(\mathtt{r}) = \mu_0(\sigma^1 = \mathtt{r}) = \frac{\Theta(1)}{z_0}\sum_{\vartheta \in \{\mathtt{r}\}\times\{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}}x(\vartheta) = \frac{\Theta(1)x(\mathtt{rw})}{z_0} + \frac{O(1)}{2^{k(1+\epsilon)}},$$

which shows that $x(\mathtt{rw}) = \Theta(2^{-k})$. It follows from this that

$$\mu_0(\sigma = \vartheta) = \frac{\Theta(1)}{2^k} \quad \text{for all } \vartheta \in \{\mathtt{ry}, \mathtt{rc}, \mathtt{yr}, \mathtt{cr}\}.$$

Lastly, we note that (9.94) directly implies $\mu_0(\sigma = \mathtt{rr}) = \zeta_0(\mathtt{rr}) \leq 2^{-k(1+\epsilon)}$. In summary we have seen that the assumptions of the lemma imply

$$\mu_0(\sigma = \vartheta) = \begin{cases} \Theta(1) & \text{if } \mathtt{r}[\vartheta] = 0, \\ \Theta(2^{-k}) & \text{if } \mathtt{r}[\vartheta] = 1, \\ O(2^{-k(1+\epsilon)}) & \text{if } \mathtt{r}[\vartheta] = 2, \end{cases} \tag{9.98}$$

for $\vartheta \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$.

Part 1. Description of update procedure. We will define a sequence of weights $(b_t)^i$, x_t ending with the desired weights at $t = 3/4$. For each t denote the corresponding measure

$$\mu_t(\sigma, \varsigma) = \frac{\mathbf{1}\{\sigma \sim \varsigma\}}{z_t} \dot{q}(\sigma)(b_t)^1(\sigma^1)(b_t)^2(\sigma^2)x_t(\varsigma), \tag{9.99}$$

with z_t the normalizing constant. Let ζ_t denote the marginal law of ς under μ_t.

(a) *Time $t = 0$ to $t = 1/4$.* First adjust the weight on $\varsigma = \mathtt{yw}$: set

$$\frac{(b_{1/4})^1(\mathtt{y})}{(b_0)^1(\mathtt{y})} = \frac{\mu_0(\sigma^1 = \mathtt{y}) - \zeta(\mathtt{yw})}{\mu_0(\sigma^1 = \mathtt{y}) - \mu_0(\varsigma = \mathtt{yw})}, \tag{9.100}$$

$$\frac{x_{1/4}(\mathtt{yw})}{x_0(\mathtt{yw})} = \frac{\mu_0(\sigma^1 = \mathtt{y}) - \mu_0(\varsigma = \mathtt{yw})}{\mu_0(\sigma^1 = \mathtt{y}) - \zeta(\mathtt{yw})} \frac{\zeta(\mathtt{yw})}{\mu_0(\varsigma = \mathtt{yw})}. \tag{9.101}$$

We call (9.100) and (9.101) the $\mathtt{yw}$-**update**. At the same time also make the updates (9.100) and (9.101) with $\mathtt{c}$ in place of $\mathtt{y}$, to define new weights $(b_{1/4})^1(\mathtt{c})$ and $x_{1/4}(\mathtt{cw})$ — we call this the $\mathtt{cw}$-**update**. Keep all the other weights unchanged, that is, $(b_{1/4})^1(\mathtt{r}) = (b_0)^1(\mathtt{r})$, $(b_{1/4})^2(\tau) = (b_0)^1(\tau)$ for all $\tau \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$, and $x_{1/4}(\varsigma) = x_0(\varsigma)$ for all $\varsigma \notin \{\mathtt{yw}, \mathtt{cw}\}$.

(b) *Time $t = 1/4$ to $t = 1/2$.* Perform the analogous update in the second copy: for the $\mathtt{wy}$-**update**, set

$$\frac{(b_{1/2})^2(\mathtt{y})}{(b_{1/4})^2(\mathtt{y})} = \frac{\mu_{1/4}(\sigma^2 = \mathtt{y}) - \zeta(\mathtt{wy})}{\mu_{1/4}(\sigma^2 = \mathtt{y}) - \mu_{1/4}(\varsigma = \mathtt{wy})}, $$

$$\frac{x_{1/2}(\mathtt{wy})}{x_{1/4}(\mathtt{wy})} = \frac{\mu_{1/4}(\sigma^2 = \mathtt{y}) - \mu_{1/4}(\varsigma = \mathtt{wy})}{\mu_{1/4}(\sigma^2 = \mathtt{y}) - \zeta(\mathtt{wy})} \frac{\zeta(\mathtt{wy})}{\mu_{1/4}(\varsigma = \mathtt{wy})}. $$

Likewise make the $\mathtt{wc}$-**update**, to define new weights $(b_{1/2})^2(\mathtt{c})$ and $x_{1/2}(\mathtt{wc})$. Again, keep all the other weights unchanged.

(c) *Time $t = 1/2$ to $t = 3/4$.* Finally, for all $\varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2$, update

$$\frac{x_{3/4}(\varsigma)}{x_{1/2}(\varsigma)} = \frac{\zeta(\varsigma)}{\zeta_{1/2}(\varsigma)}. $$

Leave the b-weights unchanged, that is, $(b_{3/4})^i(\tau) = (b_{1/2})^i(\tau)$ for $i = 1, 2$ and all $\tau \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$.

We begin with a simple observation about the first update, from $t = 0$ to $t = 1/4$. At time $t = 0$, the non-normalized weight on the event $\{\sigma^1 = \mathtt{y}\}$ is given by $z_0\mu_0(\sigma^1 = \mathtt{y})$. At time $t = 1/4$, the non-normalized weight on the same event is

$$z_{1/4}\mu_{1/4}(\sigma^1 = \mathtt{y}) = z_0 \frac{(b_{1/4})^1(\mathtt{y})}{(b_0)^1(\mathtt{y})} \left(\mu_0(\sigma^1 = \mathtt{y}, \varsigma \neq \mathtt{yw}) + \mu_0(\varsigma = \mathtt{yw}) \frac{x_{1/4}(\mathtt{yw})}{x_0(\mathtt{yw})} \right)$$

$$= z_0 \left\{ \left(\mu_0(\sigma^1 = \mathtt{y}) - \zeta(\mathtt{yw}) \right) + \zeta(\mathtt{yw}) \right\} = z_0\mu_0(\sigma^1 = \mathtt{y}), \tag{9.102}$$

and likewise $z_{1/4}\mu_{1/4}(\sigma^1 = \mathtt{c}) = z_0\mu_0(\sigma^1 = \mathtt{c})$. The weight of the event $\{\sigma^1 = \mathtt{r}\}$ remains unchanged, so altogether we have $z_{1/4} = z_0$. For the same reason $z_{1/2} = z_{1/4}$. We now turn to analyzing the effect of these updates on the σ-marginal.

Part 2. Effect of update (a) on σ-marginal. The first change in the b-weights can be bounded as

$$\left| \frac{(b_{1/4})^1(\mathtt{y})}{(b_0)^1(\mathtt{y})} - 1 \right| \overset{(9.100)}{=} \left| \frac{\mu_0(\sigma^1 = \mathtt{y}) - \zeta(\mathtt{yw})}{\mu_0(\sigma^1 = \mathtt{y}) - \mu_0(\varsigma = \mathtt{yw})} - 1 \right| \overset{(9.95)}{=} \left| \frac{\mu_0(\sigma^1 = \mathtt{y}) - \zeta(\mathtt{yw})}{\mu_0(\sigma^1 = \mathtt{y}) - \zeta(\mathtt{yw}) - k^{O(1)}\delta/2^k} - 1 \right|$$

$$\overset{(9.93)}{=} \left| \frac{\mu_0(\sigma^1 = \mathtt{y}) - k^{O(1)}\delta/2^k}{\mu_0(\sigma^1 = \mathtt{y}) - k^{O(1)}\delta/2^k} - 1 \right| \leq \frac{k^{O(1)}\delta}{2^k} \tag{9.103}$$

— the last estimate above uses the assumption that $_\star\pi_e$ is nice, which implies that $_\star\pi_e(\mathtt{y}) = \Theta(1)$. Now, note it follows from (9.102) that it does not make a difference if the $\mathtt{yw}$-update and $\mathtt{cw}$-update are done simultaneously or in sequence. Suppose for this part of the proof that we make only the $\mathtt{yw}$-update (i.e., only (9.100) and (9.101)), without

making the cw-update. Let $\mu_{1/8}$ denote the measure that results, and define the signed measure $\Gamma \equiv \Gamma_{1/8} \equiv \mu_{1/8} - \mu_0$; we now proceed to bound Γ. The event $\{\varsigma = \mathtt{yw}\}$ is affected by both the b- and x-updates:

$$\left|\Gamma(\varsigma = \mathtt{yw})\right| \equiv \left|\zeta_{1/8}(\mathtt{yw}) - \zeta_0(\mathtt{yw})\right| = \zeta_0(\mathtt{yw})\left|\frac{(b_{1/4})^1(\mathtt{y})}{(b_0)^1(\mathtt{y})}\frac{x_{1/4}(\mathtt{yw})}{x_0(\mathtt{yw})} - 1\right| \overset{(9.101)}{=} \left|\zeta(\mathtt{yw})\frac{(b_{1/4})^1(\mathtt{y})}{(b_0)^1(\mathtt{y})} - \zeta_0(\mathtt{yw})\right|$$

$$\overset{(9.101)}{=} \left|\frac{\mu_0(\sigma^1 = \mathtt{y}) - \mu_0(\varsigma = \mathtt{yw})}{\mu_0(\sigma^1 = \mathtt{y}) - \zeta(\mathtt{yw})}\zeta(\mathtt{yw}) - \zeta_0(\mathtt{yw})\right| \overset{(9.103)}{=} \left|\left(1 + \frac{k^{O(1)}\delta}{2^k}\right)\zeta(\mathtt{yw}) - \zeta_0(\mathtt{yw})\right|$$

$$\overset{(9.93)}{=} \left|\zeta(\mathtt{yw}) - \zeta_0(\mathtt{yw}) + \frac{k^{O(1)}\delta}{4^k}\right| \overset{(9.95)}{\leq} \frac{k^{O(1)}\delta}{2^k} . \tag{9.104}$$

Next let us consider events that are only affected by the b-update: using (9.103) gives

$$\left|\Gamma\left(\begin{matrix}\sigma = \mathtt{yy}, \\ \varsigma = \mathtt{yy}\end{matrix}\right)\right| = \mu_0\left(\begin{matrix}\sigma = \mathtt{yy}, \\ \varsigma = \mathtt{yy}\end{matrix}\right)\cdot\left|\frac{(b_{1/4})^1(\mathtt{y})}{(b_0)^1(\mathtt{y})} - 1\right| \overset{(9.94)}{\leq} \frac{k^{O(1)}\hat{\imath}\delta}{2^k} ,$$

$$\left|\Gamma\left(\begin{matrix}\sigma = \mathtt{yy}, \\ \varsigma = \mathtt{wy}\end{matrix}\right)\right| = \mu_0\left(\begin{matrix}\sigma = \mathtt{yy}, \\ \varsigma = \mathtt{wy}\end{matrix}\right)\cdot\left|\frac{(b_{1/4})^1(\mathtt{y})}{(b_0)^1(\mathtt{y})} - 1\right| \overset{(9.93)}{\leq} \frac{k^{O(1)}\delta}{4^k} .$$

Similarly, for any $\vartheta \in \{\mathtt{y}, \mathtt{w}\} \times \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$ we have, again using (9.103), that

$$\left|\Gamma\left(\begin{matrix}\sigma^1 = \mathtt{y}, \\ \varsigma = \vartheta\end{matrix}\right)\right| = \mu_0\left(\begin{matrix}\sigma^1 = \mathtt{y}, \\ \varsigma = \vartheta\end{matrix}\right)\cdot\left|\frac{(b_{1/4})^1(\mathtt{y})}{(b_0)^1(\mathtt{y})} - 1\right| \overset{(9.93)}{\leq} \frac{k^{O(1)}\delta}{4^k} . \tag{9.105}$$

As a consequence, for any $\vartheta \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2 \setminus \{\mathtt{yw}, \mathtt{ww}\}$ we have

$$\left|\Gamma(\varsigma = \vartheta)\right| = \left|\Gamma\left(\begin{matrix}\sigma^1 = \mathtt{y}, \\ \varsigma = \vartheta\end{matrix}\right) + \overbrace{\Gamma\left(\begin{matrix}\sigma^1 \neq \mathtt{y}, \\ \varsigma = \vartheta\end{matrix}\right)}^{\text{zero}}\right| \overset{(9.105)}{\leq} \frac{k^{O(1)}\delta}{4^k} . \tag{9.106}$$

Similarly we also have the bound

$$\left|\Gamma\left(\begin{matrix}\sigma^1 = \mathtt{y}, \\ \varsigma \in \{\mathtt{ww}, \mathtt{yw}\}\end{matrix}\right)\right| = \left|\overbrace{\Gamma(\sigma^1 = \mathtt{y})}^{\text{zero by (9.102)}} - \Gamma\left(\begin{matrix}\sigma^1 = \mathtt{y}, \\ \varsigma \in \{\mathtt{y}, \mathtt{w}\} \times \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}\end{matrix}\right)\right| \overset{(9.105)}{\leq} \frac{k^{O(1)}\delta}{4^k} . \tag{9.107}$$

Next we note that under both μ_0 and $\mu_{1/4}$, the conditional probability of σ^2 given $\sigma^1 = \mathtt{y}$ and ς is given by

$$\mu_t\left(\sigma^2 = \mathtt{y}\,\middle|\,\begin{matrix}\sigma^1 = \mathtt{y}, \\ \varsigma = \vartheta\end{matrix}\right) = \begin{cases} 1 & \text{if } \vartheta \in \{\mathtt{yy}, \mathtt{wy}\}, \\ \dfrac{\dot{q}(\mathtt{yy})(b_0)^2(\mathtt{y})}{\dot{q}(\mathtt{yy})(b_0)^2(\mathtt{y}) + \dot{q}(\mathtt{yc})(b_0)^2(\mathtt{c})} \leq O(1) & \text{if } \vartheta \in \{\mathtt{yw}, \mathtt{ww}\}, \end{cases}$$

for both $t = 0$ and $t = 1/4$. It follows that

$$\left|\Gamma(\sigma = \mathtt{yy})\right| = \left|\underbrace{\Gamma\left(\begin{matrix}\sigma^1 = \mathtt{y}, \\ \varsigma \in \{\mathtt{yy}, \mathtt{wy}\}\end{matrix}\right)}_{\text{bounded by (9.105)}}\cdot 1 + \underbrace{\Gamma\left(\begin{matrix}\sigma^1 = \mathtt{y}, \\ \varsigma \in \{\mathtt{ww}, \mathtt{yw}\}\end{matrix}\right)}_{\text{bounded by (9.107)}}\cdot O(1)\right| \leq \frac{k^{O(1)}\delta}{4^k} .$$

Similar arguments (details omitted) can be used to bound $\Gamma(\sigma = \vartheta)$ for all $\vartheta \in \{\mathtt{y}, \mathtt{c}\}^2$, so altogether we have

$$\max\left\{\left|\Gamma(\sigma = \vartheta)\right| : \vartheta \in \{\mathtt{y}, \mathtt{c}\}^2\right\} \leq \frac{k^{O(1)}\delta}{4^k} . \tag{9.108}$$

Lastly we note that $\Gamma(\sigma^1 = \mathtt{r}) = 0$, while

$$\left|\Gamma(\sigma^2 = \mathtt{r})\right| = \left|\Gamma\left(\begin{matrix}\sigma^1 = \mathtt{y}, \\ \sigma^2 = \mathtt{r}\end{matrix}\right) + \underbrace{\Gamma\left(\begin{matrix}\sigma^1 \neq \mathtt{y}, \\ \sigma^2 = \mathtt{r}\end{matrix}\right)}_{\text{zero}}\right| \overset{(9.106)}{\leq} \frac{k^{O(1)}\delta}{4^k} . \tag{9.109}$$

This concludes our analysis of the effect of (9.100) and (9.101) alone on the σ-marginal.

Part 3. Effect of updates (a) **and** (b) **on** σ-**marginal.** Recall that $\mu_{1/2}$ is the measure that results after completing updates (a) and (b), and denote the signed measure $\Delta \equiv \mu_{1/2} - \mu_0$. Recall that (9.108) bounds $|(\mu_{1/8} - \mu_0)(\sigma = \vartheta)|$ for all $\vartheta \in \{y, c\}^2$, where from time $t = 0$ to time $t = 1/8$ we perform only the yw-update. From time $t = 1/8$ to time $t = 1/2$ we perform analogously the cw-, wy-, and wc-updates, for which the analogous estimate holds. Therefore

$$\max\left\{\left|\Delta(\sigma = \vartheta)\right| : \vartheta \in \{y, c\}^2\right\} \le \frac{k^{O(1)}\delta}{4^k}.$$

On the other hand, the $\{\sigma = rr\}$ event is completely unaffected by updates (a) and (b), so $\Delta(\sigma = rr) = 0$. It remains to estimate $\Delta(\sigma = \vartheta)$ for $\vartheta \in \{r, y, c\}^2$ with $r[\vartheta] = 1$. To this end we note that for any $\vartheta \in \{r, y, c\}^2$, the event $\{\varsigma = \vartheta\}$ is affected only by the b-updates, so

$$\Delta(\varsigma = \vartheta) \overset{(9.103)}{=} \frac{k^{O(1)}\delta}{2^k}\zeta_0(\vartheta) \overset{(9.94)}{\le} \frac{k^{O(1)}\hat{\imath}\delta}{2^k} \overset{(9.94)}{\le} \frac{k^{O(1)}\delta}{2^{k(2+\epsilon)}}. \tag{9.110}$$

By combining (9.110) with (9.109) we obtain

$$\left|\Delta(\varsigma = rw)\right| = \left|\Delta(\sigma^1 = r) - \Delta(\varsigma \in \{r\} \times \{r, y, c\})\right| \le \left(\frac{1}{2^k} + \hat{\imath}\right)\frac{k^{O(1)}\delta}{2^k} \overset{(9.94)}{\le} \frac{k^{O(1)}\delta}{4^k}. \tag{9.111}$$

The probability of $\sigma = ry$ conditional on $\varsigma = rw$ is given by

$$\mu_{1/2}\left(\sigma = ry \,\middle|\, \varsigma = rw\right) = \frac{\dot{q}(ry)(b_{1/2})^2(y)}{\dot{q}(ry)(b_{1/2})^2(y) + \dot{q}(rc)(b_{1/2})^2(c)} \overset{(9.103)}{=} \mu_0\left(\sigma = ry \,\middle|\, \varsigma = rw\right)\left\{1 + \frac{k^{O(1)}\delta}{2^k}\right\}. \tag{9.112}$$

Meanwhile $\mu_{1/2}(\sigma = ry \,|\, \varsigma = ry) = 1$. Combining the last few estimates gives

$$\mu_{1/2}(\sigma = ry) = \overbrace{\left\{\mu_{1/2}(\varsigma = rw) \cdot \mu_{1/2}\left(\sigma = ry \,\middle|\, \varsigma = rw\right)\right\}}^{\text{estimate by (9.111) and (9.112)}} + \overbrace{\left\{\mu_{1/2}(\varsigma = ry) \cdot 1\right\}}^{\text{estimate by (9.110)}}$$

$$= \left\{\mu_0(\varsigma = rw) + \frac{k^{O(1)}\delta}{4^k}\right\}\mu_0(\sigma = ry \,|\, \varsigma = rw)\left\{1 + \frac{k^{O(1)}\delta}{2^k}\right\} + \left\{\mu_0(\varsigma = ry) + \frac{k^{O(1)}\hat{\imath}\delta}{2^k}\right\}$$

$$\overset{(9.93)}{=} \mu_0(\sigma = ry) + \left(\frac{1}{2^k} + \hat{\imath}\right)\frac{k^{O(1)}\delta}{2^k} \overset{(9.94)}{=} \mu_0(\sigma = ry) + \frac{k^{O(1)}\delta}{4^k}.$$

The analogous estimate holds for the event $\{\sigma = \vartheta\}$ for all $\vartheta \in \{r, y, c\}^2$ with $r[\vartheta] = 1$. Altogether, if we write ω_t for the marginal law of σ under μ_t, then we have

$$\widehat{\mathrm{ERR}}(\omega_0, \omega_{1/2}) \le k^{O(1)}\begin{pmatrix} 4^{-k} \\ 2^{-k} \\ 0 \end{pmatrix}\delta, \tag{9.113}$$

having made use of (9.98).

Part 4. Effect of updates (a) **and** (b) **on** ς-**marginal.** Recall that ζ_t denotes the marginal law of ς under μ_t. Update (a) results in $\zeta_{1/4}(\varsigma) = \zeta(\varsigma)$ for $\varsigma \in \{yw, cw\}$. Update (b) results in $\zeta_{1/2}(\varsigma) = \zeta(\varsigma)$ for $\varsigma \in \{wy, wc\}$, but it need not hold that $\zeta_{1/2}(\varsigma) = \zeta(\varsigma)$ for $\varsigma \in \{yw, cw\}$. However we claim that the discrepancy is very small. Recall that as a

consequence of (9.102) we have $z_0 = z_{1/4} = z_{1/2}$. Therefore

$$\mu_{1/2}(\varsigma = \mathtt{yw}) \overset{(9.99)}{=} \frac{1}{z_{1/2}} x_{1/2}(\mathtt{yw})(b_{1/2})^1(\mathtt{y})\left\{\dot{q}(\mathtt{yy})(b_{1/2})^2(\mathtt{y}) + \dot{q}(\mathtt{yc})(b_{1/2})^2(\mathtt{c})\right\}$$

$$\overset{(9.103)}{=} \frac{1}{z_{1/4}} x_{1/4}(\mathtt{yw})(b_{1/4})^1(\mathtt{y})\left\{\dot{q}(\mathtt{yy})(b_{1/4})^2(\mathtt{y}) + \dot{q}(\mathtt{yc})(b_{1/4})^2(\mathtt{c})\right\}\left\{1 + \frac{k^{O(1)}\delta}{2^k}\right\}$$

$$\overset{(9.99)}{=} \mu_{1/4}(\varsigma = \mathtt{yw})\left\{1 + \frac{k^{O(1)}\delta}{2^k}\right\} \overset{(9.93)}{=} \zeta(\mathtt{yw}) + \frac{k^{O(1)}\delta}{4^k}.$$

The analogous estimate holds for $\varsigma = \mathtt{cw}$. Denote the signed measure $\Upsilon \equiv \zeta_{1/2} - \zeta$; the above can be rewritten as

$$\max\left\{\left|\Upsilon(\mathtt{yw})\right|, \left|\Upsilon(\mathtt{cw})\right|\right\} = \left|(\zeta_{1/2} - \zeta)(\mathtt{yw})\right| \le \frac{k^{O(1)}\delta}{4^k}. \tag{9.114}$$

Next, for all $\vartheta \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$ we have

$$\left|\Upsilon(\vartheta)\right| = \left|(\zeta_{1/2} - \zeta_0)(\vartheta) + (\zeta_0 - \zeta)(\vartheta)\right| \overset{(9.110)}{\le} \frac{k^{O(1)}\hat{\imath}\delta}{2^k} + \left|(\zeta_0 - \zeta)(\vartheta)\right| \overset{(9.96)}{\le} O(\hat{\imath}\delta). \tag{9.115}$$

Recall that $\mu_0(\sigma^i = \mathtt{r}) = {}_\star\pi_e(\mathtt{r}) = \zeta^i(\mathtt{r})$ for $i = 1, 2$. Therefore, using (9.111) and (9.115), we have

$$\Upsilon(\mathtt{rw}) = \left|(\mu_{1/2} - \mu_0)(\sigma^1 = \mathtt{r}) - \sum_{\vartheta \in \{\mathtt{r}\} \times \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}} \Upsilon(\vartheta)\right| \le \left(\frac{1}{4^k} + \hat{\imath}\right)k^{O(1)}\delta \overset{(9.94)}{\le} k^{O(1)}\hat{\imath}\delta. \tag{9.116}$$

Combining (9.114), (9.115), and (9.116) gives

$$\left|\Upsilon(\mathtt{ww})\right| = \left|-\Upsilon\left(\begin{array}{c}(\{\mathtt{w}\} \times \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}) \\ \cup (\{\mathtt{r}, \mathtt{y}, \mathtt{c}\} \times \{\mathtt{w}\})\end{array}\right) - \Upsilon\left(\{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2\right)\right| \le \left(\frac{1}{4^k} + \hat{\imath}\right)k^{O(1)}\delta \overset{(9.94)}{\le} k^{O(1)}\hat{\imath}\delta. \tag{9.117}$$

This concludes our analysis of updates (a) and (b).

