

## MIT Open Access Articles

*Satisfiability Threshold for Random Regular nae-sat*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Ding, Jian, Allan Sly, and Nike Sun. "Satisfiability Threshold for Random Regular Nae-Sat." *Communications in Mathematical Physics* 341, no. 2 (November 26, 2015): 435–489. © 2015 Springer-Verlag

**As Published:** <http://dx.doi.org/10.1007/s00220-015-2492-8>

**Publisher:** Springer-Verlag

**Persistent URL:** <http://hdl.handle.net/1721.1/109568>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Communicated by F. Toninelli  
Receipt: 6/9/14  
Accept: 8/25/15

## SATISFIABILITY THRESHOLD FOR RANDOM REGULAR NAE-SAT

\*JIAN DING, †ALLAN SLY, AND ‡NIKE SUN

ABSTRACT. We consider the random regular  $k$ -NAE-SAT problem with  $n$  variables each appearing in exactly  $d$  clauses. For all  $k$  exceeding an absolute constant  $k_0$ , we establish explicitly the satisfiability threshold  $d_* \equiv d_*(k)$ . We prove that for  $d < d_*$  the problem is satisfiable with high probability, while for  $d > d_*$  the problem is unsatisfiable with high probability. If the threshold  $d_*$  lands exactly on an integer, we show that the problem is satisfiable with probability bounded away from both zero and one. This is the first result to locate the exact satisfiability threshold in a random constraint satisfaction problem exhibiting the condensation phenomenon identified by Krzakala et al. (2007). Our proof verifies the one-step replica symmetry breaking formalism for this model. We expect our methods to be applicable to a broad range of random constraint satisfaction problems and combinatorial problems on random graphs.

### 1. INTRODUCTION

Given a conjunctive normal form (CNF) boolean formula (an “AND of ORs”), a *not-all-equal-SAT* (NAE-SAT) solution is an assignment  $\underline{x}$  of literals to variables so that both  $\underline{x}$  and its negation  $\neg\underline{x}$  evaluate to TRUE — equivalently, so that no clause evaluates its variables to all TRUE or all FALSE. A  $k$ -NAE-SAT problem is defined by a  $k$ -CNF formula, in which each clause has exactly  $k$  literals.

A major direction of research on boolean satisfiability has concerned the large-system limit for *random problem instances*, seeking to establish typical behavior and phase transitions. A random  $k$ -satisfiability problem is given by choosing uniformly random clause literals, then assigning variables to clauses so that the constraint structure is that of a random  $k$ -uniform hypergraph. In the relevant asymptotic scaling the clause density  $\alpha$  (clause-to-variable ratio) is of constant order, that is to say the graph is sparse. Much effort has been directed towards locating the *satisfiability transition*: the critical density  $\alpha_*$  where solutions cease to exist.

The (sparse) random satisfiability problems belong to a broad universality class of sparse random *constraint satisfaction problems* (CSPs) — which includes also the coloring and independent set problems on sparse random graphs — that has been intensively studied in theoretical computer science, statistical physics, and combinatorics. A common feature of these problems is that in a non-trivial regime below  $\alpha_*$ , *the number of solutions fails to concentrate about its mean*, preventing standard (first and second) moment methods from locating the exact transition.

Statistical physicists ([MP85, MZK+99, MPZ02] and references therein) have described these problems via a deep but non-rigorous theory of *replica symmetry breaking* (RSB) which posits a breakup of the solution space into well-separated *clusters*. The non-concentration of the number of solutions below  $\alpha_*$  is understood in terms of a regime of *condensation* [KMR+07, MRS08] in which most solutions are contained in a bounded number of large clusters. Inference on clusters yields an explicit prediction of the satisfiability threshold, the *one-step replica symmetry breaking* (1RSB) solution, for a wide range of models including random  $k$ -SAT, coloring, and independent set. However, no such prediction has been rigorously verified in the presence of a condensation regime, with all previous bounds leaving a constant gap.

---

*Date:* September 9, 2015.

Research supported by \*NSF grant DMS-1313596; †Sloan Research Fellowship; ‡NDSEG and NSF GRF.



In this paper we consider *random  $d$ -regular  $k$ -NAE-SAT*, in which each variable is involved in exactly  $d$  clauses, and clause literals are chosen uniformly at random. We establish the following sharp satisfiability threshold, the first of its kind among this class of CSPs:

**Theorem 1.** *For  $k \geq k_0$  there is a threshold  $d_* \equiv d_*(k)$ , given by the largest zero of the explicit function (1), such that the probability for a random  $d$ -regular  $k$ -NAE-SAT instance to be satisfiable tends to one for  $d < d_*$ , to zero for  $d > d_*$ , and stays bounded away from zero and one for  $d = d_*$ .*

Throughout we consider the large-system limit in which the number  $n$  of variables tends to infinity while  $d, k$  remain fixed. Explicitly,  $d_*$  is given by the largest zero of the function

$$*\Phi(d) = -\log(1-x) - d(1-k^{-1}-d^{-1})\log(1-2x^k) + (d-1)\log(1-x^{k-1}), \quad (1)$$

where  $x = x(d)$  is the unique solution of the equation

$$d = 1 + \left( \log \frac{1-2x}{1-x} \right) / \log \left( \frac{1-2x^{k-1}}{1-x^{k-1}} \right) \quad \text{on the interval } \frac{1}{2} - 2^{-k} \leq x \leq \frac{1}{2}.$$

Furthermore,  $d_*$  is the unique zero of  $*\Phi$  on the interval  $(2^{k-1}-2)k \log 2 \leq d \leq 2^{k-1}k \log 2$ .

It is possible for the root  $d_*$  to be integer-valued, though we have no reason to believe that this ever occurs. Nevertheless, we address this hypothetical possibility by showing that if  $d = d_*$  then the probability for the NAE-SAT instance to be satisfiable is asymptotically bounded away from both zero and one. This gives a full characterization of the satisfiability transition in this problem.

The threshold  $d_*$  has been predicted before using 1RSB methods [CNRZ03, DRZ08], and arises from solving *survey propagation* recursions, a special case of the 1RSB cavity recursions — for further background we refer the reader to [MPZ02, AGK04, BZ04, BMZ05, MMW07, MM09]. The main challenge overcome in this paper is the non-concentration of the number of solutions occurring before the satisfiability transition. Our approach is based on the 1RSB intuition that while the number of solutions is not well concentrated, *the number of clusters is*. Indeed, a key innovation in our proof is to rigorously establish a simple combinatorial description of clusters, with which we show the necessary concentration to locate the exact threshold  $d_*$ . Our method of proof thus gives rigorous validation of the 1RSB heuristics for this problem.

We believe that the methods developed in this paper are flexible to the model specification, and offer a robust new approach to establishing exact thresholds for other problems within this class of sparse random CSPs. Indeed, in a companion paper [DSS13] we consider the maximum independent set (MAX-IND-SET) problem on random regular graphs, and apply similar methods to determine the explicit MAX-IND-SET threshold, and furthermore show tight concentration of the maximum independent set size about the threshold value.

Previous work has identified sharp satisfiability transitions in models not exhibiting condensation. The 2-SAT transition can be identified by a branching process argument [CR92, Goe96, FdlV01], and even the finite-size scaling has been characterized [BBC<sup>+</sup>01]. The satisfiability transitions in the 1-in- $k$ -SAT [ACIM01] and XOR-SAT [MRTZ03, PS12] problems have also been exactly determined. See [MM09] for more detailed discussions. The study of random  $k$ -satisfiability problems has focused on the Erdős-Rényi version, where variables are included in clauses independently at random. For random  $k$ -CNF formulas with  $k \geq 3$ , successively improving bounds on the location of the satisfiability transition have been proved ([KKKS98, AP04, CP13, Coj14] for random  $k$ -SAT, and [AM06, CZ12, CP12] for random  $k$ -NAE-SAT); in each case a gap remains to be closed.

After we announced this result, A. Coja-Oghlan posted a paper [Coj13] on a different symmetrization of regular  $k$ -SAT in which a 2-clause joins each consecutive pair of variables, forcing them to take opposite literals. While not establishing a satisfiability threshold, he establishes a 1RSB-type formula for the existence of solutions that satisfy all but  $o(n)$  clauses, via an approach of modeling clusters which is similar to ours. In subsequent work [Coj14] he demonstrated how to apply the method to establish an improved bound on the threshold for random  $k$ -SAT, but removing the gap to establish an exact threshold remains an open problem.

In the remainder of this introductory section, we review the concept of RSB and explain why it obstructs standard approaches for locating thresholds, then present a brief overview of our solution. A more detailed overview is available in conference proceedings [DSS14].

**1.1. Moments and non-concentration.** Throughout we let  $\mathcal{G}_{n,d,k}$  denote a uniformly random  $d$ -regular  $k$ -NAE-SAT problem instance, on  $n$  boolean variables constrained by  $m$  clauses where  $nd = mk$  (see §2.1 for the formal definition). A natural approach to locating the satisfiability transition is via the (first and second) moment method, as follows. Let  $Z \equiv Z_n$  count the number of solutions for the random NAE-SAT instance  $\mathcal{G}_{n,d,k}$ . The probability of satisfiability is upper bounded by Markov's inequality,  $\mathbb{P}(Z > 0) \leq \mathbb{E}Z$ . Given any fixed  $\underline{x} \in \{0, 1\}^n$ , each clause has probability  $2/2^k$  to evaluate its  $k$  variables to all 0 or all 1, thereby invalidating the solution; otherwise the clause is satisfied. The clauses are satisfied independently of one another, thus

$$\begin{aligned} \mathbb{E}Z &= 2^n(1 - 2/2^k)^m = \exp\{n\Phi(d)\} \\ &\text{with } \Phi(d) \equiv \log 2 + (d/k) \log(1 - 2/2^k). \end{aligned} \quad (2)$$

For fixed  $k$  the rate function  $\Phi$  is clearly decreasing in  $d$ , with unique zero at

$$d_\circ \equiv \frac{k \log 2}{-\log(1 - 2/2^k)} = [2^{k-1} - 1/2 - (6 \cdot 2^k)^{-1} - O(4^{-k})] k \log 2 \leq 2^{k-1} k \log 2 \equiv d_{\text{ubd}} \quad (3)$$

Above  $d_\circ$  the first moment  $\mathbb{E}Z$  is exponentially small in  $n$ , so the random NAE-SAT instance is unsatisfiable *with high probability* (with probability tending to one as  $n \rightarrow \infty$ ).

On the other hand, the probability to be satisfiable can be lower bounded by the Cauchy–Schwarz inequality,  $(\mathbb{E}Z)^2 \leq \mathbb{P}(Z > 0) \mathbb{E}[Z^2]$ . If the second moment remains within a bounded factor of the first moment squared as  $n \rightarrow \infty$ , then  $Z > 0$  *with positive probability* (with probability bounded uniformly away from zero in the limit  $n \rightarrow \infty$ ). We show that

**Proposition 1.1.** *For  $k \geq k_0$  and  $d \leq d_{\text{ibd}} \equiv (2^{k-1} - 2) k \log 2$ , the number  $Z$  of NAE-SAT solutions on  $\mathcal{G}_{n,d,k}$  satisfies  $\mathbb{E}[Z^2] \lesssim_k (\mathbb{E}Z)^2$ , implying that the problem is satisfiable with positive probability.*

(We write  $f \lesssim_k g$  to indicate  $\limsup_{n \rightarrow \infty} (f/g) \leq C_k$  for a constant  $C_k$  depending only on  $k$ . If  $f \lesssim_k g$  and  $g \lesssim_k f$  then we write  $f \asymp_k g$ . We drop the subscript  $k$  to indicate that the constant can be chosen uniformly over all  $k \geq k_0$ .) Propn. 1.1 is entirely analogous to the second moment estimates of [AM06] for the Erdős–Rényi version of random NAE-SAT.

**1.2. Non-concentration due to clustering.** The above shows that if the threshold  $d_*$  exists, then it must lie between  $d_{\text{ibd}}$  and  $d_\circ \leq d_{\text{ubd}}$ . However, the basic moment approach cannot locate the exact satisfiability transition: there is a non-trivial regime  $d \leq d_\circ$  in which the first moment of  $Z$  is exponentially large, but the second moment is exponentially larger than the first moment squared, yielding no conclusions about the typical behavior of  $Z$ .

We will now review a heuristic explanation for the failure of the second moment. We refer also to detailed discussions in prior work on satisfiability lower bounds [AM06, AP04]. Observe that, due to the sparsity of the graph, a typical NAE-SAT solution will have a non-negligible fraction of “free” variables, whose states can be flipped without violating any clauses. Indeed, suppose variable  $v$  is incident to clause  $a$ : if the other  $k - 1$  variables incident to  $a$  have the same evaluation (which occurs with probability roughly  $4/2^k$ ), then  $v$  will be forced to a particular literal to satisfy  $a$ ; otherwise,  $a$  is satisfied regardless of the literal on  $v$ . The probability that  $v$  can be flipped without violating *any* of its  $d$  neighboring clauses is then

$$\epsilon \approx \frac{(1 - 4/2^k)^d}{2(1 - 2/2^k)^d - (1 - 4/2^k)^d}, \quad (4)$$

where the denominator accounts for the fact that we have conditioned on a valid solution, so different clauses cannot force the variable to conflicting literals. (Let us emphasize again that this discussion is only meant as a heuristic guide; in particular, the calculation incorrectly assumes that



neighboring variables are roughly independent, when in fact the model has long-range correlations.) Thus the existence of a single solution implies the existence of a *cluster* of at least  $2^{nc}$  nearby solutions, related by flips on the free variables, where  $c$  is a positive constant depending on  $k$  and  $d$ . Observe moreover that  $\epsilon$  remains positive (at least on the order of  $2^{-k}$ ) *up to and beyond* the first moment threshold, and consequently  $c$  can be chosen uniformly over  $d \leq d_0$ . Thus there is a regime  $d \leq d_0$  where

$$\mathbb{E}[Z^2] \geq 2^{nc} \mathbb{E}Z \gg (\mathbb{E}Z)^2 \quad (5)$$

— that is, the amount of non-rigidity in a typical solution overcomes the total expected number of solutions, causing the second moment method to fail.

Statistical physicists have developed a sophisticated but non-rigorous theory which explains how this non-concentration fits within a more comprehensive picture. We briefly summarize here the (conjectural) phase diagram of a CSP, referring the reader to [KMR<sup>+</sup>07, MRS08] for the details. For small  $d$ , almost all of the solution space belongs to a single well-connected cluster. As  $d$  surpasses a certain *clustering* threshold  $d_s$ , the solution space undergoes an abrupt structural transition: for  $d$  above  $d_s$ , most of the mass in the solution space becomes divided into exponentially many well-separated clusters, each contributing an exponentially small fraction of the total mass. This geometry persists up to a further (conjectured) *condensation* transition  $d_c$ , above which most of the solution space becomes concentrated within a bounded number of large clusters. In the non-trivial regime between the condensation and satisfiability transitions, the within-cluster correlation dominates the moment calculation, causing the failure of the second moment method. In a sense this issue characterizes this class of CSPs.

**1.3. Combinatorial representation of clusters.** Given a random NAE-SAT instance, define a graph on the solution space by putting an edge between any two solutions differing in a single variable. Each connected component of this graph constitutes a *cluster*. Note that changing the state of a free variable can potentially free some of its neighboring variables, so the path joining two solutions in the same cluster may be complicated. Nevertheless, we have the following simple combinatorial description of clusters:

**Definition 1.2** (frozen model). On a given NAE-SAT instance, a *frozen configuration* is a vector  $\eta \in \{0, 1, \mathbf{f}\}^V$  satisfying the following properties:

- (i) No clause is invalidated by having  $k$  variables evaluating to all 0 or all 1 (f evaluates to f, so a clause involving any f-variables is automatically satisfied);
- (ii) Variables take value 0 or 1 if and only if forced to do so, that is,  $\eta_v$  takes value  $x \in \{0, 1\}$  if and only if setting  $\eta_v = \neg x$  invalidates a clause.

Variables with spin f are *free* while the rest are *rigid* or *forced*.

We claim that frozen configurations effectively encode NAE-SAT solution clusters. Indeed, the following *coarsening algorithm* projects clusters to frozen configurations:

**Definition 1.3** (coarsening algorithm). Starting from a valid NAE-SAT solution, whenever a 0/1 variable is seen to be unforced (that is, can be flipped without invalidating any clause), change its state to f. Iterate until no more variables can be set to f.

Natural analogues of this procedure can be defined for many CSPs of interest. Coarsening was originally introduced in the context of the coloring model, where it was called *whitening* [Par02]. We refer to [MMW07] for a study of the combinatorics of the coarsening procedure.

In the regime  $d_{\text{libd}} \leq d \leq d_{\text{ubd}}$ , we see from (4) that the density of variables which are already free in the initial NAE-SAT solution is very low, on the order of  $2^{-k}$ . Setting a variable to free can cause more of its neighbors to become free, so in principle the coarsening algorithm may terminate in the configuration of all frees (which clearly does satisfy the conditions for a frozen configuration). However, the following heuristic calculation suggests that this propagation of frees

will be extremely subcritical: suppose variables  $u$  and  $v$  are joined through clause  $a$ . Freeing  $u$  can change the state of  $v$  from forced to free only if (i) the other  $k - 2$  variables in  $a$  have the same evaluation, which has probability  $8/2^k$ ; and (ii)  $v$  is not forced by any of its neighboring clauses other than  $a$ , which also has probability around  $2^{-k}$  (roughly, (4) with  $d - 1$  in place of  $d$ ). Altogether this has probability  $\lesssim 4^{-k}$ , while each variable has  $d(k - 1) \ll 4^k$  neighboring variables: thus the frees propagate through the graph in the manner of a subcritical branching process. Based on this intuition we prove

**Proposition 1.4.** *In the regime  $d_{\text{ld}} \leq d \leq d_{\text{ubd}}$ , with high probability the coarsening algorithm projects all NAE-SAT solutions to frozen configurations with density of frees  $\leq 7/2^k$ .*

In reverse, if we start from a frozen configuration and try to recover a valid NAE-SAT solution, setting a free variable to 0 or 1 can force the states of neighboring variables, and cycles in the graph can create conflicts making it impossible to recover any solution. However, a simple argument shows that a valid NAE-SAT solution can be recovered as long as each connected component in this propagation contains at most one cycle. But by the same token as above, if the initial density of frees in the frozen configuration is sufficiently low, then the propagation of this effect is subcritical. As a result

**Proposition 1.5.** *Conditioned on the event that there exist frozen configurations with density of frees  $\leq 7/2^k$ , with high probability there also exist NAE-SAT solutions.*

**1.4. Sharp threshold for clusters.** In view of Propns. 1.4 and 1.5, Thm. 1 will follow by establishing a sharp threshold  $d_*$  for positivity of the “cluster partition function”

$$Z \equiv Z_n \equiv \text{number of frozen configurations on } \mathcal{G}_{n,d,k} \text{ with density of frees } \leq 7/2^k. \quad (6)$$

The sharp threshold for  $Z$ , which comprises the majority of our proof, is proved via the moment method followed by a variance reduction argument. The principle is that the frozen model does not suffer the non-rigidity (5) present in the original model, and consequently its partition function  $Z$  can be expected to have good concentration. Guided by this intuition, in §3-5 we prove

**Theorem 2.** *For  $k \geq k_0$ ,  $d_{\text{ld}} \leq d \leq d_{\text{ubd}}$ , and  ${}^*\Phi \equiv {}^*\Phi(d)$  as in (1),*

$$\mathbb{E}Z \asymp_k \exp\{n {}^*\Phi\} \quad \text{and} \quad \mathbb{E}[Z^2] \lesssim_k (\mathbb{E}Z)^2 + \mathbb{E}Z.$$

The proof of Thm. 2 comprises a large portion of the present paper. The first moment  $\mathbb{E}Z$  is calculated in §3. We refer to [MS08] for prior work upper bounding  $\mathbb{E}Z$  in random (Erdős-Rényi) 3-SAT, where the definition of  $Z$  differs but is closely analogous. In this work, in order to obtain the precise asymptotic order of  $\mathbb{E}Z$ , we identify the exact local neighborhood profile that gives the maximal contribution to the expectation. This is done by a Bethe variational principle which relates stationary points of the rate function to fixed points of certain tree recursions. A major technical difficulty is the high dimensionality of the maximization problem, and the possibility of multiple stationary points which must be ruled out. This is done by delicate *a priori* estimates which allow us to reduce the dimensionality by certain symmetry conditions.

The second moment can be understood in the same framework by regarding it as the first moment of the *pair model*: on a given NAE-SAT instance, a valid *pair configuration* is a pair  $(\eta^1, \eta^2)$  where each  $\eta^i$  is a valid frozen configuration for the same underlying instance. The dimensionality is substantially increased in the pair model compared with the original (single-copy) model, so the analysis becomes more difficult. We show in §4 that the dominant contribution comes from two local maximizers: one corresponding to pairs whose overlap distribution looks like product measure, and the other corresponding to perfectly correlated pairs — in each case, with marginals given by the first moment maximizer. The results of §3 and 4 control the moments up to polynomial prefactors, which are determined in §5 by establishing negative-definiteness of the Hessians for the first- and second-moment rate functions at their maximizers.



Recall  $(\mathbb{E}Z)^2/\mathbb{E}[Z^2] \leq \mathbb{P}(Z > 0) \leq \mathbb{E}Z$ . It follows from Thm. 2 (with Propns. 1.4 and 1.5) that for  $d > d_*$  the NAE-SAT instance  $\mathcal{G}_{n,d,k}$  is with high probability unsatisfiable, while for  $d \leq d_*$  it is satisfiable with positive probability. For  $d$  strictly below  $d_*$  we improve the latter statement from positive probability to high probability to establish the sharp transition:

**Theorem 3.** *For  $k \geq k_0$  and  $d_{\text{lb}} \leq d < d_*$ , the cluster partition function  $Z$  is positive with high probability. For  $d \leq d_{\text{lb}}$  the original NAE-SAT partition function  $Z$  is positive with high probability.*

Thm. 3 is proved in §6 by a variance reduction argument. The variance issue arises commonly in applications of the moment method, and is most often dealt with by a somewhat standard machinery known as the *subgraph conditioning method* ([RW92, RW94]; see also [JLR00, Wor99]) which “explains” the variance in terms of the short cycles in the graph. However, applying this method is technically demanding, though we refer to [GSV14] for recent advances on this front. It seems to us potentially intractable in our models due to the dimensionality of the problem.