Part 5. Effect of update (c). It is clear that we will have $\zeta_{3/4} = \zeta$, so it remains to understand the effect of update (c) on the σ-marginal. Note that since the b-weights remain the same during update (c), the conditional probabilities of σ given ς remain unchanged between times $t = 1/2$ and $t = 3/4$. Thus we have

$$\mu_{3/4}(\sigma = \mathtt{yy}) = \sum_{\vartheta \in \{\mathtt{y}, \mathtt{w}\}^2} \mu_{3/4}(\varsigma = \vartheta)\mu_{1/2}(\sigma = \mathtt{yy} \mid \varsigma = \vartheta) = \mu_{1/2}(\sigma = \mathtt{yy}) + k^{O(1)}\hat{\imath}\delta,$$

where the last estimate uses (9.114), (9.115), and (9.117). The analogous bound holds for $\{\sigma = \vartheta\}$ for each $\vartheta \in \{\mathtt{y}, \mathtt{c}\}^2$. Similarly we also have

$$\mu_{3/4}(\sigma = \mathtt{ry}) = \left\{\mu_{3/4}(\varsigma = \mathtt{rw})\mu_{1/2}\left(\sigma = \mathtt{ry} \mid \varsigma = \mathtt{rw}\right)\right\} + \mu_{3/4}(\varsigma = \mathtt{ry}) \overset{(9.116)}{=} \mu_{3/4}(\sigma = \mathtt{ry}) + k^{O(1)}\hat{\imath}\delta,$$

and the analogous bound holds for all $\sigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$ with $\mathtt{r}[\sigma] = 1$. Lastly, for the case $\sigma = \mathtt{rr}$, we note that

$$\frac{\mu_{3/4}(\sigma = \mathtt{rr})}{\mu_{1/2}(\sigma = \mathtt{rr})} = \frac{\mu_{3/4}(\varsigma = \mathtt{rr})}{\mu_0(\varsigma = \mathtt{rr})} = \frac{\zeta(\mathtt{rr})}{\zeta_0(\mathtt{rr})} \overset{(9.97)}{=} 1 + O(\ddot{\delta}).$$

Combining the above estimates with (9.98) and (9.113) gives

$$\widehat{\mathrm{ERR}}(\omega_0, \omega_{3/4}) \le k^{O(1)}\begin{pmatrix} \hat{\imath} & 0 \\ 2^k\hat{\imath} & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \delta \\ \ddot{\delta} \end{pmatrix}.$$

This can be translated to a similar error bound for the clause-to-variable messages via the relation

$$\hat{q}_t(\sigma) = \frac{\omega_t(\sigma)/\dot{q}(\sigma)}{\sum_{\sigma'} \omega_t(\sigma')/\dot{q}(\sigma')}.$$

The claimed result follows. □

Proof of Proposition 9.13. We shall apply the result of Lemma 9.14. Condition (9.73) of the lemma is satisfied due to Lemma 9.5. Condition (9.74) of the lemma is satisfied due to Lemma 9.7, using the assumption that $\not{D}(v) = \not{L}(v) = 0$. The claimed result immediately follows from the bounds (9.75) of Lemma 9.14.   □

9.6. Non-defective variables neighboring mostly diverse light clauses. In §9.5 we proved Proposition 9.13 under the assumption that v is a non-defective variable with $\not{D}(v) = \not{L}(v) = 0$. In the current subsection we state and prove Proposition 9.16, which holds under the weaker assumption that $\not{D}(v)$ and $\not{L}(v)$ are small but not necessarily zero. (The estimates given by Proposition 9.16 are also weaker than those given by Proposition 9.13.)

Proposition 9.16 (used only in proof of Proposition 9.17). *For any $\epsilon_\infty > 0$ there exists $k_0 < \infty$ large enough (depending only on ϵ_∞) such that the following holds for all $k \geq k_0$. Suppose v is a non-defective variable such that (in the notation of Lemma 9.11)*

$$\max\left\{\not{D}(v), \not{L}(v)\right\} \leq \frac{2^k}{2^{k\epsilon_\infty}} . \tag{9.118}$$

Then v is diverse (Definition 9.3), and furthermore π_v satisfies the estimate

$$\max\left\{\left|\left|\pi_v(x) - \frac{1}{4}\right|\right| : x \in \{-,+\}^2\right\} \leq \frac{1}{100} . \tag{9.119}$$

Proof. Without loss of generality we can assume that ϵ_∞ is much smaller than the other constants $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_\circ$ appearing in this section.

We will prove the result by contradiction, so let v be a variable of type T which satisfies the conditions of the proposition, but for which the estimate (9.119) fails to hold. It will be used repeatedly in the proof that since v is non-defective (Definition 3.10), it must be nice, and all clause types neighboring to v must also be nice (Definition 3.8). As in (9.76) and (9.77) we let $X_e \equiv (\sigma_e, \varsigma_e)$. We also denote $\mathcal{X}_e \equiv (X_e, L_e)$. We consider the law $\mathbf{P}$ of the random variable

$$\mathcal{X} \equiv \mathcal{X}_v \equiv (\mathcal{X}_e)_{e \in \delta v} \equiv (\underline{X}_{\delta v}, \underline{L}_{\delta v}) \equiv \left((X_e)_{e \in \delta v}, (L_e)_{e \in \delta v}\right).$$

The measure $\mathbf{P}$ must maximize entropy subject to the following constraints (cf. (9.78)–(9.80)): for all $e \in \delta v$,

$$\mathbf{P}(L_e = L) = \pi_{\mathscr{D}}(L \mid t_e) \quad \text{for all } L, \tag{9.120}$$

$$\mathbf{P}(\sigma_e^i = \sigma^i \mid L_e = L) = {}_\star\pi_e(\sigma^i) \qquad \text{for all } L \text{ and all } \sigma^i \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}, \tag{9.121}$$

$$\mathbf{P}(\varsigma_e = \varsigma \mid L_e = L) = \zeta_e(\varsigma \mid L) \qquad \text{for all } L \text{ and all } \varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2. \tag{9.122}$$

Via a series of transformations of the measure $\mathbf{P}$ we will construct another distribution $\mathbf{P}^f$ satisfying the same constraints (9.120)–(9.122) but with higher entropy, yielding the required contradiction. The remainder of the proof is outlined as follows:

– In Part 1 we define a measure $\mathbf{P}^a$ satisfying constraints (9.120), (9.121), and (9.123), where (9.123) is a relaxation of (9.122). We show that $\mathbf{P}^a$ has a simple explicit form, and has substantially larger entropy than $\mathbf{P}$ (see (9.132) below).

– In Part 2 we prove probabilistic estimates on various quantities under the measure $\mathbf{P}^a$. Let $\mathcal{X}^a$ denote a sample from $\mathbf{P}^a$; in the rest of the proof we transform $\mathcal{X}^a$ without changing its frozen spin $x = x^a \in \{-,+,\mathtt{f}\}^2$.

– In Part 3 we transform $\mathcal{X}^a$ into $\mathcal{X}^b$, and in Part 4 we transform $\mathcal{X}^b$ into $\mathcal{X}^c$. In these transformations, $(\varsigma_e)^i$ can only change between red and white, so any $(\varsigma_e)^i \in \{\mathtt{y}, \mathtt{c}\}$ remains unchanged. The transformations combined will guarantee (see (9.160) below) that in the configuration $\mathcal{X}^c$, if $x = x^c$ has $x^i = z \in \{-,+\}$, then some $e \in \delta v(z)$ will have $(\sigma^{c,i})_e = \mathtt{r}$, $(\sigma^c)_e \neq \mathtt{rr}$, and $(L^c)_e \in \mathbb{D} \cap \mathbb{L}$. The condition (9.160) will allow more flexibility to change spins from non-white to white in the following step. The law $\mathbf{P}^c$ of $\mathcal{X}^c$ satisfies (9.120), but need not satisfy (9.121) or (9.122), in particular, it can have less than the correct density of red spins (see (9.172) and (9.173)).

– In Part 5 we transform $\mathcal{X}^c$ to $\mathcal{X}^d$ by changing $(\varsigma^{c,i})_e \neq \mathtt{w}$ to $(\varsigma^{d,i})_e = \mathtt{w}$ in some cases — the condition (9.160) from the previous step allows more flexibility to do this without invalidating the configuration. The goal of this step is to create enough white spins which can be altered in the next step to ensure that the constraints (9.121) or (9.122) will be satisfied. At the same time, we cannot change too many spins to white, since this could result in too much decrease in entropy.

– In Part 6 we transform $\mathcal{X}^d$ to $\mathcal{X}^f$, where we use $\mathcal{X}^f$ rather than $\mathcal{X}^e$ to avoid confusion with edge labels $e \in \delta v$. In this final step we change $(\varsigma^{d,i})_e = \mathtt{w}$ to $(\varsigma^{f,i})_e \neq \mathtt{w}$ in some cases, such that the law $\mathbf{P}^f$ of $\mathcal{X}^f$ satisfies the original constraints (9.120)–(9.122) (see (9.182) below). Finally we show that $\mathbf{P}^f$ does not have much smaller entropy than $\mathbf{P}^a$, and as a result must have substantially larger entropy than $\mathbf{P}$, giving the contradiction.

We now turn to the details of the transformation.

Part 1. Product solution $\mathbf{P}^a$ for relaxed constraints. Let $\mathbf{P}^a$ be the probability measure over configurations $\mathcal{X}$ which maximizes entropy subject to the constraints (9.120), (9.121), and (in place of (9.122))

$$\mathbf{P}^a(\varsigma_e^i = \varsigma^i \mid L_e = L) = \zeta_e(\varsigma^i \mid L) \quad \text{for all } L \text{ and all } \varsigma^i \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}. \tag{9.123}$$

Let us note that the law $\mathbf{P}^a$ has a quite simple form: by the method of Lagrange multipliers, it must be expressible as (cf. (9.81))

$$\mathbf{P}^a\left(\underline{\sigma}_{\delta v}, \underline{\varsigma}_{\delta v}, \underline{L}_{\delta v}\right) \cong \psi_v(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \left\{ \mathbf{1}\{\sigma_e \sim \varsigma_e\} \psi_e(L_e) \prod_{i=1,2} \left[(\beta_e)^i((\sigma_e)^i \mid L_e)(\chi_e)^i((\varsigma_e)^i \mid L_e) \right] \right\}, \tag{9.124}$$

where $\psi_v(\underline{\sigma}_{\delta v})$ is parametrized as in (7.55) from Definition 7.24. A valid solution is given by simply setting $\psi_v \equiv 1$, $\psi_e(L) = \pi_{\mathcal{D}}(L \mid t_e)$, and (cf. (9.84) and (9.85))

$$\begin{pmatrix} (\beta_e)^i(\mathtt{r} \mid L) \\ (\beta_e)^i(\mathtt{y} \mid L) \\ (\beta_e)^i(\mathtt{g} \mid L) = (\beta_e)^i(\mathtt{b} \mid L) \equiv (\beta_e)^i(\mathtt{c} \mid L) \end{pmatrix} = \begin{pmatrix} 1 \\ [\star\pi_e(\mathtt{y}) - (\zeta_e)^i(\mathtt{y} \mid L)]/\dot\star{q}_e(\mathtt{y}) \\ [\star\pi_e(\mathtt{c}) - (\zeta_e)^i(\mathtt{c} \mid L)]/\dot\star{q}_e(\mathtt{c}) \end{pmatrix} \overset{(9.82)}{=} \begin{pmatrix} \Theta(1) \\ \Theta(1) \\ \Theta(1) \end{pmatrix},$$

$$\begin{pmatrix} (\chi_e)^i(\mathtt{r} \mid L) \\ (\chi_e)^i(\mathtt{y} \mid L) \\ (\chi_e)^i(\mathtt{c} \mid L) \\ (\chi_e)^i(\mathtt{w} \mid L) \end{pmatrix} = \begin{pmatrix} \star\pi_e(\mathtt{r})/\dot\star{q}_e(\mathtt{r}) \\ (\zeta_e)^i(\mathtt{y} \mid L)/[\star\pi_e(\mathtt{y}) - (\zeta_e)^i(\mathtt{y} \mid L)] \\ (\zeta_e)^i(\mathtt{c} \mid L)/[\star\pi_e(\mathtt{c}) - (\zeta_e)^i(\mathtt{c} \mid L)] \\ 1 \end{pmatrix} \overset{(9.82)}{=} \begin{pmatrix} \Theta(1/2^k) \\ O(k^2/2^k) \\ O(k^2/2^k) \\ 1 \end{pmatrix},$$

where the estimates on $(\zeta_e)^i(\tau \mid L)$ for $\tau \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}$ come from Lemma 9.5. Note that for $e \in \delta v$, $L = L_e$, $i = 1, 2$, and $\sigma^i \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}$ we have

$$(\hat{q}_e)^i(\sigma^i \mid L) \equiv (\beta_e)^i(\sigma^i \mid L) \sum_{\varsigma^i} \mathbf{1}\{\varsigma^i \sim \sigma^i\}(\chi_e)^i(\varsigma^i \mid L) = \frac{\star\pi_e(\sigma^i)}{\dot\star{q}_e(\sigma^i)}, \tag{9.125}$$

where we emphasize that the right-hand side of (9.125) does not depend on L. Indeed,

$$(\hat{q}_e)^i(\mathtt{r} \mid L) = (\beta_e)^i(\mathtt{r} \mid L)(\chi_e)^i(\mathtt{r} \mid L) = \frac{\star\pi_e(\mathtt{r})}{\dot\star{q}_e(\mathtt{r})},$$

which verifies (9.125) for the case $\sigma^i = \mathtt{r}$. Next,

$$(\hat{q}_e)^i(\mathtt{y} \mid L) = (\beta_e)^i(\mathtt{y} \mid L)\left((\chi_e)^i(\mathtt{y} \mid L) + (\chi_e)^i(\mathtt{w} \mid L) \right)$$

$$= \frac{\star\pi_e(\mathtt{y}) - (\zeta_e)^i(\mathtt{y} \mid L)}{\dot\star{q}_e(\mathtt{y})} \cdot \left\{ \frac{(\zeta_e)^i(\mathtt{y} \mid L)}{\star\pi_e(\mathtt{y}) - (\zeta_e)^i(\mathtt{y} \mid L)} + 1 \right\} = \frac{\star\pi_e(\mathtt{y})}{\dot\star{q}_e(\mathtt{y})},$$

which verifies (9.125) for the case $\sigma^i = \mathtt{y}$. A similar calculation gives (9.125) for $\sigma^i \in \{\mathtt{g}, \mathtt{b}\}$. It follows from (9.125), together with the identity $\star\pi_e(\tau)/\dot\star{q}_e(\tau) \cong \star\hat{q}_e(\tau)$, that the marginal of (9.124) over $(\underline{\varsigma}_{\delta v}, \underline{L}_{\delta v})$ is given by

$$\mathbf{P}^a\left(\underline{\sigma}_{\delta v}, \underline{L}_{\delta v}\right) \cong \varphi_v(\underline{\sigma}_{\delta v}) \prod_{e \in \delta v} \left\{ \pi_{\mathcal{D}}(L_e \mid t_e) \prod_{i=1,2} \star\hat{q}_e((\sigma_e)^i) \right\}$$

$$\cong \star\dot{v}_v((\underline{\sigma}_{\delta v})^1)\star\dot{v}_v((\underline{\sigma}_{\delta v})^2) \prod_{e \in \delta v} \pi_{\mathcal{D}}(L_e \mid t_e), \tag{9.126}$$

where $\star\dot{v}_v$ is as in (3.56). In particular, we see that under $\mathbf{P}^a$, the random variables $(\underline{\sigma}_{\delta v})^1$, $(\underline{\sigma}_{\delta v})^2$, and $\underline{L}_e$ (for $e \in \delta v$) are mutually independent. Moreover, conditional on $(\underline{\sigma}, \underline{L})_{\delta v}$, the $(\varsigma_e)^i$ are all independent from one another, with

conditional laws depending only on $(L_e, (\sigma_e)^i)$. If $(\sigma_e)^i = \mathtt{r}$ then $(\varsigma_e)^i = \mathtt{r}$ also. If $(\sigma_e)^i = \mathtt{y}$ then $(\varsigma_e)^i \in \{\mathtt{y}, \mathtt{w}\}$; and the conditional probability of $(\varsigma_e)^i = \mathtt{y}$ is given by

$$\mathbf{P}^a\left((\varsigma_e)^i = \mathtt{y} \,\middle|\, L_e = L, (\sigma_e)^i = \mathtt{y}\right) = \frac{(\chi_e)^i(\mathtt{y} \mid L)}{(\chi_e)^i(\mathtt{y} \mid L) + (\chi_e)^i(\mathtt{w} \mid L)} = \frac{(\zeta_e)^i(\mathtt{y} \mid L)}{{}_\star\pi_e(\mathtt{y})} \le O\left(\frac{k^2}{2^k}\right), \qquad (9.127)$$

where the last inequality follows by Lemma 9.5. Similarly, if $(\sigma_e)^i = \tau \in \{\mathtt{g}, \mathtt{b}\}$ then $(\varsigma_e)^i \in \{\mathtt{c}, \mathtt{w}\}$; and the conditional probability of $(\varsigma_e)^i = \mathtt{c}$ is given by

$$\mathbf{P}^a\left((\varsigma_e)^i = \mathtt{c} \,\middle|\, L_e = L, (\sigma_e)^i = \tau\right) = \frac{(\zeta_e)^i(\mathtt{c} \mid L)}{{}_\star\pi_e(\mathtt{b}) + {}_\star\pi_e(\mathtt{g})} \le O\left(\frac{k^2}{2^k}\right) \qquad (9.128)$$

for $\tau \in \{\mathtt{g}, \mathtt{b}\}$. The independence between the two copies $i = 1, 2$ gives furthermore

$$\mathbf{P}^a\left(\varsigma_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2 \,\middle|\, L_e = L\right) = \prod_{i=1}^{2} \mathbf{P}^a\left((\varsigma_e)^i \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\} \,\middle|\, L_e = L\right) \le O\left(\frac{k^4}{4^k}\right). \qquad (9.129)$$

Note also that if $x = x_v \in \{\mathtt{-}, \mathtt{+}, \mathtt{f}\}^2$ is the frozen spin, then x^1 and x^2 are independent under $\mathbf{P}^a$, and the marginal law of each x^i is determined by the ${}_\star\pi_e$ for $e \in \delta v$. Since we assumed that v is non-defective, it follows that

$$\max\left\{\left|\mathbf{P}^a\left(x = z\right) - \frac{1}{4}\right| : z \in \{\mathtt{-}, \mathtt{+}\}^2\right\} = o_k(1). \qquad (9.130)$$

If $\mathbf{P}'$ is any other probability measure satisfying the same constraints as $\mathbf{P}^a$ (namely (9.120), (9.121), and (9.123)), then it follows by the same derivation as for (8.9) that the relative entropy between $\mathbf{P}'$ and $\mathbf{P}^a$ satisfies

$$\mathcal{H}(\mathbf{P}' \mid \mathbf{P}^a) = \left\langle \mathbf{P}', \ln \frac{\mathbf{P}'}{\mathbf{P}^a} \right\rangle = \mathcal{H}(\mathbf{P}^a) - \mathcal{H}(\mathbf{P}'). \qquad (9.131)$$

Now suppose that $x \in \{\mathtt{-}, \mathtt{+}, \mathtt{f}\}^2$ is the frozen spin corresponding to $X \sim \mathbf{P}$ (meaning that its law fails condition (9.119)), while $x^a \equiv (x^{a,1}, x^{a,2}) \in \{\mathtt{-}, \mathtt{+}, \mathtt{f}\}^2$ is the frozen spin corresponding to $X^a \sim \mathbf{P}^a$ (so that its law satisfies condition (9.130)). Therefore we obtain

$$\mathcal{H}(\mathbf{P} \mid \mathbf{P}^a) \ge \mathcal{D}_{\mathrm{KL}}(x \mid x^a) \ge \frac{1}{10^6}. \qquad (9.132)$$

It follows by combining the last few calculations that

$$\mathcal{H}(\mathbf{P}^a) \stackrel{(9.131)}{=} \mathcal{H}(\mathbf{P}) + \mathcal{H}(\mathbf{P} \mid \mathbf{P}^a) \stackrel{(9.132)}{\ge} \mathcal{H}(\mathbf{P}) + \frac{1}{10^6}. \qquad (9.133)$$

In the remainder of the proof we will make a sequence of transformations, from $\mathbf{P}^a$ to $\mathbf{P}^b$ and so on, until we arrive at a measure $\mathbf{P}^f$ that satisfies the original constraints (9.120)–(9.122). We will show that none of the transformations substantially reduce the entropy. Thus the final measure $\mathbf{P}^f$ will have similar entropy as $\mathbf{P}^a$, and hence will have greater entropy than $\mathbf{P}$ by (9.132). But we assumed $\mathbf{P}$ to be the measure of maximal entropy satisfying constraints (9.120)–(9.122), so this will give the desired contradiction. The measures $\mathbf{P}^a$ through $\mathbf{P}^f$ will all be coupled together, and the frozen spin x will be left unchanged throughout the transformations.