We develop instead a novel approach of taking a certain log-transform of the partition function, and bounding the incremental fluctuations of its Doob martingale with respect to the edge-revealing filtration; each increment amounts to the effect of adding a clause. We control the variance by discrete Fourier analysis applied on the spins at the boundary of a large local neighborhood of the added clause, and we show that the main contribution comes from the degree-two Fourier coefficients which correspond to the formation of short cycles in the graph. We expect this approach to be applicable within a broad range of models.

**Acknowledgements.** We are grateful to Amir Dembo, Elchanan Mossel, Andrea Montanari, David Wilson, and Lenka Zdeborová for helpful conversations. We thank the anonymous referee for helpful comments on the manuscript.

## 2. SATISFYING ASSIGNMENTS

**2.1. Preliminaries.** The constraint structure of a  $d$ -regular  $k$ -NAE-SAT problem corresponds to a  $d$ -regular  $k$ -uniform hypergraph — equivalently, a  $(d, k)$ -regular bipartite graph  $G = (V, F, E)$ , where  $(V, F)$  gives the bipartition of the vertex set into  $n$  variables  $V$  (each degree  $d$ ) and  $m$  clauses  $F$  (each degree  $k$ ), and  $E$  is the set of edges joining clauses to variables, with  $|E| = nd = mk$  (Fig. 1). Multi-edges are permitted, i.e. a clause may include the same variable more than once.

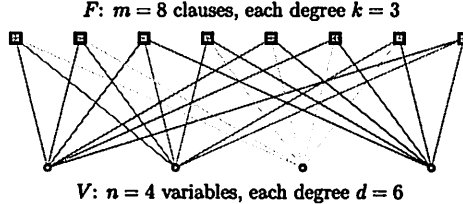


FIGURE 1.  $(d, k)$ -regular bipartite factor graph

Write  $0 \equiv \text{TRUE}$ ,  $1 \equiv \text{FALSE}$ , and  $\neg x \equiv x \oplus 1$ . A *variable assignment* is a vector  $\underline{x} \in \{0, 1\}^V$ , and a *literal assignment* is a vector  $\underline{L} \in \{0, 1\}^E$ . Use  $\partial a \equiv (v_1, \dots, v_k)$  to denote the  $k$  variables adjacent to clause  $a$ , with repetition if the graph has multi-edges. More generally, for any subset  $U$  of  $V \cup F$ , we use  $\partial U$  to denote the external boundary of  $U$ , that is to say, the vertices in  $(V \cup F) \setminus U$  which are neighboring to  $U$ . For a variable  $v \in V$  participating in a clause  $a \in F$ , the *value* of  $v$  is  $x_v$ , while the *evaluation* of  $v$  in  $a$  is  $L_{av} \oplus x_v$ . Let  $(\underline{L}\underline{x})_a \equiv (L_{av} \oplus x_v)_{v \in \partial a} \in \{0, 1\}^k$  denote the vector of evaluations for the variables incident to  $a$ ; the evaluation of  $\underline{x}$  by clause  $a$  is the boolean OR of this vector. The assignment  $\underline{x}$  is SAT if all clauses evaluate to TRUE, and is NAE-SAT if both  $\underline{x}$  and its negation  $\neg \underline{x}$  are SAT.

To generate a random problem instance, sample  $G$  according to the usual “configuration model”: starting with  $V \cup F$  as a collection of  $n + m$  isolated vertices, equip each  $v \in V$  with  $d$  half-edges, and each  $a \in F$  with  $k$  half-edges. The edge set  $E$  is given by a uniformly random matching between clause and variable half-edges. (Equivalently, take a uniformly random permutation  $\pi$  on  $[nd]$ , and for each  $1 \leq i \leq nd$  put an edge between variable  $\lfloor i/d \rfloor$  and clause  $\lfloor \pi(i)/k \rfloor$ .) Each edge  $e \in E$  is equipped with an independent random literal  $L_e$ , 0 or 1 with equal probability. The resulting random graph  $\mathcal{G}_{n,d,k}$  constitutes a random  $d$ -regular  $k$ -NAE-SAT problem instance.

For non-negative functions  $f(k, d, n)$  and  $g(k, d, n)$  we shall use any of the equivalent notations  $f = O_k(g)$ ,  $g = \Omega_k(f)$ ,  $f \lesssim_k g$ ,  $g \gtrsim_k f$  to indicate  $f \leq C(k)g$  for a finite constant  $C(k)$  depending on  $k$  but not on  $d, n$ . (In this paper, if  $f \leq C(k, d)g$  then  $f \leq C(k)g$  simply by taking the maximum of  $C(k, d)$  over the finitely many integers  $d \leq d_{\text{ubd}}(k)$ .) We drop the subscript  $k$  to indicate when we can take the same constant  $C(k) \equiv C$  for all  $k \geq k_0$ .

**2.2. Satisfiability below critical regime.** We now prove Propn. 1.1 by applying the second moment method to the NAE-SAT partition function. The analysis here is similar to that of [AM06], but due to the slightly different combinatorics of the random regular model we are not able to directly apply their result. We begin by fixing some notation. For any measures  $p, q$  defined on a discrete space  $\mathcal{S}$  we denote the entropy by  $H(p)$ , and the relative entropy by  $H(q|p)$ :

$$H(p) = - \sum_{x \in \mathcal{S}} p(x) \log p(x), \quad H(q|p) = \sum_{x \in \mathcal{S}} q(x) \log \frac{q(x)}{p(x)}.$$

If  $p, q$  are measures on the binary set  $\{0, 1\}$  then we may abuse notation and represent the measures  $p, q$  by the scalars  $x = p(1), y = q(1)$ :

$$H(x) = -x \log x - (1-x) \log(1-x), \quad H(y|x) = y \log \frac{y}{x} + (1-y) \log \frac{1-y}{1-x}.$$

Write  $\text{bin}_{n,p}(j) \equiv \binom{n}{j} p^j (1-p)^{n-j}$ .

*Proof of Propn. 1.1.* Assume throughout that  $d \leq d_{\text{ibd}}$ . By definition,  $\mathbb{E}[Z^2]$  is the sum over pairs  $\underline{x}^1, \underline{x}^2 \in \{0, 1\}^V$  of the probability that both  $\underline{x}^i$  are valid NAE-SAT solutions. By symmetry, the sum over  $\underline{x}^2$  is the same for all  $\underline{x}^1$ ; further, conditioned on  $\underline{x}^1$  being a valid solution, the probability that  $\underline{x}^2$  is also valid depends only on the number  $n\alpha$  of vertices in which the  $\underline{x}^i$  agree. Therefore

$$\mathbb{E}[Z^2] = (\mathbb{E}Z) \sum_{\alpha} \binom{n}{n\alpha} \sum_{\gamma} \mathbf{p}_{\gamma}^{\alpha} (1-\vartheta)^{m\gamma}$$

where  $\vartheta \equiv 2/(2^k - 2)$ , and  $\mathbf{p}_{\gamma}^{\alpha}$  is the probability, given vectors  $\underline{x}^1, \underline{x}^2 \in \{0, 1\}^V$  which agree in  $n\alpha$  coordinates, that there are exactly  $m\gamma$  clauses  $a \in F$  for which  $(\underline{x}^1 \oplus \underline{x}^2)_{\partial a}$  is not identically 0 or identically 1. Let  $D_1, \dots, D_n$  be i.i.d.  $\text{Bin}(k, \alpha)$  random variables: then

$$\begin{aligned} \mathbf{p}_{\gamma}^{\alpha} &= \mathbb{P}(\sum_{a=1}^m \mathbf{1}\{D_a \notin \{0, k\}\} = m\gamma \mid \sum_{a=1}^m D_a = m\alpha k) \\ &\leq \frac{\mathbb{P}(\sum_{a=1}^m \mathbf{1}\{D_a \notin \{0, k\}\} = m\gamma)}{\mathbb{P}(\sum_{a=1}^m D_a = m\alpha k)} = n^{O(1)} \exp\{-m H(\gamma|\gamma_0)\}, \end{aligned} \quad (7)$$

where  $\gamma_0 \equiv \gamma_0(\alpha) \equiv 1 - \alpha^k - (1 - \alpha)^k$ . Thus we conclude

$$\begin{aligned} \mathbb{E}[Z^2] &\leq n^{O(1)} (\mathbb{E}Z) \exp\{n \sup_{\alpha, \gamma} \mathbf{a}(\alpha, \gamma)\} \\ &\quad \text{with } \mathbf{a}(\alpha, \gamma) \equiv H(\alpha) + (d/k)[-H(\gamma|\gamma_0) + \gamma \log(1-\vartheta)]. \end{aligned}$$

For fixed  $\alpha$ ,  $\mathbf{a}$  is strictly concave in  $\gamma$  with second derivative  $-\frac{1}{\gamma(1-\gamma)} \leq -4$ , and is uniquely maximized at  $\gamma^*(\alpha) = \gamma_0(1-\vartheta)/(1-\vartheta\gamma_0)$  with optimal value

$$\begin{aligned} \bar{\mathbf{a}}(\alpha) &= H(\alpha) + (d/k) \log(1 - \gamma_0\vartheta) \\ &= (d/k) \log(1 - \vartheta) + H(\alpha) + (d/k) \log(1 + \frac{\vartheta}{1-\vartheta}[\alpha^k + (1-\alpha)^k]). \end{aligned} \quad (8)$$



The function  $\bar{a}$  is symmetric in  $\alpha$  with  $\bar{a}(1/2) = \Phi$ , the first-moment exponent of (2). We now show that  $\alpha = 1/2$  is the global maximizer for  $d \leq d_{\text{bd}}$ . Since  $\bar{a}(\alpha) - \Phi$  is nondecreasing in  $d$ , it suffices to show this for  $d = d_{\text{bd}}$ . Since  $(d/k)[\vartheta/(1-\vartheta)] = \log 2$  for  $d = d_{\text{bd}}$  we find

$$\bar{a}(\alpha) - \Phi \leq H(\alpha) + [-1 + \alpha^k + (1-\alpha)^k - 2/2^k] \log 2 + O(8^{-k}).$$

It is straightforward to calculate that for  $k^{-1}(\log k)^2 \leq \alpha \leq 1 - k^{-1}(\log k)^2$ ,

$$(\bar{a})''(\alpha) = H''(\alpha) + O(k^{-(\log k)/2}) < -3,$$

so clearly  $\alpha = 1/2$  is the unique maximizer on this interval. For  $0 \leq \alpha \leq 1/2$ ,  $H(\alpha)$  is increasing while  $\alpha^k + (1-\alpha)^k$  is decreasing, and we use this to bound

$$\sup \{[\bar{a}(\alpha) - \Phi] : k^{-3/2} \leq \alpha \leq k^{-4/5}\} \leq O(k^{-4/5} \log k) - k^{-1/2} \log 2 < 0.$$

For  $\alpha \leq k^{-3/2}$  we have  $(1-\alpha)^k = 1 - k\alpha + O(k^{1/2}\alpha)$ , therefore

$$\sup \{[\bar{a}(\alpha) - \Phi] : 2^{-3k/4} \leq \alpha \leq k^{-3/2}\} \leq \alpha(3k/4) \log 2 - \alpha k \log 2 + O(k^{1/2}\alpha) < 0.$$

Lastly, recalling  $H(x) + x \log c \leq \log(1+c) \leq c$  gives

$$\begin{aligned} \sup \{[\bar{a}(\alpha) - \Phi] : \alpha \leq 2^{-3k/4}\} &\leq -(2/2^k) \log 2 + O(k^2/2^{3k/2}) + \sup_{\alpha \leq 2^{-3k/4}} [H(\alpha) - \alpha k \log 2] \\ &\leq 2^{-k} [1 - 2 \log 2] + O(k^2/2^{3k/2}) < 0. \end{aligned}$$

Therefore  $\bar{a}$  is uniquely maximized at  $(\alpha^*, \gamma^*(\alpha^*)) = (1/2, 1-\vartheta)$  with maximal value  $\Phi$ , which proves  $\mathbb{E}[Z^2] \leq n^{O(1)} (\mathbb{E}Z)^2$ .

To remove the polynomial factor we give a more precise calculation of the probabilities  $\mathbf{p}_\gamma^\alpha$  of (7). Let  $D_1, \dots, D_m$  be i.i.d.  $\text{Bin}(k, \alpha)$  as before, and for  $0 \leq j \leq k$  let  $p_j(\alpha) \equiv (1-\vartheta)^{\mathbf{1}\{j \neq 0, k\}} \text{bin}_{k, \alpha}(j)$ . Then, since  $\bar{a}(\alpha)$  is uniquely maximized at  $\alpha = 1/2$ ,

$$\mathbb{E}[Z^2]/(\mathbb{E}Z)^2 = o(1) + \sum_{|\alpha-1/2| \leq 1/3} \frac{\binom{n}{n\alpha} \sum_{\nu} \mathbf{1}\{\sum_j j\nu_j = k\alpha\} \binom{m}{m\nu} \prod_j p_j(\alpha)^{m\nu_j}}{(\mathbb{E}Z) e^{O((mk)^{-1})/\sqrt{2\pi mk\alpha(1-\alpha)}}},$$

where the inner sum is taken over probability measures  $\nu$  on  $\{0, \dots, k\}$  such that  $m\nu$  is integer-valued. By Stirling's approximation,

$$\mathbb{E}[Z^2]/(\mathbb{E}Z)^2 = o(1) + \sum_{|\alpha-1/2| \leq 1/4} \sum_{\nu} \frac{\mathbf{1}\{\sum_j j\nu_j = k\alpha\} \exp\{n \mathbf{b}(\alpha, \nu)\}}{\mathcal{P}(\alpha, \nu) \mathbb{E}Z}$$

where  $\mathbf{b}(\alpha, \nu) \equiv H(\alpha) - (d/k) \sum_j \nu_j \log[\nu_j/p_j(\alpha)]$  is strictly concave in  $(\alpha, \nu)$ , and the correction term  $\mathcal{P}(\alpha, \nu)$  is  $n^{O(1)}$  in general, and is  $\asymp_k n^{(k+1)/2}$  for  $\nu$  satisfying  $\max_j 1/\nu_j \lesssim_k 1$ . It is easily seen that this is indeed satisfied by  $\arg \max_{\nu} \mathbf{b}(\alpha, \nu)$  for  $1/4 \leq \alpha \leq 3/4$ , so it follows using the strict concavity of  $\mathbf{b}$  that

$$\mathbb{E}[Z^2]/(\mathbb{E}Z)^2 = o(1) + O_k(1) \sum_{|\alpha-1/2| \leq 1/4} \frac{\exp\{n \sup_{\nu} \mathbf{b}(\alpha, \nu)\}}{n^{1/2} \mathbb{E}Z}.$$

Of course  $\sup_{\nu} \mathbf{b}(\alpha, \nu)$  need not be concave in  $\alpha$ , however, since we previously took an upper bound on  $\mathbf{p}_\gamma^\alpha$ ,  $\sup_{\nu} \mathbf{b}(\alpha, \nu) \leq \bar{a}(\alpha)$  which is strictly concave near  $\alpha = 1/2$  with global maximum  $\bar{a}(1/2) = \Phi$ . This proves  $\mathbb{E}[Z^2] \lesssim_k (\mathbb{E}Z)^2$  for  $d \leq d_{\text{bd}}$ .  $\square$

**2.3. Coarsening algorithm and frozen model.** In view of Propn. 1.1 we hereafter assume unless indicated otherwise that  $k \geq k_0$  where  $k_0$  is a large absolute constant,

$$d_{\text{bd}} \leq d = (2^{k-1} - \rho)k \log 2 \leq d_{\text{ubd}} \quad (0 \leq \rho \leq 2), \quad \text{so } \Phi = 2^{-k}(2\rho - 1) \log 2 + O(4^{-k}). \quad (9)$$

In this regime, we define the following algorithm to map a satisfying variable assignment  $\underline{x} \in \{0, 1\}^V$  to a *coarsened* configuration  $\underline{\eta} \equiv \underline{\eta}(\underline{x}) \in \{0, 1, \mathbf{f}\}^V$ . In the coarsened model, 0 and 1 indicate variables which are ‘‘rigid’’ or ‘‘forced’’ while  $\mathbf{f}$  indicates variables which are ‘‘free,’’ as follows:

**Definition 2.1.** Given  $\eta \in \{0, 1, \mathbf{f}\}^V$ , a clause-variable edge  $(av)$  is said to be  $\eta$ -forcing (or simply forcing if  $\eta$  is unambiguous) if  $\eta_{\partial a} \in \{0, 1\}^k$  with  $L_{av} \oplus \eta_v = \neg L_{aw} \oplus \eta_w$  for all  $w \in \partial a \setminus v$ . We also say that  $a$  is  $\eta$ -forcing and  $v$   $\eta$ -forced; note each clause can have at most one  $\eta$ -forcing edge.<sup>1</sup> Given  $\eta_{\partial a} \in \{0, 1\}^k$ , of the  $2^k - 2$  valid configurations of  $\underline{L}_a$  there are exactly two which are  $(av_i)$ -forcing for each  $1 \leq i \leq k$ , with the remaining  $2^k - 2 - 2k$  configurations not forcing to any  $v_i$ . A variable which is not  $\eta$ -forced is said to be  $\eta$ -free.

**Coarsening algorithm.**

Set  $\eta^0 \equiv \underline{x}$ . For  $t \geq 0$ , if there exists  $v \in V$  which has  $\eta_v^t \neq \mathbf{f}$  but which is not  $\eta^t$ -forced, then take the first<sup>2</sup> such  $v$  and set  $\eta_v^{t+1} = \mathbf{f}$ . Set  $\eta_w^{t+1} = \eta_w^t$  for all  $w \neq v$ .

Iterate until the first time  $t_1$  that no such vertex  $v$  remains.

Denote the terminal configuration  $\eta \equiv \eta(\underline{x}) \equiv \eta^{t_1}$ .<sup>3</sup> Let  $Z_{\geq n\beta}$  denote the contribution to  $Z$  from assignments  $\underline{x} \in \{0, 1\}^V$  such that the coarsened configuration  $\eta(\underline{x}) \in \{0, 1, \mathbf{f}\}^V$  has more than  $n\beta$  free variables.

**Proposition 2.2.** *In regime (9),  $\mathbb{E}Z_{\geq n\beta}$  is exponentially small in  $n$  for  $\beta = 7/2^k$ .*

*Proof.* By symmetry,  $\mathbb{E}Z_{\geq n\beta} = (\mathbb{E}Z) \mathbf{f}_{n\beta}$  where  $\mathbf{f}_{n\beta}$  denotes the probability, conditioned on  $\underline{x} = \underline{0}$  being a valid NAE-SAT solution, that its coarsening  $\eta$  has at least  $n\beta$  free variables.

We simulate the coarsening algorithm as follows: of the  $nd$  half-edges incident to variables, choose  $e_1, \dots, e_m$  uniformly at random (with random ordering) to be *potentially forcing*. Edge  $e_a$  corresponds to clause  $a$ , though here the clauses are not explicitly formed. Conditioned on  $\underline{x}$  being a valid solution, each clause independently has probability  $\vartheta \equiv 2k/(2^k - 2)$  to be  $\underline{x}$ -forcing (cf. Defn. 2.1): therefore set each  $e_a$  to be *initially forcing* with probability  $\vartheta$ , independently over  $a$ . Then, for each  $t \geq 0$ , if there exists  $v \in V$  which is incident to no (remaining) initially forcing half-edge, then take the first such  $v$  and

- (i) Delete all  $d_v^t$  remaining potentially forcing half-edges incident to  $v$ ; and
- (ii) Delete the first  $d - d_v^t$  potentially forcing half-edges among all those remaining.

The interpretation is that the coarsening algorithm sets  $v$  to be a free variable at stage  $t$ . Thus the  $d - d_v^t$  clauses incident to  $v$  and potentially forcing to other variables can no longer be forcing, so we remove these clauses from consideration (step (ii)).<sup>4</sup>

Say a variable  $v$  is  $t$ -free if it has no initially forcing half-edges remaining after  $nt$  iterations of the above procedure. Since initially forcing edges are deleted in order,  $v$  must avoid the set  $E_t$  of initially forcing edges  $e_a$  with index  $a > ndt$ . If there are  $\geq nt$  free variables in the coarsened configuration  $\eta$ , then the above process must survive at least  $nt$  iterations. The law of  $|E_t|$  is  $\text{Bin}(m - ndt, \vartheta)$ , so (by a union bound)

$$\begin{aligned} \mathbf{f}_{nt} &\leq \binom{n}{nt} \mathbb{E} \left[ \binom{nd(1-t)}{|E_t|} / \binom{nd}{|E_t|} \right] \leq \binom{n}{nt} \mathbb{E}[(1-t)^{|E_t|}] \\ &= n^{O(1)} \exp\{n[H(t) + d(k^{-1} - t) \log(1 - \vartheta t)]\}. \end{aligned}$$

If  $t = C/2^k$  with  $C \asymp 1$  then  $\mathbf{f}_{nt} \leq n^{O(1)} \exp\{n(C/2^k)[1 - \log C + O(k^2/2^k)]\}$ . Then recalling (9) we have  $\mathbb{E}Z_{\geq nt} \leq e^{n\Phi} \mathbf{f}_{nt}$  exponentially small in  $n$  for  $C = 7$ .  $\square$

**Definition 2.3.** We say  $\eta \in \{0, 1, \mathbf{f}\}^V$  is a  $0/1/\mathbf{f}$  frozen model configuration on  $(G, \underline{L})$  if

- (a) No clause  $a \in F$  is *unsatisfied* (meaning  $\eta_{\partial a} \in \{0, 1\}^k$  with  $(\underline{L}\eta)_a$  identically 0 or 1);

<sup>1</sup>Recall that  $\partial a$  indicates the neighbors of  $a$  with multiplicity; for example, if  $\partial a = (v, v, v, w, \dots, w)$  with  $\eta_v \neq \eta_w$  and  $\underline{L}_a = (0^k)$ , the clause is not considered  $\eta$ -forcing.

<sup>2</sup>First with respect to the ordering on  $V = [n]$ . This choice is useful for the analysis of the coarsening algorithm, but in fact the terminal configuration does not depend on the order in which variables are set to free.

<sup>3</sup>We could define a *cluster* of NAE-SAT solutions to be the pre-image of any  $\eta$  under the coarsening algorithm.

<sup>4</sup>Step (i) does not delete any initially forcing half-edges, but step (ii) can.



- (b) Each variable  $v \in V$  has  $\eta_v \neq \mathbf{f}$  if and only if there is a clause  $a \in \partial v$  with  $\eta_{\partial a} \in \{0, 1\}^k$  and  $L_{av} \oplus \eta_v = -L_{aw} \oplus \eta_w$  for all  $w \in \partial a \setminus v$  (cf. Defn. 2.1).

Some of our computations are simplified by working with the image of the 0/1/f frozen model under the projection  $\{0, 1\} \mapsto \mathbf{r}$ , hereafter  $\mathbf{r}/\mathbf{f}$  frozen model.

Let  $Z_{n\beta}$  denote the frozen model partition function on  $(G, \underline{L})$  restricted to configurations with exactly  $n\beta$  f-vertices. In view of Propn. 2.2, in regime (9) we hereafter restrict all consideration to the truncated 0/1/f frozen model partition function

$$Z \equiv \sum_{t \leq \beta_{\max}} Z_{nt}, \quad \beta_{\max} \equiv 7/2^k. \quad (10)$$

We will show in §7 that restricted frozen model solutions indeed correspond to true NAE-solutions.<sup>5</sup>

### 3. FIRST MOMENT OF FROZEN MODEL

In this section we identify the leading exponential order  $^*\Phi = \lim_{n \rightarrow \infty} n^{-1} \log \mathbb{E}Z$  of the first moment of the (truncated) frozen model partition function (10). The random  $(d, k)$ -regular bipartite factor graph  $G \equiv \mathcal{G}_{n,d,k}$  converges locally weakly (in the sense of [BS01, AL07]) to the infinite  $(d, k)$ -regular tree  $T_{d,k}$  — the infinite tree with levels indexed by  $\mathbb{Z}_{\geq 0}$  such that all vertices at even integer levels are of degree  $d$  (variables) and all vertices at odd integer levels are of degree  $k$  (clauses). Our calculation is based on a variational principle which relates the exponent  $^*\Phi$  to a certain class of Gibbs measures for the frozen model on  $T_{d,k}$  which are characterized by fixed-point recursions. In fact the recursions can have multiple solutions, and much of the work goes into identifying (via *a priori* estimates) the unique fixed point which gives rise to  $^*\Phi$ . We begin by introducing the Gibbs measures which will be relevant for the variational principle.

**3.1. Frozen model tree recursions.** We shall specify a Gibbs measure  $\nu$  on  $T_{d,k}$  by defining a consistent family of finite-dimensional distributions  $\nu_U$  on finite subtrees  $U \subseteq T_{d,k}$ . A typical manner of specifying  $\nu_U$  is to specify a “boundary law” for the configuration on  $\partial U$ , and then to define  $\nu_U$  as an appropriate finite-volume Gibbs measure on  $U$  conditioned on the  $\partial U$ -configuration. The family  $(\nu_U)_U$  is consistent if  $\nu_S$  is a marginal  $\nu_U$  whenever  $S \subseteq U$ .

In our setting some difficulty is imposed by the fact that the frozen model is not a factor model (or Markov random field) in the conventional sense that  $\eta|_A$  and  $\eta|_B$  are conditionally independent given the configuration  $\eta|_C$  on any subset  $C$  separating  $A$  from  $B$  — in particular, given the variable spins at level  $2t$  of  $T_{d,k}$ , whether a variable at level  $2(t-1)$  is permitted to take spin  $\mathbf{f}$  depends on whether its neighboring 0’s and 1’s in level  $2t$  are forced by clauses in level  $2t+1$ .

We shall instead specify Gibbs measures for the frozen model via a message-passing system, as follows. First sample uniformly random literals  $\underline{L}$  on the edges of the tree. Given the literals, each variable  $v$  will send a message  $\sigma_{v \rightarrow a}$  to each neighboring clause  $a \in \partial v$  which represents the “state of  $v$  ignoring  $a$ ”, and will receive in return a message  $\sigma_{a \rightarrow v}$  representing the “state of  $a$  ignoring  $v$ .” That is,  $\sigma_{v \rightarrow a}$  will be a function  $\hat{m}_{d-1}$  of  $d-1$  incoming messages  $(\sigma_{b \rightarrow v})_{b \in \partial v \setminus a}$ , and likewise  $\sigma_{a \rightarrow v}$  will be a function  $\hat{m}_{kv}$  (which will involve the literals at  $a$ ) of  $k-1$  incoming messages  $(\sigma_{w \rightarrow a})_{w \in \partial a \setminus v}$ . The actual state  $\eta_v$  of  $v$  is then a function  $\hat{m}_d$  of all its incoming messages  $\underline{\sigma}_{\partial v \rightarrow v} \equiv (\sigma_{u \rightarrow v})_{u \in \partial v}$ ; the configuration may be invalidated if any variable receives conflicting incoming messages.

Say at the boundary of the subtree  $U$  we are given a vector  $\eta^\dagger \equiv (\eta_x^\dagger)_x$  of incoming messages, where  $x$  runs over the external boundary  $\partial U$  and  $\eta_x^\dagger$  stands for the message  $\sigma_{x \rightarrow y}$  from  $x$  to its unique neighbor  $y \in U$ . Then there is *at most one* valid completion of  $\eta^\dagger$  to a bi-directional message configuration on  $U$ , which is obtained simply by iterating the maps  $\hat{m}_{d-1}, \hat{m}_{kv}$  from the leaves inward.

<sup>5</sup>Some truncation is indeed necessary: the identically-f vector is a valid configuration of the unrestricted frozen model, and in fact it turns out that the dominant contribution to the partition function of the unrestricted 0/1/f frozen model comes from configurations with much higher density of free variables (roughly  $\approx (\log k)/k$ ) — hence not corresponding to NAE-solutions.

The measure  $\nu_U$  can then be specified by giving the law of the boundary messages  $\eta^\dagger$ : our choice will be to take the  $\eta_x^\dagger$  mutually independent, distributed according to law  $\hat{q}$  if  $x$  is a variable, and law  $\hat{q}$  if  $x$  is a clause. Consistency of the family  $(\nu_U)_U$  will then amount to fixed-point relations on  $\hat{q}, \hat{q}$ , as we see below.

The message-passing rules for our frozen model are as follows:

1. *Vertex message-passing rule*  $\hat{m}_D : \{0, 1, \mathbf{f}\}^D \rightarrow \{0, 1, \mathbf{f}\}$ : output

$$\left\{ \begin{array}{l} \mathbf{f} \text{ if all } D \text{ incoming messages are } \mathbf{f}; \\ 0 \text{ if at least one } 0 \text{ but no } 1\text{'s incoming}; \\ 1 \text{ if at least one } 1 \text{ but no } 0\text{'s incoming}; \\ \text{UNSAT otherwise (i.e. both } 0, 1 \text{ incoming)}. \end{array} \right.$$

2. *Clause message-passing rule*  $\hat{m}_{av} : \underline{\sigma}_{\partial a \setminus v \rightarrow a} \mapsto \sigma_{a \rightarrow v}$ : output

$$\left\{ \begin{array}{l} 0 \text{ if } L_{av} \oplus 0 = \neg L_{aw} \oplus \sigma_{w \rightarrow a} \text{ for all } w \in \partial a \setminus v; \\ 1 \text{ if } L_{av} \oplus 1 = \neg L_{aw} \oplus \sigma_{w \rightarrow a} \text{ for all } w \in \partial a \setminus v; \\ \mathbf{f} \text{ otherwise.} \end{array} \right.$$

We then define

$$Z_U \nu_U(\underline{L}_U, \underline{\eta}_U, \underline{\eta}^\dagger) = I(\underline{L}_U, \underline{\eta}_U, \underline{\eta}^\dagger) \prod_{v \in V \cap \partial U} \hat{q}(\eta_v^\dagger) \prod_{a \in F \cap \partial U} \hat{q}(\eta_a^\dagger) \quad (11)$$

where  $I(\underline{L}_U, \underline{\eta}_U, \underline{\eta}^\dagger)$  is the indicator that, under the literals  $\underline{L}(t)$ , the boundary messages  $\eta^\dagger$  can be completed (no UNSAT messages) to a valid message configuration on  $U$ , which in turn corresponds to the frozen configuration  $\underline{\eta}_U$ . The marginal at the root vertex  $o$  is then given by

$$\nu_o(\eta_o = x) = \frac{(\hat{q}_x + \hat{q}_x)^d - (\hat{q}_x)^d}{(\hat{q}_x + \hat{q}_0)^d + (\hat{q}_x + \hat{q}_1)^d - (\hat{q}_x)^d} \quad \text{for } x = 0 \text{ or } 1,$$

with the remaining probability going to  $\eta_o = \mathbf{f}$ . The measures  $(\nu_U)_U$  are consistent if and only if  $q \equiv (\hat{q}, \hat{q})$  satisfies the *frozen model recursions*

$$\begin{aligned} \hat{q}_0 &= \hat{q}_1 = (2/2^k)(\hat{q}_0 + \hat{q}_1)^{k-1} = 1 - \hat{q}_x/2, \\ \hat{q}_0 &= \hat{q}_1 = \frac{(\hat{q}_x + \hat{q}_0)^{d-1} - (\hat{q}_x)^{d-1}}{(\hat{q}_x + \hat{q}_0)^{d-1} + (\hat{q}_x + \hat{q}_1)^{d-1} - (\hat{q}_x)^{d-1}} = 1 - \hat{q}_x/2. \end{aligned} \quad (12)$$

For example, the equations for  $\hat{q}$  can be obtained by comparing  $\nu_o$  and  $\nu_U$  where  $U$  is the subgraph induced by  $\{o, a\}$  for some clause  $a \in \partial v$ . In particular one sees that  $\hat{q}_0 = \hat{q}_1$  due to the randomness of the literals in clause  $a$ .

**Lemma 3.1.** *In the regime  $\hat{q}_x \lesssim 2^{-k}$ , the recursion (12) has a unique solution  $q^*$ , which furthermore satisfies  $2^k(\hat{q}_x)^* = 1/2 + O(k^2/2^k)$ .*

*Proof.* Writing  $q \equiv 1 - \hat{q}_x$  and  $v \equiv \hat{q}_x/(\hat{q}_0 + \hat{q}_1)$ , we see that a solution of (12) corresponds to a solution of the equations

$$q = q_{d-1}(v) \equiv \frac{2 - 2v^{d-1}}{2 - v^{d-1}}, \quad v = v_{k-1}(q) \equiv \frac{1 - 2(q/2)^{k-1}}{1 - (q/2)^{k-1}}. \quad (13)$$

If  $1 - q \lesssim 2^{-k}$  then  $v_{k-1}(q) = 1 - 2/2^k + O(k/4^k)$ , therefore  $v_{k-1}(q)^{d-1} = 2^{-k} + O(k^2/4^k)$  and then  $q_{d-1} \circ v_{k-1}(q) = 1 - 2^{-k-1} + O(k^2/4^k)$ . In this regime we also calculate

$$v'_{k-1}(q) = -\frac{(k-1)(q/2)^{k-1}}{q[1 - (q/2)^{k-1}]^2} \approx k2^{-k}, \quad q'_{d-1}(v) = -\frac{2(d-1)v^{d-1}}{v(2 - v^{d-1})^2} \approx k,$$

thus  $(q_{d-1} \circ v_{k-1})' \approx k^2/2^k$  so in this regime (13) must have the unique solution as claimed.  $\square$

**Remark 3.2.** Note that if  $\nu$  is the Gibbs measure on  $T_{d,k}$  corresponding to a solution  $q^*$  of (12), then  $\nu(\sigma_o \neq \mathbf{f})$  is a fixed point of  $q_d \circ v_{k-1}$ . In the regime of Lem. 3.1 the fixed points of  $q_d \circ v_{k-1}$  and  $q_{d-1} \circ v_{k-1}$  are nearly identical, so in view of Propn. 2.2 we are justified in restricting attention to fixed points with  $\hat{q}_t \lesssim 2^{-k}$ .

**3.2. Auxiliary model.** On the tree  $T_{d,k}$ , the frozen model configuration  $\eta$  can be uniquely recovered from the configuration  $\underline{\sigma}$  of messages on all the directed edges: each vertex spin  $\eta_v$  is determined by applying  $\hat{m}_d$  to the incoming messages. We refer to  $\underline{\sigma}$  as the *auxiliary* configuration, and we now observe that we can define a model on auxiliary configurations on  $(d, k)$ -regular bipartite graphs which is in bijection with the frozen model but has the advantage of being a *factor model* in a relatively simple sense.

The spins of the auxiliary model on the bipartite factor graph are the bidirectional messages  $\sigma_{va} \equiv \sigma_{av} \equiv (\sigma_{v \rightarrow a}, \sigma_{a \rightarrow v})$ , taking values in the alphabet  $\mathcal{M} \equiv \{0, 1, \mathbf{f}\}^2 \setminus \{01, 10\}$ . Write  $\underline{\sigma}_v$  for the  $d$ -tuple of spins on the edges incident to variable  $v \in V$ , and write  $\underline{\hat{\sigma}}_a$  for the pair of spins on the edges incident to clause  $a \in F$ .

In the *auxiliary model*, each configuration  $\underline{\sigma} \in \mathcal{M}^E$  receives the factor model weight

$$\Psi(\underline{\sigma}) \equiv \Psi_{G, \underline{L}}(\underline{\sigma}) \equiv \prod_{v \in V} \hat{\varphi}(\underline{\hat{\sigma}}_v) \prod_{a \in F} \hat{\varphi}^a(\underline{\hat{\sigma}}_a) \quad (14)$$

where the variable factor weight  $\hat{\varphi}(\underline{\hat{\sigma}}_v)$  is simply the indicator that each outgoing message  $\sigma_{v \rightarrow a}$  is determined by the message-passing rule  $\hat{m}_{d-1}$  from the incoming messages  $\sigma_{b \rightarrow v}$ ,  $b \in \partial v \setminus a$ ; and likewise the clause factor weight  $\hat{\varphi}^a(\underline{\hat{\sigma}}_a)$  is the indicator that each outgoing message  $\sigma_{a \rightarrow v}$  is determined by the message-passing rule  $\hat{m}_{av}$  from the incoming messages  $\sigma_{w \rightarrow a}$ ,  $w \in \partial a \setminus v$ . Then, with  $-\mathbf{f} \equiv \mathbf{f}$ , we have  $\hat{\varphi}^a(\underline{\hat{\sigma}}_a) = \hat{\varphi}^\circ(\underline{\hat{\sigma}}_a \oplus \underline{L}_a)$  where  $\hat{\varphi}$  and  $\hat{\varphi}^\circ$  are given explicitly by

$$\hat{\varphi}(\underline{\hat{\sigma}}) \equiv \begin{cases} 1, & \underline{\hat{\sigma}} = (\mathbf{f}\mathbf{f}^d), \\ 1, & \underline{\hat{\sigma}} \in \text{Per}[(\mathbf{f}0, 0\mathbf{f}^{d-1})], \\ 1, & \underline{\hat{\sigma}} \in \text{Per}[(\mathbf{f}1, 1\mathbf{f}^{d-1})], \\ 1, & \underline{\hat{\sigma}} \in \text{Per}[(00^j, 0\mathbf{f}^{d-j})_{j \geq 2}], \\ 1, & \underline{\hat{\sigma}} \in \text{Per}[(11^j, 1\mathbf{f}^{d-j})_{j \geq 2}], \\ 0, & \text{else;} \end{cases}, \quad \hat{\varphi}^\circ(\underline{\hat{\sigma}}) \equiv \begin{cases} 1, & \underline{\hat{\sigma}} \in \text{Per}[(00 \text{ or } \mathbf{f}0, 1\mathbf{f}^{k-1})] \\ 1, & \underline{\hat{\sigma}} \in \text{Per}[(11 \text{ or } \mathbf{f}1, 0\mathbf{f}^{k-1})] \\ 1, & \underline{\hat{\sigma}} \in \text{Per}[(0\mathbf{f}^j, 1\mathbf{f}^{k-j})_{2 \leq j \leq k-2}], \\ 1, & \underline{\hat{\sigma}} \in \text{Per}[(\mathbf{f}\mathbf{f}, 0\mathbf{f}^j, 1\mathbf{f}^{k-1-j})_{1 \leq j \leq k-2}], \\ 1, & \underline{\hat{\sigma}} \in \text{Per}[(\mathbf{f}\mathbf{f}^j, 0\mathbf{f}^s, 1\mathbf{f}^{k-1-s})_{j \geq 2}], \\ 0, & \text{else;} \end{cases}, \quad (15)$$

with  $\text{Per}(\underline{\sigma})$  the set of permutations of  $\underline{\sigma}$ . We refer to this as the *factor model with specification*  $\varphi \equiv (\hat{\varphi}, \hat{\varphi}^\circ)$ .

**Remark 3.3.** The frozen model is in exact bijection with the auxiliary model. Given an auxiliary configuration  $\underline{\sigma} \in \mathcal{M}^E$ , the corresponding frozen configuration  $\eta$  is given by coordinate-wise application of  $\hat{m}_d$ . The inverse mapping  $\eta \mapsto \underline{\sigma}$  can be defined as follows: first determine the clause-to-variable messages by setting  $\sigma_{a \rightarrow v}$  to be  $\eta_v$  if  $(av)$  is  $\eta$ -forcing and  $\mathbf{f}$  otherwise, equivalently  $\sigma_{a \rightarrow v} = \hat{m}_{av}(\eta_{\partial a \setminus v})$ . Then determine the variable-to-clause messages  $\sigma_{v \rightarrow a}$  by applying  $\hat{m}_{d-1}$  (since we assumed  $\eta$  is a valid frozen model configuration,  $v$  cannot receive conflicting incoming messages  $\sigma_{a \rightarrow v} = 0$  and  $\sigma_{b \rightarrow v} = 1$ ).

**Definition 3.4.** The  $0/1/\mathbf{f}$  *auxiliary model* on  $G$  is defined to be the average of the auxiliary model (14) over all literal configurations  $\underline{L}$ . The  $r/\mathbf{f}$  *auxiliary model* is the image of the  $0/1/\mathbf{f}$  auxiliary model under the projection  $\Pi : \{0, 1\} \mapsto r, \mathbf{f} \mapsto \mathbf{f}$ .

It is easily seen that the  $0/1/\mathbf{f}$  auxiliary model is again a factor model on  $G$ , with variable factor  $\hat{\varphi}$  as before and clause factor  $\hat{\varphi}(\underline{\hat{\sigma}}) \equiv 2^{-k} \sum_{\underline{L}} \hat{\varphi}^\circ(\underline{\hat{\sigma}} \oplus \underline{L})$ . Further,  $\hat{\varphi}(\underline{\hat{\sigma}})$  and  $\hat{\varphi}(\underline{\hat{\sigma}})$  depend on  $\underline{\hat{\sigma}}$  and  $\underline{\hat{\sigma}}$  only through their projections under  $\Pi$ , so we conclude that  $r/\mathbf{f}$  auxiliary model on  $G$  is a factor

model with specification

$$\hat{\varphi}(\hat{\sigma}) = \begin{cases} 1, & \hat{\sigma} = (\text{ff}^d), \\ 2, & \hat{\sigma} \in \text{Per}(\text{fr}, \text{rf}^{d-1}), \\ 2, & \hat{\sigma} \in \text{Per}[(\text{rr}^j, \text{rf}^{d-j})_{j \geq 2}], \\ 0, & \text{else;} \end{cases} \quad \hat{\varphi}(\hat{\sigma}) \equiv 2^{-k} \begin{cases} 2, & \hat{\sigma} \in \text{Per}[(\text{rr} \text{ or } \text{fr}, \text{rf}^{k-1})], \\ 2^k - 2 - 2k, & \hat{\sigma} = (\text{rf}^k), \\ 2^k - 4, & \hat{\sigma} \in \text{Per}(\text{ff}, \text{rf}^{k-1}), \\ 2^k, & \hat{\sigma} \in \text{Per}[(\text{ff}^j, \text{rf}^{k-j})_{j \geq 2}], \\ 0, & \text{else.} \end{cases} \quad (16)$$

**3.3. Bethe variational principle.** The primary purpose of defining the auxiliary model is that it gives us the following approach for calculating  $\mathbb{E}Z$ . Given an auxiliary configuration  $\underline{\sigma}$ , consider the normalized empirical measures

$$\begin{aligned} \hat{h}(\hat{\sigma}) &\equiv n^{-1} \sum_{v \in V} \mathbf{1}\{\hat{\sigma}_v = \hat{\sigma}\} & (\hat{\sigma} \in \mathcal{M}^d) & \quad \text{variable empirical measure;} \\ \hat{h}(\hat{\sigma}) &\equiv m^{-1} \sum_{a \in F} \mathbf{1}\{\hat{\sigma}_a = \hat{\sigma}\} & (\hat{\sigma} \in \mathcal{M}^k) & \quad \text{clause empirical measure.} \end{aligned}$$

We regard  $\mathbf{h} \equiv (\hat{h}, \hat{h})$  as a vector indexed by  $\text{supp } \varphi \equiv (\text{supp } \hat{\varphi}, \text{supp } \hat{\varphi})$ . For  $\sigma \in \mathcal{M}$  and  $\hat{\sigma} \in \text{supp } \hat{\varphi}$  let  $\hat{H}_{\sigma, \hat{\sigma}}$  denote the number of appearances of  $\sigma$  in  $\hat{\sigma}$ , and similarly write  $\hat{H}_{\sigma, \hat{\sigma}}$  for the number of appearances of  $\sigma$  in  $\hat{\sigma}$ . For  $\mathbf{h}$  to correspond to a valid configuration  $\underline{\sigma}$ , the variable and clause empirical measures must give rise to the same edge marginals

$$\bar{h} = d^{-1} \hat{H} \hat{h} = k^{-1} \hat{H} \hat{h}, \quad \bar{h}(\sigma) \equiv (nd)^{-1} \sum_{(va) \in H} \mathbf{1}\{\sigma_{va} = \sigma\}.$$

**Definition 3.5.** Given  $\varphi \equiv (\hat{\varphi}, \hat{\varphi})$  let  $\Delta$  denote the space of probability measures  $\mathbf{h} \equiv (\hat{h}, \hat{h})$  on  $\text{supp } \varphi$  (that is,  $\hat{h}$  is a probability measure on  $\text{supp } \hat{\varphi}$  while  $\hat{h}$  is a probability measure on  $\text{supp } \hat{\varphi}$ ) such that

- (i)  $(\hat{h}, (d/k)\hat{h})$  lies in the kernel of matrix  $H_{\Delta} \equiv (\hat{H} \quad -\hat{H})$ , and
- (ii)  $\mathbf{h}(\hat{m}_d(\hat{\sigma}) = \mathbf{f}) \leq \beta_{\max}$  (cf. (10)).