Part 2. Bounds on $\mathbf{P}^a$. Using the assumed bound (9.118) on $\not{D}(v)$ and $\not{L}(v)$, we can select a subset of four distinct edges incident to v,

$$E \equiv \left\{e^{+,1}, e^{+,2}, e^{-,1}, e^{-,2}\right\} \subset \delta v, \qquad (9.134)$$

where $e^{+,i} \in \delta v(\mathtt{+})$, and $e^{-,i} \in \delta v(\mathtt{-})$, such that each $e \in E$ satisfies

$$\sum_L \pi_{\mathscr{D}}(L \mid t_e)\mathbf{1}\{L \in \mathbb{D} \cap \mathbb{L}\} \ge \frac{1}{2}. \qquad (9.135)$$

For a configuration $\mathcal{X}$ as in (9.77), we define the following quantities:

$$\mathfrak{R}^i \equiv \mathfrak{R}^i(\mathcal{X}) \equiv \sum_{e \in \delta v \setminus E} \mathbf{1}\left\{ (\sigma_e)^i = \mathbf{r} \right\},$$

$$\mathfrak{H} \equiv \mathfrak{H}(\mathcal{X}) \equiv \sum_{e \in \delta v \setminus E} \sum_{i=1,2} \mathbf{1}\left\{ (\varsigma_e)^i \in \{\mathbf{r}, \mathbf{y}, \mathbf{c}\} \right\},$$

$$\mathfrak{W} \equiv \mathfrak{W}(\mathcal{X}) \equiv \sum_{e \in \delta v \setminus E} \mathbf{1}\left\{ \varsigma_e = \mathbf{ww}, L_e \in \mathbb{D} \cap \mathbb{L} \right\},$$

$$T \equiv T(\mathcal{X}) \equiv \mathbf{1}\left\{ \mathfrak{W}(\mathcal{X}) \le \left(1 - \frac{1}{k^{10}}\right) |\delta v| \right\}.$$

We now prove some straightforward probabilistic bounds (under the measure $\mathbf{P}^a$) for the above quantities. For $e \in \delta v$ let R_e denote independent Bernoulli random variables with

$$\mathbb{P}(R_e = 1) = \frac{\star \hat{q}_e(\mathbf{r})}{\star \hat{q}_e(\{\mathbf{r}, \mathbf{b}\})} = \Theta\left(\frac{1}{2^k}\right). \tag{9.136}$$

We see from (9.126) that under $\mathbf{P}^a(\cdot \mid x^i = +)$, the total number of edges $e \in \delta v$ with $(\sigma_e)^i = \mathbf{r}$ is equidistributed as the sum of R_e over $e \in \delta v(+)$, conditioned on that sum being strictly positive. Therefore

$$\mathbf{P}^a\left(\sum_{i=1,2} \mathfrak{R}^i \ge \frac{k^3}{2}\right) \le \mathbf{P}^a\left(\sum_{i=1,2} \sum_{e \in \delta v} \mathbf{1}\{(\sigma_e)^i = \mathbf{r}\} \ge \frac{k^3}{2}\right) \le O(1) \max_{z \in \{-,+\}} \left\{ \frac{\mathbb{P}(\sum_{e \in \delta v(z)} R_e \ge k^3/4)}{\mathbb{P}(\sum_{e \in \delta v(z)} R_e \ge 1)} \right\}$$

$$\le \frac{\exp(-\Omega(k^3 \ln k))}{1 - o_k(1)} \le \frac{1}{\exp(\Omega(k^3 \ln k))}, \tag{9.137}$$

where the second-to-last inequality follows by a Chernoff bound. Next recall from (9.127) and (9.128) that under $\mathbf{P}^a$, conditional on $(\sigma_e)^i \in \{\mathbf{y}, \mathbf{g}, \mathbf{b}\}$, the chance for $(\varsigma_e)^i \in \{\mathbf{y}, \mathbf{c}\}$ is $O(k^2/2^k)$. Thus, under $\mathbf{P}^a$, the total number of edges e with $(\varsigma_e)^i \in \{\mathbf{y}, \mathbf{c}\}$ for either $i = 1, 2$ is stochastically dominated by a binomial random variable with $\Theta(k 2^k)$ trials and success probability $\Theta(1/2^k)$. It follows by another Chernoff bound, and by combining with (9.137), that

$$\mathbf{P}^a\left(\mathfrak{H} \ge k^3\right) \le \mathbf{P}^a\left(\sum_{e \in \delta v} \sum_{i=1,2} \mathbf{1}\left\{ (\varsigma_e)^i \in \{\mathbf{r}, \mathbf{y}, \mathbf{c}\} \right\} \ge k^3\right) \le \frac{1}{\exp(\Omega(k^3 \ln k))}. \tag{9.138}$$

Next recall from (9.126) that the L_e for $e \in \delta v$ are mutually independent under $\mathbf{P}^a$. The assumption (9.118) bounds the expected number (under $\mathbf{P}^a$) of edges $e \in \delta v$ with $L_e \notin \mathbb{D} \cap \mathbb{L}$. It follows by Azuma's inequality that

$$\mathbf{P}^a\left(\sum_{e \in \delta v} \mathbf{1}\{L_e \notin \mathbb{D} \cap \mathbb{L}\} \ge \frac{k 2^k}{2^{k \epsilon_\infty}}\right) \le \frac{1}{\exp(\Omega(2^{k(1 - 2\epsilon_\infty)} k))}. \tag{9.139}$$

Combining with (9.138), and recalling the definitions of $\mathfrak{W}$ and T, we obtain

$$\mathbf{P}^a(T = 1) = \mathbf{P}^a\left(\mathfrak{W} \le \left(1 - \frac{1}{k^{10}}\right)|\delta v|\right)$$

$$\le \mathbf{P}^a\left(\max\left\{\mathfrak{H}, \sum_{e \in \delta v} \mathbf{1}\{L_e \notin \mathbb{D} \cap \mathbb{L}\}\right\} \ge \frac{|\delta v|}{3 k^{10}}\right) \le \frac{1}{\exp(\Omega(k^3 \ln k))}. \tag{9.140}$$

Recall from (9.134) the choice of four distinct edges E from δv. For each $e \in E$ define

$$\Gamma(e) \equiv \begin{cases} 1 & \text{if } T(\mathcal{X}^a) = 1, e = e^{+,i}, \text{ and } x^{a,i} = +, \\ 1 & \text{if } T(\mathcal{X}^a) = 1, e = e^{-,i}, \text{ and } x^{a,i} = -, \\ 0 & \text{otherwise.} \end{cases} \tag{9.141}$$

In the first transformation (described below), given $\mathcal{X}^a$ we will construct a modified configuration $\mathcal{X}^b$ in which $(\sigma_e)^i = \mathbf{r}$ and $L_e \in \mathbb{D} \cap \mathbb{L}$ whenever $e = e^{\pm,i}$ with $\Gamma(e) = 1$. (This ensures that a frozen spin $x^i \in \{-, +\}$ has at least

one incident red edge in $\delta v(x^i)$, and gives more flexibility to adjust the colors on the other incident edges.) For $e \in E$ and $z \in \{-, +, \mathfrak{f}\}^2$ define

$$p_z(e) \equiv \mathbf{P}^a(\Gamma(e) = 1, x = z) \le \mathbf{P}^a(T = 1) \overset{(9.140)}{\le} \frac{1}{\exp(\Omega(k^3 \ln k))}. \tag{9.142}$$

Note the definition of $\Gamma(e)$ implies $p_z(e^{+,i}) = 0$ if $z^i \neq +$, and likewise $p_z(e^{-,i}) = 0$ if $z^i \neq -$. For $e \in E$ we also let

$$r_z(e) \equiv \mathbf{P}^a \left(L_e \in \mathbb{D} \cap \mathbb{L}, x = z, (\sigma_e)^i = \mathtt{r}, \sigma_e \neq \mathtt{rr}, \mathfrak{R}^i \ge 2, T = 0 \right).$$

Note having $(\sigma_e)^i = \mathtt{r}$ necessitates $x^i = +$, so $r_z(e^{+,i}) = 0$ if $z^i \neq +$ and $r_z(e^{-,i}) = 0$ if $z^i \neq -$. In all other cases we claim that $r_z(e) \ge \Omega(1)/4^k$. Without loss of generality it suffices to consider the case $e = e^{+,1}$ and $z \in \{+\} \times \{-, +, \mathfrak{f}\}$. In this case, recalling the bound (9.140) gives

$$r_z(e) \ge \mathbf{P}^a \left(L_e \in \mathbb{D} \cap \mathbb{L}, x^2 = z^2, \sigma_e \in \{\mathtt{r}\} \times \{\mathtt{y}, \mathtt{g}, \mathtt{b}\}, \mathfrak{R}^i \ge 2 \right) - \frac{1}{\exp(\Omega(k^3 \ln k))}.$$

Next recall from (9.126) that $(\underline{\sigma}_{\delta v})^1$, $(\underline{\sigma}_{\delta v})^2$, and L_e ($e \in \delta v$) are mutually independent under $\mathbf{P}^a$. It follows by straightforward calculations that for any L we have

$$\mathbf{P}^a \left(\mathfrak{R}^i \ge 2 \,\middle|\, L_e = L, x^2 = z^2, \sigma_e \in \{\mathtt{r}\} \times \{\mathtt{y}, \mathtt{g}, \mathtt{b}\} \right) \ge 1 - o_k(1).$$

Combining with (9.135) gives

$$r_z(e) \ge \frac{1 - o_k(1)}{2} \mathbf{P}^a \left((\sigma_e)^1 = \mathtt{r} \right) \mathbf{P}^a \left(x^2 = z^2, (\sigma_e)^2 \neq \mathtt{r} \right) - \frac{1}{\exp(\Omega(k^3 \ln k))} \ge \frac{\Omega(1)}{4^k}, \tag{9.143}$$

as claimed. It follows by comparing (9.142) with (9.143) that $p_z(e)/r_z(e) = o_k(1)$. For $e \in E$, we define Bernoulli random variables $\Xi(e)$ such that if $r_z(e) = 0$ then $\Xi(e) \equiv 0$, and otherwise if $e = e^{\pm,i}$ with $r_z(e) > 0$ then

$$\mathbf{P}^a \left(\Xi(e) = 1 \,\middle|\, x = z, \mathcal{X}^a = \mathcal{X} \right) = \frac{p_z(e)}{r_z(e)} \mathbf{1} \left\{ L_e \in \mathbb{D} \cap \mathbb{L}, x = z, (\sigma_e)^i = \mathtt{r}, \sigma_e \neq \mathtt{rr}, \mathfrak{R}^i \ge 2, T = 0 \right\} \tag{9.144}$$

(for all possible values of $z, \mathcal{X}$). It then follows from the definition of $r_z(e)$ that for all $z \in \{-, +, \mathfrak{f}\}^2$ we have

$$\mathbf{P}^a \left(\Xi(e) = 1, x = z \right) = p_z(e) = \mathbf{P}^a \left(\Gamma(e) = 1, x = z \right), \tag{9.145}$$

and consequently $\mathbf{P}^a(\Xi(e) = 1) = \mathbf{P}^a(\Gamma(e) = 1)$. Further, by the independence properties of $\mathbf{P}^a$ that we see from the expression (9.126), we have

$$\mathbf{P}^a \left(L_e = L \,\middle|\, \Xi(e) = 1, \left(\sigma_e, (\sigma_{e'}, \varsigma_{e'}, L_{e'})_{e' \in \delta v \setminus e} \right) \right) = \pi_{\mathscr{D}} \left(L \,\middle|\, t_e, L_e \in \mathbb{D} \cap \mathbb{L} \right). \tag{9.146}$$

This concludes our estimates on $\mathbf{P}^a$, and we now turn to the construction of $\mathbf{P}^b$.

Part 3. Construction of $\mathbf{P}^b$. Recall (9.134) that E is a subset of four distinct edges in δv. In the above we defined Bernoulli random variables $\Gamma(e)$ and $\Xi(e)$ for $e \in E$. Recall moreover that $\Gamma(e) = 1$ can only occur on the event $T = T(\mathcal{X}^a) = 1$, while $\Xi(e) = 1$ can only occur on the event $T = T(\mathcal{X}^a) = 0$, so in particular we can never have $\Gamma(e) = \Xi(e) = 1$. In the construction we will essentially "swap" the events $\Gamma(e) = 1$ and $\Xi(e) = 1$. To make this precise, recall that we write $\mathcal{X}_e \equiv (\sigma_e, \varsigma_e, L_e)$. Given $\mathcal{X}^a \sim \mathbf{P}^a$, we let $\mathcal{X}^b$ be defined as follows:

- If $e \in \delta v \setminus E$ then set $(\mathcal{X}^b)_e = (\mathcal{X}^a)_e$.
- If $e \in E$ with $\Gamma(e) = \Xi(e) = 0$ then we also set $(\mathcal{X}^b)_e = (\mathcal{X}^a)_e$.
- If $e \in E$ with $\Gamma(e) = 1$ (hence $\Xi(e) = 0$), and $x_v = z$, then we let $(\mathcal{X}^b)_e$ be sampled from the law

$$\mathbf{P}^b \left(\mathcal{X}_e \in \cdot \,\middle|\, x_v = z, \Gamma(e) = 1 \right) = \mathbf{P}^a \left(\mathcal{X}_e \in \cdot \,\middle|\, x_v = z, \Xi(e) = 1 \right). \tag{9.147}$$

- If $e \in E$ with $\Xi(e) = 1$ (hence $\Gamma(e) = 0$), and $x_v = z$, then we let $(\mathcal{X}^b)_e$ be sampled from the law

$$\mathbf{P}^b \left(\mathcal{X}_e \in \cdot \,\middle|\, x_v = z, \Xi(e) = 1 \right) = \mathbf{P}^a \left(\mathcal{X}_e \in \cdot \,\middle|\, x_v = z, \Gamma(e) = 1 \right). \tag{9.148}$$

This results in a valid configuration $\mathcal{X}^b$ which has the same frozen spin as $\mathcal{X}^a$, that is, $x^a = x^b \in \{\text{-}, \text{+}, \text{f}\}^2$. For the sake of concreteness, we give two examples of the above construction:

(a) *Example on event $T = 1$.* Suppose we have $\mathcal{X}^a \sim \mathbf{P}^a$ with frozen spin $x^a = \text{++}$, such that $T = T(\mathcal{X}^a) = 1$. In this case it follows from the definition (9.141) of $\Gamma(e)$ that $\Gamma(e^{+,1}) = \Gamma(e^{+,2}) = 1$. Since $x^a = \text{++}$, we must have $(\sigma^a)_e \in \{\text{r}, \text{b}\}^2$ for $e \in \{e^{+,1}, e^{+,2}\}$. The resampling step (9.147) results in a modified configuration $\mathcal{X}^b$. It follows from the definition (9.144) of $\Xi(e)$ that in the modified configuration,

$$(\sigma^b)_e = \text{rb for } e = e^{+,1},$$
$$(\sigma^b)_e = \text{br for } e = e^{+,2}.$$

Thus, for $e = e^{\pm,i}$ with $\Gamma(e) = 1$, the procedure (9.147) makes $(\sigma_e)^i$ marginally more likely to be red, but makes σ_e less likely to be rr.

(b) *Example on event $T = 0$.* Suppose we have $\mathcal{X}^a \sim \mathbf{P}^a$ with frozen spin $x^a = \text{++}$, such that $T = T(\mathcal{X}^a) = 0$. Recalling (9.144), suppose that we have $\Xi(e^{+,1}) = 1$ while $\Xi(e^{+,2}) = 0$. In particular, this means for $e = e^{+,1}$ we must have $(\sigma^a)_e = \text{rb}$. The resampling step (9.148) results in a modified configuration $\mathcal{X}^b$. It follows from the definition (9.141) of $\Gamma(e)$ that in the modified configuration we must have $(\sigma^b)_e \in \{\text{r}, \text{b}\}^2$ for $e = e^{+,1}$. Thus, for $e = e^{\pm,i}$ with $\Xi(e) = 1$, the procedure (9.148) makes $(\sigma_e)^i$ marginally less likely to be red, but makes σ_e more likely to be rr.

We let $\mathbb{P}$ denote the joint law of $(\mathcal{X}^a, \mathcal{X}^b)$, and let $\mathbf{P}^b$ denote the marginal law of $\mathcal{X}^b$. Note that $\mathbf{P}^a$ and $\mathbf{P}^b$ have the same edge marginals: for $e \in \delta v \setminus E$ it is clear that $(\mathcal{X}^a)_e$ and $(\mathcal{X}^b)_e$ have the same marginal law, since in the coupling we set $(\mathcal{X}^b)_e = (\mathcal{X}^a)_e$. For $e \in E$, we have

$$\mathbf{P}^b(\mathcal{X}_e \in \cdot \mid x_v = z, \Gamma(e) + \Xi(e) = 1)$$
$$= \mathbf{P}^a(\Gamma(e) = 1 \mid x_v = z)\mathbf{P}^b(\mathcal{X}_e \in \cdot \mid x_v = z, \Gamma(e) = 1) + \mathbf{P}^a(\Xi(e) = 1 \mid x_v = z)\mathbf{P}^b(\mathcal{X}_e \in \cdot \mid x_v = z, \Xi(e) = 1)$$
$$= \mathbf{P}^a(\Xi(e) = 1 \mid x_v = z)\mathbf{P}^a(\mathcal{X}_e \in \cdot \mid x_v = z, \Xi(e) = 1) + \mathbf{P}^a(\Gamma(e) = 1 \mid x_v = z)\mathbf{P}^a(\mathcal{X}_e \in \cdot \mid x_v = z, \Gamma(e) = 1)$$
$$= \mathbf{P}^a(\mathcal{X}_e \in \cdot \mid x_v = z, \Gamma(e) + \Xi(e) = 1), \tag{9.149}$$

where the transition to the third line uses (9.145), (9.147), and (9.148). This verifies that $\mathbf{P}^b(\mathcal{X}_e \in \cdot) = \mathbf{P}^a(\mathcal{X}_e \in \cdot)$ for all $e \in \delta v$, so the measure $\mathbf{P}^b$ again satisfies the constraints (9.120), (9.121), and (9.123). We now compare the entropy of $\mathbf{P}^a$ and $\mathbf{P}^b$. As above, denote $T = T(\mathcal{X}^a)$. Write $E_a \equiv \{e \in E : \Gamma(e) + \Xi(e) = 1\}$. Denote $\mathcal{X} \equiv (\mathcal{Y}, \mathcal{Z})$ where

$$\mathcal{Y} \equiv (\mathcal{X}_e : e \in E_a),$$
$$\mathcal{Z} \equiv (\mathcal{X}_e : e \in \delta v \setminus E_a).$$

If $E_a = \varnothing$ then $\mathcal{Y}$ is the null vector. Likewise denote $\underline{X}_{\delta v} \equiv (Y, Z)$ where $Y \equiv (X_e : e \in E_a)$. We then have

$$\mathcal{H}(\mathcal{X}^b) \geq \mathcal{H}(T, E_a, \mathcal{Z}^b) + \mathcal{H}(\mathcal{Y}^b \mid T, E_a, \mathcal{Z}^b) - \mathcal{H}(T) - \mathcal{H}(E_a)$$
$$= \mathcal{H}(T, E_a, \mathcal{Z}^b) + \mathcal{H}((\underline{L}^b)_{E_a} \mid T, E_a, \mathcal{Z}^b) + \mathcal{H}(Y^b \mid T, E_a, \mathcal{Z}^b, (\underline{L}^b)_{E_a}) - \mathcal{H}(T) - \mathcal{H}(E_a), \tag{9.150}$$
$$\mathcal{H}(\mathcal{X}^a) \leq \mathcal{H}(T, E_a, \mathcal{Z}^a) + \mathcal{H}(\mathcal{Y}^a \mid T, E_a, \mathcal{Z}^a)$$
$$= \mathcal{H}(T, E_a, \mathcal{Z}^a) + \mathcal{H}((\underline{L}^a)_{E_a} \mid T, E_a, \mathcal{Z}^a) + \mathcal{H}(Y^a \mid T, E_a, \mathcal{Z}^a, (\underline{L}^a)_{E_a}). \tag{9.151}$$

Recall that the edges outside E_a are left unchanged by the transformation from $\mathcal{X}^a$ to $\mathcal{X}^b$, so

$$\mathcal{H}\left(T, E_a, \mathcal{Z}^b\right) = \mathcal{H}\left(T, E_a, \mathcal{Z}^a\right). \tag{9.152}$$

Next, (9.140) and (9.145) together imply that

$$\mathbf{P}^a(E_a \neq \varnothing) \leq O\left(\mathbf{P}^a(T = 1)\right) \leq O\left(\frac{1}{\exp(\Omega(k^3 \ln k))}\right),$$

from which it follows that

$$\max\left\{\mathcal{H}(T), \mathcal{H}(E_a), \mathcal{H}(Y^a \mid E_a), \mathcal{H}(Y^b \mid E_a)\right\} = o_k(1). \tag{9.153}$$

We emphasize that in (9.153) we must put Y^a and Y^b rather than $\mathcal{Y}^a$ and $\mathcal{Y}^b$: this is because Y consists of spins X_e which take only $O(1)$ possibilities, while $\mathcal{Y}$ consists of spins $\mathcal{X}_e = (X_e, L_e)$ which take a large number of possibilities

(indeed, growing with the neighborhood radius R of (3.1)) and therefore can have large conditional entropy given E_a. To deal with the clause types L_e we note that

$$\mathcal{H}\left((\underline{L}^a)_{E_a}\,\Big|\,E_a, T, \mathcal{Z}^a\right)$$

$$\overset{(9.146)}{=} \sum_{e \in E} \left\{ \mathbf{P}^a(\Gamma(e) = 1)\mathcal{H}\left(\pi_\mathscr{D}(\cdot \mid t_e)\right) + \mathbf{P}^a(\Xi(e) = 1)\mathcal{H}\left(\pi_\mathscr{D}(\cdot \mid t_e, L_e \in \mathbb{D} \cap \mathbb{L})\right) \right.$$

$$\overset{(9.145)}{=} \sum_{e \in E} \left\{ \mathbf{P}^b(\Gamma(e) = 1)\mathcal{H}\left(\pi_\mathscr{D}(\cdot \mid t_e)\right) + \mathbf{P}^b(\Xi(e) = 1)\mathcal{H}\left(\pi_\mathscr{D}(\cdot \mid t_e, L_e \in \mathbb{D} \cap \mathbb{L})\right) \right\}$$

$$\overset{(9.146)}{=} \mathcal{H}\left((\underline{L}^b)_{E_a}\,\Big|\,E_a, T, \mathcal{Z}^b\right). \tag{9.154}$$

Combining (9.153) and (9.154) gives

$$\left| \mathcal{H}\left(\mathcal{Y}^a\,\Big|\,E_a, T, \mathcal{Z}^a\right) - \mathcal{H}\left(\mathcal{Y}^b\,\Big|\,E_a, T, \mathcal{Z}^b\right) \right|$$

$$= \left| \mathcal{H}\left((\underline{L}^a)_{E_a}\,\Big|\,E_a, T, \mathcal{Z}^a\right) - \mathcal{H}\left((\underline{L}^b)_{E_a}\,\Big|\,E_a, T, \mathcal{Z}^b\right) \right.$$

$$\left. + \mathcal{H}\left(Y^a\,\Big|\,E_a, T, \mathcal{Z}^a, (\underline{L}^a)_{E_a}\right) - \mathcal{H}\left(Y^b\,\Big|\,E_a, T, \mathcal{Z}^b, (\underline{L}^b)_{E_a}\right) \right|$$

$$\overset{(9.154)}{=} \left| \mathcal{H}\left(Y^a\,\Big|\,E_a, T, \mathcal{Z}^a, (\underline{L}^a)_{E_a}\right) - \mathcal{H}\left(Y^b\,\Big|\,E_a, T, \mathcal{Z}^b, (\underline{L}^b)_{E_a}\right) \right| \overset{(9.153)}{=} o_k(1). \tag{9.155}$$

Finally, combining (9.150), (9.151), (9.152), (9.153), and (9.155) gives

$$\mathcal{H}(\mathbf{P}^b) - \mathcal{H}(\mathbf{P}^a) = \mathcal{H}(\mathcal{X}^b) - \mathcal{H}(\mathcal{X}^a) \geq o_k(1). \tag{9.156}$$

This concludes our analysis of the measure $\mathbf{P}^b$.