Let  $\hat{s} \equiv |\text{supp } \hat{\varphi}|$ ,  $\hat{s} \equiv |\text{supp } \hat{\varphi}|$ , and  $\bar{s} \equiv |\text{supp } \hat{\varphi}| = |\mathcal{M}|$ : we shall show (Lem. 6.6) that  $H_{\Delta}$  is surjective, therefore  $\Delta$  is an  $(\hat{s} + \hat{s} - \bar{s} - 1)$ -dimensional space.

The expected number of auxiliary configurations on  $\mathcal{G}_{n,d,k}$  with empirical measure  $\mathbf{h}$  is

$$\mathbb{E}Z(\mathbf{h}) = \frac{\binom{n}{n\hat{h}} \binom{m}{m\hat{h}}}{\binom{nd}{nd\hat{h}}} \varphi^{n\hat{h}} \hat{\varphi}^{m\hat{h}} \equiv \frac{n!m! \prod_{\sigma} (nd\bar{h}(\sigma))!}{(nd)!} \prod_{\hat{\sigma}} \frac{\hat{\varphi}(\hat{\sigma})^{n\hat{h}(\hat{\sigma})}}{(n\hat{h}(\hat{\sigma}))!} \prod_{\hat{\sigma}} \frac{\hat{\varphi}(\hat{\sigma})^{m\hat{h}(\hat{\sigma})}}{(m\hat{h}(\hat{\sigma}))!}.$$

Stirling's formula gives  $\mathbb{E}Z(\mathbf{h}) = n^{O(1)} \exp\{n\Phi(\mathbf{h})\}$  where

$$\Phi(\mathbf{h}) \equiv \sum_{\hat{\sigma}} \hat{h}(\hat{\sigma}) \log \frac{\hat{\varphi}(\hat{\sigma})}{\hat{h}(\hat{\sigma})} + (d/k) \sum_{\hat{\sigma}} \hat{h}(\hat{\sigma}) \log \frac{\hat{\varphi}(\hat{\sigma})}{\hat{h}(\hat{\sigma})} - d \sum_{\sigma} \bar{h}(\sigma) \log \frac{1}{\bar{h}(\sigma)}. \quad (17)$$

If further  $\min \bar{h} \gtrsim_k 1$  as  $n \rightarrow \infty$ , then

$$\mathbb{E}Z(\mathbf{h}) = \frac{e^{O_k(n^{-1})}}{(2\pi n)^{(\hat{s} + \hat{s} - \bar{s} - 1)/2}} \underbrace{\left[ \frac{\prod_{\sigma} d\bar{h}(\sigma)}{k \prod_{\hat{\sigma}} \hat{h}(\hat{\sigma}) \prod_{\hat{\sigma}} \left(\frac{d}{k}\hat{h}(\hat{\sigma})\right)} \right]^{1/2}}_{\mathcal{D}(\mathbf{h})} \exp\{n\Phi(\mathbf{h})\} \quad (18)$$

The first moment of frozen model configurations is  $\mathbb{E}Z = \sum_{\mathbf{h} \in \Delta} \mathbb{E}Z(\mathbf{h})$ . The aim of this section is to compute the exponent  ${}^* \Phi = \lim_n n^{-1} \log \mathbb{E}Z$  by determining the maximizer  ${}^* \mathbf{h} \equiv ({}^* \hat{h}, {}^* \hat{h})$  of  $\Phi$  on  $\Delta$ . Observe it is clear from the functional form of  $\Phi$  that  ${}^* \hat{h}$  and  ${}^* \hat{h}$  must be *symmetric* functions on  $\mathcal{M}^d$  and  $\mathcal{M}^k$  respectively.

If  ${}^* \mathbf{h}$  lies in the interior  $\Delta^\circ$  of  $\Delta$  then it must be a stationary point for  $\Phi$ . Such points correspond to a generalization of the tree Gibbs measures considered in §3.1, where the boundary conditions are specified by a law on incoming *and outgoing* messages, as follows. Given a finite subtree  $U$  of the infinite tree  $T_{d,k} = (V, F, E)$ , sample uniformly random literals  $\underline{L}_U$  on the edges of  $U$  as



before. Let  $U^\circ$  denote the interior vertices of  $U$  (the vertices with no neighbors in  $T \setminus U$ ). Let  $\delta U$  denote the internal edge boundary of  $U$ , that is, the edges  $(xy)$  where  $x \in U^\circ$  and  $y \in U \setminus U^\circ$ . Let  $\delta_V U \subseteq \delta U$  denote the subset of such  $(xy)$  where  $x \in V$ ; and let  $\delta_F U$  denote the rest, with  $x \in F$ . For probability measures  $\dot{h}, \hat{h}$  on  $\mathcal{M}$  we define the measures

$$Z_U \nu_U(\underline{L}_U, \underline{\sigma}_U) = \prod_{v \in V \cap U^\circ} \dot{\varphi}(\dot{\sigma}_v) \prod_{a \in F \cap U^\circ} \dot{\varphi}^a(\dot{\sigma}_a) \prod_{e \in \delta_V U} \dot{h}_{\sigma_e} \prod_{e \in \delta_F U} \hat{h}_{\sigma_e}, \quad (19)$$

with  $Z_U$  the normalizing constant which makes  $\nu_U$  a probability measure. This generalizes the definition of  $\nu_U$  in (11) by taking  $\dot{h}_{\eta\eta'}$  proportional to  $\dot{q}_\eta$  and  $\hat{h}_{\eta\eta'}$  proportional to  $\hat{q}_{\eta'}$ , i.e.

$$\dot{h}_{\eta\eta'} = \dot{q}_\eta / (2 + \dot{q}_\eta), \quad \hat{h}_{\eta\eta'} = \hat{q}_{\eta'} / (2 + \hat{q}_{\eta'}). \quad (20)$$

The family  $(\nu_U)_U$  is consistent if and only if  $h \equiv (\dot{h}, \hat{h})$  satisfies the *Bethe recursions*

$$\dot{z}_h \dot{h}_\sigma = \sum_{\dot{\sigma}: \sigma_1 = \sigma} \dot{\varphi}(\dot{\sigma}) \prod_{i=2}^d \dot{h}_{\sigma_i}, \quad \hat{z}_h \hat{h}_\sigma = \sum_{\hat{\sigma}: \sigma_1 = \sigma} \hat{\varphi}(\hat{\sigma}) \prod_{i=2}^k \hat{h}_{\sigma_i} \quad (21)$$

(with  $\dot{z}_h, \hat{z}_h$  the normalizing constants); these generalize the frozen model recursions (12), as we shall see explicitly below. Thus a solution  $h$  of (21) specifies a Gibbs measure  $\nu$  for the auxiliary model on  $T_{d,k}$  which generalizes the measures  $\nu$  described in §3.1.

It is clear from the 0/1 symmetries of  $\dot{\varphi}$  that any solution  $h$  of the 0/1/f Bethe recursions must also have the 0/1 symmetry, and as a consequence must correspond to a solution  $g$  of the r/f Bethe recursions via

$$\dot{h}_{\eta\eta'} = \hat{g}_{\Pi\eta, \Pi\eta'} / (2 - \hat{g}_{\eta\eta'}), \quad \hat{h}_{\eta\eta'} = 2^{1\{\eta=\eta'=f\}} \dot{g}_{\Pi\eta, \Pi\eta'} / 2. \quad (22)$$

The r/f Bethe recursions read explicitly as follows:

$$\begin{aligned} \dot{z}_g \hat{g}_{\mathbf{r}\mathbf{r}} &= \dot{z}_g \hat{g}_{\mathbf{f}\mathbf{r}} = (2/2^k)(\dot{g}_{\mathbf{r}\mathbf{f}})^{k-1}, \\ \dot{z}_g \hat{g}_{\mathbf{f}\mathbf{f}} &= (\dot{g}_{\mathbf{r}\mathbf{f}} + \dot{g}_{\mathbf{f}\mathbf{f}})^{k-1} - (4/2^k)(\dot{g}_{\mathbf{r}\mathbf{f}})^{k-1}, \\ \dot{z}_g \hat{g}_{\mathbf{r}\mathbf{f}} &= (\dot{g}_{\mathbf{r}\mathbf{f}} + \dot{g}_{\mathbf{f}\mathbf{f}})^{k-1} - (2/2^k)(k+1)(\dot{g}_{\mathbf{r}\mathbf{f}})^{k-1} + (2/2^k)(k-1)(\dot{g}_{\mathbf{r}\mathbf{f}})^{k-2}(\dot{g}_{\mathbf{r}\mathbf{r}} + \dot{g}_{\mathbf{f}\mathbf{r}} - 2\dot{g}_{\mathbf{f}\mathbf{f}}), \\ \dot{z}_g \hat{g}_{\mathbf{f}\mathbf{f}} &= (\hat{g}_{\mathbf{f}\mathbf{f}})^{d-1}, \quad \dot{z}_g \hat{g}_{\mathbf{f}\mathbf{r}} = 2(\hat{g}_{\mathbf{r}\mathbf{f}})^{d-1}, \\ \dot{z}_g \hat{g}_{\mathbf{r}\mathbf{r}} &= \dot{z}_g \hat{g}_{\mathbf{r}\mathbf{f}} = 2[(\hat{g}_{\mathbf{r}\mathbf{r}} + \hat{g}_{\mathbf{r}\mathbf{f}})^{d-1} - (\hat{g}_{\mathbf{r}\mathbf{f}})^{d-1}], \end{aligned}$$

where  $\dot{g}_{\mathbf{r}\mathbf{f}}$  was simplified using  $\hat{g}_{\mathbf{r}\mathbf{r}} = \hat{g}_{\mathbf{f}\mathbf{r}}$ . The recursion for  $\hat{g}_{\mathbf{r}\mathbf{f}}$  then simplifies to

$$\dot{z}_g \hat{g}_{\mathbf{r}\mathbf{f}} = \dot{z}_g \hat{g}_{\mathbf{f}\mathbf{f}} + (2/2^k)(k-1)(\dot{g}_{\mathbf{r}\mathbf{f}})^{k-2}(\dot{g}_{\mathbf{r}\mathbf{f}} - 2\dot{g}_{\mathbf{f}\mathbf{f}}),$$

so we see that  $\dot{g}_{\mathbf{r}\mathbf{f}} = 2\dot{g}_{\mathbf{f}\mathbf{f}}$  if and only if  $\hat{g}_{\mathbf{r}\mathbf{f}} = \hat{g}_{\mathbf{f}\mathbf{f}}$ , in which case the corresponding solution  $h$  of the 0/1/f Bethe recursions satisfies the symmetries (20). A fixed point of the recursion (13) is given by  $q = \hat{g}_{\mathbf{r}\mathbf{f}} / (\hat{g}_{\mathbf{r}\mathbf{f}} + \hat{g}_{\mathbf{f}\mathbf{f}})$  and  $v = \hat{g}_{\mathbf{r}\mathbf{f}} / (\hat{g}_{\mathbf{r}\mathbf{r}} + \hat{g}_{\mathbf{r}\mathbf{f}}) = \hat{h}_{\mathbf{f}\mathbf{f}} / (\hat{h}_{11} + \hat{h}_{\mathbf{f}\mathbf{f}}) = 4\hat{h}_{\mathbf{f}\mathbf{f}} / (1 + \hat{h}_{\mathbf{f}\mathbf{f}})$ , using the relation  $4\hat{h}_{11} + 3\hat{h}_{\mathbf{f}\mathbf{f}} = 1$ . In the reverse direction, any solution  $q, v$  of (13) gives rise to a Bethe solution via

$$\begin{aligned} \hat{g}_{\mathbf{r}\mathbf{f}} = \hat{g}_{\mathbf{f}\mathbf{f}} &= 2\hat{h}_{\mathbf{f}\mathbf{f}} / (1 + \hat{h}_{\mathbf{f}\mathbf{f}}) = v/2, & \hat{g}_{\mathbf{r}\mathbf{r}} = \hat{g}_{\mathbf{f}\mathbf{r}} &= 2\hat{h}_{11} / (1 + \hat{h}_{\mathbf{f}\mathbf{f}}) = (1-v)/2, \\ \dot{g}_{\mathbf{f}\mathbf{r}} = 2\dot{g}_{\mathbf{f}\mathbf{f}} &= 2\hat{h}_{\mathbf{f}\mathbf{f}} = 2\hat{q}_{\mathbf{f}} / (2 + \hat{q}_{\mathbf{f}}), & \dot{g}_{\mathbf{r}\mathbf{r}} = \dot{g}_{\mathbf{r}\mathbf{f}} &= 2\hat{h}_{1\mathbf{f}} = q / (2 + \hat{q}_{\mathbf{f}}). \end{aligned} \quad (23)$$

This proves our claim that the measures  $\nu$  generalize the measures  $\nu$  of §3.1.

The connection between these Gibbs measures and the rate function  $\Phi$  is given by the following variational principle. Versions of this principle have appeared in many places in the prior literature; we refer the reader to the bibliographic notes of [MM09, Ch. 14].

**Lemma 3.6.** *If  $\varphi \equiv (\dot{\varphi}, \hat{\varphi})$  is such that both  $\dot{H}$  and  $\hat{H}$  are surjective, then any stationary point  $h$  of  $\Phi$  belonging to  $\Delta^\circ$  corresponds to a Bethe fixed point solving (21) via*

$$\dot{z}_h \dot{h}(\dot{\sigma}) = \dot{\varphi}(\dot{\sigma}) \prod_{i=1}^d \dot{h}_{\sigma_i}, \quad \hat{z}_h \hat{h}(\hat{\sigma}) = \hat{\varphi}(\hat{\sigma}) \prod_{i=1}^k \hat{h}_{\sigma_i}, \quad \bar{z}_h \bar{h}(\sigma) = \dot{h}_\sigma \hat{h}_\sigma \quad (24)$$

with  $\dot{z}_h, \hat{z}_h, \bar{z}_h$  normalizing constants satisfying  $\bar{z}_h = \dot{z}_h / \hat{z}_h = \hat{z}_h / \dot{z}_h$  for  $\dot{z}_h, \hat{z}_h$  as in (21).

*Proof.* At an interior stationary point  $\mathbf{h}$ , consider differentiating  $\Phi$  in direction  $\delta \equiv (\hat{\delta}, 0)$  with  $\hat{H}\hat{\delta} = 0$ , so that  $\mathbf{h} + s\delta \in \Delta^\circ$  for  $|s|$  small. Writing  $\hat{\mathbf{a}} \equiv \log[\hat{\varphi}(\hat{\underline{x}})/\hat{h}(\hat{\underline{x}})]$ ,

$$0 = \partial_s \Phi(\mathbf{h} + s\delta)|_{s=0} = \sum_{\hat{\underline{x}}} \hat{\delta}(\hat{\underline{x}}) \hat{\mathbf{a}}(\hat{\underline{x}}) = \sum_{\hat{\underline{x}}} \hat{\delta}(\hat{\underline{x}}) \dot{\hat{\mathbf{a}}}(\hat{\underline{x}}),$$

$$\dot{\hat{\mathbf{a}}}(\hat{\underline{x}}) \equiv \hat{\mathbf{a}}(\hat{\underline{x}}) + \sum_{i=1}^d \dot{\lambda}(\sigma_i) \text{ with } \dot{\lambda} : \mathcal{M} \rightarrow \mathbb{R} \text{ arbitrary.}$$

We claim it is possible to choose  $\dot{\lambda}$  such that  $\dot{\hat{\mathbf{a}}}$  has marginals  $\bar{\varepsilon} \equiv 0$ : in vector notation  $\dot{\hat{\mathbf{a}}} = \hat{\mathbf{a}} + \hat{H}^t \dot{\lambda}$ , so this amounts to solving  $\hat{H}\hat{\mathbf{a}} + \hat{H}\hat{H}^t \dot{\lambda} = 0$ , which has a unique solution  $\dot{\lambda}$  by surjectivity of  $\hat{H}$ . Taking  $\hat{\delta} = \dot{\hat{\mathbf{a}}}$  with this value of  $\dot{\lambda}$  in the above derivative gives

$$\dot{h}(\hat{\underline{x}}) = \dot{\varphi}(\hat{\underline{x}}) \prod_{i=1}^d e^{\dot{\lambda}(\sigma_i)}, \quad \text{likewise} \quad \dot{h}(\hat{\underline{x}}) = \dot{\varphi}(\hat{\underline{x}}) \prod_{i=1}^k e^{\dot{\lambda}(\sigma_i)}. \quad (25)$$

Now differentiate in the direction of general  $\delta$  with  $d^{-1}\hat{H}\hat{\delta} = \bar{\delta} = k^{-1}\hat{H}\hat{\delta}$ , so  $\mathbf{h} + s\delta \in \Delta$  for small  $|s|$ . Applying (25) and simplifying gives

$$0 = \partial_s \Phi(\mathbf{h} + s\delta)|_{s=0} = d \langle \bar{\delta}, \bar{\rho} \rangle \quad \text{where } \bar{\rho}(\sigma) \equiv \log \bar{h}(\sigma) - \dot{\lambda}(\sigma) - \hat{\lambda}(\sigma).$$

By surjectivity we may choose  $\bar{\delta}$  with  $\bar{\delta}(\sigma) = \bar{\rho}(\sigma) - |\mathcal{M}|^{-1} \sum_{\sigma'} \bar{\rho}(\sigma')$ , and then substituting into the above we find that  $\log \bar{h} - \dot{\lambda} - \hat{\lambda}$  is a constant function of  $\sigma$ , that is,

$$\bar{h}(\sigma) \text{ equals } e^{\dot{\lambda}(\sigma)} e^{\hat{\lambda}(\sigma)} \text{ up to normalizing constant.}$$

On the other hand, the marginal of (25) reads

$$\bar{h}(\sigma) = e^{\dot{\lambda}(\sigma)} \sum_{\hat{\underline{x}}: \sigma_1 = \sigma} \dot{\varphi}(\hat{\underline{x}}) \prod_{i=2}^d e^{\dot{\lambda}(\sigma_i)} = e^{\dot{\lambda}(\sigma)} \sum_{\hat{\underline{x}}: \sigma_1 = \sigma} \dot{\varphi}(\hat{\underline{x}}) \prod_{i=2}^k e^{\dot{\lambda}(\sigma_i)}.$$

Comparing the expressions for  $\bar{h}(\sigma)$  shows that the probability measures  $\dot{h}$  and  $\hat{h}$  on  $\mathcal{M}$  obtained by normalizing respectively  $e^{\dot{\lambda}(\sigma)}$  and  $e^{\hat{\lambda}(\sigma)}$  must solve the Bethe recursions (21). Lastly (25) shows that  $\mathbf{h}$  corresponds to  $h \equiv (\dot{h}, \hat{h})$  via (24), concluding the proof.  $\square$

**Theorem 3.7.** *In the 0/1/f auxiliary model, let  ${}^*h$  denote the unique stationary point of  $\Phi$  which corresponds — via (24) and (20) — to the solution  $q^*$  of the frozen model recursions (12) which was identified in Lem. 3.1. The unique maximizer of  $\Phi$  on  $\Delta$  is given by  ${}^*h$ .*

In view of Lem. 3.6 and our preceding discussion of Gibbs measures, Thm. 3.7 will follow by showing

1. Any global maximizer  $\mathbf{h}$  of  $\Phi$  on  $\Delta$  must lie in the interior  $\Delta^\circ$ , and so corresponds via (24) to a solution  $h$  of the Bethe recursions (21). (For the required surjectivity of  $\hat{H}, \hat{H}$  see Lem. 6.6.)
2. Any such Bethe solution  $h$  satisfies the symmetries (21), therefore reduces to a solution  $q$  of the frozen model recursions (12). Further  $q$  is in the regime of Lem. 3.1, which uniquely identifies  $\mathbf{h} = {}^*h$ .

**3.4. Boundary maximizers.** In this section we verify (by *a priori* estimates) that  $\Phi$  has no maximizers on the boundary of  $\Delta$ . By Rmk. 3.3 we may work interchangeably with the frozen and auxiliary models.

We begin with a preliminary calculation. For a vector  $\underline{\ell} \in \mathbb{Z}^n$  let  $\mathbf{g}_r^{\underline{\ell}}(n, E)$  denote the probability, with respect to a uniformly random assignment of  $E$  forcing half-edges to  $n$  degree- $d$  variables, that variable  $i$  receives at least  $\ell_i$  of the  $E$  edges for each  $1 \leq i \leq n$ . If  $\underline{\ell}$  is the constant vector  $(l, \dots, l)$  we write  $\mathbf{g}_r^{\underline{\ell}}(n, E) \equiv \mathbf{g}_r^l(n, E)$ .