Part 4. Construction of $\mathbf{P}^c$. Given a configuration $\mathcal{X}$, for $z \in \{-, +\}$ and $i = 1, 2$ let

$$\hat{E}^{z,i} \equiv \hat{E}^{z,i}(\mathcal{X}) \equiv \left\{ e \in \delta v(z) : (\sigma_e)^i = \mathtt{r}, \sigma_e \neq \mathtt{rr}, L_e \in \mathbb{D} \cap \mathbb{L} \right\}. \tag{9.157}$$

Write $\hat{E}^i \equiv \hat{E}^{-,i} \cup \hat{E}^{+,i}$. Note that $\hat{E}^1 \cap \hat{E}^2 = \varnothing$. Recall from (9.134) the definition of E, and for $z \in \{-, +\}$ let

$$\acute{E}^z \equiv \acute{E}^z(\mathcal{X}) \equiv \left\{ e \in \delta v(z) \setminus E : \varsigma_e = \mathtt{ww}, L_e \in \mathbb{D} \cap \mathbb{L} \right\}.$$

(The sets $\acute{E}^\pm(\mathcal{X}^a)$ are disjoint from E, but the sets $\hat{E}^i$ can intersect E.) We construct $\mathcal{X}^c$ as follows:

(I-1) On the event $\{\mathcal{X}^a = \mathcal{X}^b, T(\mathcal{X}^a) = 0, |\hat{E}^1(\mathcal{X}^a)| \geq 2\}$, choose $\hat{e}^1$ uniformly at random from $\hat{E}^1(\mathcal{X}^a)$. For $e = \hat{e}^1$, set $(\mathcal{X}^c)_e = (\mathcal{X}^b)_e$ with probability $1 - 2^{-k\epsilon_\infty/2}$. With the remaining probability $2^{-k\epsilon_\infty/2}$ let $(\mathcal{X}^c)_e$ for $e = \hat{e}^1$ be defined by

$$(\sigma^{c,1}, \varsigma^{c,1})_e = (\mathtt{b}, \mathtt{w}),$$
$$(\sigma^{c,2}, \varsigma^{c,2}, L^c)_e = (\sigma^{b,2}, \varsigma^{b,2}, L^b)_e$$

Let $p^{\mathrm{I},1}(e, L)$ denote the probability that edge e has clause type L and is changed by the above:

$$p^{\mathrm{I},1}(e, L) = \frac{1}{2^{k\epsilon_\infty/2}} \mathbb{E}\left(\frac{\mathbf{1}\{e \in \hat{E}^1(\mathcal{X}^a)\}}{|\hat{E}^1(\mathcal{X}^a)|}; (L^a)_e = L, \mathcal{X}^a = \mathcal{X}^b, T(\mathcal{X}^a) = 0, |\hat{E}^1(\mathcal{X}^a)| \geq 2 \right), \tag{9.158}$$

where $\mathbb{E}$ denotes expectation with respect to the joint law $\mathbb{P}$ of $(\mathcal{X}^a, \mathcal{X}^b)$.

(I-2) On the event $\{\mathcal{X}^a = \mathcal{X}^b, T(\mathcal{X}^a) = 0, |\hat{E}^2(\mathcal{X}^a)| \geq 2\}$, perform the analogue of (I-1) switching the roles of the two copies, and define the corresponding probability $p^{\mathrm{I},2}(e, L)$.

(II-1) On the event $\{x^{a,1} \in \{-,+\}, T(\mathcal{X}^a) = 0, |\hat{E}^1(\mathcal{X}^a)| \leq 1\}$, choose $\acute{e}^1$ uniformly at random from $\acute{E}^{x^1}(\mathcal{X}^a)$. (Note the condition $T(\mathcal{X}^a) = 0$ guarantees that the sets $\acute{E}^{\pm}(\mathcal{X}^a)$ are both large, so it is possible to choose $\acute{e}^1$.) Let $(\mathcal{X}^c)_e$ for $e = \acute{e}^1$ be defined by

$$(\sigma^{c,1}, \varsigma^{c,1})_e = (\mathtt{r}, \mathtt{r}),$$
$$(\sigma^{c,2}, \varsigma^{c,2}, L^c)_e = (\sigma^{b,2}, \varsigma^{b,2}, L^b)_e.$$

Let $p^{\text{II},1}(e, L)$ denote the probability that edge e has clause type L and is changed by the above:

$$p^{\text{II},1}(e, L) = \sum_{z \in \{-,+\}} \mathbb{E}\left(\frac{1\{e \in \acute{E}^z(\mathcal{X}^a)\}}{|\acute{E}^z(\mathcal{X}^a)|}; (L^a)_e = L, x^{a,1} = z, T(\mathcal{X}^a) = 0, |\hat{E}^1(\mathcal{X}^a)| \leq 1 \right), \tag{9.159}$$

where again $\mathbb{E}$ denotes expectation with respect to the joint law $\mathbb{P}$ of $(\mathcal{X}^a, \mathcal{X}^b)$.

(II-2) Lastly, on the event $\{x^{a,2} \in \{-,+\}, T(\mathcal{X}^a) = 0, |\hat{E}^2(\mathcal{X}^a)| \leq 1\}$, perform the analogue of (II-1) switching the roles of the two copies, with the modification that if (II-1) already occurred, then $\acute{e}^2$ is chosen uniformly at random from $\acute{E}(\mathcal{X}^a)^2 \setminus \{\acute{e}^2\}$. Define the corresponding probability $p^{\text{II},2}(e, L)$.

We point out that steps (I-1) and (I-2) can only occur if $\mathcal{X}^a = \mathcal{X}^b$, but steps (II-1) and (II-2) can occur even in the case $\mathcal{X}^a \neq \mathcal{X}^b$. Note also that (I-1) and (II-1) cannot both occur, and likewise (I-2) and (II-2) cannot both occur. To finish the construction, for all edges $e \in \delta v$ which were not chosen as $\hat{e}^i$ or $\acute{e}^i$ in any of the above steps, set $(\mathcal{X}^c)_e = (\mathcal{X}^b)_e$. Note that the frozen spin has not changed, $x = x^a = x^b = x^c \in \{-,+,\mathtt{f}\}^2$. Moreover we have $(L^b)_e = (L^c)_e$ for all $e \in \delta v$. It is straightforward to verify that the resulting $\mathcal{X}^c$ is a valid configuration with the same frozen spin, $x^b = x^c \in \{-,+,\mathtt{f}\}^2$. Moreover, we observe that the construction so far guarantees

$$\hat{E}^{z,i}(\mathcal{X}^c) \neq \varnothing \quad \text{whenever} \quad x^{c,i} = z \in \{-,+\}, \tag{9.160}$$

for $\hat{E}^{z,i}$ as defined by (9.157). Indeed, if $x = x^c$ has in the first copy $x^1 = z \in \{-,+\}$, we have these cases:

– Suppose $T(\mathcal{X}^a) = 1$. It follows from (9.141) that we will have $\Gamma(e) = 1$ for $e = e^{z,1} \in E$. In this case the resampling step (9.147) and the definition (9.144) of $\Xi(e)$ results in a modified configuration $\mathcal{X}^b$ such that e belongs to $\hat{E}^{z,1}(\mathcal{X}^b)$. (If $x^2 = y \in \{-,+\}$ then we may also have $\Gamma(e^{y,2}) = 1$, in which case the set $\hat{E}^{y,2}$ also increases going from $\mathcal{X}^a$ to $\mathcal{X}^b$, but this is not relevant to the current considerations.) Then steps (I-1) and (I-2) do not occur because $\mathcal{X}^a \neq \mathcal{X}^b$, and steps (II-1) and (II-2) do not occur because $T(\mathcal{X}^a) = 1$. It follows that

$$|\hat{E}^{z,1}(\mathcal{X}^c)| = |\hat{E}^{z,1}(\mathcal{X}^b)| \geq |\hat{E}^{z,1}(\mathcal{X}^a)| + 1 \geq 1,$$

since $e = e^{z,1} \in \hat{E}^{z,1}(\mathcal{X}^b) \setminus \hat{E}^{z,1}(\mathcal{X}^a)$.

– Suppose $T(\mathcal{X}^a) = 0, \mathcal{X}^a = \mathcal{X}^b$, and $|\hat{E}^1(\mathcal{X}^a)| \geq 2$. In this case, the only way for an edge e to belong to $\hat{E}^1(\mathcal{X}^a) = \hat{E}^1(\mathcal{X}^b)$ and not to $\hat{E}^1(\mathcal{X}^c)$ is if e is chosen as $\hat{e}^1$ in step (I-1). It follows that

$$|\hat{E}^1(\mathcal{X}^c)| \geq |\hat{E}^1(\mathcal{X}^b)| - 1 = |\hat{E}^1(\mathcal{X}^a)| - 1 \geq 1.$$

– Suppose $T(\mathcal{X}^a) = 0, \mathcal{X}^a \neq \mathcal{X}^b$, and $|\hat{E}^1(\mathcal{X}^a)| \geq 2$. In this case, we see from (9.144) that we may have $\Xi(e) = 1$ for $e = e^{z,1}$. This means $e \in \hat{E}^1(\mathcal{X}^a)$, and the resampling step (9.148) results in $\mathcal{X}^b$ such that $e \notin \hat{E}^1(\mathcal{X}^b)$. (If $x^2 = y \in \{-,+\}$ then we may also have $\Xi(e^{y,2}) = 1$, in which case the set $\hat{E}^{y,2}$ also decreases going from $\mathcal{X}^a$ to $\mathcal{X}^b$, but this is not relevant to the current considerations since the $\hat{E}^i$ are disjoint.) Then steps (I-1) and (I-2) do not occur because $\mathcal{X}^a \neq \mathcal{X}^b$, and steps (II-1) and (II-2) do not occur because $|\hat{E}^1(\mathcal{X}^a)| \geq 2$. It follows that

$$|\hat{E}^1(\mathcal{X}^c)| = |\hat{E}^1(\mathcal{X}^b)| \geq |\hat{E}^1(\mathcal{X}^a)| - 1 \geq 1.$$

– Suppose $T(\mathcal{X}^a) = 0$ and $|\hat{E}^1(\mathcal{X}^a)| \leq 1$. In this case, step (II-1) occurs and results in a configuration $\mathcal{X}^c$ such that $\hat{E}^1(\mathcal{X}^c)$ contains $\acute{e}^1$, and so is nonempty.

This verifies the property (9.160), which will be used in later steps of the proof. We now let $\mathbb{P}$ denote the joint law of $(\mathcal{X}^a, \mathcal{X}^b, \mathcal{X}^c)$, and let $\mathbf{P}^c$ denote the marginal law of $\mathcal{X}^c$.

We first estimate the probability $p^{\text{I},1}(e, L)$ from (9.158). For an upper bound, note in order for $e \in \hat{E}^1(\mathcal{X})$ we must have $(\sigma_e)^1 = \mathtt{r}$, so for all clause types L we have

$$p^{\text{I},1}(e, L) \leq \frac{\mathbf{P}^a(L_e = L, (\sigma_e)^1 = \mathtt{r})}{2^{k\epsilon_\infty/2}} = \frac{\pi_{\mathscr{D}}(L \mid t_e) \cdot {}_\star \pi_e(\mathtt{r})}{2^{k\epsilon_\infty/2}}. \tag{9.161}$$

Next, since the total probability that (I-1) occurs is at most $2^{-k\epsilon_\infty/2}$, we can also bound

$$\sum_{e \in \delta v} \sum_L p^{\mathrm{I},1}(e, L) \leq \frac{1}{2^{k\epsilon_\infty/2}} . \tag{9.162}$$

Next we note that $p^{\mathrm{I},1}(e, L) = 0$ if $L \notin \mathbb{D} \cap \mathbb{L}$. In the case $L \in \mathbb{D} \cap \mathbb{L}$ we will derive a lower bound on $p^{\mathrm{I},1}(e, L)$. Recall R_e from (9.136), and let S_e be an independent copy of R_e. By similar considerations as for (9.137), we have

$$\mathbf{P}^a\left(\ddot{\mathfrak{r}}(X) \equiv \sum_{e \in \delta v} \mathbf{1}\{\sigma_e = \mathrm{rr}\} \geq k^{1/2}\right) \leq O(1) \max_{z \in \{-,+\}} \left\{ \frac{\mathbb{P}(\sum_{e \in \delta v(z)} R_e S_e \geq k^{1/2})}{\mathbb{P}(\sum_{e \in \delta v(z)} R_e \geq 1)^2} \right\} \leq \frac{1}{\exp(\Omega(k^{3/2}))} . \tag{9.163}$$

Next, for $e \in \delta v$, let G_e denote independent Bernoulli random variables with

$$\mathbb{P}(G_e = 1) = \mathbf{P}^a(L_e \in \mathbb{D} \cap \mathbb{L}) = \sum_L \pi_{\mathscr{D}}(L \,|\, t_e)\mathbf{1}\{L \in \mathbb{D} \cap \mathbb{L}\} .$$

For $z \in \{-,+\}$ and any fixed $e \in \delta v(z)$, on the event $\{(\sigma_e)^1 = \mathrm{r}\}$ we have

$$(\mathfrak{r}_e)^1(X) \equiv \sum_{e' \in \delta v(z)\backslash e} \mathbf{1}\left\{(\sigma_{e'})^1 = \mathrm{r}, L_{e'} \in \mathbb{D} \cap \mathbb{L}\right\} \geq \left|\hat{E}^1(X) \setminus \{e\}\right| \geq (\mathfrak{r}_e)^1(X) - \ddot{\mathfrak{r}}(X) , \tag{9.164}$$

where $\ddot{\mathfrak{r}}(X)$ was defined on the left-hand side of (9.163). By similar considerations as for (9.137) and (9.163), we can express

$$\mathbf{P}^a\left((\mathfrak{r}_e)^1 \leq \frac{k}{3} \,\middle|\, (\sigma_e)^1 = \mathrm{r}\right) = \mathbb{P}\left(\sum_{e' \in \delta v(z)\backslash e} R_{e'} G_{e'} \leq \frac{k}{3}\right).$$

To bound this last probability, note that for both $z \in \{-,+\}$ we have

$$\mathrm{Var}\left(\sum_{e' \in \delta v(z)\backslash e} R_{e'} G_{e'}\right) \leq \mathbb{E}\left(\sum_{e' \in \delta v(z)\backslash e} R_{e'} G_{e'}\right) = \sum_{e' \in \delta v(z)\backslash e} \mathbb{E}R_{e'} - \sum_{e' \in \delta v(z)\backslash e} (\mathbb{E}R_{e'})\mathbb{E}(1 - G_{e'})$$

$$= \left(1 - o_k(1)\right)\frac{k \ln 2}{2} - \sum_{e' \in \delta v(z)\backslash e} (\mathbb{E}R_{e'})\mathbb{E}(1 - G_{e'}) \overset{(9.118)}{=} \left(1 - o_k(1)\right)\frac{k \ln 2}{2} .$$

We can use this last estimate together with Bernstein's inequality to bound

$$\mathbf{P}^a\left((\mathfrak{r}_e)^1 \leq \frac{k}{3} \,\middle|\, (\sigma_e)^1 = \mathrm{r}\right) = \mathbb{P}\left(\sum_{e' \in \delta v(z)\backslash e} R_{e'} G_{e'} \leq \frac{k}{3}\right) \leq \frac{1}{\exp(\Omega(k))} . \tag{9.165}$$

On the other hand, by similar arguments as for (9.137), we can bound

$$\mathbf{P}^a\left((\mathfrak{r}_e)^1 \geq k \,\middle|\, (\sigma_e)^1 = \mathrm{r}\right) \leq \mathbb{P}\left(\sum_{e' \in \delta v(z)\backslash e} R_{e'} \geq k\right) \leq \frac{1}{\exp(\Omega(k \ln k))} .$$

Combining the last two bounds gives

$$\mathbf{P}^a\left((\mathfrak{r}_e)^1 \in \left[\frac{k}{3}, k\right] \,\middle|\, (\sigma_e)^1 = \mathrm{r}\right) \geq 1 - \frac{1}{\exp(\Omega(k))} . \tag{9.166}$$

Returning to (9.158), for $L \in \mathbb{D} \cap \mathbb{L}$ we can lower bound

$$p^{\mathrm{I},1}(e, L) \geq \frac{1}{2^{k\epsilon_\infty/2} \cdot 2k}\mathbf{P}^a\left(L_e = L, (\sigma_e)^1 = \mathrm{r}, (\sigma_e)^2 \neq \mathrm{r}, (\mathfrak{r}_e)^1 \in \left[\frac{k}{3}, k\right]\right)$$

$$- \left\{\mathbf{P}^a(T = 1) + \mathbb{P}(X^a \neq X^b) + \mathbf{P}^a(\ddot{\mathfrak{r}} \geq k^{1/2})\right\}.$$

In the above we used the assumption that $L \in \mathbb{D} \cap \mathbb{L}$, so if $(\sigma_e)^1 = \mathrm{r}$ and $(\sigma_e)^2 \neq \mathrm{r}$ then $e \in \hat{E}^1$. We also used that if $(\mathfrak{r}_e)^1 \geq k/3$ and $\ddot{\mathfrak{r}} \leq k^{1/2}$, then certainly (9.164) implies $|\hat{E}^1| \geq 2$, as is required for step (I-1) to occur. Lastly we used that if $(\mathfrak{r}_e)^1 \leq k$, then it follows from (9.164) that $|\hat{E}^1| \leq k + 1$, so edge e has probability at least $1/(k+1) \geq 1/(2k)$ to

be chosen as $\hat{e}^1$. Now recall from the previous construction that $\mathbb{P}(X^a \neq X^b) \leq O(\mathbf{P}^a(T = 1))$, so applying (9.140) and (9.163) gives

$$p^{\mathrm{I},1}(e, L) \geq \frac{1}{2^{k\epsilon_\infty/2} \cdot 2k} \mathbf{P}^a\left(L_e = L, (\sigma_e)^1 = \mathtt{r}, (\sigma_e)^2 \neq \mathtt{r}, (\mathfrak{r}_e)^1 \in \left[\frac{k}{3}, k\right]\right) - \frac{1}{\exp(\Omega(k^{3/2}))} .$$

It then follows (again using the independence properties of $\mathbf{P}^a$) that

$$p^{\mathrm{I},1}(e, L) \overset{(9.126)}{\geq} \frac{\pi_\mathscr{D}(L \mid t_e)}{2^{k\epsilon_\infty/2} \cdot 2k} \mathbf{P}^a\left((\sigma_e)^1 = \mathtt{r}\right) \mathbf{P}^a\left((\sigma_e)^2 \neq \mathtt{r}\right) \mathbf{P}^a\left((\mathfrak{r}_e)^1 \in \left[\frac{k}{3}, k\right] \,\middle|\, (\sigma_e)^1 = \mathtt{r}\right)$$

$$\overset{(9.166)}{\geq} \frac{\pi_\mathscr{D}(L \mid t_e)}{2^{k\epsilon_\infty/2} \cdot 2k} \cdot \frac{1 - o_k(1)}{2^k} , \tag{9.167}$$

as long as $L \in \mathbb{D} \cap \mathbb{L}$. This concludes our analysis of $p^{\mathrm{I},1}(e, L)$. Clearly, (9.162) and (9.167) also hold for $p^{\mathrm{I},2}(e, L)$.

We next prove an upper bound on the probability $p^{\mathrm{II},1}(e, L)$ from (9.159). Recall that the condition $T(X^a) = 0$ implies $|\acute{E}^z(X^a)| = \Theta(k2^k)$ for both $z \in \{-, +\}$. It follows that we can upper bound (9.159) as

$$p^{\mathrm{II},1}(e, L) \leq O\left(\frac{1}{k2^k}\right) \cdot \mathbf{P}^a\left(L_e = L, |\hat{E}^1(X^a)| \leq 1\right) \leq O\left(\frac{\pi_\mathscr{D}(L \mid t_e)}{k2^k}\right) \cdot \mathbf{P}^a\left(|\hat{E}^1(X^a) \setminus \{e\}| \leq 1\right).$$

Recall from (9.164) that $|\hat{E}^1(X^a) \setminus \{e\}| \leq (\mathfrak{r}_e)^1(X^a)$. Then, arguing similarly as for (9.165), for each $z \in \{-, +\}$ we can apply Bernstein's inequality to upper bound

$$\mathbf{P}^a\left((\mathfrak{r}_e)^1 = \sum_{e' \in \delta v(z) \setminus e} \mathbf{1}\left\{(\sigma_{e'})^1 = \mathtt{r}, L_{e'} \in \mathbb{D} \cap \mathbb{L}\right\} \leq 1 \,\middle|\, x^1 = z\right)$$

$$= \frac{\mathbb{P}(\sum_{e' \in \delta v(z) \setminus e} R_{e'} G_{e'} \leq 1)}{\mathbb{P}(\sum_{e' \in \delta v(z)} R_{e'} \geq 1)} \leq \frac{1}{\exp(kc_0)} ,$$

for a positive absolute constant c_0. On the other hand, if $(\mathfrak{r}_e)^1 \geq 2$, in order for $|\hat{E}^1(X^a) \setminus \{e\}| \leq 1$ it must be that $(\sigma^{a,2})_{e'} = \mathtt{r}$ for all but at most one of the edges e' contributing to $(\mathfrak{r}_e)^1$, which has chance $O(2^{-k})$. Then

$$\mathbf{P}^a\left(\hat{E}^1(X^a) \setminus \{e\} = \varnothing\right) \leq \mathbf{P}^a\left((\mathfrak{r}_e)^1 \leq 1\right) + \underbrace{\mathbf{P}^a\left((\mathfrak{r}_e)^1 \geq 2, |\hat{E}^1(X^a) \setminus \{e\}| \leq 1\right)}_{O(1/2^k)} \leq \frac{1}{\exp(kc_0)} .$$

Substituting into the above expression for $p^{\mathrm{II},1}(e, L)$ gives

$$p^{\mathrm{II},1}(e, L) \leq \frac{\pi_\mathscr{D}(L \mid t_e)}{2^{k(1+c_0)}} . \tag{9.168}$$

A similar bound holds for $p^{\mathrm{II},2}(e, L)$, and this concludes our analysis of $p^{\mathrm{II},1}(e, L)$.

We now turn to estimating the entropy of $\mathbf{P}^c$. Let E_b be the (random) subset of edges $e \in \delta v$ for which $(X^b)_e \neq (X^c)_e$, or equivalently $(X^b)_e \neq (X^c)_e$. It follows by combining (9.162) and (9.168) that

$$\mathbb{E}(|E_b|) \leq \sum_{e \in \delta v} \sum_L \sum_{i=1,2} \left(p^{\mathrm{I},i}(e, L) + p^{\mathrm{II},i}(e, L)\right) \leq O\left(\frac{1}{2^{k\epsilon_\infty/2}}\right). \tag{9.169}$$

We write $\mathcal{H}(E_b)$ to denote the entropy of the random variable E_b, and we write $H(p) \equiv -p \ln p - (1-p) \ln(1-p)$ to denote the binary entropy function evaluated at $p \in [0, 1]$. Then

$$\mathcal{H}(E_b) \leq \sum_{e \in \delta v} \mathcal{H}(\mathbf{1}\{e \in E_b\}) \leq |\delta v| H\left(\frac{\mathbb{E}(|E_b|)}{|\delta v|}\right) \overset{(9.169)}{=} o_k(1) , \tag{9.170}$$

where the second step used Jensen's inequality and the concavity of the binary entropy function. Now decompose $\underline{X}_{\delta v} \equiv (U, W)$ where $U \equiv (X_e : e \in E_b)$. Note if E_b is given, then U^b takes at most $O(1)$ distinct values, so it follows

from (9.170) that the conditional entropy of U^b given E_b is small, that is, $\mathcal{H}(U^b \mid E_b) = o_k(1)$. Then, since $\mathcal{X}^b$ can be determined as a function of $(\mathcal{X}^c, E_b, U^b)$, we have

$$\mathcal{H}(\mathcal{X}^b) \leq \mathcal{H}(\mathcal{X}^c) + \Big(\mathcal{H}(E_b) + \mathcal{H}(U^b \mid E_b)\Big) \overset{(9.170)}{\leq} \mathcal{H}(\mathcal{X}^c) + o_k(1). \tag{9.171}$$

This verifies that the entropy of $\mathbf{P}^c$ is not much smaller than that of $\mathbf{P}^b$.

We finally estimate how the transformation from $\mathbf{P}^b$ to $\mathbf{P}^c$ affects the edge marginals. Recall that steps (I-1) and (I-2) can change $(\varsigma_e)^i$ from red to white, and the probability of such a change is lower bounded by (9.167). Meanwhile, steps (II-1) and (II-2) can change $(\varsigma_e)^i$ from white to red, and the probability of such a change is upper bounded by (9.168). It follows that for $L \in \mathbb{D} \cap \mathbb{L}$,

$$\mathbf{P}^c\Big((\sigma_e)^i = \mathtt{r} \,\Big|\, L_e = L\Big) \leq \mathbf{P}^b\Big((\sigma_e)^i = \mathtt{r} \,\Big|\, L_e = L\Big) + \sum_{i=1,2} \frac{p^{\mathrm{II},i}(e, L) - p^{\mathrm{I},i}(e, L)}{\pi_{\mathscr{D}}(L \mid t_e)}$$

$$\leq \mathbf{P}^b\Big((\sigma_e)^i = \mathtt{r} \,\Big|\, L_e = L\Big) + O\Big(\frac{1}{2^{k(1+c_0)}}\Big) - \Omega\Big(\frac{1}{k 2^{k(1+\epsilon_\infty/2)}}\Big) \leq {}_\star\pi_e(\mathtt{r}) - \Omega\Big(\frac{1}{k 2^{k(1+\epsilon_\infty/2)}}\Big), \tag{9.172}$$

where the last step also used our earlier observation (9.149) that $\mathbf{P}^b$ has the same edge marginals as $\mathbf{P}^a$. On the other hand, since the probability of (I-1) or (I-2) is upper bounded by (9.161), we also have

$$\mathbf{P}^c\Big((\sigma_e)^i = \mathtt{r} \,\Big|\, L_e = L\Big) \geq \mathbf{P}^b\Big((\sigma_e)^i = \mathtt{r} \,\Big|\, L_e = L\Big) - \sum_{i=1,2} \frac{p^{\mathrm{I},i}(e, L)}{\pi_{\mathscr{D}}(L \mid t_e)} \geq {}_\star\pi_e(\mathtt{r}) - O\Big(\frac{1}{2^{k(1+\epsilon_\infty/2)}}\Big) \tag{9.173}$$

for all $L \in \mathbb{D} \cap \mathbb{L}$. In the case $(L^b) \notin \mathbb{D} \cap \mathbb{L}$ we always set $(\mathcal{X}^c)_e = (\mathcal{X}^b)_e$, so

$$\mathbf{P}^c\Big((\sigma_e)^i = \mathtt{r} \,\Big|\, L_e = L\Big) = \mathbf{P}^b\Big((\sigma_e)^i = \mathtt{r} \,\Big|\, L_e = L\Big) \tag{9.174}$$

for all $L \notin \mathbb{D} \cap \mathbb{L}$. Finally, since we did not yet change any spin with $(\varsigma_e)^i \in \{\mathtt{y}, \mathtt{c}\}$ or $\varsigma_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2$, we have

$$\mathbf{P}^c((\varsigma_e)^i = \tau \mid L_e = L) = (\zeta_e)^i(\tau \mid L) \text{ for all } \tau \in \{\mathtt{y}, \mathtt{c}\}, \tag{9.175}$$

$$\mathbf{P}^c(\varsigma_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2 \mid L_e = L) = \mathbf{P}^a(\varsigma_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2 \mid L_e = L) \leq O(k^4/4^k) \tag{9.176}$$

for all L, where the final bound uses (9.129).

Part 5. Construction of $\mathbf{P}^d$. We now construct $\mathcal{X}^d$ from $\mathcal{X}^c$. We will leave the clause type unchanged, $(L^d)_e = (L^c)_e$ for all $e \in \delta v$. We change some edges from non-white to white, according to the following rules.

(A) If $L \notin \mathbb{D} \cap \mathbb{L}$ and $(\varsigma^c)_e \neq \mathtt{ww}$, then set $(\varsigma^d)_e = \mathtt{ww}$, and for $i = 1, 2$ set

$$(\sigma^{d,i})_e = \begin{cases} \mathtt{b} & \text{if } (\sigma^{c,i})_e = \mathtt{r}, \\ (\sigma^{c,i})_e & \text{otherwise.} \end{cases}$$

(B) If $L \in \mathbb{D} \cap \mathbb{L}$ then modify the spin on e as follows:
(i) If $(\varsigma^c)_e = \mathtt{rr}$ then set $(\sigma^d, \varsigma^d)_e = (\mathtt{bb}, \mathtt{ww})$.
(ii) If $(\varsigma^c)_e \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}\}^2 \setminus \{\mathtt{rr}\}$ then set $(\sigma^d)_e = (\sigma^c)_e$, and for $i = 1, 2$ set

$$(\varsigma^{d,i})_e = \begin{cases} \mathtt{r} & \text{if } (\sigma^{c,i})_e = \mathtt{r}, \\ \mathtt{w} & \text{otherwise.} \end{cases}$$

(iii) If $(\varsigma^c)_e \in \{\mathtt{y}, \mathtt{c}\} \times \{\mathtt{w}\}$ then set $(\sigma^d)_e = (\sigma^c)_e$. For $\tau = (\varsigma^{c,1})_e \in \{\mathtt{y}, \mathtt{c}\}$ define the quantity

$$\theta = (\theta_e)^1(\tau \mid L) \equiv \min\Big\{1, \frac{2^{-k(1+\epsilon_\infty/2)}}{(\zeta_e)^1(\tau \mid L)}\Big\}.$$

Set $(\varsigma^d)_e = \mathtt{ww}$ with probability θ. With the remaining probability $1 - \theta$ leave $(\varsigma^d)_e$ unchanged. If $(\varsigma^c)_e \in \{\mathtt{w}\} \times \{\mathtt{y}, \mathtt{c}\}$ then perform the analogous procedure with the roles of the two copies switched.
(iv) Lastly, if $(\varsigma^c)_e = \mathtt{ww}$ then set $(X^d)_e = (\sigma^d, \varsigma^d)_e = (\sigma^c, \varsigma^c)_e = (X^c)_e$.

Note that property (9.160) of $\mathcal{X}^c$ ensures that the above transformation results in a valid configuration $\mathcal{X}^d$ with the same frozen spin, $x^c = x^d \in \{\texttt{-}, \texttt{+}, \texttt{f}\}^2$. Since the transformation does not add any red spins, for $L \in \mathbb{D} \cap \mathbb{L}$ we can bound

$$\mathbf{P}^d\left(\varsigma_e = \texttt{rw} \,\middle|\, L_e = L\right) \leq \mathbf{P}^c\left(\varsigma_e = \texttt{rw} \,\middle|\, L_e = L\right) \overset{(9.172)}{\leq} {}_{\star}\pi_e(\texttt{r}) - \Omega\left(\frac{1}{k2^{k(1+\epsilon_\infty/2)}}\right)$$

$$= \zeta_e(\texttt{rw} \,|\, L) + \zeta_e(\{\texttt{r}\} \times \{\texttt{r},\texttt{y},\texttt{c}\} \,|\, L) - \Omega\left(\frac{1}{k2^{k(1+\epsilon_\infty/2)}}\right)$$

$$\overset{\odot}{\leq} \zeta_e(\texttt{rw} \,|\, L) + O\left(\frac{1}{2^{k(1+\epsilon_\circ)}}\right) - \Omega\left(\frac{1}{k2^{k(1+\epsilon_\infty/2)}}\right) \leq \zeta_e(\texttt{rw} \,|\, L).$$

The inequality marked $\odot$ above follows from Lemma 9.7 and the assumption $L \in \mathbb{D} \cap \mathbb{L}$, with $\epsilon_\circ$ as in Definition 9.4, and recalling that we took $\epsilon_\infty \leq \epsilon_\circ$. For a lower bound, we note that for the case $L \in \mathbb{D} \cap \mathbb{L}$, the above transformation removes red spins only in the case $(\sigma^c)_e = \texttt{rr}$. It follows that, for all $L \in \mathbb{D} \cap \mathbb{L}$, we have

$$\mathbf{P}^d\left((\varsigma_e)^1 = \texttt{r} \,\middle|\, L_e = L\right) \geq \mathbf{P}^c\left((\varsigma_e)^1 = \texttt{r} \,\middle|\, L_e = L\right) - \mathbf{P}^c\left(\sigma_e = \texttt{rr} \,\middle|\, L_e = L\right)$$

$$\overset{(9.129)}{\geq} \mathbf{P}^c\left((\varsigma_e)^1 = \texttt{r} \,\middle|\, L_e = L\right) - O\left(\frac{k^4}{4^k}\right) \overset{(9.173)}{\geq} {}_{\star}\pi_e(\texttt{r}) - O\left(\frac{1}{2^{k(1+\epsilon_\infty/2)}}\right) - O\left(\frac{k^4}{4^k}\right)$$

$$\geq \zeta_e(\varsigma = \texttt{rw} \,|\, L) - O\left(\frac{1}{2^{k(1+\epsilon_\infty/2)}}\right).$$

Similar bounds hold if we exchange the roles of the two copies, so we also have

$$\zeta_e(\texttt{wr} \,|\, L) - O\left(\frac{1}{2^{k(1+\epsilon_\infty/2)}}\right) \leq \mathbf{P}^d\left(\varsigma_e = \texttt{wr} \,\middle|\, L_e = L\right) \leq \zeta_e(\texttt{wr} \,|\, L)$$

for all $L \in \mathbb{D} \cap \mathbb{L}$. For $L \in \mathbb{D} \cap \mathbb{L}$ and $\tau \in \{\texttt{y},\texttt{c}\}$ we also have

$$\mathbf{P}^d\left(\varsigma_e = \tau\texttt{w} \,\middle|\, L_e = L\right) \leq \left(1 - (\theta_e)^1(\tau)\right)\mathbf{P}^c\left((\varsigma_e)^1 = \tau \,\middle|\, L_e = L\right) \overset{(9.175)}{=} \left(1 - (\theta_e)^1(\tau)\right)(\zeta_e)^i(\tau \,|\, L)$$

$$= \max\left\{0, (\zeta_e)^i(\tau \,|\, L) - \frac{1}{2^{k(1+\epsilon_\infty/2)}}\right\} = \max\left\{0, \zeta_e(\tau\texttt{w} \,|\, L) + \zeta_e(\tau \times \{\texttt{r},\texttt{y},\texttt{c}\} \,|\, L) - \frac{1}{2^{k(1+\epsilon_\infty/2)}}\right\}$$

$$\leq \zeta_e(\tau\texttt{w} \,|\, L) + O\left(\frac{1}{2^{k(1+\epsilon_\circ)}}\right) - \frac{1}{2^{k(1+\epsilon_\infty/2)}} \leq \zeta_e(\tau\texttt{w} \,|\, L) - \Theta\left(\frac{1}{2^{k(1+\epsilon_\infty/2)}}\right),$$

where the transition to the last line is by another application of Lemma 9.7, again using that $L \in \mathbb{D} \cap \mathbb{L}$. For a lower bound on the same quantity, we have for all $L \in \mathbb{D} \cap \mathbb{L}$ and $\tau \in \{\texttt{y},\texttt{c}\}$ that

$$\mathbf{P}^d\left(\varsigma_e = \tau\texttt{w} \,\middle|\, L_e = L\right) \geq \left(1 - (\theta_e)^1(\tau)\right)\left\{\mathbf{P}^c\left((\varsigma_e)^1 = \tau \,\middle|\, L_e = L\right) - \mathbf{P}^c\left(\varsigma_e \in \{\tau\} \times \{\texttt{r},\texttt{y},\texttt{c}\} \,\middle|\, L_e = L\right)\right\}$$

$$\overset{(9.176)}{\geq} \left(1 - (\theta_e)^1(\tau)\right)\mathbf{P}^c\left((\varsigma_e)^1 = \tau \,\middle|\, L_e = L\right) - O\left(\frac{k^4}{4^k}\right)$$

$$\overset{(9.175)}{=} \max\left\{0, \zeta_e(\tau\texttt{w} \,|\, L) + \zeta_e(\tau \times \{\texttt{r},\texttt{y},\texttt{c}\} \,|\, L) - \frac{1}{2^{k(1+\epsilon_\infty/2)}}\right\} - O\left(\frac{k^4}{4^k}\right) \geq \zeta_e(\tau\texttt{w} \,|\, L) - \Theta\left(\frac{1}{2^{k(1+\epsilon_\infty/2)}}\right).$$

Finally, the above procedure never yields $(\varsigma^d)_e \in \{\texttt{r},\texttt{y},\texttt{c}\}^2$, so for all $L \in \mathbb{D} \cap \mathbb{L}$ and all $\varsigma \in \{\texttt{r},\texttt{y},\texttt{c}\}^2$ we have

$$\zeta_e(\varsigma \,|\, L) \geq 0 = \mathbf{P}^d\left(\varsigma_e = \varsigma \,\middle|\, L_e = L\right) \geq \zeta_e(\varsigma \,|\, L) - \Theta\left(\frac{1}{2^{k(1+\epsilon_\infty/2)}}\right),$$

where the last inequality again uses Lemma 9.7. In summary, in the above we have shown

$$\zeta_e(\varsigma \,|\, L) \geq \mathbf{P}^d\left(\varsigma_e = \varsigma \,\middle|\, L_e = L\right) \geq \zeta_e(\varsigma \,|\, L) - \Theta\left(\frac{1}{2^{k(1+\epsilon_\infty/2)}}\right) \tag{9.177}$$

for all $L \in \mathbb{D} \cap \mathbb{L}$ and all $\varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2 \setminus \{\mathtt{ww}\}$. On the other hand, for $L \notin \mathbb{D} \cap \mathbb{L}$ we have

$$\zeta_e(\varsigma \mid L) \geq 0 = \mathbf{P}^d\left(\varsigma_e = \varsigma \,\Big|\, L_e = L\right) \geq \zeta_e(\varsigma \mid L) - O\left(\frac{k^2}{2^k}\right) \tag{9.178}$$

for all $\varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2 \setminus \{\mathtt{ww}\}$, where the last inequality is by Lemma 9.5.

We now turn to estimating the entropy of $\mathbf{P}^d$. To this end, let E_c denote the subset of edges $e \in \delta v$ for which $(\mathcal{X}^d)_e \neq (\mathcal{X}^c)_e$, or equivalently $(X^d)_e \neq (X^c)_e$. Note that if $(\varsigma^{c,i})_e \neq \mathtt{w}$ and $(\varsigma^{d,i})_e \neq \mathtt{w}$, then $(\varsigma^{c,i})_e = (\varsigma^{d,i})_e$ so $e \notin E_c$. If $e \in E_c$ then the above procedure must transform $(\varsigma^{c,i})_e \neq \mathtt{w}$ into $(\varsigma^{d,i})_e = \mathtt{w}$ for at least one index $i = 1, 2$, so we can bound

$$\mathbb{P}(e \in E_c \mid L_e = L) \leq \sum_{i=1,2} \left\{ \mathbf{P}^c\left((\varsigma_e)^i \neq \mathtt{w} \,\Big|\, L_e = L\right) - \mathbf{P}^d\left((\varsigma_e)^i \neq \mathtt{w} \,\Big|\, L_e = L\right) \right\}$$

$$\leq \sum_{i=1,2} \left\{ (\zeta_e)^i\left(\{\mathtt{r}, \mathtt{y}, \mathtt{c}\} \,\Big|\, L\right) - \mathbf{P}^d\left((\varsigma_e)^i \neq \mathtt{w} \,\Big|\, L\right) \right\} \leq O(1) \sum_{\varsigma \neq \mathtt{ww}} \left| \zeta_e(\varsigma \mid L) - \mathbf{P}^d\left(\varsigma_e = \varsigma \,\Big|\, L_e = L\right) \right|,$$

where the transition to the second line follows by combining (9.172) and (9.175). Applying (9.177) and (9.178) gives

$$\mathbb{E}(|E_c|) \leq O(1) \sum_{e \in \delta v} \sum_{\varsigma \neq \mathtt{ww}} \left\{ \frac{\mathbf{P}^c(L_e \in \mathbb{D} \cap \mathbb{L})}{2^{k(1+\epsilon_\infty/2)}} + \frac{\mathbf{P}^c(L_e \notin \mathbb{D} \cap \mathbb{L})}{2^k/k^2} \right\}$$

$$\leq O\left(\frac{k}{2^{k\epsilon_\infty/2}} + \frac{\mathbb{D}(v) + \mathbb{L}(v)}{2^k/k^2} \right) \overset{(9.118)}{\leq} O\left(\frac{k^2}{2^{k\epsilon_\infty/2}} \right), \tag{9.179}$$

similarly to the estimate (9.169) from the previous transformation. It follows by the same argument as for (9.171) that

$$\mathcal{H}(\mathbf{P}^c) = \mathcal{H}(\mathcal{X}^c) \leq \mathcal{H}(\mathcal{X}^d) + o_k(1) = \mathcal{H}(\mathbf{P}^d) + o_k(1). \tag{9.180}$$

This concludes our analysis of the measure $\mathbf{P}^d$.

Part 6. Construction of $\mathbf{P}^f$. In the final step of the proof we construct $\mathcal{X}^f$ from $\mathcal{X}^d$, where we have chosen the notation $\mathcal{X}^f$ rather than $\mathcal{X}^e$ to avoid confusion with edge labels $e \in \delta v$. For $\sigma \in \{\mathtt{y}, \mathtt{b}\}^2$ we define subsets A^σ of $\{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2 \times \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2$ as follows:

$$A^{\mathtt{bb}} = \{(\mathtt{rr}, \mathtt{rr}), (\mathtt{br}, \mathtt{wr}), (\mathtt{br}, \mathtt{cr}), (\mathtt{rb}, \mathtt{rw}), (\mathtt{rb}, \mathtt{rc}), (\mathtt{bb}, \mathtt{wc}), (\mathtt{bb}, \mathtt{cw}), (\mathtt{bb}, \mathtt{cc})\},$$

$$A^{\mathtt{by}} = \{(\mathtt{ry}, \mathtt{ry}), (\mathtt{by}, \mathtt{wy}), (\mathtt{by}, \mathtt{cy})\},$$

$$A^{\mathtt{yb}} = \{(\mathtt{yr}, \mathtt{yr}), (\mathtt{yb}, \mathtt{yw}), (\mathtt{yb}, \mathtt{yc})\},$$

$$A^{\mathtt{yy}} = \{(\mathtt{yy}, \mathtt{yy})\}.$$

Note that for each $\sigma \in \{\mathtt{y}, \mathtt{b}\}^2$ and every $(\sigma', \varsigma') \in A^\sigma$ we have the compatibility relation $\sigma' \sim \varsigma'$; and the spins σ, σ' correspond to the same frozen spin $x \in \{-, +\}^2$. Moreover, for each $\varsigma \in \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2 \setminus \{\mathtt{ww}\}$ there is exactly one spin $\sigma \equiv \sigma^f(\varsigma) \in \{\mathtt{y}, \mathtt{b}\}^2$ such that $(\sigma', \varsigma) \in A^\sigma$ for some σ'. We then define distributions $\xi^\sigma(\cdot \mid L)$ over pairs $X \equiv (\sigma, \varsigma) \in \{\mathtt{r}, \mathtt{y}, \mathtt{g}, \mathtt{b}\}^2 \times \{\mathtt{r}, \mathtt{y}, \mathtt{c}, \mathtt{w}\}^2$ by setting

$$\xi^\sigma((\sigma, \varsigma) \mid L) = \frac{\zeta_e(\varsigma \mid L) - \mathbf{P}^d(\varsigma_e = \varsigma \mid L_e = L)}{\mathbf{P}^d(\sigma_e = \sigma, \varsigma_e = \mathtt{ww} \mid L_e = L)} \tag{9.181}$$

for all $(\sigma, \varsigma) \in A^\sigma$, and assigning the remaining probability to the event $\{(\sigma, \varsigma) = (\sigma, \mathtt{ww})\}$. To see that these distributions are well-defined, note that on the right-hand side of (9.181), the denominator is $\Theta(1)$, while (9.177) and (9.178) together imply that the numerator is nonnegative and small. It follows that

$$\xi^\sigma((\sigma, \mathtt{ww}) \mid L) = 1 - \sum_{X' \in A^\sigma} \xi^\sigma(X' \mid L) \in [0, 1],$$

so $\xi^\sigma(\cdot \mid L)$ is a valid probability measure over $A^\sigma \cup \{(\sigma, \mathtt{ww})\}$. In the final transformation, given $\mathcal{X}^d$, for each $e \in \delta v$ we set $(L^f)_e = (L^d)_e$. If $(\sigma^d)_e \in \{\mathtt{y}, \mathtt{b}\}^2$ and $(\varsigma^d)_e = \mathtt{ww}$, then we let $(X^f)_e$ be a sample from $\xi^\sigma(\cdot \mid (L^d)_e)$. In all other

cases we let $(X^f)_e = (X^d)_e$. It is straightforward to verify that this results in a valid configuration $\mathcal{X}^f$ with the same frozen spin, $x^f = x^d \in \{\text{-},\text{+},\text{f}\}^2$. For each $\varsigma \in \{\text{r},\text{y},\text{c},\text{w}\}^2 \setminus \{\text{ww}\}$,

$$\mathbf{P}^f\left(\varsigma_e = \varsigma \,\middle|\, L_e = L\right) = \mathbf{P}^d\left(\varsigma_e = \varsigma \,\middle|\, L_e = L\right) + \sum_{\sigma : \sigma = \sigma^f(\varsigma)} \mathbf{P}^d\left(\sigma_e = \sigma, \varsigma_e = \text{ww} \,\middle|\, L_e = L\right)\xi^\sigma((\sigma,\varsigma)\,|\,L)$$

$$\overset{(9.181)}{=} \mathbf{P}^d\left(\varsigma_e = \varsigma \,\middle|\, L_e = L\right) + \left\{ \zeta_e(\varsigma\,|\,L) - \mathbf{P}^d\left(\varsigma_e = \varsigma \,\middle|\, L_e = L\right) \right\} = \zeta_e(\varsigma\,|\,L)\,, \qquad (9.182)$$

so $\mathbf{P}^f$ satisfies condition (9.122). On the other hand, since $x^a = x^f \in \{\text{-},\text{+},\text{f}\}^2$, we have

$$\mathbf{P}^f\left(x^i = z \,\middle|\, L_e = L\right) = \mathbf{P}^a\left(x^i = z \,\middle|\, L_e = L\right) \overset{(9.126)}{=} \begin{cases} \star\pi_e(\{\text{r},\text{b}\}) & \text{if } z \in \{\text{-},\text{+}\} \text{ and } e \in \delta v(z), \\ \star\pi_e(\text{y}) & \text{if } z \in \{\text{-},\text{+}\} \text{ and } e \in \delta v(\text{-}z), \\ \star\pi_e(\text{g}) & \text{if } z = \text{f}. \end{cases}$$

Moreover, since $(\varsigma_e)^i = \text{r}$ if and only if $(\sigma_e)^i = \text{r}$, it follows that

$$\mathbf{P}^f\left((\sigma_e)^i = \text{r} \,\middle|\, L_e = L\right) \overset{(9.182)}{=} (\zeta_e)^i(\text{r}\,|\,L) = \star\pi_e(\text{r})\,.$$

It follows from the last two displays combined that $\mathbf{P}^f$ satisfies condition (9.121). Lastly, since $(L^a)_e = (L^f)_e$ for all e, and $\mathbf{P}^a$ satisfied condition (9.120) by construction, we can conclude that $\mathbf{P}^f$ also satisfies condition (9.120).