**Lemma 3.8.** *For  $\zeta = y(\log d)/d$  with  $y \approx 1$  and  $\underline{\ell}$  upper bounded by  $l \lesssim 1$  (uniformly in  $d$ ),*

$$\mathbf{g}_r^{\underline{\ell}}(n, nd\zeta) \asymp \exp\{O(nd^{-2y}(\log d)^{2l-1})\} \prod_{i=1}^n \mathbb{P}_\zeta(X_i \geq \ell_i).$$

*Proof.* Let  $X_1, X_1, \dots, X_n$  be i.i.d.  $\text{Bin}(d, \theta)$  random variables, with joint law  $\mathbb{P}_\theta$ : then

$$\frac{\mathfrak{g}_r^\ell(n, nd\zeta)}{\prod_{i=1}^n \mathbb{P}_\zeta(X_i \geq \ell_i)} = \frac{\mathbb{P}_\theta(\sum_{i=1}^n X_i = nd\zeta \mid X_i \geq \ell_i \forall i)}{\mathbb{P}_\theta(\sum_{i=1}^n X_i = nd\zeta)}.$$

For any  $\ell$  the conditional mean  $\mathbb{E}_\theta[X \mid X \geq \ell]$  is increasing in  $\theta$  (the derivative is the variance of a certain random variable), thus there is a unique value  $\theta = [1 + O(d^{-y}(\log d)^{l-1})]\zeta$  such that  $\mathbb{E}_\theta[\sum_{i=1}^n X_i \mid X_i \geq \ell_i \forall i] = nd\zeta$ . For this value of  $\theta$ , the local CLT (see [DM95]) combined with Stirling's approximation gives

$$\frac{\mathfrak{g}_r^\ell(n, nd\zeta)}{\prod_{i=1}^n \mathbb{P}_\zeta(X_i \geq \ell_i)} \asymp \exp\{nd H(\zeta \mid \theta)\} = \exp\{O(nd^{-2y}(\log d)^{2l-1})\},$$

concluding the proof.  $\square$

**Lemma 3.9.** *For  $k \geq k_0$  and  $d_{\text{ibd}} \leq d \leq d_{\text{ubd}}$ , the contribution to  $\mathbb{E}Z$  (see (10)) from all  $\beta \leq \beta_{\text{max}}$  with  $|2^{k+1}\beta - 1| \geq 2^{-k/8}$  is exponentially small in  $n$  compared with  $\mathbb{E}Z$ . Further*

$$e^{O(n/2^k)} = \mathbb{E}Z \geq \mathbb{E}Z_{1/2^{k+1}} \geq \exp\{n[\Phi - 1/2^{k+1} + O(k^{O(1)}/2^{4k/3})]\}, \quad \text{with } \Phi \text{ as in (2)}. \quad (26)$$

*Proof.* Recall that  $Z_{n\beta}$  denotes the contribution to the frozen model partition function from configurations with  $n\beta$  free variables.

*Upper bound ignoring forcing constraints.*

Let  $Y_{n\beta}$  denote the partition function of 0/1/f frozen configurations with  $n\beta$  frees where we *ignore* the requirement that rigid variables be forced, so clearly  $Y_{n\beta} \geq Z_{n\beta}$ . In a given frozen configuration let  $m\nu_j$  ( $0 \leq j \leq k$ ) count the number of clauses incident to exactly  $j$  free variables; and let  $\mathbf{p}_\nu^\beta$  denote the probability of empirical measure  $\nu$  of clauses with respect to a uniformly random matching between clause half-edges and variable half-edges with density  $\beta$  of frees. Then

$$\mathbb{E}Z_{n\beta} \leq \mathbb{E}Y_{n\beta} = 2^{n(1-\beta)} \binom{n}{n\beta} \sum_{\nu} \mathbf{p}_\nu^\beta (1 - 2/2^k)^{m\nu_0} (1 - 4/2^k)^{m\nu_1}.$$

Similarly to the calculation in the proof of Propn. 1.1, let  $D_1, \dots, D_m \sim \text{Bin}(k, \beta)$ , and calculate  $\mathbf{p}_\nu^\beta = \mathbb{P}(\sum_a \mathbf{1}\{D_a = j\} = m\nu_j \text{ for all } 0 \leq j \leq k \mid \sum_a D_a = mk\beta)$ : since the local CLT implies  $\mathbb{P}(\sum_a D_a = mk\beta) = n^{O(1)}$ , we find

$$\mathbb{E}Y_{n\beta} = n^{O(1)} 2^{n(1-\beta)} \binom{n}{n\beta} \sum_{\nu} \mathbf{1}\{\sum_j j\nu_j = k\beta\} \binom{m}{m\nu} \prod_j p_j^{m\nu_j} \quad \text{where}$$

$$p_0 \equiv (1 - 2/2^k) \text{bin}_{k,\beta}(0), \quad p_1 \equiv (1 - 4/2^k) \text{bin}_{k,\beta}(1), \quad p_j \equiv \text{bin}_{k,\beta}(j) \text{ for } 2 \leq j \leq k.$$

The above is optimized at  $\nu_j = p_j u^j / c$  where  $c \equiv \sum_j p_j u^j$  and  $u$  is chosen such that  $k\beta$  matches  $\sum_j j\nu_j = (\sum_j j p_j u^j) / (\sum_j p_j u^j)$ . The latter is increasing in  $u$ , and it is straightforward to check that it has a unique solution  $u = 1 + 2/2^k + O(k/4^k)$ . This implies  $c = 1 - 2/2^k + O(k^2/8^k)$ , thus

$$\mathbb{E}Y_{n\beta} = n^{O(1)} \exp\{n y(\beta)\} \text{ with}$$

$$y(\beta) \equiv (1 - \beta) \log 2 + H(\beta) + (d/k)[\log c - k\beta \log u]$$

$$= -\log 2 + (d/k) \log(1 - 2/2^k) - \beta \log 2 + H(\beta) - d\beta(2/2^k) + O(k^2/4^k)$$

$$= \Phi + \beta[\log(e/\beta) - \log 2^{k+1}] + O(k^2/4^k),$$

with  $\Phi$  as in (2) (not depending on  $\beta$ ).

*Bounds with forcing constraints.*

Suppose we condition on an assignment of edges such that every clause is satisfied, and no  $f$ -variables are illegally forced. Each of the  $m\nu_0$  fully rigid clauses is forcing with probability  $\vartheta \equiv 2k/(2^k - 2)$ , and  $m\nu_0$  is clearly sandwiched between  $m$  and  $\hat{m} \equiv m(1 - k\beta)$ , therefore

$$\mathbb{E}Z_{n\beta} \begin{cases} \leq \mathbb{E}Y_{n\beta} \sum_{\alpha} \text{bin}_{m,\vartheta}(m\alpha) \mathfrak{g}_r(n(1-\beta), m\alpha); \\ \geq \mathbb{E}Y_{n\beta} \sum_{\alpha} \text{bin}_{\hat{m},\vartheta}(\hat{m}\alpha) \mathfrak{g}_r(n(1-\beta), \hat{m}\alpha). \end{cases} \quad (27)$$

For  $\alpha = [1 + O(2^{-k/3})]^\vartheta$  we have

$$\frac{m\alpha}{n(1-\beta)d} = [1 + O(2^{-k/3})] 2/2^k = \frac{y \log d}{d} \quad \text{with } y = 1 - \frac{\log k}{k \log 2} + O(1/k);$$

the same estimate holds with  $\acute{m}, \acute{y}$  in place of  $m, y$ . Applying Lem. 3.8 then gives

$$\begin{aligned} \mathbf{g}_r(n(1-\beta), m\alpha) &= [1 - (1 - 2/2^k)^d \exp\{O(k/2^{k/3})\}]^{n(1-\beta)} \exp\{O(nk^{O(1)}/4^k)\} \\ &= \exp\{(n/2^k)[-1 + O(k/2^{k/3})]\} = \mathbf{g}_r(n(1-\beta), \acute{m}\alpha). \end{aligned}$$

For  $|1 - \alpha/\vartheta| \geq 2^{-k/3}$  we have  $\text{bin}_{m,\vartheta}(m\alpha) \leq \exp\{-n/2^{k/2}\}$ ; consequently, in each of the two sums on the right-hand side of (27), the total contribution from such  $\alpha$  is an exponentially small fraction of the sum. We therefore conclude

$$\begin{aligned} \mathbb{E}Z_{n\beta} &= (\mathbb{E}Y_{n\beta}) \exp\{(n/2^k)[-1 + O(k/2^{k/3})]\} \\ &= \exp\{n[\Phi + \beta[\log(e/\beta) - \log 2^{k+1}] - 2^{-k} + O(k/2^{4k/3})]\}. \end{aligned}$$

This is clearly optimized with  $2^{k+1}\beta \approx 1$ , and estimating the second derivative of the exponent with respect to  $\beta$  implies the result.  $\square$

**Proposition 3.10.** *The maximum of the 0/1/f auxiliary model exponent  $\Phi$  on  $\Delta$  is not attained on the boundary  $\partial\Delta$ .*

*Proof.* Lem. 3.9 shows that the maximum cannot be obtained on the boundary  $\beta = \beta_{\max}$ , so it remains to show that the maximizer must be a strictly positive measure on  $\text{supp } \varphi$ . For  $\delta \equiv (\delta, \hat{\delta})$  such that  $\mathbf{h} + t\delta$  lies in  $\Delta$  for  $t \geq 0$  small, consider

$$\mathbf{T}\Phi(\mathbf{h}; \delta) \equiv \lim_{t \downarrow 0} \frac{\Phi(\mathbf{h} + t\delta) - \Phi(\mathbf{h})}{t \log(1/t)} = \dot{\delta}[(\text{supp } \dot{\mathbf{h}})^c] + (d/k)\hat{\delta}[(\text{supp } \hat{\mathbf{h}})^c] - d\bar{h}[(\text{supp } \bar{\mathbf{h}})^c].$$

To show that  $\mathbf{h} \in \Delta$  is not a maximizer it suffices to exhibit  $\mathbf{T}\Phi(\mathbf{h}; \delta) > 0$  for some  $\delta$ . In particular, it follows by convexity that for any  $\mathbf{h} \in \Delta$ ,  $\mathbf{h} + t(\mathbf{*h} - \mathbf{h}) \in \Delta^\circ$  for  $t > 0$  small and  $\mathbf{*h}$  as in the statement of Thm. 3.7. Therefore, if  $\mathbf{h}$  is a maximizer such that the edge marginal has full support  $\text{supp } \bar{\mathbf{h}} = \mathcal{M}$ , then necessarily  $\text{supp } \dot{\mathbf{h}} = \text{supp } \varphi$ , since otherwise  $\mathbf{T}\Phi(\mathbf{h}; \mathbf{*h} - \mathbf{h}) > 0$ .

Suppose  $\mathbf{h}$  is a maximizer for  $\Phi$  on  $\Delta$ ; recall  $\dot{\mathbf{h}}, \hat{\mathbf{h}}$  must be symmetric functions. By Lem. 3.9, almost all variables are rigid except for  $\asymp n2^{-k}$  free variables; so some but not all edges are forcing. It is also clear that the rigid variables will be divided roughly evenly between 0's and 1's, so we obtain  $\{ff, 0f, 1f\} \subseteq \text{supp } \bar{\mathbf{h}}$  as well as  $\bar{h}(\{fx, xx\}) > 0$  for  $x = 0, 1$ .

1. *Case  $\bar{h}(f0) > 0 = \bar{h}(00)$ .*

By symmetry of  $\mathbf{h}$ ,  $(f0, 0f^{d-1}) \in \text{supp } \dot{\mathbf{h}}$  and  $(f0, 0f^{k-1}) \in \text{supp } \hat{\mathbf{h}}$ .

Further  $(0f^k) \in \text{supp } \hat{\mathbf{h}}$ , else  $\mathbf{T}\Phi(\mathbf{h}; \mathbf{h}' - \mathbf{h}) > 0$  for  $\mathbf{h}' \equiv (\hat{\mathbf{h}}', \hat{\mathbf{h}}')$  defined by

$$\dot{\mathbf{h}}' = \mathbf{1}_{(f0, 0f^{d-1})}, \quad (d/k)\hat{\mathbf{h}}' = \mathbf{1}_{(f0, 0f^{k-1})} + (d/k - 1)\mathbf{1}_{(0f^k)}.$$

If  $\bar{h}(00) = 0$  then consider

$$\dot{\delta} = \mathbf{1}_{(00^2, 0f^{d-2})} - \mathbf{1}_{(f0, 0f^{d-1})}, \quad (d/k)\hat{\delta} = 2 \cdot \mathbf{1}_{(00, 0f^{k-1})} - \mathbf{1}_{(f0, 0f^{k-1})} - \mathbf{1}_{(0f^k)};$$

this has marginal  $d\bar{\delta} = 2 \cdot \mathbf{1}_{00} - \mathbf{1}_{f0} - \mathbf{1}_{0f}$  so we find  $\mathbf{T}\Phi(\mathbf{h}; \delta) = 1 + 2 - 2 > 0$ .

2. *Case  $\bar{h}(00) > 0 = \bar{h}(f0)$ .*

By symmetry of  $\mathbf{h}$ ,  $(00, 0f^{k-1}) \in \text{supp } \dot{\mathbf{h}}$ .

Further  $(00^2, 0f^{d-2}) \in \text{supp } \dot{\mathbf{h}}$ , else  $\mathbf{T}\Phi(\mathbf{h}; \mathbf{h}' - \mathbf{h}) > 0$  for  $\mathbf{h}' \equiv (\hat{\mathbf{h}}', \hat{\mathbf{h}}')$  defined by

$$\dot{\mathbf{h}}' = \mathbf{1}_{(00^2, 0f^{d-2})}, \quad (d/k)\hat{\mathbf{h}}' = 2 \cdot \mathbf{1}_{(00, 0f^{k-1})} + (d/k - 2)\mathbf{1}_{(0f^k)}.$$

If  $\bar{h}(f0) = 0$  then for  $\delta$  as above we find  $\mathbf{T}\Phi(\mathbf{h}; -\delta) = 1 + 1 - 1 > 0$ .

Clearly the same argument applies replacing 0 with 1. In each case the conclusion contradicts the assumption that  $h$  is a maximizer, concluding the proof.<sup>6</sup>  $\square$

**3.5. Bethe recursion symmetries.** Suppose  $h$  is an interior maximizer for  $\Phi$  on  $\Delta$ , and so corresponds to a Bethe solution  $h$ . Let  $\hat{T}_{d,k}$  denote  $T_{d,k}$  with a subtree incident to the root removed, leaving an unmatched half-edge  $\epsilon$  incident to  $o$  (Fig. 2). Consider defining a Gibbs measure on  $\hat{T}_{d,k}$  in the manner of (??), with boundary law given by the Bethe solution  $h$ . Then the marginal law of  $\sigma_\epsilon$  will be  $\hat{h}$ , and the marginal law of the tuple of spins incident to any given vertex will be  $\hat{h}$  if the vertex is a variable,  $\hat{h}$  if it is a clause. Further, the Gibbs measure on  $\hat{T}_{d,k}$  can be generated in Markovian fashion, starting with spin  $\sigma_\epsilon$  distributed according to  $\hat{h}$ , generating the messages on the other  $d-1$  edges incident to  $o$  according to the conditional measure  $\hat{h}(\underline{\sigma} | \sigma_1 = \sigma_\epsilon)$ , and continuing iteratively down the tree.

Write  $\sigma_\epsilon \equiv \hat{u}\hat{v}$  where  $\hat{u}$  is the variable-to-clause message and  $\hat{v}$  the clause-to-variable message (in Fig. 2,  $\hat{u}$  is directed upwards,  $\hat{v}$  downwards). Given any valid auxiliary configuration  $\underline{\sigma}$  on the edges of  $\hat{T}_{d,k}$ , changing  $\hat{v}$  and passing the changed message through the tree (via  $\hat{m}_{d-1}, \hat{m}_{av}$ ) produces a new auxiliary configuration  $\underline{\sigma}'$  (Fig. 2). The symmetries (20) will follow by showing that for any fixed  $\hat{u}$ , the effect of changing  $\hat{v}$  is *measure-preserving* under the Gibbs measure  $\nu$  corresponding to  $h$ . From our definition of the Gibbs measure via the boundary law, the measure-preserving property will follow by showing that the effect of changing  $\hat{v}$  almost surely does not percolate down the tree.

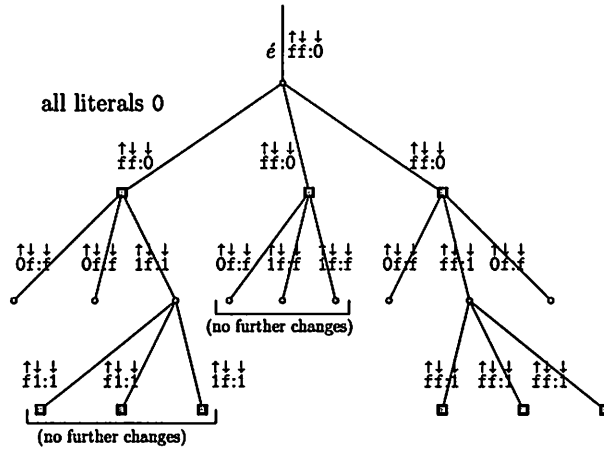


FIGURE 2. Change of message incoming down to  $\epsilon$  is passed down  $\hat{T}_{d,k}$   
 $(\eta\hat{\eta}:\hat{\eta}$  means message  $\eta$  up, message  $\hat{\eta}$  down in  $\underline{\sigma}$ , message  $\hat{\eta}$  down in  $\underline{\sigma}'$ )

Indeed, recall that we already saw directly from the Bethe recursions that  $\hat{h}_{00} = \hat{h}_{f0}$ : this came from the observation that  $\hat{\varphi}$  does not distinguish between 00 and f0, which corresponds to the fact that changing the message incoming to a clause along a *forcing* edge has no effect on the other  $k-1$  edges. We also saw that  $\hat{h}_{00} = \hat{h}_{f0}$  implies  $\hat{h}_{00} = \hat{h}_{0f}$ : this corresponds to the fact that if  $\hat{u} = 0$ , changing  $\hat{v}$  at most can change messages incoming to clauses in  $\partial o$  along forcing edges, so the effect terminates before the second level of the tree.

*Proof of Thm. 3.7.* By Propn. 3.10, any maximizer  $h$  for  $\Phi$  on  $\Delta$  must lie in the interior  $\Delta^\circ$ , and so corresponds to a solution  $h$  of the Bethe recursions (21). From the above discussion it remains to

<sup>6</sup>In our setting we have checked  $\text{supp } \hat{h} = \mathcal{M}$  in a rather ad hoc manner. A simpler argument applies generally to any specification  $\varphi$  which is *everywhere* positive on  $\mathcal{M}^d, \mathcal{M}^k$ : if  $\sigma \notin \text{supp } \hat{h}$  then take  $\hat{\sigma} \in \text{supp } \hat{h}$ , and observe that  $\mathbf{T}\Phi(h; \delta) > 0$  for  $\delta$  defined by  $\delta = \mathbf{1}_{(\sigma, \delta^{d-1})} - \mathbf{1}_{(\delta^d)}$ ,  $(d/k)\delta = \mathbf{1}_{(\sigma, \delta^{k-1})} - \mathbf{1}_{(\delta^k)}$ .



show that  $h$  satisfies  $\dot{h}_{\mathbf{ff}} = \dot{h}_{\mathbf{f0}}$ : meaning that in the Gibbs measure  $\nu$  corresponding to  $h$ , changing  $\dot{v}$  with  $\dot{u} = \mathbf{f}$  fixed has a finite-range effect. Let  $g$  correspond to  $h$  via (22).

The effect of changing  $\dot{u}\dot{v}$  from  $\mathbf{ff}$  to  $\mathbf{f0}$  can only propagate through clauses in which the parent variable and exactly one descendant variable send message  $\mathbf{f}$ , and the evaluation of the remaining  $k-2$  messages under the clause literals is identically 0 or 1. The vertex-preceding edges of  $\hat{T}_{d,k}$  whose spins will be affected by changing  $\dot{v}$  from  $\mathbf{f}$  to  $\mathbf{0}$  form a branching process with mean

$$(d-1)(k-1)(8/2^k)\hat{h}(\hat{\sigma} \in (\mathbf{ff}^2, \mathbf{rf}^{k-2}) \mid \sigma_1 = \mathbf{ff}) \leq dk(8/2^k) \frac{(\dot{g}_{\mathbf{ff}})^2(\dot{g}_{\mathbf{rf}})^{k-2}}{(1-4/2^k)\dot{g}_{\mathbf{ff}}(\dot{g}_{\mathbf{rf}})^{k-1}} \lesssim k^2\dot{g}_{\mathbf{ff}}/\dot{g}_{\mathbf{rf}}, \quad (28)$$

where the intermediate step follows from (24). Similarly, the effect of changing  $\dot{u}\dot{v}$  from  $\mathbf{f0}$  to  $\mathbf{ff}$  can only propagate through clauses in which exactly one descendant variable sends message  $\mathbf{f}$ , and the evaluation of the remaining  $k-1$  messages under the clause literals is identically 0 or 1. This forms a branching process with mean

$$(d-1)(k-1)(4/2^k)\hat{h}(\hat{\sigma} \in (\mathbf{rf}^{k-1}, \mathbf{fr}) \mid \sigma_1 \in \mathbf{rf}) \leq dk(4/2^k) \frac{(2^k-4)\dot{g}_{\mathbf{ff}}\dot{g}_{\mathbf{rf}}^{k-1}}{(2^k-2-2k)(\dot{g}_{\mathbf{rf}})^k} \lesssim k^2\dot{g}_{\mathbf{ff}}/\dot{g}_{\mathbf{rf}}. \quad (29)$$

To show that both processes are subcritical, we estimate the ratio  $\dot{u} \equiv \dot{g}_{\mathbf{ff}}/\dot{g}_{\mathbf{rf}}$ . Recall from the proof of Lem. 3.9 that  $m\hat{h}(\mathbf{rf}^k)$ , the number of all-rigid non-forcing clauses, is  $m\nu_0\alpha = m(1 - O(k/2^k))$  (otherwise the contribution to the partition function is an exponentially small fraction of the whole). Applying (24) again we have

$$\frac{nd\beta}{m\nu_0\alpha} = \hat{h}(\mathbf{rf}^k)^{-1} \sum_{j \geq 1} j \hat{h}(\text{Per}(\mathbf{ff}^j, \mathbf{rf}^{k-j})) = \sum_{j \geq 1} j \binom{k}{j} \dot{u}^j = k\dot{u}(1 + \dot{u})^{k-1},$$

so we conclude  $\dot{u} = \dot{g}_{\mathbf{ff}}/\dot{g}_{\mathbf{rf}} = \beta[1 + O(k/2^k)] \lesssim 2^{-k}$ , which clearly shows that the effect of changing  $\dot{v}$  given  $\dot{u} = \mathbf{f}$  does not percolate. Therefore  $h$  satisfies the symmetries (20), and so corresponds to a solution  $\dot{q}$  of the frozen model recursions (13). Further  $1 - q = \dot{g}_{\mathbf{ff}}/(\dot{g}_{\mathbf{ff}} + \dot{g}_{\mathbf{rf}}) \lesssim 2^{-k}$ , so Lem. 3.1 implies  $\dot{q} = \dot{q}^*$  as claimed.  $\square$

**3.6. Explicit form of first moment exponent.** We conclude this section by giving the explicit form of  $^*\Phi \equiv ^*\Phi_k(d)$ .

**Proposition 3.11.** *For  $k \geq k_0$ ,  $d_{\text{ibd}} \leq d \leq d_{\text{ubd}}$ ,  $^*\Phi \equiv ^*\Phi_k(d)$  is given by*

$$^*\Phi = \log 2 - \log(2 - q) - d(1 - k^{-1} - d^{-1}) \log[1 - 2(q/2)^k] + (d-1) \log[1 - (q/2)^{k-1}] \quad (30)$$

where  $q$  is the unique solution of

$$d = 1 + \left( \log \frac{2(1-q)}{2-q} \right) / \left( \log \frac{1-2(q/2)^{k-1}}{1-(q/2)^{k-1}} \right) \quad \text{with } 0 \leq 1-q \leq 1/2^k. \quad (31)$$

The function  $^*\Phi$  is strictly decreasing in  $d$  with  $2^k[\Phi - ^*\Phi] = 1/2 + O(k^2/2^k)$ , and so has a unique zero  $d_{\text{ibd}} < d_* < d_{\text{ubd}}$  satisfying

$$d_* = \left( 2^{k-1} - \frac{1}{2} - \frac{1}{4 \log 2} \right) k \log 2 + O\left(\frac{k^3}{2^k}\right) = d_o - \left( \frac{1}{4 \log 2} - \frac{1}{6} \right) k \log 2 + O\left(\frac{k^3}{2^k}\right), \quad (32)$$

with  $d_o$  the first moment threshold of the original NAE-SAT partition function (3).

*Proof.* The equation (31) is a rewriting of the frozen model recursions (13), which by Lem. 3.1 has a unique solution with  $0 \leq 1 - q^* \leq 2^{-k}$ . Write  $q \equiv q^*$ ,  $q_{\mathbf{f}} \equiv 1 - q = 1/2^{k+1} + O(k^2/4^k)$ , and  $v \equiv v(q) \equiv v_{k-1}(q)$  as in (13). We also abbreviate

$$\begin{aligned} Q &\equiv Q(q) \equiv (q/2)^{k-1} = (1-v)/v \quad \text{and} \\ v_{\mathbf{r}} &\equiv 1 - v = Q/(1-Q) = 2^{-k}[2 - 2(k-1)q_{\mathbf{f}} + 4/2^k + O(k^2/4^k)]. \end{aligned} \quad (33)$$

Clearly, Lem. 3.6 applies for both the 0/1/f and r/f auxiliary models. Substituting (24) into (17) and rearranging gives

$$*\Phi = \Phi(*h) = \log \dot{z}_h + (d/k) \log \hat{z}_h - d \log \bar{z}_h = \log \dot{z}_g + (d/k) \log \hat{z}_g - d \log \bar{z}_g. \quad (34)$$

We use (23) to calculate

$$\begin{aligned} \dot{z}_g &= 2(\hat{g}_{rr} + \hat{g}_{ff})^d - (\hat{g}_{ff})^d = 2^{-d}(2 - v^d) = (2/2^d)(1 + q_f v_r)/(1 + q_f), \\ \hat{z}_g &= (\hat{g}_{rf} + \hat{g}_{ff})^k - 2(\hat{g}_{rf}/2)^k = (2 + q_f)^{-k}(1 + q_f v_r)/(1 + v_r), \\ \bar{z}_g &= (\hat{g}_{rr} + \hat{g}_{ff})(\hat{g}_{rr} + \hat{g}_{ff}) + \hat{g}_{ff} \hat{g}_{rr} = (1/2)(1 + q_f v_r)/(2 + q_f). \end{aligned} \quad (35)$$

(From the Bethe recursions (21) we see that  $\dot{z}_g \hat{g}_{ff} = (\hat{g}_{ff})^{d-1}$  and  $\hat{z}_g \hat{g}_{rr} = (\hat{g}_{rf}/2)^{k-1}$ , so we can use (23) again to express  $\dot{z}_g, \hat{z}_g$  in terms of  $q_f, v_r$  and confirm that the relations  $\bar{z}_g = \dot{z}_g/\hat{z}_g = \hat{z}_g/\bar{z}_g$  of Lem. 3.6 are indeed satisfied.) Then

$$\begin{aligned} *\Phi &= \log 2 - (d/k) \log(1 + v_r) - d(1 - k^{-1} - d^{-1}) \log(1 + q_f v_r) - \log(1 + q_f) \\ &= \Phi - (d/k) \log[(1 - 2/2^k)(1 + v_r)] - d(1 - k^{-1} - d^{-1}) \log(1 + q_f v_r) - \log(1 + q_f) \end{aligned}$$

with  $\Phi \equiv \log 2 + (d/k) \log(1 - 2/2^k)$  the first-moment exponent for the original NAE-SAT partition function (2). From (33) we have

$$\begin{aligned} (1 - 2/2^k)(1 + v_r) &= 1 - 2(k-1)q_f/2^k + O(k^2/8^k), \quad \text{therefore} \\ *\Phi - \Phi &= d(1 - k^{-1})q_f(2/2^k - v_r) - \log(1 + q_f) + O(k^2/4^k) = -q_f + O(k^2/4^k). \end{aligned} \quad (36)$$

Let us now see that  $*\Phi$  is strictly decreasing in  $d$ . Recalling (33) that  $v = Q/(1 - Q)$ , we find

$$*\Phi = \log 2 + (d/k) \log(1 - qQ) + (d-1) \log \frac{1-Q}{1-qQ} - \log(2-q),$$

and rearranging gives (30). This can be expressed as a function of  $q$  alone by taking  $Q = Q(q)$  as in (33) and  $d = d(q)$  as in (31). With  $D_q$  denoting differentiation in  $q$ , we calculate

$$\begin{aligned} D_q Q &= \frac{(k-1)Q}{q} = \frac{2(k-1)}{2^k} [1 + O(k/2^k)], \quad D_q v = -\frac{D_q Q}{(1-Q)^2} = -\frac{2(k-1)}{2^k} [1 + O(k/2^k)], \\ D_q d &= -[q_f(1 + q_f) \log v]^{-1} + \frac{(d-1)D_q v}{-v \log v} = (q_f v_r)^{-1} + O(k^2/2^k) = 4^k [1 + O(k^2/2^k)]. \end{aligned}$$

The total derivative of  $*\Phi \equiv *\Phi(d(q))$  with respect to  $q$  is then straightforward to calculate: the main contribution comes from

$$\begin{aligned} D_q [(d/k) \log(1 - qQ)] &= -(dkQ)/(1 - qQ) + k^{-1}(D_q d) \log(1 - qQ) \\ &= -k^{-1}Q(D_q d) + O(k) = -(2/k)2^k [1 + O(k^3/2^k)], \\ \text{while } D_q [(d-1) \log[(1-Q)/(1-qQ)] - \log(2-q)] &= O(k^2) \end{aligned}$$

Thus  $*\Phi \equiv *\Phi_k$  is strictly decreasing on the interval  $d_{\text{lb}d} \leq d \leq d_{\text{ub}d}$  with derivative

$$*\Phi'(d) = \frac{D_q *\Phi}{D_q d} = -\frac{2}{2^k k} [1 + O(k^3/2^k)],$$

so it must have a unique zero  $d_{\text{lb}d} \leq d_* \leq d_{\text{ub}d}$ . We further estimate from (9), (33), and (36) that  $d_*$  must satisfy (32), concluding the proof.  $\square$

#### 4. SECOND MOMENT OF AUXILIARY MODEL

In this section we compute the exponential growth rate  ${}_2\Phi = \lim_{n \rightarrow \infty} n^{-1} \log \mathbb{E}[Z^2]$  of the second moment of the (truncated) frozen model partition function (10). This is done in the same framework as introduced in §3, regarding the second moment as the first moment of the partition function of *pair* frozen model configurations  $\underline{\omega} \equiv (\eta^1, \eta^2)$  on the same underlying graph. The corresponding model of pair auxiliary configurations  $\underline{\tau} \equiv (\sigma^1, \sigma^2)$  has variable factors  ${}_2\dot{\varphi} \equiv \dot{\varphi} \otimes \dot{\varphi}$ , clause factors  ${}_2\hat{\varphi}^\alpha(\hat{\tau}) \equiv \hat{\varphi}^\circ(\hat{\tau} \oplus \underline{L}) \equiv \hat{\varphi}^\circ(\hat{\sigma}^1 \oplus \underline{L}) \hat{\varphi}^\circ(\hat{\sigma}^2 \oplus \underline{L})$ , and rate function  ${}_2\Phi$  on the space  ${}_2\Delta$  of empirical measures  ${}_2h$  having both marginals in  $\Delta$ . We again average over literals to define the 0/1/f auxiliary

model on  $G$ ; note however that the pair auxiliary model does *not* have a simple r/f projection as was found in (16). In this section we prove

**Theorem 4.1.** *The rate function  ${}_2\Phi$  on  ${}_2\Delta$  attains its maximum only at the product measure  ${}^*\mathfrak{h} \equiv {}^*\mathfrak{h} \otimes {}^*\mathfrak{h}$ , or at the measures  ${}^x\mathfrak{h}$  ( $x = 0, 1$ ) with marginals  ${}^*\mathfrak{h}$  supported on pair configurations  $\mathfrak{I} = (\underline{\mathfrak{a}}, x \oplus \underline{\mathfrak{a}})$ .*

We begin with an *a priori* estimate in the frozen model. As before  $\mathfrak{r}$  denotes  $\{0, 1\}$ . We partition  $\mathfrak{rr} \equiv \{0, 1\}^2$  into  $\mathfrak{rr}^- \equiv \{00, 11\}$  and  $\mathfrak{rr}^+ \equiv \{01, 10\}$ , and we decompose  $Z^2 = \sum_{\pi} Z_{\pi}^2$  where  $Z_{\pi}^2$  denotes the partition function of pair frozen configurations with associated empirical measure  $\pi$  on  $\mathcal{P} \equiv \{\mathfrak{rr}^-, \mathfrak{rr}^+, \mathfrak{rf}, \mathfrak{fr}, \mathfrak{ff}\}$ . We use  $\nu$  to denote  $\{0, 1, \mathfrak{f}\}$ , so for example  $\pi_{\mathfrak{f}\nu} \equiv \pi_{\mathfrak{f}\mathfrak{r}} + \pi_{\mathfrak{f}\mathfrak{f}}$ , etc.; in view of (10) we always assume  $\pi_{\mathfrak{f}\nu}, \pi_{\nu\mathfrak{f}} \leq \beta_{\max}$ . Throughout the following we write  $\alpha \equiv \pi_{\mathfrak{rr}^-} / \pi_{\mathfrak{rr}}$  for the fraction of  $\mathfrak{rr}$ -vertices taking the same spin in both coordinates.

**Lemma 4.2.** *With  $\alpha \equiv \pi_{\mathfrak{rr}^-} / \pi_{\mathfrak{rr}}$ , the function  ${}_2\Phi$  can only attain its global maximum on  $\Delta$  either in the near-independent regime  ${}^*\Delta \subseteq \Delta$  of measures with  $|2\alpha - 1| \leq k/2^{k/2}$ , or in the near-identical regime  ${}^{\dagger}\Delta \subseteq \Delta$  of measures with  $\min\{\alpha, 1 - \alpha\} \leq 2^{-3k/4}$ .*

*Proof.* Given empirical measure  $\pi$  on  $\mathcal{P}$ , let  $\mathfrak{p}_{\nu}^{\pi}$  denote the probability, with respect to a uniformly random matching between variable and clause half-edges, that there are exactly  $m\nu$  clauses which are incident to only  $\mathfrak{rr}^-$  or only  $\mathfrak{rr}^+$  variables: such clauses have two invalid literal assignments. Of the remaining  $m(1 - \nu)$  clauses, all but  $O(mk^2/4^k)$  must have fewer than two frees in at least one of the two coordinates, hence must have at least *four* invalid literal assignments. Therefore

$$\mathbb{E}[Z_{\pi}^2] \leq 2^{n(1-\pi_{\mathfrak{rr}})} \binom{n}{n\pi} \sum_{\nu} \mathfrak{p}_{\nu}^{\pi} (1 - 2/2^k)^{m\nu} (1 - 4/2^k)^{m(1-\nu)} \exp\{O(nk^2/4^k)\}.$$

The typical value of  $\nu$  given  $\pi$  is  $\bar{\nu} \equiv \pi_{\mathfrak{rr}}^k [\alpha^k + (1 - \alpha)^k]$ ; conditioning and applying the local CLT (see (7) or the proof of Lem. 3.9) gives  $\mathfrak{p}_{\nu}^{\pi} \leq n^{O(1)} \exp\{-m H(\nu | \bar{\nu})\}$ . The optimal contribution to the summation above comes from

$$\frac{\nu}{1 - \nu} = \frac{\bar{\nu}}{(1 - \bar{\nu})(1 - \vartheta)} \quad \text{where } \vartheta \equiv \frac{2}{2^k - 2}.$$

Since  $\pi_{\mathfrak{f}\nu}, \pi_{\nu\mathfrak{f}} \leq \beta_{\max}$  we find  $2^{n(1-\pi_{\mathfrak{rr}})} \binom{n}{n\pi} = \exp\{n[\log 2 + H(\alpha) + O(k/2^k)]\}$ . Combining and recalling (2), (8) gives

$$\mathbb{E}[Z_{\pi}^2] \leq \exp\{n[\Phi + \bar{\mathfrak{a}}(\alpha) + O(k/2^k)]\}.$$

Recall from (26) that  $\mathbb{E}Z = e^{O(n/2^k)}$ ; it therefore follows from the estimates in the proof of Propn. 1.1 that  $\mathbb{E}[Z_{\pi}^2]/(\mathbb{E}Z)^2$  is exponentially small in  $n$  for  $2^{-3k/4} \leq \min\{\alpha, 1 - \alpha\} \leq \frac{1}{2}(1 - k/2^{k/2})$ .  $\square$

**4.1. Near-independence regime.** We now complete our analysis of the near-independent regime  ${}^*\Delta$  to prove

**Proposition 4.3.** *The unique global maximizer of the restriction of  ${}_2\Phi$  to  ${}^*\Delta$  is  ${}^*\mathfrak{h}$ .*

**Lemma 4.4.** *The contribution to  $\mathbb{E}[Z^2]$  from frozen configurations with  $|2\alpha - 1| \leq k/2^{k/2}$  and  $(2/3)\beta_{\max} \leq \max\{\pi_{\mathfrak{f}\nu}, \pi_{\nu\mathfrak{f}}\} \leq \beta_{\max}$  is exponentially small in  $n$  compared with  $(\mathbb{E}Z)^2$ .*

*Proof.* Write  $\pi_{\mathfrak{f}} \equiv 1 - \pi_{\mathfrak{rr}}$ . Let  $m\nu_j$  count the number of clauses with exactly  $2j$  invalid literal assignments, and let  $\bar{\nu}_j$  denote the typical value of  $\nu_j$  given  $\pi$ :

$$\begin{aligned} \bar{\nu}_1 &\geq (\pi_{\mathfrak{rr}})^k [\alpha^k + (1 - \alpha)^k] = 2/2^k + O(k^2/2^{3k/2}), \\ \bar{\nu}_2 &\geq (\pi_{\mathfrak{rr}})^k [1 - \alpha^k - (1 - \alpha)^k] = 1 - 2/2^k - k\pi_{\mathfrak{f}} + O(k^2/2^{3k/2}), \\ \bar{\nu}_3 &\geq k(\pi_{\mathfrak{rf}} + \pi_{\mathfrak{fr}})(\pi_{\mathfrak{rr}})^{k-1} [1 - \alpha^{k-1} - (1 - \alpha)^{k-1}] = k(\pi_{\mathfrak{rf}} + \pi_{\mathfrak{fr}}) + O(k/4^k), \\ \bar{\nu}_4 &\geq k\pi_{\mathfrak{ff}}(\pi_{\mathfrak{rr}})^{k-1} [1 - \alpha^{k-1} - (1 - \alpha)^{k-1}] = k\pi_{\mathfrak{ff}} + O(k/4^k). \end{aligned}$$

By the argument of Lem. 4.2,

$$\begin{aligned} \mathbb{E}[Z_\pi^2] &\leq 2^{n(1-\pi_{\mathbf{rf}})} \binom{n}{n\pi} \sum_{\nu} \mathbf{P}_\nu^\pi \prod_{j \geq 0} (1 - 2j/2^k)^{m\nu_j} \\ &\leq \exp\{n[(1 - \pi_{\mathbf{rf}}) \log 2 + H(\pi) + (d/k) \log(1 - (2/2^k) \sum_{j=1}^3 j\bar{\nu}_j)]\}. \end{aligned}$$

Note that  $H(\pi)$  is maximized at  $\alpha = 1/2$ , therefore

$$\begin{aligned} H(\pi) &\leq -\sum_{\omega \in \mathcal{P} \setminus \mathbf{rf}} \pi_\omega \log \pi_\omega - \pi_{\mathbf{rf}} \log \pi_{\mathbf{rf}} + \pi_{\mathbf{rf}} \log 2 \\ &\leq \log 2 - \sum_{\omega \in \mathcal{P} \setminus \mathbf{rf}} \pi_\omega \log \pi_\omega + \pi_{\mathbf{f}} \log(e/2). \end{aligned}$$

From the above estimates on the  $\bar{\nu}_j$  we find

$$1 - (2/2^k) \sum_{j=1}^3 j\bar{\nu}_j = (1 - 2/2^k)^2 - (2k/2^k)(\pi_{\mathbf{rf}} + \pi_{\mathbf{fr}} + 2\pi_{\mathbf{ff}}) + O(k^2/2^{5k/2}).$$

Combining these estimates and recalling  $\Phi = \log 2 + (d/k) \log(1 - 2/2^k)$  from (2) gives

$$n^{-1} \log \frac{\mathbb{E}[Z_\pi^2]}{e^{2n\Phi}} \leq \underbrace{-\pi_{\mathbf{rf}} \log \left( \frac{\pi_{\mathbf{rf}} 2^{k+1}}{e} \right)}_{\geq 1/2^{k+1}} - \underbrace{\pi_{\mathbf{fr}} \log \left( \frac{\pi_{\mathbf{fr}} 2^{k+1}}{e} \right)}_{\geq 1/2^{k+1}} - \underbrace{\pi_{\mathbf{ff}} \log \left( \frac{\pi_{\mathbf{ff}} 4^{k+1}}{e} \right)}_{\geq 1/4^{k+1}} + O\left(\frac{k^2}{2^{3k/2}}\right).$$

Recalling (26) gives the two upper bounds

$$n^{-1} \log \frac{\mathbb{E}[Z_\pi^2]}{(\mathbb{E}Z)^2} \leq O\left(\frac{k^{O(1)}}{2^{4k/3}}\right) + \begin{cases} -\pi_{\mathbf{rf}} \log(\pi_{\mathbf{rf}} 2^{k+1}/e) + 3/2^{k+1}; & \text{(A)} \\ -\pi_{\mathbf{ff}} \log(\pi_{\mathbf{ff}} 4^{k+1}/e) + 2/2^k. & \text{(B)} \end{cases}$$

Recall (10) that  $\beta_{\max} = 7/2^k$ ; thus (A) implies that  $\mathbb{E}[Z_\pi^2]/(\mathbb{E}Z)^2$  is exponentially small in  $n$  for  $2\pi_{\mathbf{rf}} \geq \beta_{\max}$ , or symmetrically for  $2\pi_{\mathbf{fr}} \geq \beta_{\max}$ . However (B) implies that  $\mathbb{E}[Z_\pi^2]/(\mathbb{E}Z)^2$  is exponentially small in  $n$  for  $\pi_{\mathbf{ff}} \geq 4/(k2^k)$ , and combining gives the result.  $\square$

**Proposition 4.5.** *Any global maximizer of  ${}_2\Phi$  on  ${}_2\Delta$  must be an interior stationary point.*

*Proof.* Lem. 4.4 shows that the maximum cannot be obtained on the boundary where density of frees in either coordinate is  $\beta_{\max}$ , so it remains to show that the maximizer must be a strictly positive measure on  $\text{supp } \varphi$ . For this we argue similarly as in the proof of Propn. 3.10. If  ${}_2\mathbf{h} \equiv ({}_2\hat{h}, {}_2\check{h}) \in {}_2\Delta$  is defined by  ${}_2\hat{h}(\hat{x}) = \hat{h}(\hat{x}^1) \mathbf{1}\{\hat{x}^1 = x \oplus \hat{x}^2\}$  and  ${}_2\check{h}(\hat{x}^1) \mathbf{1}\{\hat{x}^1 = x \oplus \hat{x}^2\}$  for  $\mathbf{h} \in \Delta$  and  $x \in \{0, 1\}$ , then clearly  ${}_2\Phi({}_2\mathbf{h}) = \Phi(\mathbf{h})$ , so Propn. 3.10 implies

$$\{(\hat{x}, x \oplus \hat{x}) : \hat{x} \in \text{supp } \hat{h}\} \subseteq \text{supp } \check{h}, \quad \{(\hat{x}, x \oplus \hat{x}) : \hat{x} \in \text{supp } \check{h}\} \subseteq \text{supp } \hat{h} \quad \text{for } x = 0, 1.$$

In the following we write  $r, s, x, y$  for elements of  $\{0, 1\}$ .