We now estimate the entropy of $\mathbf{P}^f$. Let E_d denote the subset of edges $e \in \delta v$ for which $(\mathcal{X}^d)_e \neq (\mathcal{X}^f)_d$. For any edge in $e \in E_d$ we must have $(\varsigma^d)_e = \text{ww}$ and $(\varsigma^f)_e \neq \text{ww}$. As a result we can bound

$$\mathbb{E}(|E_d|) \leq \sum_{e \in \delta v} \left(\mathbf{P}^f(\varsigma_e \neq \text{ww}) - \mathbf{P}^d(\varsigma_e \neq \text{ww}) \right) = \sum_{e \in \delta v} \sum_{\varsigma \neq \text{ww}} \left(\zeta_e(\varsigma) - \mathbf{P}^d(\varsigma_e = \varsigma) \right) \leq O\left(\frac{k^2}{2^{k\epsilon_\infty/2}} \right),$$

using the same reasoning as for (9.179). It then follows by the argument of (9.171) that

$$\mathcal{H}(\mathbf{P}^d) \leq \mathcal{H}(\mathbf{P}^f) + o_k(1)\,. \qquad (9.183)$$

Combining equations (9.156), (9.171), (9.180),(9.183), and (9.133) gives

$$\mathcal{H}(\mathbf{P}^f) \geq \mathcal{H}(\mathbf{P}^a) - o_k(1) \overset{(9.133)}{\geq} \mathcal{H}(\mathbf{P}) + \frac{1}{10^6} - o_k(1)\,.$$

This contradicts the assumption that $\mathbf{P}$ is the maximal-entropy measure satisfying constraints (9.120)–(9.122), thereby concluding the proof of the proposition. $\qquad\square$

9.7. **Conclusion of *a priori* estimates.** In this subsection we complete the proof of Proposition 8.4. We first prove the expansion result, Lemma 8.3, which was stated in §8.1.

Proof of Lemma 8.3. We first show that in the original k-SAT graph $\mathcal{G}' = (V', F', E') \sim \mathbf{P} = \mathbb{P}_{n,m}$, **every** subset of variables $S \subseteq V'$ satisfies the bound

$$|F_\bullet(S)| \leq \begin{cases} n2^{29k/30}/k & \text{if } 0 \leq |S|/n \leq 4/5, \\ |S| & \text{if } c_1/2 \leq |S|/n \leq 1/4, \end{cases} \qquad (9.184)$$

where c_1 is as in Proposition 3.23. (Note that (8.5) refers to the number $|V|$ of variables in the processed graph, while in (9.184) and throughout this proof we use $n \equiv |V'| \geq |V|$ to refer to the number of variables in the original graph. Thus in (9.184) we crudely divided the first bound by a factor k, which we will use below to account for the discrepancy in the number of variables in $\mathcal{G}'$ versus in $\mathcal{G}$.) Towards the proof of (9.184), recall that $F_\bullet(S)$ denotes the subset of clauses $a \in F$ with $|S \cap \partial a| \geq 9k/10$. Note that for any fixed subset $S \subseteq V'$ of size $|S| = ns$, the probability that a particular clause $a \in F'$ belongs to $F_\bullet(S)$ is given by

$$p_s = \mathbb{P}\left(\text{Bin}(k,s) \geq \frac{9k}{10}\right) \leq \exp\left\{ -k\mathcal{H}\left(\frac{9}{10}\,\middle|\,s\right) \right\} \leq \begin{cases} 2^{-k/20} & \text{for all } 0 \leq s \leq 4/5, \\ s^{2k/3} & \text{for all } 0 \leq s \leq 1/4, \end{cases}$$

where the first inequality is from the Chernoff bound and the second is by direct comparison. If $\mathbb{E}$ denotes expectation over the law $\mathbb{P}$ of $\mathscr{G}'$, then another application of the Chernoff bound gives

$$E_n(s,\gamma) \equiv \mathbb{E}\left|\left\{ S : |S| = ns \text{ and } |F_\bullet(S)| \geq n\gamma \right\}\right| \leq \binom{n}{ns}\mathbb{P}\left(\text{Bin}(m,p_s) \geq n\gamma\right)$$

$$\leq \exp\left\{ n\left[\mathcal{H}(s) - \alpha\mathcal{H}\left(\frac{\gamma}{\alpha}\,\Big|\,p_s\right)\right]\right\}.$$

For $0 \leq s \leq 4/5$, taking $\gamma = 2^{29k/30}/k$ gives

$$\alpha\mathcal{H}\left(\frac{\gamma}{\alpha}\,\Big|\,p_s\right) = \gamma\ln\frac{\gamma/\alpha}{p_s} + (\alpha - \gamma)\ln\frac{1 - \gamma/\alpha}{1 - p_s} \geq \Omega(\gamma k) - O(\gamma + \alpha p_s) \geq \Omega(\gamma k)\,,$$

so we see that $E_n(s,\gamma)$ is exponentially small in n. For $c_1/2 \leq s \leq 1/4$, taking $\gamma = s$ gives

$$\alpha\mathcal{H}\left(\frac{\gamma}{\alpha}\,\Big|\,p_s\right) \geq s\ln\frac{s/\alpha}{s^{2k/3}} - O(s + \alpha p_s) \geq \Omega\left(ks\ln\frac{1}{s}\right) \geq \Omega\left(k\mathcal{H}(s)\right),$$

so again $E_n(s,\gamma)$ is exponentially small in n. It follows that for the original k-SAT instance, we have

$$\mathbb{P}\left(\text{any subset } S' \subseteq V' \text{ violates (9.184)}\right) \leq \exp\left\{ -n\Omega\left(k\Omega\left(\mathcal{H}\left(\frac{c_1}{2}\right)\right)\right)\right\} \leq \exp\{-nc'\}\,, \tag{9.185}$$

where c' depends only on k and c_1 (and c_1 in turn depends only on k, α, R). If in the processed k-SAT instance $\mathscr{G} = (V, F, E)$ we have any subset $S \subseteq V$ violating (8.5), then in the original instance $\mathscr{G}' = (V', F', E')$ it must be the case that either some subset $S' \subseteq V'$ violates (9.184), or $|V| \leq |V'|/k$. It follows by combining with (9.185) and Proposition 3.22 that

$$\mathbb{E}\left[\mathbb{P}_{\mathscr{D}}\left(\text{any subset } S \subseteq V \text{ violates (8.5)}\right)\right]$$

$$\leq \mathbb{P}\left(|V| \leq \frac{|V'|}{k}\right) + \mathbb{P}\left(\text{any subset } S' \subseteq V' \text{ violates (9.184)}\right) \leq o_n(1)\,,$$

and the claimed result follows by Markov's inequality. $\qquad\qquad\square$

Proposition 9.17 (used only in proof of Proposition 8.4). *On the event that $\mathscr{D}$ expands on type-subsets (Definition 8.2), all non-defective variables are diverse (Definition 9.3); and all strongly non-defective clauses are light (Definition 9.4).*

Proof. Since we have restricted to $\omega \in I_0$ (as has been the case throughout this section), the total number of variables v with frozen spin $x_v \in \{\texttt{++}, \texttt{--}\}$ is close to $n/2$, in any pair frozen configuration that is consistent with ω. Let S' denote the set of all non-diverse variables: then

$$|S'| = \sum_{v \in V}\mathbf{1}\left\{\pi_v(\{\texttt{++}, \texttt{--}\}) > \frac{3}{4}\right\} \leq \frac{4}{3}\sum_{v \in V}\pi_v(\{\texttt{++}, \texttt{--}\}) \leq \frac{3n}{4} \tag{9.186}$$

(where the last inequality uses the above observation on the number of variables with $x_v \in \{\texttt{++}, \texttt{--}\}$).

Let $\not{D} \subseteq F$ denote the set of non-diverse clauses; it follows from Definitions 9.3 and 8.2 that $\not{D} \subseteq F_\bullet(S')$. Now let $S \subseteq S'$ denote the subset of non-diverse variables that are non-defective (Definition 3.10), so we see from the above that $|S| \leq |S'| \leq 3n/4$. Recall from Remark 6.5 that a clause is termed strongly non-defective if it neighbors only non-defective variables. Let $N \subseteq \not{D}$ denote the set of non-diverse clauses that are strongly non-defective; it follows that $N \subseteq F_\bullet(S)$. Then, on the event that $\mathscr{D}$ expands on type-subsets, it follows from Definition 8.2 that

$$|N| \leq \begin{cases} n2^{29k/30} & \text{if } |S|/n \leq 4/5\,. \\ |S| & \text{if } |S|/n \leq 1/16\,. \end{cases} \tag{9.187}$$

Next let $\not{L} \subseteq F$ denote the set of all heavy clauses, and let $H \subseteq \not{L}$ denote the subset of heavy clauses that are strongly non-defective. By Proposition 9.16, if v is non-defective but not diverse, then

$$\frac{2^k}{2^{k\epsilon_\infty}} < \max\left\{\not{D}(v), \not{L}(v)\right\} \leq \not{D}(v) + \not{L}(v)\,,$$

where we can take ϵ_∞ to be much smaller than all the other constants ϵ in this section. Summing the above over all $v \in S$, and recalling the definitions of $\mathbb{D}(v)$ and $\mathbb{L}(v)$ from Definitions 9.3 and 9.4, we obtain

$$\frac{2^k|S|}{2^{k\epsilon_\infty}} < \sum_{v \in S} \sum_{e \in \delta v} \sum_{L} \pi_\mathscr{D}(L \mid t_e)\left(\mathbf{1}\{L \in \mathbb{D} \cup \mathbb{L}\}\right) \leq \left(|N| + |H|\right)k + |S|, \tag{9.188}$$

where the right-hand side of (9.188) is obtained as follows: the contribution from the case of strongly non-defective clause types L is bounded by the term $(|N| + |H|)k$, where the factor k accounts for the fact that one clause can neighbor up to k variables. As for the case where L fails to be strongly non-defective, we recall from Remark 6.5 that each $v \in S$ neighbors at most one such clause, so the contribution from this case is bounded by the term $|S|$.

Now note that (9.187) bounds $|N|$ in terms of $|S|$, while (9.188) bounds $|S|$ in terms of $|N|$ and $|H|$. We will close the loop by bounding $|H|$ in terms of $|N|$. In fact we will bound $|E_H|$ where E_H is a subset of edges defined as follows: for any edge $e = (av)$ where $L_a = L$ and $j = j(t_e)$, write $\omega_e \equiv \omega_{L,j}$. Then let

$$E_H \equiv \left\{ e = (av) : a \text{ is a strongly non-defective clause and } \omega_e(\mathrm{rr}) \geq \frac{1}{2^{k(1+\epsilon_\circ)}} \right\}.$$

Note that a clause belongs to H if and only if it lies incident to E_H, so $|H| \leq |E_H|$. We will bound

$$|E_H| \leq \sum_{i=0}^{2} |E_H(U_i)|$$

where $E_H(U_i)$ denotes the subset of edges in E_H that are incident to U_i, and (U_0, U_1, U_2) gives a partition of the set of all non-defective variables:

$$U_0 \equiv \{\text{non-defective variables } v : \mathbb{D}(v) \leq 2^{k(1-\epsilon_\infty)}, \mathbb{L}(v) \leq 2^{k(1-\epsilon_\infty)}\},$$
$$U_1 \equiv \{\text{non-defective variables } v : \mathbb{D}(v) \leq 2^{k(1-\epsilon_\infty)}, \mathbb{L}(v) > 2^{k(1-\epsilon_\infty)}\},$$
$$U_2 \equiv \{\text{non-defective variables } v : \mathbb{D}(v) > 2^{k(1-\epsilon_\infty)}\}.$$

It follows from Proposition 9.16 that if $v \in U_0$ then v is diverse, and moreover that $\pi_v(x)$ is close to $1/4$ for each $x \in \{-,+\}^2$. Thus each $v \in U_0$ satisfies the conditions of Lemma 9.11 — in particular, Lemma 9.11 has a condition $\pi_v(++) \geq 2^{-k/16}$ which is certainly guaranteed if $\pi_v(++)$ is close to $1/4$. It follows by Lemma 9.11 that an edge $e = (av)$ with $v \in U_0$ cannot belong to E_H unless the clause a is non-diverse, which means $a \in N$. Therefore

$$|E_H(U_0)| \leq \left|\left\{(av) \in E_H : a \in N\right\}\right| \leq |N|k.$$

Next, similarly to (9.188) we have the bound

$$\frac{2^k|U_1|}{2^{k\epsilon_\infty}} < \sum_{v \in U_1} \mathbb{L}(v) \leq |H|k + |U_1|. \tag{9.189}$$

For $v \in U_1$, applying Corollary 9.12 gives

$$\sum_{e \in \delta v} \sum_{L \in \mathbb{D}} \pi(L|t_e) \mathbf{1}\left\{\omega_{L,j}(\mathrm{rr}) \geq \frac{1}{2^{k(1+\epsilon_\circ)}}\right\} \leq \frac{k^2 2^{k(1+\epsilon_\circ)}}{2^{k\epsilon_4}} \leq \frac{2^k}{2^{k\epsilon_\circ}}, \tag{9.190}$$

where the last bound holds because we took $\epsilon_\circ$ to be much smaller than ϵ_4. Combining with the preceding upper bound on $|U_1|$ gives

$$|E_H(U_1)| \leq \left|\left\{(av) : a \in N\right\}\right| + \left|\left\{(av) \in E_H : v \in U_1, a \in \mathbb{D}\right\}\right|$$
$$\overset{(9.190)}{\leq} |N|k + |U_1|\frac{2^k}{2^{k\epsilon_\circ}} \overset{(9.189)}{\leq} |N|k + |H|\frac{2^{k\epsilon_\infty}}{2^{k\epsilon_\circ}}k\left(1 + o_k(1)\right).$$

Finally, similarly to (9.188) and (9.189) we have

$$\frac{2^k|U_2|}{2^{k\epsilon_\infty}} < \sum_{v \in U_2} \mathbb{D}(v) \leq |N|k + |U_2|, \tag{9.191}$$

from which it follows that

$$|E_H(U_2)| \leq \left|\left\{(av) : v \in U_2\right\}\right| \leq |U_2| 2^k k^2 \overset{(9.191)}{\leq} |N| 2^{k\epsilon_\infty} k^3 \Big(1 + o_k(1)\Big).$$

Combining the above bounds gives

$$|H| \leq \sum_{i=0}^{2} |E_H(U_i)| \leq |N| 2^{k\epsilon_\infty} k^4 + |H| \frac{2^{k\epsilon_\infty}}{2^{k\epsilon_\circ}} k^2 \leq |N| 2^{k\epsilon_\infty} k^4 + \frac{|H|}{2^{k\epsilon_\circ/2}},$$

where the last bound holds since we took ϵ_∞ to be much smaller than $\epsilon_\circ$, as discussed above. Rearranging gives $|H| \leq O(1)|N|2^{k\epsilon_\infty}k^4$, and combining with (9.188) gives

$$|S| \overset{(9.188)}{\leq} O(1)\Big(|N| + |H|\Big)\frac{2^{k\epsilon_\infty}}{2^k}k \leq \frac{|N|k^{O(1)}}{2^{k(1-2\epsilon_\infty)}}.$$

Together with the expansion bound (9.187) (which can be applied, since we saw earlier that $|S| \leq 3n/4$), we see that the only possibility is for $|S| = |N| = 0$. This implies $|H| = 0$ and concludes the proof. □

Proof of Proposition 8.4. If $\mathscr{D}$ expands on type-subsets, then Proposition 9.17 yields that all non-defective variables are diverse, and all strongly non-defective clauses are light. Moreover, since a non-defective clause can neighbor at most one defective variable, it follows that every non-defective clause is diverse in the sense of Definition 9.3. Thus, if a variable v is strongly non-defective, then all the clauses around it must be both diverse and light, meaning $\not\!\!\mathscr{D}(v) = \not\!\!\mathscr{L}(v) = 0$. Proposition 9.13 then gives the desired estimate (9.72) (by taking $\zeta_* = \epsilon_\circ$) for $\omega_{L,j}$ whenever $L(j)$ is a strongly non-defective edge type. □

10. Monotonicity of the 1-rsb free energy

In this section we prove Proposition 1.2, which says that the 1-rsb free energy is strictly decreasing with respect to α in the interval (1.7). Throughout this section we let $\alpha_{\mathrm{lbd}} \leq \dot{\alpha} \leq \ddot{\alpha} \leq \alpha_{\mathrm{ubd}}$.

Definition 10.1 (monotone coupling of trees). Recall from Definition 4.3 that $\mathrm{PGW} \equiv \mathrm{PGW}(\alpha)$ denotes the law of the variable-to-clause tree $\mathscr{T}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$. We now define a pair of such trees as follows. First let $\ddot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}} \sim \mathrm{PGW}(\ddot{\alpha})$. Then delete each clause in $\ddot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}} \sim \mathrm{PGW}(\ddot{\alpha}) \setminus \{a_{\mathrm{RT}}\}$ independently with probability $1 - \dot{\alpha}/\ddot{\alpha}$, and let $\dot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$ be the connected component containing v_{RT}, so that $\dot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}} \sim \mathrm{PGW}(\dot{\alpha})$. Let $\mathrm{PGW}(\dot{\alpha}, \ddot{\alpha})$ denote the law of the pair $(\dot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}, \ddot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}})$. This is the **monotone coupling** between the measures $\mathrm{PGW}(\dot{\alpha})$ and $\mathrm{PGW}(\ddot{\alpha})$. Recalling Definition 4.11, let

$$(\dot{\boldsymbol{\eta}}^\ell, \ddot{\boldsymbol{\eta}}^\ell) \equiv \left(\mathbf{F}_\ell(\dot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}), \mathbf{F}_\ell(\ddot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}})\right),$$

for all $\ell \geq 0$. Recall that $\dot{\boldsymbol{\eta}}^\ell$ and $\ddot{\boldsymbol{\eta}}^\ell$ are (random) probability measures over $\{\text{-}, \text{+}, \text{f}\}$; similarly as before we will denote $\dot{\eta}^\ell \equiv \dot{\boldsymbol{\eta}}^\ell(\text{-})$ and $\ddot{\eta}^\ell \equiv \ddot{\boldsymbol{\eta}}^\ell(\text{-})$. In the notation of Proposition 1.1, the marginal law of $\dot{\eta}^\ell$ is the measure $\mu^\ell(\dot{\alpha})$, while that of $\ddot{\eta}^\ell$ is $\mu^\ell(\ddot{\alpha})$. We hereafter write $\mu^\ell(\dot{\alpha}, \ddot{\alpha})$ for the law of the pair $(\dot{\eta}^\ell, \ddot{\eta}^\ell)$.

Lemma 10.2. *Let $0 \leq \ddot{\alpha} - \dot{\alpha} \leq \exp(-2^k)$, and let $(\dot{\eta}^\ell, \ddot{\eta}^\ell) \sim \mu^\ell(\dot{\alpha}, \ddot{\alpha})$ as described in Definition 10.1. Then*

$$\mathbb{E}\left(\mathbb{E}\left[(\dot{\eta}^\ell - \ddot{\eta}^\ell)^2 \,\Big|\, \dot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}\right]^2\right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}}$$

for all $\ell \geq 0$. (On the left-hand side above, in the inner expectation, $\dot{\eta}^\ell$ is determined as a measurable function of $\dot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}}$. The outer expectation is with respect to the law of $\dot{\mathscr{T}}_{v_{\mathrm{RT}}a_{\mathrm{RT}}} \sim \mathrm{PGW}(\dot{\alpha})$.)

Proof. The bound certainly holds for $\ell = 0$, where we have $\dot{\eta}^\ell = \ddot{\eta}^\ell = 1/2$ with probability one. Suppose inductively that the bound holds for $\ell - 1 \geq 0$. Define the i.i.d. array

$$\underline{t} \equiv \left((t_{ij}^+, t_{ij}^-)_{i,j \geq 1}\right)$$

where each entry $t = (\dot{t}, \ddot{t})$ is an independent sample from the monotone coupling $\mathrm{PGW}(\dot{\alpha}, \ddot{\alpha})$. Let

$$\underline{a} \equiv \left((a_{ij}^+, a_{ij}^-)_{i,j \geq 1}\right) \tag{10.1}$$

where each entry $a = (\dot\eta, \ddot\eta)$ is obtained by applying the map $F_{\ell-1}$ to the corresponding entry in $\underline{t}$: that is,

$$a_{ij}^+ \equiv (\dot\eta_{ij}^+, \ddot\eta_{ij}^+) \equiv \left(\left[F_{\ell-1}(\dot t_{ij}^+) \right](-), \left[F_{\ell-1}(\ddot t_{ij}^+) \right](-) \right) \in [0,1)^2 \, .$$

Next, let $\dot d^\pm, \delta^\pm$ be independent random variables with

$$\dot d^\pm \sim \mathrm{Pois}\!\left(\frac{\dot\alpha k}{2} \right), \quad \delta^\pm \sim \mathrm{Pois}\!\left(\frac{(\ddot\alpha - \dot\alpha)k}{2} \right).$$

Let $\ddot d^\pm \equiv \dot d^\pm + \delta^\pm$. Let $\dot T \equiv \dot{\mathscr{T}}_{v_{\mathrm{RT}} a_{\mathrm{RT}}}$ be the variable-to-clause tree with $|\partial v_{\mathrm{RT}}(\pm) \setminus a_{\mathrm{RT}}| = \dot d^\pm$, such that the i-th clause in $\partial v_{\mathrm{RT}}(\pm) \setminus a_{\mathrm{RT}}$ has child subtrees $\dot t_{ij}$ for $1 \le j \le k-1$. Likewise, let $\ddot T \equiv \ddot{\mathscr{T}}_{v_{\mathrm{RT}} a_{\mathrm{RT}}}$ be the variable-to-clause tree with $|\partial v_{\mathrm{RT}}(\pm) \setminus a_{\mathrm{RT}}| = \ddot d^\pm$, such that the i-th clause in $\partial v_{\mathrm{RT}}(\pm) \setminus a_{\mathrm{RT}}$ has child subtrees $\ddot t_{ij}$ for $1 \le j \le k-1$. Then

$$(\dot T, \ddot T) \equiv \left(\dot{\mathscr{T}}_{v_{\mathrm{RT}} a_{\mathrm{RT}}}, \ddot{\mathscr{T}}_{v_{\mathrm{RT}} a_{\mathrm{RT}}} \right) \sim \mathrm{PGW}(\dot\alpha, \ddot\alpha) \, .$$

With $F(x,y)$ as in (4.28), we can express

$$\dot\eta^\ell = F(\dot\Pi^-, \dot\Sigma) = \frac{(1 - \dot\Pi^-)\exp(\dot\Sigma)}{1 + (1 - \dot\Pi^-)\exp(\dot\Sigma)}$$

where we use similar notation as in Definition 4.9, (4.13), and (4.38):

$$\dot\Pi^\pm \equiv \prod_{i=1}^{\dot d^\pm} \dot s_i^\pm, \quad \dot s_i^\pm \equiv 1 - \prod_{j=1}^{k-1} \dot\eta_{ij}^\pm, \quad \dot\Sigma_i^\pm \equiv -\ln \dot\Pi^\pm, \quad \dot\Sigma \equiv \ln \frac{\dot\Pi^+}{\dot\Pi^-} \, .$$

We make all the analogous definitions to express $\ddot\eta^\ell = F(\ddot\Pi^-, \ddot\Sigma)$. As an intermediary between $\dot T$ and $\ddot T$, we define quantities that use the random messages $\ddot\eta_{ij}^\pm$ with the random degrees $\dot d^\pm$:

$$\tilde\Pi^\pm \equiv \prod_{i=1}^{\dot d^\pm} \ddot s_i^\pm, \quad \ddot s_i^\pm \equiv 1 - \prod_{j=1}^{k-1} \ddot\eta_{ij}^\pm, \quad \tilde\Sigma_i^\pm \equiv -\ln \tilde\Pi^\pm, \quad \tilde\Sigma \equiv \ln \frac{\tilde\Pi^+}{\tilde\Pi^-} \, .$$

With this notation, we can decompose

$$\ddot\eta^\ell - \dot\eta^\ell = \overbrace{\left(F(\ddot\Pi^-, \ddot\Sigma) - F(\tilde\Pi^-, \tilde\Sigma) \right)}^{\text{denote this } \mathbf{A}} + \overbrace{\left(F(\tilde\Pi^-, \tilde\Sigma) - F(\dot\Pi^-, \dot\Sigma) \right)}^{\text{denote this } \mathbf{a}} \tag{10.2}$$