1. If  $\text{supp } \bar{h}$  does not contain  $\begin{smallmatrix} r \\ f \\ s \end{smallmatrix}$  or  $\begin{smallmatrix} r \\ s \\ f \end{smallmatrix}$  then  $\mathbf{T}({}_2\Phi)(\mathbf{h}; \delta) > 0$  for

$$\begin{aligned} \hat{\delta} &= \sum_{x, y \neq f} \mathbf{1}\left\{ \left( \begin{smallmatrix} xx & xx & xf^{d-2} \\ fy & yf & yf^{d-2} \end{smallmatrix} \right) - \mathbf{1}\left\{ \left( \begin{smallmatrix} xx^2 & xf^{d-2} \\ yy^2 & yf^{d-2} \end{smallmatrix} \right) \right\} \right\}, \\ (d/k)\hat{\delta} &= \sum_{x, y \neq f} \mathbf{1}\left\{ \left( \begin{smallmatrix} xx & xf^{k-1} \\ fy & yf^{k-1} \end{smallmatrix} \right), \left( \begin{smallmatrix} xx & xf^j & xf^{k-1-j} \\ yf & yf^j & -yf^{k-1-j} \end{smallmatrix} \right) \right\} - 2 \cdot \mathbf{1}\left\{ \left( \begin{smallmatrix} xx & xf^{k-1} \\ yy & yf^{k-1} \end{smallmatrix} \right) \right\} \quad \text{with } j = \lfloor \frac{k-1}{2} \rfloor, \\ d\bar{\delta} &= \sum_{x, y \neq f} \mathbf{1}\left\{ \left( \begin{smallmatrix} xx & xx \\ fy & yf \end{smallmatrix} \right) - 2 \cdot \mathbf{1}\left\{ \left( \begin{smallmatrix} xx \\ yy \end{smallmatrix} \right) \right\} \right\}. \end{aligned}$$

2. If  $\text{supp } \bar{h}$  does not contain  $\begin{smallmatrix} r \\ f \\ s \end{smallmatrix}$  or  $\begin{smallmatrix} r \\ s \\ f \end{smallmatrix}$  then  $\mathbf{T}({}_2\Phi)(\mathbf{h}; \delta) > 0$  for

$$\begin{aligned} \hat{\delta} &= \sum_{x, y \neq f} \mathbf{1}\left\{ \left( \begin{smallmatrix} xf & fx & xf^{d-2} \\ fy & yf & yf^{d-2} \end{smallmatrix} \right) - \mathbf{1}\left\{ \left( \begin{smallmatrix} fx & xf^{d-1} \\ fy & yf^{d-1} \end{smallmatrix} \right) \right\} \right\}, \\ (d/k)\hat{\delta} &= \sum_{x, y \neq f} \mathbf{1}\left\{ \left( \begin{smallmatrix} -xf & fx & -xf^{k-2} \\ fy & -yf & -yf^{k-2} \end{smallmatrix} \right) - \mathbf{1}\left\{ \left( \begin{smallmatrix} fx & xf^{k-1} \\ fy & yf^{k-1} \end{smallmatrix} \right) \right\} \right\}, \\ d\bar{\delta} &= \sum_{x, y \neq f} \mathbf{1}\left\{ \left( \begin{smallmatrix} xf & fx \\ fy & yf \end{smallmatrix} \right) - \mathbf{1}\left\{ \left( \begin{smallmatrix} fx & xf \\ fy & yf \end{smallmatrix} \right) \right\} \right\}. \end{aligned}$$

3. If  $\text{supp } \bar{h}$  does not contain any of  $\begin{smallmatrix} r \\ f \\ f \end{smallmatrix}$ ,  $\begin{smallmatrix} r \\ f \\ f \end{smallmatrix}$ ,  $\begin{smallmatrix} f \\ r \\ f \end{smallmatrix}$  then  $\mathbf{T}({}_2\Phi)(\mathbf{h}; \delta) > 0$  for

$$\begin{aligned} \hat{\delta} &= \mathbf{1}\left\{ \left( \begin{smallmatrix} fx & xf^{d-1} \\ ff & ff^{d-1} \end{smallmatrix} \right), \left( \begin{smallmatrix} xx^2 & xf^{d-2} \\ ff^2 & ff^{d-2} \end{smallmatrix} \right) \right\} - 2 \cdot \mathbf{1}\left\{ \left( \begin{smallmatrix} xf^d \\ ff^d \end{smallmatrix} \right) \right\}, \\ (d/k)\hat{\delta} &= \mathbf{1}\left\{ \left( \begin{smallmatrix} fx & xf^{k-1} \\ ff & ff^{k-1} \end{smallmatrix} \right) \right\} + 2 \cdot \mathbf{1}\left\{ \left( \begin{smallmatrix} xx & xf^{k-1} \\ ff & ff^{k-1} \end{smallmatrix} \right) \right\} + ((2d/k) - 3) \mathbf{1}\left\{ \left( \begin{smallmatrix} xf^k \\ ff^k \end{smallmatrix} \right) \right\} - (2d/k) \cdot \mathbf{1}\left\{ \left( \begin{smallmatrix} xf^k \\ ff^k \end{smallmatrix} \right) \right\}, \\ d\bar{\delta} &= \mathbf{1}\left\{ \left( \begin{smallmatrix} fx \\ ff \end{smallmatrix} \right) \right\} + 2 \cdot \mathbf{1}\left\{ \left( \begin{smallmatrix} xx \\ ff \end{smallmatrix} \right) \right\} + (2d - 3) \mathbf{1}\left\{ \left( \begin{smallmatrix} xf \\ ff \end{smallmatrix} \right) \right\} - 2d \cdot \mathbf{1}\left\{ \left( \begin{smallmatrix} ff \\ ff \end{smallmatrix} \right) \right\}. \end{aligned}$$

It follows by symmetry considerations that  $\text{supp } \bar{h} = \mathcal{M}^2$ , hence any maximizer  $h$  of  ${}_2\Phi$  must be positive on  $\text{supp } \varphi$  since otherwise  $\mathbf{T}_2\Phi(h; \frac{1}{2}h - h)$  would be positive.  $\square$

**Lemma 4.6.** *The pair frozen model tree recursions on measures  $q \equiv (\dot{q}, \hat{q})$  have the unique solution  $q = q^* \otimes q^*$  in the regime  $\dot{q}(\{\text{ff}, \text{rf}, \text{fr}\}) \lesssim 2^{-k}$ ,  $(\dot{q}_{00} + \dot{q}_{11})/(\dot{q}_{01} + \dot{q}_{10}) = 1 + O(k/2^{k/2})$ .*

*Proof.* The pair frozen model tree recursions are as follows. Write  $\dot{q}_= \equiv \dot{q}_{00} + \dot{q}_{11}$ ,  $\dot{q}_\neq \equiv \dot{q}_{01} + \dot{q}_{10}$ . By assumption,  $\dot{q}_= = (1 + \epsilon)\dot{q}_\neq$  with  $|\epsilon| \lesssim k/2^{k/2}$ . The clause recursions are

$$\begin{aligned} \hat{q}_= &\equiv \hat{q}_{00} = \hat{q}_{11} = (2/2^k) (\dot{q}_=)^{k-1}, & \hat{q}_\neq &\equiv \hat{q}_{01} = \hat{q}_{10} = (2/2^k) (\dot{q}_\neq)^{k-1}, \\ \hat{q}_{\text{rf}} &\equiv \hat{q}_{0\text{f}} = \hat{q}_{1\text{f}} = (2/2^k) [(\dot{q}_{\text{r}\neq})^{k-1} - (\dot{q}_=)^{k-1} - (\dot{q}_\neq)^{k-1}], \\ \hat{q}_{\text{fr}} &\equiv \hat{q}_{\text{f}0} = \hat{q}_{\text{f}1} = (2/2^k) [(\dot{q}_{\text{r}\neq})^{k-1} - (\dot{q}_=)^{k-1} - (\dot{q}_\neq)^{k-1}], \\ \hat{q}_{\text{ff}} &\equiv 1 - 2[\hat{q}_= + \hat{q}_\neq + \hat{q}_{\text{rf}} + \hat{q}_{\text{fr}}] = 1 - 2[\hat{q}_{00} + \hat{q}_{01} + \hat{q}_{10} + \hat{q}_{11}] = 1 - O(2^{-k}). \end{aligned}$$

The variable recursions are (with  $\dot{c}$  the normalizing constant)

$$\begin{aligned} \dot{c}\hat{q}_{\text{ff}} &= (\hat{q}_{\text{ff}})^{d-1}, \\ \dot{c}\hat{q}_{\text{f}0} &= \dot{c}\hat{q}_{\text{f}1} = (\hat{q}_{\text{ff}} + \hat{q}_{\text{fr}})^{d-1} - (\hat{q}_{\text{ff}})^{d-1}, & \dot{c}\hat{q}_{0\text{f}} &= \dot{c}\hat{q}_{1\text{f}} = (\hat{q}_{\text{ff}} + \hat{q}_{\text{rf}})^{d-1} - (\hat{q}_{\text{ff}})^{d-1}, \\ \dot{c}\hat{q}_{00} &= \dot{c}\hat{q}_{11} = (\hat{q}_= + \hat{q}_{\text{rf}} + \hat{q}_{\text{fr}} + \hat{q}_{\text{ff}})^{d-1} - (\hat{q}_{\text{rf}} + \hat{q}_{\text{ff}})^{d-1} - (\hat{q}_{\text{fr}} + \hat{q}_{\text{ff}})^{d-1} + (\hat{q}_{\text{ff}})^{d-1} \\ \dot{c}\hat{q}_{01} &= \dot{c}\hat{q}_{10} = (\hat{q}_\neq + \hat{q}_{\text{rf}} + \hat{q}_{\text{fr}} + \hat{q}_{\text{ff}})^{d-1} - (\hat{q}_{\text{rf}} + \hat{q}_{\text{ff}})^{d-1} - (\hat{q}_{\text{fr}} + \hat{q}_{\text{ff}})^{d-1} + (\hat{q}_{\text{ff}})^{d-1}. \end{aligned}$$

By the assumption that  $|\epsilon| \lesssim k/2^{k/2}$ , the clause recursions give  $\hat{q}_{\text{rf}} = (2/2^k)[1 + O(k/2^k)] = \hat{q}_{\text{fr}}$ , and therefore from the variable recursions we must have  $\hat{q}_{\text{rf}} = (1 + \delta)\hat{q}_{\text{fr}}$  with  $|\delta| \lesssim 2^{-k}$ . But then the clause recursions give  $|\hat{q}_{\text{rf}} - \hat{q}_{\text{fr}}| \lesssim \delta k/4^k$ , consequently

$$\frac{\hat{q}_{\text{fr}}}{\hat{q}_{\text{rf}}} = \frac{1 + O(\delta dk/4^k) - (\hat{q}_{\text{ff}})^{d-1}/(\hat{q}_{\text{ff}} + \hat{q}_{\text{rf}})^{d-1}}{1 - (\hat{q}_{\text{ff}})^{d-1}/(\hat{q}_{\text{ff}} + \hat{q}_{\text{rf}})^{d-1}} = 1 + O(\delta k^2/2^k),$$

proving that the recursion contracts to  $\delta = 0$ , i.e.  $\hat{q}_{\text{rf}} = \hat{q}_{\text{fr}}$  and  $\hat{q}_{\text{rf}} = \hat{q}_{\text{fr}}$ . Similarly, the clause recursions give  $\hat{q}_= = \hat{q}_\neq = O(\epsilon k/4^k)$ , and substituting this into the variable recursions then gives  $\hat{q}_=/\hat{q}_\neq = 1 + O(\epsilon k^2/2^k)$ , so we also have contraction to  $\epsilon = 0$ ,  $\dot{q}_= = \dot{q}_\neq \equiv \dot{q}_{\text{rr}}/2$ .

It remains to show  $(\dot{q}_{\text{rr}}, \hat{q}_{\text{rf}}, \hat{q}_{\text{ff}}) = (\dot{q}^2, \dot{q}(1 - \dot{q}), (1 - \dot{q})^2)$  with  $\dot{q} = \dot{q}^*$ . To this end write

$$q \equiv \dot{q}_{\text{rr}} + \hat{q}_{\text{rf}} \equiv 4\dot{q}_{00} + 2\hat{q}_{0\text{f}}, \quad \dot{q} \equiv \dot{q}_{\text{rr}}/q \equiv \dot{q}_{00}/(2\dot{q}_{00} + \hat{q}_{0\text{f}}).$$

Writing  $Q \equiv (q/2)^{k-1}$ ,  $\dot{Q} \equiv (\dot{q}/2)^{k-1}$ , and  $Q^* \equiv (\dot{q}^*/2)^{k-1}$ , we have

$$1 - q = \left. \begin{aligned} &\frac{\langle (0, 2, -1), W^{d-1} \rangle}{\langle (4, -4, 1), W^{d-1} \rangle} \\ \text{and } 1 - \dot{q} &= \frac{\langle (0, 1, -1), W^{d-1} \rangle}{\langle (2, -3, 1), W^{d-1} \rangle} \end{aligned} \right\} \text{ where } \begin{cases} W_1 \equiv \hat{q}_{\text{ff}} + 2\hat{q}_{0\text{f}} + \hat{q}_{00} = 1 - 2Q + Q\dot{Q}, \\ W_2 \equiv \hat{q}_{\text{ff}} + \hat{q}_{0\text{f}} = 1 - 3Q + 2Q\dot{Q}, \\ W_3 \equiv \hat{q}_{\text{ff}} = 1 - 4Q + 4Q\dot{Q}. \end{cases}$$

By assumption,  $x \equiv 2^k(|q - \dot{q}^*| + |\dot{q} - \dot{q}^*|) \lesssim 1$ , so  $Q + O(xk/4^k) = Q^* = \dot{Q} + O(xk/4^k)$  and consequently  $W^{d-1}[1 + O(xdk/4^k)] = (W^*)^{d-1}$ . It follows that

$$\frac{1 - q}{1 + O(xk^2/2^k)} = \frac{1 - \dot{q}}{1 + O(xk^2/2^k)} = \frac{\langle (0, 2, -1), (W^*)^{d-1} \rangle}{\langle (4, -4, 1), (W^*)^{d-1} \rangle} = 1 - \dot{q}^*$$

implying that the recursion contracts to  $x = 0$ ,  $q = \dot{q} = \dot{q}^*$  as claimed.  $\square$

*Proof of Propn. 4.3.* By Propn. 4.5 and Lem. 3.6, any maximizer  $h$  of  ${}_2\Phi$  on  ${}^*\Delta$  corresponds to some solution  $h$  of the Bethe recursions for the pair 0/1/f auxiliary model. We now show that  $h$  must satisfy the symmetries (20): that is,  $\hat{h}(\sigma_i) = \hat{h}(\sigma_i')$ , where  $\sigma$  now indicates the outgoing pair of variable-to-clause messages, and  $i$  or  $i'$  indicates the incoming pair of clause-to-variable messages. Let  $\mathcal{R}_= \equiv \left\{ \begin{smallmatrix} \sigma_{\text{f}}^{\text{f}} \\ \sigma_{\text{f}}^{\text{f}} \end{smallmatrix} \right\}$  and  $\mathcal{R}_\neq \equiv \left\{ \begin{smallmatrix} \sigma_{\text{f}}^{\text{f}} \\ \sigma_{\text{f}}^{\text{f}} \end{smallmatrix} \right\}$ , and write  $\mathcal{D}_j \equiv \mathcal{R}_=^j \cup \mathcal{R}_\neq^j$  and  $\mathcal{R} \equiv \mathcal{D}_1$ . Analogously to the first-moment symmetries that were seen directly from the Bethe recursions, in the pair model it is easily seen that

$$\hat{h}(\sigma_{xx}) = \hat{h}(\sigma_{\text{f}x}) \text{ and } \hat{h}(\sigma_{xx}) = \hat{h}(\sigma_{\text{f}x}) \text{ for all } \sigma \in \mathcal{M}, x \in \{0, 1\}.$$



Further, the 0/1 symmetry in the clause factors implies  $\hat{h}(\tau) = \hat{h}(-\tau)$  and  $\dot{h}(\tau) = \dot{h}(-\tau)$  for any  $\tau \in \mathcal{M}^2$ . It remains to prove

$$\dot{h}\left(\begin{smallmatrix} \sigma \\ \mathbf{f}x \end{smallmatrix}\right) = \dot{h}\left(\begin{smallmatrix} \sigma \\ \mathbf{f}\mathbf{f} \end{smallmatrix}\right) \quad \text{for } x \in \{0, 1\}.$$

*Estimates on messages.* The number of clauses incident to any variable which are free in either coordinate is  $\lesssim mk/2^k$ , while an easy *a priori* estimate implies that the number of fully-rigid clauses which are non-forcing is  $\asymp m$ . Recalling (24) then gives

$$k/2^k \gtrsim \frac{k \hat{h}(\{\begin{smallmatrix} \mathbf{f}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix}, \begin{smallmatrix} \mathbf{r}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix}, \begin{smallmatrix} \mathbf{f}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix}\}, \mathcal{D}^{k-1})}{\hat{h}(\mathcal{D}^k)} \geq [1 - O(k/2^k)] \frac{k [\dot{h}(\begin{smallmatrix} \mathbf{f}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix}) + \dot{h}(\begin{smallmatrix} \mathbf{r}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix}) + \dot{h}(\begin{smallmatrix} \mathbf{f}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix})]}{\dot{h}(\begin{smallmatrix} \mathbf{r}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix})}, \quad (37)$$

where the last inequality follows because all the  $\mathcal{Z}^{\hat{\phi}}$  factor weights involved in the application of (24) are  $1 - O(k/2^k)$ .

We now estimate the ratio  $\dot{\gamma} \equiv \dot{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \sigma\mathbf{f} \end{smallmatrix})/\dot{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix})$ . Another application of (24) gives

$$\dot{\gamma} \equiv \frac{\alpha}{1 - \alpha} = \frac{\pi_{\mathbf{r}\mathbf{r}^=} }{\pi_{\mathbf{r}\mathbf{r}^{\neq}}} = \frac{\hat{h}(\begin{smallmatrix} \{00, \sigma\mathbf{f}\} \\ \{00, \sigma\mathbf{f}\} \end{smallmatrix})^d - \hat{h}(\begin{smallmatrix} \{00, \sigma\mathbf{f}\} \\ \sigma\mathbf{f} \end{smallmatrix})^d - \hat{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \{00, \sigma\mathbf{f}\} \end{smallmatrix})^d + \hat{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \sigma\mathbf{f} \end{smallmatrix})^d}{\hat{h}(\begin{smallmatrix} \{00, \sigma\mathbf{f}\} \\ \{11, \mathbf{f}\mathbf{f}\} \end{smallmatrix})^d - \hat{h}(\begin{smallmatrix} \{00, \sigma\mathbf{f}\} \\ \mathbf{f}\mathbf{f} \end{smallmatrix})^d - \hat{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \{11, \mathbf{f}\mathbf{f}\} \end{smallmatrix})^d + \hat{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix})^d}$$

where we have used the symmetry  $\hat{h}(\begin{smallmatrix} \sigma \\ \mathbf{f}x \end{smallmatrix}) = \hat{h}(\begin{smallmatrix} \sigma \\ \mathbf{f}\mathbf{f} \end{smallmatrix})$  noted above. The ratio  $\dot{\gamma}$  is given by the same expression with  $d - 1$  in place of  $d$ . Writing  $\hat{h}^{\otimes}$  for the product measure with marginals  $\dot{h}$ , the Bethe recursions give

$$\hat{z}_h \hat{h}(\begin{smallmatrix} 00 \\ 00 \end{smallmatrix}) = \frac{2}{2^k} \hat{h}^{\otimes}(\mathcal{D}_{k-1}), \quad \hat{z}_h \hat{h}(\begin{smallmatrix} 00 \\ \sigma\mathbf{f} \end{smallmatrix}) = \frac{2}{2^k} \hat{h}^{\otimes}(\mathcal{D}^{k-1} \setminus \mathcal{D}_{k-1}), \quad \hat{z}_h \hat{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \sigma\mathbf{f} \end{smallmatrix}) = (1 - O(k/2^k)) \dot{h}^{\otimes}(\mathcal{D}^{k-1}),$$

where the last estimate uses (37). Thus  $\hat{h}(\begin{smallmatrix} \{00, \sigma\mathbf{f}\} \\ \{00, \sigma\mathbf{f}\} \end{smallmatrix}) = [1 + \frac{2}{2^k} + O(\frac{k}{4^k})] \hat{h}(\begin{smallmatrix} \{00, \sigma\mathbf{f}\} \\ \sigma\mathbf{f} \end{smallmatrix})$ , and so

$$\dot{\gamma} \equiv \frac{\dot{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \sigma\mathbf{f} \end{smallmatrix})}{\dot{h}(\begin{smallmatrix} \sigma\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix})} = e^{O(1/2^k)} \frac{\hat{h}(\begin{smallmatrix} \{00, \sigma\mathbf{f}\} \\ \{00, \sigma\mathbf{f}\} \end{smallmatrix})^{d-1}}{\hat{h}(\begin{smallmatrix} \{00, \sigma\mathbf{f}\} \\ \{00, \sigma\mathbf{f}\} \end{smallmatrix})^{d-1}} = e^{O(1/2^k)} (\gamma e^{O(1/2^k)})^{(d-1)/d} = 1 + O(k/2^{k/2}), \quad (38)$$

where the last step uses the assumption that  $\hat{h}$  lies in  $\mathfrak{A}^*$ .