Let $\dot X_i^\pm \equiv -\ln \dot s_i^\pm$ and $\ddot X_i^\pm \equiv -\ln \ddot s_i^\pm$. Using the bound (4.29) from Lemma 4.24, we have

$$\mathbb{E}\!\left[\mathbf{A}^2 \,\Big|\, \dot T \right] \le 2 \,\mathbb{E}\!\left[\left(\ddot\Pi^- - \tilde\Pi^- \right)^2 + \left(\ddot\Sigma - \tilde\Sigma \right)^2 \,\Big|\, \dot T \right]$$

$$= 2 \,\mathbb{E}\!\left[(\tilde\Pi^-)^2 \,\Big|\, \dot T \right] \mathbb{E}\!\left[\left(1 - \prod_{i=\dot d^- +1}^{\dot d^- + \delta^-} \ddot s_i^- \right)^2 \,\Big|\, \dot T \right] + 2 \,\mathbb{E}\!\left[\left(\sum_{i=\dot d^- +1}^{\dot d^- + \delta^-} \ddot X_i^- - \sum_{i=\dot d^+ +1}^{\dot d^+ + \delta^+} \ddot X_i^+ \right)^2 \,\Big|\, \dot T \right]$$

$$\le 2 \,\mathbb{E}\!\left[(\tilde\Pi^-)^2 \,\Big|\, \dot T \right] \mathbb{P}\!\left(\mathrm{Pois}\!\left(\frac{k(\ddot\alpha - \dot\alpha)}{2} \right) > 0 \right) + 2 \,\mathbb{E}\!\left[\mathrm{Pois}(k(\ddot\alpha - \dot\alpha)) \right] \mathbb{E}\!\left[(\ddot X_i^-)^2 \right].$$

Recall from Lemma 4.13 that $\ddot X_i^-$ is well concentrated around roughly $1/2^{k-1}$, from which it follows that

$$\tilde\Pi^- = \frac{1}{\exp(\Sigma^-)} = \exp\!\left\{ -\sum_{i=1}^{\dot d^-} \ddot X_i^- \right\}$$

is well concentrated around $1/2^k$. It follows that

$$\mathbb{E}\left(\mathbb{E}\left[\mathbf{A}^2 \,\middle|\, \dot{T}\right]^2\right) \le \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}}.$$

As for **a**, again using the bound (4.29) from Lemma 4.24, we have

$$\mathbb{E}\left[\mathbf{a}^2 \,\middle|\, \dot{T}\right] \le 2\,\mathbb{E}\left[(\dot{\Pi}^- - \tilde{\Pi}^-)^2 + (\dot{\Sigma} - \tilde{\Sigma})^2 \,\middle|\, \dot{T}\right]. \tag{10.3}$$

For the first term on the right-hand side of (10.3) we can bound

$$(\dot{\Pi}^- - \tilde{\Pi}^-)^2 \le \dot{d}^- \sum_{i=1}^{\dot{d}^-}(\Pi^-[i])^2(\dot{s}_i^- - \ddot{s}_i^-)^2, \quad \Pi^-[i] \equiv \prod_{i'=1}^{i-1}\dot{s}_{i'}^- \prod_{i'=i+1}^{\dot{d}^-}\ddot{s}_{i'}^-.$$

Then, using that $\dot{d}^\pm$ are measurable functions of $\dot{T}$, we have

$$\mathbb{E}\left(\mathbb{E}\left[(\dot{\Pi}^- - \tilde{\Pi}^-)^2 \,\middle|\, \dot{T}\right]^2\right) \le \overbrace{\mathbb{E}\left[(\dot{d}^-)^3 \sum_{i=1}^{\dot{d}^-}(\Pi^-[i])^4\right]}^{\le k^{O(1)}} \mathbb{E}\left(\mathbb{E}\left[(\dot{s}_1^- - \ddot{s}_1^-)^2 \,\middle|\, \dot{T}\right]^2\right),$$

where the first factor is $\le k^{O(1)}$ because $\dot{d}^-$ is well concentrated around $k\dot{\alpha}/2$ while each $\Pi^-[i]$ is well concentrated around $1/2^k$ (using Lemma 4.13 again). For the second factor, we can bound

$$(\dot{s}_i^- - \ddot{s}_i^-)^2 \le k \sum_{j=1}^{k-1} u_i^-[j]^2(\dot{\eta}_{ij}^- - \ddot{\eta}_{ij}^-)^2, \quad u_i^-[j] \equiv \prod_{j'=1}^{j-1}\dot{\eta}_{ij'}^- \prod_{j'=j+1}^{k-1}\ddot{\eta}_{ij'}^-.$$

Combining with the inductive hypothesis gives

$$\mathbb{E}\left(\mathbb{E}\left[(\dot{s}_i^- - \ddot{s}_i^-)^2 \,\middle|\, \dot{T}\right]^2\right) \le \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}}\left\{\max_{1 \le j \le k-1}\mathbb{E}[u_i^-[j]^4]\right\} \le \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{8k}/k^{O(1)}}.$$

Turning to the second term on the right-hand side of (10.3), we have

$$\mathbb{E}\left(\mathbb{E}\left[(\dot{\Sigma} - \tilde{\Sigma})^2 \,\middle|\, \dot{T}\right]^2\right) = \mathbb{E}\left(\mathbb{E}\left[\left(\sum_{i=1}^{\dot{d}^-}(\dot{X}_i^- - \ddot{X}_i^-)\right)^2 \,\middle|\, \dot{T}\right]^2\right) \le \mathbb{E}\left[(\dot{d}^-)^2\right]\mathbb{E}\left(\mathbb{E}\left[(\dot{X}_i^- - \ddot{X}_i^-)^2 \,\middle|\, \dot{T}\right]^2\right).$$

By a similar derivation as for (4.47), we can bound

$$\mathbb{E}\left(\mathbb{E}\left[(\dot{X}_1^- - \ddot{X}_1^-)^2 \,\middle|\, \dot{T}\right]^2\right) \le k^3 \sum_{j=1}^{k-1}\mathbb{E}\left(\left(\frac{u[j]}{1 - u[j]}\right)^4\right)\mathbb{E}\left(\mathbb{E}\left[(\eta_j - \tilde{\eta}_j)^2 \,\middle|\, \dot{T}\right]^2\right) \le \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{8k}/k^{O(1)}},$$

where the last step again uses the inductive hypothesis together with Lemma 4.13. Altogether we obtain

$$\mathbb{E}\left(\mathbb{E}\left[\mathbf{a}^2 \,\middle|\, \dot{T}\right]^2\right) \le \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}}.$$

Combining with the bound on **A** verifies the inductive hypothesis, concluding the proof. $\qquad\square$

Lemma 10.3. *Let* $0 \le \ddot{\alpha} - \dot{\alpha} \le \exp(-2^k)$, *and* $(\dot{\eta}^\ell, \ddot{\eta}^\ell) \sim \mu^\ell(\dot{\alpha}, \ddot{\alpha})$. *Then*

$$\mathbb{E}\left(\mathbb{E}\left[\ddot{\eta}^\ell - \dot{\eta}^\ell \,\middle|\, \dot{T}\right]^2\right) \le \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{3k}/k^{O(1)}}$$

for all $\ell \ge 0$.

Proof. We will argue by induction, using the same notation as in the proof of Lemma 10.2. The bound clearly holds for $\ell = 0$, so suppose it holds for $\ell - 1 \geq 0$. Decompose $\ddot{\eta}^\ell - \dot{\eta}^\ell = \mathbf{A} + \mathbf{a}$ as in (10.2). Applying the bound (4.30) from Lemma 4.24 gives

$$
\mathbf{a} = \overbrace{\left\{ \left(\tilde{\dot{\Pi}}^- - \dot{\Pi}^- \right) F_x(\dot{\Pi}^-, \dot{\Sigma}) \right\}}^{\text{denote this } \mathbf{a}_x} + \overbrace{\left\{ (\tilde{\Sigma} - \dot{\Sigma}) F_y(\dot{\Pi}^-, \dot{\Sigma}) \right\}}^{\text{denote this } \mathbf{a}_y} + \overbrace{O\left((\tilde{\dot{\Pi}}^- - \dot{\Pi}^-)^2 \right)}^{\text{denote this } \mathbf{a}_{xx}} + \overbrace{O\left(\left(\tilde{\Sigma} - \dot{\Sigma} \right)^2 \right)}^{\text{denote this } \mathbf{a}_{yy}} .
$$

We deal with each term separately in what follows.

Bound on $\mathbf{a}_x$. Let $\Pi^-[i]$ and $u_i^-[j]$ be as defined in the proof of Lemma 10.2. Then we have

$$
\tilde{\dot{\Pi}}^- - \dot{\Pi}^- = \sum_{i=1}^{\dot{d}^-} \Pi^-[i](\tilde{\dot{s}}_i^- - \dot{s}_i^-) = \sum_{i=1}^{\dot{d}^-} \Pi^-[i] \sum_{j=1}^{k-1} u_i^-[j](\dot{\eta}_{ij}^- - \ddot{\eta}_{ij}^-).
$$

Meanwhile, $F_x(\dot{\Pi}^-, \dot{\Sigma})$ is a measurable function of $\dot{T}$, and from Lemma 4.24 we have $|F_x(\dot{\Pi}^-, \dot{\Sigma})| \leq 1$. Combining with the inductive hypothesis, together with the independence of the random subtrees, we have

$$
\mathbb{E}\left(\mathbb{E}\left[\mathbf{a}_x \,\middle|\, \dot{T} \right]^2 \right) \leq \mathbb{E}\left(\left\{ \sum_{i=1}^{\dot{d}^-} \sum_{j=1}^{k-1} \mathbb{E}\left[\Pi^-[i] u_i^-[j] \,\middle|\, \dot{T} \right] \mathbb{E}\left[(\dot{\eta}_{ij}^- - \ddot{\eta}_{ij}^-) \,\middle|\, \dot{T} \right] \right\}^2 \right)
$$

$$
\leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{3k}/k^{O(1)}} \mathbb{E}\left(\dot{d}^- k \sum_{i=1}^{\dot{d}^-} \sum_{j=1}^{k-1} \mathbb{E}\left[\Pi^-[i] u_i^-[j] \,\middle|\, \dot{T} \right]^2 \right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{3k}/k^{O(1)}},
$$

where the last inequality uses the estimates from Lemma 4.13.

Bound on $\mathbf{a}_y$. For this part of the proof we will generate the pair $(\dot{T}, \ddot{T})$ in a slightly different way, as follows. We take the subtrees $\dot{t}_{ij}$ and $\ddot{t}_{ij}$ as before; the only difference is in the $\pm$-degrees of the root variable v_{RT}. Let $d \sim \text{Pois}(k\dot{\alpha})$, and $\delta \sim \text{Pois}(k(\ddot{\alpha} - \dot{\alpha}))$. Let S_i be i.i.d. symmetric random signs, and let

$$
\dot{d}^\pm \equiv \sum_{i=1}^{d} \mathbf{1}\{S_i = \pm\}, \quad \ddot{d}^\pm \equiv \sum_{i=1}^{d+\delta} \mathbf{1}\{S_i = \pm\}.
$$

We then let v_{RT} have $\pm$-degrees $\dot{d}^\pm$ in $\dot{T}$, and $\ddot{d}^\pm$ in $\ddot{T}$. Then, similarly to (4.47) we can decompose

$$
\tilde{\Sigma} - \dot{\Sigma} = \sum_{i=1}^{d} S_i(\ddot{X}_i - \dot{X}_i) = -\sum_{i=1}^{d} S_i \sum_{j=1}^{k-1} \ln \frac{1 - u_i[j]\ddot{\eta}_{ij}}{1 - u_i[j]\dot{\eta}_{ij}}.
$$

Note in the above that $\dot{X}_i$ stands for $\dot{X}_i^+$ or $\dot{X}_i^-$ depending on the sign S_i, but we have dropped the superscripts for notational convenience. Applying the approximation

$$
f(\eta) = \ln \frac{1}{1 - u_i[j]\eta} = f(\eta_0) + (\eta - \eta_0)f'(\eta_0) + O\left(\|f''\|_\infty (\eta - \eta_0)^2 \right)
$$

gives the following expansion for $\mathbf{a}_y$:

$$
\mathbf{a}_y = \overbrace{\left(-F_y(\dot{\Pi}^-, \dot{\Sigma}) \sum_{i=1}^{d} S_i \sum_{j=1}^{k-1} \frac{u_i[j](\ddot{\eta}_{ij} - \dot{\eta}_{ij})}{1 - u_i[j]\dot{\eta}_{ij}} \right)}^{\text{denote this } \mathbf{a}_{y,1}} + \overbrace{O\left(\sum_{i=1}^{d} \sum_{j=1}^{k-1} \left(\frac{u_i[j](\ddot{\eta}_{ij} - \dot{\eta}_{ij})}{1 - u_i[j]} \right)^2 \right)}^{\text{denote this } \mathbf{a}_{y,2}}.
$$

We bound these terms separately, beginning with the quadratic term $\mathbf{a}_{y,2}$:

$$
\mathbb{E}[\mathbf{a}_{y,2} \,|\, \dot{T}]^2 \leq dk \sum_{i=1}^{d} \sum_{j=1}^{k-1} \mathbb{E}\left[\left(\frac{u_i[j]}{1 - u_i[j]} \right)^2 \,\middle|\, \dot{T} \right]^2 \mathbb{E}\left[(\ddot{\eta}_{ij} - \dot{\eta}_{ij})^2 \,\middle|\, \dot{T}_{ij} \right]^2 .
$$

Combining with the inductive hypothesis, Lemma 4.13, and Lemma 10.2 gives

$$\mathbb{E}\left(\mathbb{E}[\mathbf{a}_{y,2}\,|\,\dot{T}]^2\right) \le \frac{(\ddot{\alpha}-\dot{\alpha})^2}{2^{4k}/k^{O(1)}}\,.$$

We next bound the contribution from the linear term $\mathbf{a}_{y,1}$, which requires some more care. Expanding $\mathbb{E}[\mathbb{E}[\mathbf{a}_{y,1}\,|\,\dot{T}]^2]$ as a double sum over indices $1 \le i_1, i_2 \le d$, we first bound the contribution from the cross terms $i_1 \ne i_2$. Here the issue is that although $S_{i_1}S_{i_2}$ certainly has mean zero, the cross term may not have mean zero because of the factor $F_y(\dot{\Pi}^-, \dot{\Sigma})^2$ which is correlated with S_{i_1}, S_{i_2}. However we shall argue that this effect is negligible. To this end, we shall approximate $\dot{\Pi}^-$ and $\dot{\Sigma}$ by

$$\dot{\Pi}^-[i_1 i_2] \equiv \prod_{\substack{i\in[\dot{d}]\setminus\{i_1,i_2\},\\ S_i=-}} \dot{s}_i\,, \quad \dot{\Sigma}[i_1 i_2] \equiv \sum_{i\in[\dot{d}]\setminus\{i_1,i_2\}} S_i X_i\,.$$

Let $H(x,y) \equiv F_y(x,y)^2$, and note that for $0 \le x \le 1$ and $xe^y \le 1$ we have

$$|H_x(x,y)| = 2\cdot\left|\frac{(1-x)e^y}{(1+(1-x)e^y)^3}\frac{1-(1-x)e^y}{1+(1-x)e^y}\frac{e^y}{(1-xe^y)+e^y}\right| \le 2\,,$$

$$|H_y(x,y)| = 2\cdot\left|\left(\frac{(1-x)e^y}{(1+(1-x)e^y)^2}\right)^2\frac{1-(1-x)e^y}{1+(1-x)e^y}\right| \le 2\,.$$

It follows that, abbreviating $H[i_1 i_2] \equiv H(\dot{\Pi}^-[i_1 i_2], \dot{\Sigma}[i_1 i_2])$, we have

$$\left|H(\dot{\Pi}^-, \dot{\Sigma}) - H[i_1 i_2]\right| \le 2\left\{\left|\dot{\Pi}^- - \dot{\Pi}^-[i_1 i_2]\right| + \left|\dot{\Sigma} - \dot{\Sigma}[i_1 i_2]\right|\right\}.$$

Returning to the expression for $\mathbf{a}_{y,1}$, let us denote

$$T(i) \equiv \sum_{j=1}^{k-1} \mathbb{E}\left[\frac{u_i[j]}{1-u_i[j]\dot{\eta}_{ij}}\,\bigg|\,\dot{T}\right]\mathbb{E}\left[(\ddot{\eta}_{ij}-\dot{\eta}_{ij})\,\bigg|\,\dot{T}\right].$$

We then use the Cauchy–Schwarz inequality to bound

$$T(i)^2 \le k\sum_{j=1}^{k-1}\left(\overbrace{\mathbb{E}\left[\frac{u_i[j]}{1-u_i[j]}\,\bigg|\,\dot{T}\right]^2\mathbb{E}\left[(\ddot{\eta}_{ij}-\dot{\eta}_{ij})\,\bigg|\,\dot{T}\right]^2}^{\text{denote this }T(ij)^2}\right).$$

Using the inductive hypothesis and Lemma 4.13, the i_1, i_2 cross term in $\mathbb{E}[\mathbb{E}[\mathbf{a}_{y,1}\,|\,\dot{T}]^2]$ can be bounded by

$$\overbrace{\mathbb{E}\left(H[i_1 i_2]\prod_{l=1}^{2}\left\{S_{i_l}T(i_l)\right\}\right)}^{\text{zero}} + \mathbb{E}\left[\left(H(\dot{\Pi}^-, \dot{\Sigma}) - H[i_1 i_2]\right)\prod_{l=1}^{2}\left\{S_{i_l}T(i_l)\right\}\right]$$

$$= O(1)\mathbb{E}\left[\left\{\left|\dot{\Pi}^- - \dot{\Pi}^-[i_1 i_2]\right| + \left|\dot{\Sigma} - \dot{\Sigma}[i_1 i_2]\right|\right\}\sum_{l=1}^{2}T(i_l)^2\right]$$

$$\le \frac{(\ddot{\alpha}-\dot{\alpha})^2}{2^{3k}/k^{O(1)}}\mathbb{E}\left[\left(\left|\dot{\Pi}^- - \dot{\Pi}^-[i_1 i_2]\right| + \left|\dot{\Sigma} - \dot{\Sigma}[i_1 i_2]\right|\right)\sum_{l=1}^{2}\sum_{j=1}^{k-1}\mathbb{E}\left[\frac{u_{i_l}[j]}{1-u_{i_l}[j]}\,\bigg|\,\dot{T}\setminus\dot{T}_{ij}\right]^2\right] \le \frac{(\ddot{\alpha}-\dot{\alpha})^2}{2^{6k}/k^{O(1)}}\,.$$

The total contribution to $\mathbb{E}[\mathbb{E}[\mathbf{a}_{y,1}\,|\,\dot{T}]^2]$ from cross terms $i_1 \neq i_2$ is thus $\leq k^{O(1)}(\ddot{\alpha} - \dot{\alpha})^2/2^{4k}$, so

$$\mathbb{E}\left(\mathbb{E}[\mathbf{a}_{y,1}\,|\,\dot{T}]^2\right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}} + \mathbb{E}\left(F_y(\dot{\Pi}^-, \dot{\Sigma})^2 \sum_{i=1}^{d} \mathbb{E}\left[\left.\left|\sum_{j=1}^{k-1} \frac{u_i[j](\ddot{\eta}_{ij} - \dot{\eta}_{ij})}{1 - u_i[j]\dot{\eta}_{ij}}\right|\,\right|\,\dot{T}\right]^2\right)$$

$$\leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}} + k\,\mathbb{E}\left(\sum_{i=1}^{d}\sum_{j=1}^{k-1}\mathbb{E}\left[\left.\frac{u_i[j]}{1 - u_i[j]}\,\right|\,\dot{T}\right]^2 \mathbb{E}\left[\left.(\ddot{\eta}_{ij} - \dot{\eta}_{ij})^2\,\right|\,\dot{T}\right]\right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}},$$

having again used the inductive hypothesis together with Lemma 4.13.

Bounds on $\mathbf{a}_{xx}$ and $\mathbf{a}_{yy}$. The quadratic terms are more straightforward to bound:

$$\mathbb{E}\left[\mathbb{E}[\mathbf{a}_{xx}\,|\,\dot{T}]^2\right] \leq k^{O(1)}\mathbb{E}\left((\dot{d}^-)^3 \sum_{i=1}^{\dot{d}^-}\sum_{j=1}^{k-1}\mathbb{E}\left[\left.\Pi^-[i]^2 u_i[j]^2\,\right|\,\dot{T}\right]^2 \mathbb{E}\left[\left.(\ddot{\eta}_{ij}^- - \dot{\eta}_{ij}^-)^2\,\right|\,\dot{T}\right]^2\right),$$

$$\mathbb{E}\left[\mathbb{E}[\mathbf{a}_{yy}\,|\,\dot{T}]^2\right] \leq k^{O(1)}\mathbb{E}\left(d^3 \sum_{k=1}^{d}\sum_{j=1}^{k-1}\mathbb{E}\left[\left.\left(\frac{u_i[j]}{1 - u_i[j]}\right)^2\,\right|\,\dot{T}\right]^2 \mathbb{E}\left[\left.(\ddot{\eta}_{ij}^- - \dot{\eta}_{ij}^-)^2\,\right|\,\dot{T}\right]^2\right).$$

Combining with Lemma 4.13 and Lemma 10.2 gives

$$\mathbb{E}\left[\mathbb{E}[\mathbf{a}_{xx}\,|\,\dot{T}]^2\right] + \mathbb{E}\left[\mathbb{E}[\mathbf{a}_{yy}\,|\,\dot{T}]^2\right] \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}}.$$

This concludes our bounds for $\mathbf{a}$.

Bound on $\mathbf{A}$. Recall that the tree $\dot{T}$ has root degree $\dot{d}$ while $\ddot{T}$ has root degree $\ddot{d} \equiv \dot{d} + \delta$. Let $\tilde{T}$ be the subtree of $\ddot{T}$ formed by deleting all subtrees descended from root neighbors with indices larger than $\dot{d}$, so that $\dot{T} \subseteq \tilde{T} \subseteq \ddot{T}$. By Jensen's inequality,

$$\mathbb{E}\left(\mathbb{E}[\mathbf{A}\,|\,\dot{T}]^2\right) \leq \mathbb{E}\left(\mathbb{E}[\mathbf{A}\,|\,\tilde{T}]^2\right).$$

If $\ddot{d} = \dot{d}$ then clearly $\mathbf{A} = 0$. Since $|\mathbf{A}| \leq 1$ almost surely, we can bound

$$\mathbb{E}\left(\mathbb{E}\left[\left.\mathbf{A}\mathbf{1}\{\delta \geq 2\}\,\right|\,\tilde{T}\right]^2\right) \leq \mathbb{P}(\delta \geq 2)^2 \leq k^{O(1)}(\ddot{\alpha} - \dot{\alpha})^4.$$

On the event $\delta = 1$, applying the bound (4.30) from Lemma 4.24 gives

$$\mathbb{E}\left[\left.\mathbf{A}\mathbf{1}\{\delta = 1\}\,\right|\,\tilde{T}\right] = \mathbb{E}\left[\overbrace{\left\{F_x(\tilde{\Pi}^-, \tilde{\Sigma})\left(\ddot{\Pi}^- - \tilde{\Pi}^-\right)\right\}}^{\text{denote this }\bar{\mathbf{A}}_x} + \overbrace{\left\{F_y(\tilde{\Pi}^-, \tilde{\Sigma})\left(\ddot{\Sigma} - \tilde{\Sigma}\right)\right\}}^{\text{denote this }\bar{\mathbf{A}}_y}\right.$$

$$\left.+ \underbrace{O\left((\ddot{\Pi}^- - \tilde{\Pi}^-)^2\right)}_{\text{denote this }\bar{\mathbf{A}}_{xx}} + \underbrace{O\left((\ddot{\Sigma}^- - \tilde{\Sigma}^-)^2\right)}_{\text{denote this }\bar{\mathbf{A}}_{yy}}\right].$$

Since $\delta = 1$ with probability $O(k)(\ddot{\alpha} - \dot{\alpha})$, we find

$$\mathbb{E}\left(\mathbb{E}\left[\left.\bar{\mathbf{A}}_x\mathbf{1}\{\delta = 1\}\,\right|\,\tilde{T}\right]^2\right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2\mathbb{E}[(\tilde{\Pi}^-)^2]}{2^{2k}/k^{O(1)}} \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}},$$

and by similar considerations

$$\mathbb{E}\left(\mathbb{E}\left[\left.\bar{\mathbf{A}}_{xx}\mathbf{1}\{\delta = 1\}\,\right|\,\tilde{T}\right]^2\right) + \mathbb{E}\left(\mathbb{E}\left[\left.\bar{\mathbf{A}}_{yy}\mathbf{1}\{\delta = 1\}\,\right|\,\tilde{T}\right]^2\right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}}.$$

Finally, by symmetry, $\mathbb{E}[\bar{\mathbf{A}}_y\,|\,\tilde{T}] = 0$. Altogether this yields

$$\mathbb{E}\left[\mathbb{E}[\mathbf{A}\,|\,\dot{T}]^2\right] \leq k^{O(1)}(\ddot{\alpha} - \dot{\alpha})^4 + \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}} \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{4k}/k^{O(1)}},$$

where the last step uses the assumed bound on $\ddot{\alpha} - \dot{\alpha}$.

Conclusion. Substituting the above estimates into (10.2) gives the claimed bound. □

Recall from Proposition 1.1 that $\mu^\ell(\alpha)$ converges weakly to $\mu(\alpha)$ as $\ell \to \infty$. Under the monotone coupling, it is clear that we also have $\mu^\ell(\dot{\alpha}, \ddot{\alpha})$ converging to a well-defined limit $\mu(\dot{\alpha}, \ddot{\alpha})$. If $(\dot{\eta}, \ddot{\eta})$ is sampled from $\mu(\dot{\alpha}, \ddot{\alpha})$, then $\dot{\eta}$ has marginal law $\mu(\dot{\alpha})$ while $\ddot{\eta}$ has marginal law $\mu(\ddot{\alpha})$.

Corollary 10.4. *Let* $0 \leq \ddot{\alpha} - \dot{\alpha} \leq \exp(-2^k)$. *Let* $\underline{a}$ *be as in* (10.1), *except that the entries are independent samples from* $\mu(\dot{\alpha}, \ddot{\alpha})$ *rather than from* $\mu^{\ell-1}(\dot{\alpha}, \ddot{\alpha})$. *Then, with the same notation as before, we have*

$$\mathbb{E}\left(\mathbb{E}[\ddot{X}_i - \dot{X}_i \mid \dot{T}]^2\right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{5k}/k^{O(1)}},$$

$$\mathbb{E}\left(\mathbb{E}\left[(\ddot{X}_i - \dot{X}_i)^2 \mid \dot{T}\right]^2\right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{8k}/k^{O(1)}}$$

for all $1 \leq i \leq \dot{d} = \dot{d}^+ + \dot{d}^-$.

Proof. With the same notation as before, we can expand

$$\ddot{X}_i - \dot{X}_i = \overbrace{\left\{ \sum_{j=1}^{k-1} \frac{u_i[j](\ddot{\eta}_{ij} - \dot{\eta}_{ij})}{1 - u_i[j]\dot{\eta}_{ij}} \right\}}^{\text{denote this } \mathbf{x}_1} + \overbrace{O\left(\sum_{j=1}^{k-1} \left(\frac{u_i[j](\ddot{\eta}_{ij} - \dot{\eta}_{ij})}{1 - u_i[j]} \right)^2 \right)}^{\text{denote this } \mathbf{x}_2}.$$

For the quadratic term, applying Lemma 10.2 (in the limit $\ell \to \infty$) gives

$$\mathbb{E}\left(\mathbb{E}[\mathbf{x}_2 \mid \dot{T}]^2\right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{8k}/k^{O(1)}}.$$

For the linear term, applying Lemma 10.3 gives

$$\mathbb{E}\left(\mathbb{E}[\mathbf{x}_1 \mid \dot{T}]^2\right) \leq k \, \mathbb{E}\left(\sum_{j=1}^{k-1} \mathbb{E}\left[\frac{u[j]}{1 - u[j]} \,\bigg|\, \dot{T} \right]^2 \mathbb{E}\left[\ddot{\eta}_j - \dot{\eta}_j \,\bigg|\, \dot{T} \right]^2 \right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{5k}/k^{O(1)}}.$$

Combining these gives the first assertion. Similar calculations give

$$\mathbb{E}\left(\mathbb{E}[(\ddot{X}_i - \dot{X}_i)^2 \mid \dot{T}]^2\right) \leq k^{O(1)} \mathbb{E}\left(\sum_{j=1}^{k-1} \mathbb{E}\left[\left(\frac{u[j]}{1 - u[j]}\right)^2 \,\bigg|\, \dot{T} \right]^2 \mathbb{E}\left[(\ddot{\eta}_j - \dot{\eta}_j)^2 \,\bigg|\, \dot{T} \right]^2 \right) \leq \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{8k}/k^{O(1)}}.$$

This gives the second assertion. □

Proof of Proposition 1.2. Let $\alpha_{\mathrm{lbd}} \leq \dot{\alpha} \leq \ddot{\alpha} \leq \alpha_{\mathrm{ubd}}$ with $\ddot{\alpha} - \dot{\alpha} \leq \exp(-2^k)$. Our goal is to upper bound the difference $\Phi(\ddot{\alpha}) - \Phi(\dot{\alpha})$ where Φ is the 1-RSB free energy defined by (1.11). To this end, define the function

$$G(x, y) \equiv \ln\left(\frac{1}{e^x} + \frac{1}{e^y} - \frac{1}{e^{x+y}} \right),$$

and note that $\Phi(\alpha) = \Phi_1(\alpha) + (k-1)\Phi_2(\alpha)$ where

$$\Phi_1(\alpha) \equiv \mathbb{E}\ln\left(\frac{1}{\exp(\Sigma^+)} + \frac{1}{\exp(\Sigma^-)} - \frac{1}{\exp(\Sigma^+ + \Sigma^-)}\right) = \mathbb{E}G(\Sigma^+, \Sigma^-),$$

$$\Phi_2(\alpha) \equiv -\alpha \, \mathbb{E}\ln\left(1 - \prod_{j=1}^k \eta_j\right) \equiv \alpha \, \mathbb{E}X'. \tag{10.4}$$

We define $\tilde{\Sigma}^\pm$ as in the proof of Lemma 10.2, and decompose (similarly to (10.2))

$$\Phi_1(\ddot{\alpha}) - \Phi_1(\dot{\alpha}) = \overbrace{\mathbb{E}\left(G(\ddot{\Sigma}^+, \ddot{\Sigma}^-) - G(\tilde{\Sigma}^+, \tilde{\Sigma}^-) \right)}^{\text{denote this } \mathbf{G}} + \overbrace{\mathbb{E}\left(G(\tilde{\Sigma}^+, \tilde{\Sigma}^-) - G(\dot{\Sigma}^+, \dot{\Sigma}^-) \right)}^{\text{denote this } \mathbf{g}}.$$

We further decompose

$$\mathbf{g} = \mathbb{E}\left(\overbrace{\left\{G(\tilde{\Sigma}^+, \tilde{\Sigma}^-) - G(\tilde{\Sigma}^+, \dot{\Sigma}^-)\right\}}^{\text{denote this } \mathbf{g}^-} + \overbrace{\left\{G(\tilde{\Sigma}^+, \dot{\Sigma}^-) - G(\dot{\Sigma}^+, \dot{\Sigma}^-)\right\}}^{\text{denote this } \mathbf{g}^+}\right).$$

It is easily checked that $|G_x| \le 1$ and $|G_{xx}| \le 1$, so

$$\mathbf{g}^+ = \overbrace{\left\{\sum_{i=1}^{\dot{d}^+} G_x\left(\Sigma^+[i] + \dot{X}_i^+, \dot{\Sigma}^-\right) \cdot \left(\ddot{X}_i^+ - \dot{X}_i^+\right)\right\}}^{\text{denote this } \mathbf{g}_x} + O\overbrace{\left(\sum_{i=1}^{\dot{d}^+} \left(\ddot{X}_i^+ - \dot{X}_i^+\right)^2\right)}^{\text{denote this } \mathbf{g}_{xx}}.$$

Applying Corollary 10.4 gives

$$(\mathbb{E}\mathbf{g}^+)^2 \le 2\,\mathbb{E}\left(\mathbb{E}[\mathbf{g}_x \mid \dot{T}]^2 + \mathbb{E}[\mathbf{g}_{xx} \mid \dot{T}]^2\right) \le \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{3k}/k^{O(1)}} + \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{6k}/k^{O(1)}} \le \frac{(\ddot{\alpha} - \dot{\alpha})^2}{2^{3k}/k^{O(1)}}.$$

Similar bounds hold for $\mathbf{g}^-$, so we conclude

$$|\mathbf{g}| = \left|\mathbb{E}\left(\mathbf{g}^- + \mathbf{g}^-\right)\right| \le \frac{\ddot{\alpha} - \dot{\alpha}}{2^{3k/2}/k^{O(1)}}.$$

It remains to estimate $\mathbf{G}$. This is similar to the analysis of $\mathbf{A}$ in the proof of Lemma 10.3: write $\delta \equiv \ddot{d} - \dot{d} = \delta^+ + \delta^-$ for the difference in root degrees between $\dot{T}$ and $\ddot{T}$. When $\delta = 0$ we have simply $\mathbf{G} = 0$. The expected contribution from the event $\delta \ge 2$ is negligible. It remains to consider the case $\delta = 1$. Without loss of generality, take the case that $\delta^+ = 1$ and $\delta^- = 0$:

$$\mathbf{G}_{1,0} \equiv \mathbb{E}\left[G(\ddot{\Sigma}^+, \ddot{\Sigma}^-) - G(\tilde{\Sigma}^+, \tilde{\Sigma}^-); (\delta^+, \delta^-) = (1, 0)\right]$$

$$= \mathbb{E}\left[G(\tilde{\Sigma}^+ + \ddot{X}_{\ddot{d}^+}, \tilde{\Sigma}^-) - G(\tilde{\Sigma}^+, \tilde{\Sigma}^-); (\delta^+, \delta^-) = (1, 0)\right]$$

$$= -\frac{1}{2} \cdot \frac{k(\ddot{\alpha} - \dot{\alpha})}{2} \cdot \frac{1 + O(2^{-k/20})}{2^{k-1}},$$

where the last estimate uses Lemma 4.13 and the fact that the partial derivative $G_x(\tilde{\Sigma}^+, \tilde{\Sigma}^-)$ is well concentrated around $-1/2$. Altogether we conclude

$$\Phi_1(\ddot{\alpha}) - \Phi_1(\dot{\alpha}) = -\frac{k(\ddot{\alpha} - \dot{\alpha})}{2^k}\left\{1 + \frac{O(1)}{2^{k/20}}\right\}.$$

Next, let $\dot{X}'$ and $\ddot{X}'$ be defined analogously to X' in (10.4), but using the messages $\dot{\eta}_j$ and $\ddot{\eta}_j$. Then

$$\Phi_2(\ddot{\alpha}) - \Phi_2(\dot{\alpha}) = \overbrace{\left\{(\ddot{\alpha} - \dot{\alpha}) \cdot \mathbb{E}\dot{X}'\right\}}^{\text{denote this } \mathbf{h}_{21}} + \overbrace{\left\{\ddot{\alpha}\,\mathbb{E}(\ddot{X}' - \dot{X}')\right\}}^{\text{denote this } \mathbf{h}_{22}}.$$

A trivial modification of Corollary 10.4 gives

$$(\mathbb{E}\mathbf{h}_{22})^2 \le \frac{\ddot{\alpha}^2(\ddot{\alpha} - \dot{\alpha})}{2^{5k}/k^{O(1)}} \le \frac{(\ddot{\alpha} - \dot{\alpha})}{2^{3k}/k^{O(1)}}.$$

On the other hand, a trivial modification of bounds from Lemma 4.13 (with k in place of $k - 1$) gives

$$\mathbb{E}\mathbf{h}_{21} = \frac{\ddot{\alpha} - \dot{\alpha}}{2^k}\left\{1 + \frac{O(1)}{2^{k/20}}\right\}.$$

Combining the above estimates gives finally

$$\Phi(\ddot{\alpha}) - \Phi(\dot{\alpha}) = \left\{ -\frac{k(\ddot{\alpha} - \dot{\alpha})}{2^k} + \frac{(k-1)(\ddot{\alpha} - \dot{\alpha})}{2^k} \right\} \left\{ 1 + \frac{O(1)}{2^{k/20}} \right\} = -\frac{\ddot{\alpha} - \dot{\alpha}}{2^k} \left\{ 1 + \frac{O(k)}{2^{k/20}} \right\}.$$

This proves that Φ is strictly decreasing on the interval $\alpha_{\mathrm{lbd}} \le \alpha \le \alpha_{\mathrm{ubd}}$, as desired. □

References

[AC08]   D. Achlioptas and A. Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proc. 49th FOCS*, pages 793–802. IEEE, 2008.

[AC15]   A. Auffinger and W.-K. Chen. The Parisi formula has a unique minimizer. *Comm. Math. Phys.*, 335(3):1429–1444, 2015.

[AC16]   A. Auffinger and W.-K. Chen. The Legendre structure of the Parisi formula. *Comm. Math. Phys.*, 348(3):751–770, 2016.

[ACIM01]   D. Achlioptas, A. Chtcherba, G. Istrate, and C. Moore. The phase transition in 1-in-k-SAT and NAE-3-SAT. In *Proc. 12th SODA*, pages 721–722. ACM–SIAM, 2001.

[ACZ20]   A. Auffinger, W.-K. Chen, and Q. Zeng. The SK model is infinite step replica symmetry breaking at zero temperature. *Comm. Pure Appl. Math.*, 73(5):921–943, 2020.

[AL07]   D. Aldous and R. Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12(54):1454–1508, 2007.

[AM02]   D. Achlioptas and C. Moore. The asymptotic order of the random k-SAT threshold. In *Proc. 43rd FOCS*, pages 779–788, 2002.

[AP03]   D. Achlioptas and Y. Peres. The threshold for random k-SAT is $2^k \ln 2 - O(k)$. In *Proc. 35th STOC*, pages 223–231 (electronic), New York, 2003. ACM.

[BBC+01]   B. Bollobás, C. Borgs, J. T. Chayes, J. H. Kim, and D. B. Wilson. The scaling window of the 2-SAT transition. *Random Structures Algorithms*, 18(3):201–256, 2001.

[BC15]   C. Bordenave and P. Caputo. Large deviations of empirical neighborhood distribution in sparse random graphs. *Probab. Theory Related Fields*, 163(1-2):149–222, 2015.

[BCH+14]   V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann, and D. Vilenchik. The condensation phase transition in random graph coloring. In *Proc. 18th RANDOM*, 2014.

[BCO16]   V. Bapst and A. Coja-Oghlan. The condensation phase transition in the regular k-SAT model. In *Approximation, randomization, and combinatorial optimization. Algorithms and techniques*, volume 60 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 22, 18. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016.

[BMZ05]   A. Braunstein, M. Mézard, and R. Zecchina. Survey propagation: an algorithm for satisfiability. *Random Struct. Algor.*, 27(2):201–226, 2005.

[Bol80]   B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[BS01]   I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6(23):13 pp. (electronic), 2001.

[CKT91]   P. Cheeseman, B. Kanefsky, and W. Taylor. Where the really hard problems are. In *Proc. 12th IJCAI*. Morgan Kaufmann, 1991.

[CO09]   A. Coja-Oghlan. A better algorithm for random k-SAT. In *Automata, languages and programming. Part I*, volume 5555 of *Lecture Notes in Comput. Sci.*, pages 292–303. Springer, Berlin, 2009.

[CP12]   A. Coja-Oghlan and K. Panagiotou. Catching the k-NAESAT threshold. In *Proc. 45th STOC*, pages 899–907. ACM, New York, 2012.

[CP13]   A. Coja-Oghlan and K. Panagiotou. Going after the k-SAT threshold. In *Proc. 45th STOC*, pages 705–714, New York, NY, USA, 2013. ACM.

[CP16]   A. Coja-Oghlan and K. Panagiotou. The asymptotic k-SAT threshold. *Adv. Math.*, 288:985–1068, 2016.

[CPS21]   W.-K. Chen, D. Panchenko, and E. Subag. The generalized TAP free energy II. *Comm. Math. Phys.*, 381(1):257–291, 2021.

[CR92]   V. Chvátal and B. Reed. Mick gets some (the odds are on his side) [satisfiability]. In *Proc. 33rd FOCS*, pages 620–627. IEEE, 1992.

[DSS13]   J. Ding, A. Sly, and N. Sun. Satisfiability threshold for random regular NAE-SAT. In *Proc. 46th STOC*. ACM, 2013.

[DSS16a]   J. Ding, A. Sly, and N. Sun. Maximum independent sets on random regular graphs. *Acta Math.*, 217(2):263–340, 2016.

[DSS16b]   J. Ding, A. Sly, and N. Sun. Satisfiability threshold for random regular NAE-SAT. *Comm. Math. Phys.*, 341(2):435–489, 2016.

[FL03]   S. Franz and M. Leone. Replica bounds for optimization problems and diluted spin systems. *J. Statist. Phys.*, 111(3-4):535–564, 2003.

[FP83]   J. Franco and M. Paull. Probabilistic analysis of the Davis–Putnam procedure for solving the satisfiability problem. *Disc. Appl. Math.*, 5(1):77–87, 1983.

[Fri99]   E. Friedgut. Sharp thresholds of graph properties, and the k-SAT problem. *J. Amer. Math. Soc.*, 12(4):1017–1054, 1999. With an appendix by Jean Bourgain.

[Goe92]   A. Goerdt. A threshold for unsatisfiability. In *Proc. 17th MFCS*, volume 629 of *Lect. Notes Comput. Sci.*, pages 264–274. Springer, Berlin, 1992.

[GS14]   D. Gamarnik and M. Sudan. Limits of local algorithms over sparse random graphs. In *Proc. 5th ITCS*, pages 369–376, New York, NY, USA, 2014. ACM.

[GT02]   F. Guerra and F. L. Toninelli. The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.*, 230(1):71–79, 2002.

[Gue03]   F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.

[JT16]   A. Jagannath and I. Tobasco. A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.*, 144(7):3135–3150, 2016.

[Kar72]   R. M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972)*, pages 85–103. Plenum, New York, 1972.

[KKKS98]   L. M. Kirousis, E. Kranakis, D. Krizanc, and Y. C. Stamatiou. Approximating the unsatisfiability threshold of random formulas. *Random Struct. Algor.*, 12(3):253–269, 1998.

[KMR+07]   F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proc. Natl. Acad. Sci. USA*, 104(25):10318–10323 (electronic), 2007.

[Lev86]   L. A. Levin. Average case complete problems. *SIAM J. Comput.*, 15(1):285–286, 1986.

[LPP95]   R. Lyons, R. Pemantle, and Y. Peres. Ergodic theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure. *Ergodic Theory Dynam. Systems*, 15(3):593–619, 1995.

[MM09]   M. Mézard and A. Montanari. *Information, physics, and computation*. Oxford Graduate Texts. Oxford University Press, Oxford, 2009.

[MMW07]   E. Maneva, E. Mossel, and M. J. Wainwright. A new look at survey propagation and its generalizations. *J. ACM*, 54(4):Art. 17, 41, 2007.

[MMZ06]   S. Mertens, M. Mézard, and R. Zecchina. Threshold values of random k-SAT from the cavity method. *Random Struct. Algor.*, 28(3):340–373, 2006.

[Mon19]   A. Montanari. Optimization of the Sherrington–Kirkpatrick Hamiltonian. In *Proc. 60th FOCS*. SIAM, 2019.

[MP85]   M. Mézard and G. Parisi. Replicas and optimization. *J. Phys. Lett.*, 46(17):771–778, 1985.

[MPV87]   M. Mézard, G. Parisi, and M. Virasoro. *Spin glass theory and beyond: An introduction to the replica method and its applications*, volume 9. World Scientific Publishing Company, 1987.

[MPZ02]   M. Mézard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297(5582):812–815, 2002.

[MRS08]   A. Montanari, F. Ricci-Tersenghi, and G. Semerjian. Clusters of solutions and replica symmetry breaking in random k-satisfiability. *J. Stat. Mech. Theory E.*, 2008(04):P04004, 2008.

[MRZ03]   M. Mézard, F. Ricci-Tersenghi, and R. Zecchina. Two solutions to diluted p-spin models and xorsat problems. *J. Statist. Phys.*, 111(3-4):505–533, 2003.

[MSL92]   D. Mitchell, B. Selman, and H. Levesque. Hard and easy distributions of SAT problems. In *Proc. 10th AAAI*, volume 92, pages 459–465. Citeseer, 1992.

[Pan13a]   D. Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013.

[Pan13b]   D. Panchenko. *The Sherrington–Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York, 2013.

[Pan13c]   D. Panchenko. Spin glass models from the point of view of spin distributions. *Ann. Probab.*, 41(3A):1315–1361, 2013.

[Pan14]   D. Panchenko. Structure of 1-RSB asymptotic Gibbs measures in the diluted p-spin models. *J. Stat. Phys.*, 155(1):1–22, 2014.

[Pan15]   D. Panchenko. Hierarchical exchangeability of pure states in mean field spin glass models. *Probab. Theory Related Fields*, 161(3-4):619–650, 2015.

[Par79]   G. Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43(23):1754, 1979.

[Par80a]   G. Parisi. The order parameter for spin glasses: A function on the interval 0-1. *J. Phys. A*, 13(3):1101, 1980.

[Par80b]   G. Parisi. A sequence of approximated solutions to the SK model for spin glasses. *J. Phys. A*, 13(4):L115, 1980.

[Par83]   G. Parisi. Order parameter for spin-glasses. *Phys. Rev. Lett.*, 50(24):1946, 1983.

[Par05]   G. Parisi. On local equilibrium equations for clustering states. arXiv:cs/0212047v2, 2005.

[PS16]   B. Pittel and G. B. Sorkin. The satisfiability threshold for k-XORSAT. *Combin. Probab. Comput.*, 25(2):236–268, 2016.

[PT04]   D. Panchenko and M. Talagrand. Bounds for diluted mean-fields spin glass models. *Probab. Theory Rel. Fields*, 130(3):319–336, 2004.

[Roc97]   R. T. Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.

[SK75]   D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Phys. Rev. Lett.*, 35(26):1792, 1975.

[SSZ16]   A. Sly, N. Sun, and Y. Zhang. The number of solutions for random regular NAE-SAT. In *Proc. 57th FOCS*, pages 724–731. IEEE, 2016.

[Sub21]   E. Subag. Following the ground states of full-RSB spherical spin glasses. *Comm. Pure Appl. Math.*, 74(5):1021–1044, 2021.

[Tal06]   M. Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006.

[Tal11]   M. Talagrand. *Mean field models for spin glasses. Volume I*, volume 54 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer-Verlag, Berlin, 2011. Basic examples.

[Wor99]   N. C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.

[YFW05]   J. S. Yedidia, W. T. Freeman, and Y. Weiss. Constructing free-energy approximations and generalized belief propagation algorithms. *IEEE Trans. Inform. Theory*, 51(7):2282–2312, 2005.

[ZK07]   L. Zdeborova and F. Krzakala. Phase transitions in the coloring of random graphs. *Phys. Rev. E*, 76(3):031131, 2007.