*Finite-range effect of changed incoming message.* We now show  $\dot{h}(\begin{smallmatrix} \sigma \\ \mathbf{f}x \end{smallmatrix}) = \dot{h}(\begin{smallmatrix} \sigma \\ \mathbf{f}\mathbf{f} \end{smallmatrix})$  for  $x \in \{0, 1\}$ . The effect propagates through clauses which in the second copy are as described in the proof of Thm. 3.7: that is, in the second copy, exactly one descendant variable sends message  $\mathbf{f}$ , and the evaluation of all the incoming 0/1 messages (of which there are  $k - 2$  or  $k - 1$  depending on whether the parent variable sends  $\mathbf{f}$  or not) under the clause literals is identically 0 or 1. The mean of the branching process is bounded as in (28) and (29) except that we must now condition on the *pair* spin  $\tau_1$  on the edge preceding the clause.

We now explain the rather delicate case where the clause is forcing to its parent variable in the *first* coordinate. Conditioned on spin  $\tau_1 = \begin{smallmatrix} 00 \\ \mathbf{f}\mathbf{f} \end{smallmatrix}$  on the preceding edge, the probability of having a clause as described above is (using (24) and (38))

$$\leq \frac{(2/2^k) \dot{h}(\begin{smallmatrix} 00 \\ \mathbf{f}\mathbf{f} \end{smallmatrix}) (k-1) \dot{h}(\begin{smallmatrix} \mathbf{r}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix}) \dot{h}^{\otimes}(\mathcal{D}_{k-2})}{(2/2^k) \dot{h}(\begin{smallmatrix} 00 \\ \mathbf{f}\mathbf{f} \end{smallmatrix}) \dot{h}^{\otimes}(\mathcal{D}^{k-1} \setminus \mathcal{D}_{k-1})} \lesssim (k/2^k) \dot{h}(\begin{smallmatrix} \mathbf{r}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix}) / \dot{h}(\begin{smallmatrix} \mathbf{r}\mathbf{f} \\ \mathbf{f}\mathbf{f} \end{smallmatrix}) \lesssim (k/4^k)$$

and this is  $\ll d^{-1}$  so the propagation through clauses started from  $\tau_1 = \begin{smallmatrix} 00 \\ \mathbf{f}\mathbf{f} \end{smallmatrix}$  is subcritical. The calculations for the remaining cases of  $\tau_1$  are similar but easier, and so are left to the reader. We therefore see that  $\hat{h}$  satisfies the symmetries (20), and so corresponds to a solution  $q \equiv (\hat{q}, \hat{q})$  of the pair frozen model recursions. By (37) and (38) this solution falls in the regime of Lem. 4.6, which uniquely identifies  $\hat{h}$  as  $\mathfrak{A}^* \hat{h}$ .  $\square$

**4.2. A priori rigidity estimate.** Recalling Lem. 4.2, let  $\mathfrak{Z}$  denote the contribution to  $Z^2$  from the near-identical regime  $\mathfrak{h} \in \mathfrak{Z}\Delta$ . In this subsection we prove

**Proposition 4.7.** *For  $k \geq k_0$ ,  $d_{\text{ld}} \leq d \leq d_{\text{ubd}}$ , and  $n \geq n_0(k)$ ,  $\mathbb{E}[\mathfrak{Z}] \lesssim \mathbb{E}Z$ .*

**Lemma 4.8.** *Given a frozen configuration  $\eta$ , for  $1 \leq j \leq k$  let  $m\nu_j$  count the number of clauses incident to exactly  $j$   $\eta$ -free variables, and write  $\nu_{\geq 2} \equiv 1 - \nu_0 - \nu_1$ . Let  $m_f$  count the number of  $\eta$ -forcing clauses, and let  $\gamma$  denote the fraction of rigid variables which are  $\eta$ -forced at most  $k^{1/2}$  times. Then for  $k \geq k_0$ ,  $n \geq n_0(k)$  it holds that*

$$\begin{aligned} \mathbb{E}[Z_{n\beta}; (\Omega_A)^c] &\leq (\mathbb{E}Z) \exp\{-10nk^2/2^k\} && \text{for } \Omega_A \equiv \{\nu_{\geq 2} \leq k^3\beta^2\}; \\ \mathbb{E}[Z_{n\beta}; (\Omega_B)^c] &\leq (\mathbb{E}Z) \exp\{-10nk^2/2^{k/2}\} && \text{for } \Omega_B \equiv \left\{ \begin{array}{l} |1 - m_f/(m \cdot 2k/2^k)| \leq 2^{-k/8} \\ \text{and } \gamma \leq k^2/2^{k/2} \end{array} \right\} \end{aligned}$$

*Proof.* As in the proof of Lem. 3.9, let  $\mathbf{p}_{\nu_0, \nu_1}^\beta$  denote the probability of  $\nu_0, \nu_1$  with respect to a uniformly random matching between clause half-edges and variable half-edges with density  $\beta$  of frees. Conditioned on all fully-rigid clauses being satisfied, the number  $m_f$  of forcing clauses is distributed  $\text{Bin}(m\nu_0, \vartheta)$  with  $\vartheta \equiv 2k/(2^k - 2)$ . Conditioned on  $m_f$ , the probability of having  $\gamma$ -fraction of the rigid variables forced  $\leq k^{1/2}$  times is

$$\mathbf{a}_{m_f}^\beta(\gamma) = \mathbb{P}(\sum_{i=1}^{np} \mathbf{1}\{F_i \leq k^{1/2}\} = np\gamma \mid \sum_{i=1}^{np} F_i = m_f), \quad F_i \sim \text{Bin}(d, \alpha)$$

(where  $0 < \alpha < 1$  may be arbitrarily chosen). We therefore bound

$$\mathbb{E}[Z_{n\beta}; \nu_{\geq 2}, m_f, \gamma] \leq 2^{n(1-\beta)} \binom{n}{n\beta} \sum_{\nu_0 + \nu_1 = 1 - \nu_{\geq 2}} \mathbf{p}_{\nu_0, \nu_1}^\beta (1 - 2/2^k)^{m\nu_0} \text{bin}_{m\nu_0, \vartheta}(m_f) \mathbf{a}_{m_f}^\beta(\gamma).$$

From the trivial bound  $\nu_0 \geq 1 - k\beta$ , together with our estimate (26) that  $\mathbb{E}Z = e^{O(n/2^k)}$ ,

$$\mathbb{E}[Z_{n\beta}; \nu_{\geq 2}, m_f, \gamma] \leq (\mathbb{E}Z) \exp\{O(nk/2^k)\} \sum_{\nu_0 + \nu_1 = 1 - \nu_{\geq 2}} \mathbf{p}_{\nu_0, \nu_1}^\beta \text{bin}_{m\nu_0, \vartheta}(m_f) \mathbf{a}_{m_f}^\beta(\gamma).$$

Summing over  $m_f, \gamma$  and simply upper bounding  $\sum_{m_f, \gamma} \text{bin}_{m\nu_0, \vartheta}(m_f) \mathbf{a}_{m_f}^\beta(\gamma) \leq 1$  yields the bound on  $\mathbb{E}[Z_{n\beta}; (\Omega_A)^c]$ , recalling that the typical value of  $m\nu_{\geq 2}$  is  $\lesssim mk^2\beta^2$ . To bound  $\mathbb{E}[Z_{n\beta}; (\Omega_B)^c]$ , we first estimate

$$\text{bin}_{m\nu_0, \vartheta}(m_f) \leq \exp\{-n2^{-k/3}\} \quad \text{on the event } |1 - m_f/(m \cdot 2k/2^k)| \geq 2^{-k/8}.$$

On the complementary event  $|1 - m_f/(m \cdot 2k/2^k)| < 2^{-k/8}$ , in the above expression for  $\mathbf{a}_{m_f}^\beta(\gamma)$  we can set  $\alpha = m_f/(ndp) = (2/2^k)[1 + O(2^{-k/8})]$ , and apply the local CLT to bound

$$\begin{aligned} \mathbf{a}_{m_f}^\beta(\gamma) &\leq \exp\{-np H(\gamma \mid \bar{\gamma})\} \quad \text{with } \bar{\gamma} \equiv \mathbb{P}(\text{Bin}(d, \alpha) \leq k^{1/2}) \leq 2^{-k} \exp\{O(k^{1/2} \log k)\} \\ &\leq \exp\{-nk^{5/2}/2^{k/2}\} \quad \text{for } \gamma \geq k^2/2^{k/2}. \end{aligned}$$

Combining these estimates gives the bound on  $\mathbb{E}[Z_{n\beta}; (\Omega_B)^c]$ .  $\square$

We now decompose  $Z^2 = \sum_\pi Z^2[\pi]$  where  $Z^2[\pi]$  denotes the contribution from empirical measure  $\pi$  on  $\{0, 1, \mathfrak{f}\}^2$ . For  $j = 1, 2$  we write  $\pi^j$  for the projection of  $\pi$  onto the  $j$ -th coordinate, e.g.  $\pi^1 \equiv (\pi_{0\cdot}, \pi_{1\cdot}, \pi_{\mathfrak{f}\cdot})$ , and we decompose  $Z = \sum_{\pi^1} Z[\pi^1]$ .

**Lemma 4.9.** *It holds for  $k \geq k_0$  and  $n \geq n_0(k)$  that for any empirical measure  $\pi$  on  $\{0, 1, \mathfrak{f}\}^2$  with  $\max\{\pi_{\mathfrak{f}}^1, \pi_{\mathfrak{f}}^2\} \leq \beta_{\text{max}}$  and  $\Delta \equiv n\pi(\eta^1 \neq \eta^2) \leq n/2^{k/2}$ ,*

$$\mathbb{E}[Z^2[\pi]] \leq e^{-n/2^{k/2}} \mathbb{E}Z + 2^{-\Delta k/10} (\mathbb{E}[Z[\pi^1]] + \mathbb{E}[Z[\pi^2]]).$$

*Proof.* Write  $p \equiv \pi_{\mathfrak{f}\cdot}$  and  $\beta \equiv \pi_{\mathfrak{f}\cdot} \equiv 1 - p$ . Given any  $\eta^1 \in \{0, 1, \mathfrak{f}\}^V$ , the number of choices for  $\eta^2$  for which  $|\{v : \eta^1 \neq \eta^2\}| \leq n/2^{k/2}$  is (crudely) upper bounded by  $\exp\{O(nk/2^{k/2})\}$  even in absence of satisfiability constraints. Combining with Lem. 4.8 gives  $\mathbb{E}[Z^2[\pi]; (\Omega_B)^c] \leq (\mathbb{E}Z) \exp\{-5nk^2/2^{k/2}\}$ , so we hereafter restrict consideration to the event  $\Omega_B$ .

For the remainder of the proof let  $\omega \equiv (\eta^1, \eta^2)$  be a fixed spin configuration with empirical measure  $\pi$ , and for  $\omega \in \{0, 1, \mathfrak{f}\}^2$  write  $V_\omega \equiv \{v \in V : \omega_v = \omega\}$ . Decompose  $\Omega_B$  as the disjoint union

of events  $\Omega_{B,x}$  where  $x \equiv (\nu_0, \nu_1, m_f, \gamma)$  is defined as in the statement of Lem. 4.8 *with respect to*  $\eta^1$ . Let  $\mathcal{R} \equiv V_{\underline{r}} \setminus V_{\underline{rr}^\#}$  and  $|\mathcal{R}| \equiv np\epsilon = n\pi_{\underline{rr}^\#} + n\pi_{\underline{rr}}$ , and write  $F_\delta$  for the event that exactly  $np\epsilon\delta$  in  $\mathcal{R}$  are  $\eta^1$ -forced  $\leq k^{1/2}$  times. We then bound

$$\mathbb{E}[Z^2[\pi]; \Omega_B] \leq \sum_x \mathbb{E}[Z[\pi^1]; \Omega_{B,x}] \sum_\delta c_r^{\pi,x,\delta} c_f^\pi \mathbb{P}(\eta^2 \text{ valid} \mid \eta^1 \text{ valid}, \Omega_{B,x}, F_\delta)$$

where  $c_r^{\pi,x,\delta} \equiv 2^{np\epsilon} \binom{np\gamma}{np\epsilon\delta} \binom{np(1-\gamma)}{np\epsilon(1-\delta)}$  and  $c_f^\pi \equiv \binom{n\beta}{n\pi_{\underline{rr}}}$   $\leq 2^{n\beta}$ ,

(the factor  $2^{np\epsilon}$  in  $c_r^{\pi,x,\delta}$  counts all the mappings  $\mathcal{R} \rightarrow \{\underline{rr}^\#, \underline{rf}\}$ ).

*Constraints on clauses incident to  $\{\underline{rr}^\#, \underline{rf}\}$ -variables.*

On the event  $F_\delta$  there must be at least  $np\epsilon[\delta + (k^{1/2} + 1)(1 - \delta)] \geq np\epsilon(2 - \delta) \equiv m_f(\mathcal{R})$  clauses  $\eta^1$ -forcing to  $\mathcal{R}$ . Let  $Q_r$  denote the event that each such clause is incident to at least one other  $\mathcal{R}$ -variable; then  $\{\eta^2 \text{ valid}\} \subseteq Q_r$ . Consider  $\mathcal{R}$ -incident half-edges (of which there are at most  $ndp\epsilon$ ) that are matched to any of the  $mk\nu_0 - m_f$  clause-incident half-edges that are non- $\eta^1$ -forcing: this matching is uniformly random and  $ndp\epsilon/(mk\nu_0 - m_f) \leq 2\epsilon$ , so

$$\begin{aligned} \mathbf{p}_r^{\pi,x,\delta} &\equiv \mathbb{P}(Q_r \mid \eta^1 \text{ valid}, \Omega_{B,x}, F_\delta) \\ &\leq \mathbb{P}(D_a > 0 \forall a \leq m_f(\mathcal{R}) \mid \sum_{a=1}^{m\nu_0} D_a = (mk\nu_0 - m_f)2\epsilon), \\ &\text{with } D_a \text{ independent random variables distributed as} \\ &\quad \text{Bin}(k-1, 2\epsilon) \text{ for } a \leq m_f, \text{ Bin}(k, 2\epsilon) \text{ for } a > m_f. \end{aligned}$$

Note that for any realization of  $(D_a : a \leq m_f(\mathcal{R}))$ ,

$$\frac{\mathbb{P}(\sum_{a=1}^{m\nu_0} D_a = (mk\nu_0 - m_f)2\epsilon \mid (D_a : a \leq m_f(\mathcal{R})))}{\mathbb{P}(\sum_{a=1}^{m\nu_0} D_a = (mk\nu_0 - m_f)2\epsilon)} \lesssim 1, \quad (39)$$

therefore  $\mathbf{p}_r^{\pi,x,\delta} \lesssim [1 - (1 - 2\epsilon)^{k-1}]^{m_f(\mathcal{R})} \leq (4k\epsilon)^{m_f(\mathcal{R})}$ . Combining with  $c_r^{\pi,x,\delta}$  and rearranging gives

$$\begin{aligned} (np)^{-1} \log(c_r^{\pi,x,\delta} \mathbf{p}_r^{\pi,x,\delta}) &\leq \gamma [H(\frac{\epsilon\delta}{\gamma}) + (\frac{\epsilon\delta}{\gamma}) \log(\frac{\epsilon}{\gamma})] + (1-\gamma) [H(\frac{\epsilon(1-\delta)}{1-\gamma}) + \frac{\epsilon(1-\delta)}{1-\gamma} \log \epsilon] \\ &\quad + \epsilon\delta \log \gamma + \epsilon(1-\delta) \log \epsilon + O(\epsilon \log k) \\ &\leq \epsilon \log(\gamma + \epsilon) + O(\epsilon \log k), \end{aligned}$$

where we have used the trivial inequality  $H(x) + x \log c \leq \log(1+c) \leq c$ . Recall  $\epsilon \lesssim 2^{-k/2}$  by assumption, and  $\gamma \leq k^2/2^{k/2}$  by the restriction to  $\Omega_B$ , therefore

$$c_r^{\pi,x,\delta} \mathbf{p}_r^{\pi,x,\delta} \leq \exp\{-np\epsilon(k/2) \log 2[1 - O(k^{-1} \log k)]\}. \quad (40)$$

Recalling  $c_f^\pi \leq 2^{n\beta}$  we see that

$$\text{if } \epsilon \geq 2\beta \text{ then } \mathbb{E}[Z^2[\pi]; \Omega_B] \leq \frac{\mathbb{E}[Z[\pi^1]; \Omega_B]}{2^{n\epsilon k/3 - n\beta}} \leq \frac{\mathbb{E}[Z[\pi^1]; \Omega_B]}{2^{n(\epsilon+\beta)k/5}} \leq \frac{\mathbb{E}[Z[\pi^1]; \Omega_B]}{2^{\Delta k/5}}.$$

*Forcing of  $\underline{fr}$ -variables.*

Now suppose  $\epsilon \lesssim \beta$ : the number of choices for  $\eta^2$  is then  $\leq \exp\{O(nk/2^k)\}$ , so combining with Lem. 4.8 gives in this case  $\mathbb{E}[Z^2[\pi]; (\Omega_A)^c] \leq (\mathbb{E}Z) \exp\{-5nk^2/2^k\}$ . Therefore we restrict consideration hereafter to the event  $\Omega_A$ . On  $\Omega_A$ , consider the event  $Q_f \supseteq \{\eta^2 \text{ valid}\}$  that every  $\underline{fr}$ -variable is  $\eta^2$ -forced, conditioned on  $\{\eta^1 \text{ valid}, \Omega_{B,x}, F_\delta, Q_r\}$ . A clause can be  $\eta^2$ -forcing to an  $\underline{fr}$ -variable in precisely one of two ways:

1. For  $v \in V_{\underline{r}}$  let  $\tilde{A}_v$  count the number of clauses  $a \in \partial v$  that are incident to no  $\eta^1$ -free variables besides  $v$ . A clause  $a \in \partial v$  of this type will be  $\eta^2$ -forcing to  $v$  for certain arrangements of literals and of spins  $\underline{rr}, \underline{rr}^\#$  among the neighbors  $u \in \partial a \setminus v$ ; since  $v \in V_{\underline{r}}$  the clause  $a$  is conditioned not to be  $\eta^1$ -forcing. Conditioning on  $\{\eta^1 \text{ valid}, \Omega_{B,x}, F_\delta, Q_r\}$  gives information about the arrangement of  $\eta^2$  only with regards to the  $\eta^1$ -forcing clauses, so the matching between  $\mathcal{R}$ -incident half-edges with half-edges incident to non- $\eta^1$ -forcing clauses remains uniformly random. Since the number of edges joining non- $\eta^1$ -forcing clauses to  $\eta^1$ -rigid variables is  $\asymp mk$ , we conclude that  $\eta^2$ -forcing arrangements of clause  $a$  will occur with conditional probability  $\lesssim \epsilon k/2^k$ . Moreover

this estimate remains valid even after conditioning on a subset of such clauses, since the total number of clauses incident to free variables is negligible compared with  $mk$ .

2. In the  $m\nu_{\geq 2}$  clauses incident to more than one  $\eta^1$ -free variable, the conditioning so far gives no information about the arrangement of the literals. Therefore, in each such clause distinguish a uniformly random edge to be *potentially*  $\eta^2$ -forcing. For  $v \in V_{\mathbf{f}_\perp}$  let  $A_v$  denote the number of such edges incident to  $v$ , and write  $m_A \equiv \sum_{v \in V_{\mathbf{f}_\perp}} A_v \leq m\nu_{\geq 2}$ . A clause  $a \in \partial v$  of this type is  $\eta^2$ -forcing to  $v$  for certain arrangements of literals and of spins  $\mathbf{r}_r, \mathbf{f}_r$  among the neighbors  $u \in \partial a \setminus v$ . In particular, at least one neighbor  $u \in \partial a \setminus v$  is  $\eta^1$ -free, therefore must have spin  $\mathbf{f}_r$  in the pair model. Thus  $\eta^2$ -forcing arrangements of  $a$  will occur with conditional probability  $\leq (k/2^k)(\pi_{\mathbf{f}_r}/\beta)$ . Moreover this estimate remains valid even after conditioning on a subset of such clauses — due to the restriction to event  $\Omega_A$  we have  $m_A \leq nd\beta^2 k^{O(1)}$ , and this is negligible compared with the total number  $nd\beta$  of half-edges leaving  $V_{\mathbf{f}_\perp}$ .

Crudely bounding  $\tilde{\Lambda}_v \leq d$ , there exists a uniform constant  $C$  such that

$$\mathbf{p}_{\mathbf{f}}^{\pi, \mathbf{x}, \delta} \equiv \mathbb{P}(Q_{\mathbf{f}} | \eta^1 \text{ valid}, \Omega_{B, \mathbf{x}}, F_{\delta}, Q_{\mathbf{r}}, \Omega_A) \leq \mathbb{E} \left[ \prod_{v \in V_{\mathbf{f}_\perp}} \left\{ 1 - \left( 1 - \frac{C\epsilon k}{2^k} \right)^d \left( 1 - \frac{Ck\pi_{\mathbf{f}_r}}{2^k\beta} \right)^{A_v} \right\} \right].$$

Let  $B, (B_v)_{v \in V_{\mathbf{f}_\perp}}$  be i.i.d.  $\text{Bin}(d, m_A/(nd\beta))$  random variables; we can compute the above by replacing  $A_v$  with  $B_v$  and conditioning on  $\sum_{v \in V_{\mathbf{f}_\perp}} B_v = m_A$ . We then claim that removing this conditioning costs at most  $e^{O(n\pi_{\mathbf{f}_r})}$ , that is to say,

$$\mathbf{p}_{\mathbf{f}}^{\pi, \mathbf{x}, \delta} \lesssim e^{O(n\pi_{\mathbf{f}_r})} \left[ 1 - \left( 1 - \frac{C\epsilon k}{2^k} \right)^d \mathbb{E} \left\{ \left( 1 - \frac{Ck\pi_{\mathbf{f}_r}}{2^k\beta} \right)^B \right\} \right]^{n\pi_{\mathbf{f}_r}}.$$

If  $\pi_{\mathbf{f}_r} \geq n/(\log n)$  this is immediate from the local CLT; otherwise it follows by arguing as in (39). Recalling  $\mathbb{E}[(1-x)^{\text{Bin}(d,p)}] = (1-px)^d$  we see that

$$\mathbf{p}_{\mathbf{f}}^{\pi, \mathbf{x}, \delta} \leq e^{O(n\pi_{\mathbf{f}_r})} \left( \frac{dk\epsilon}{2^k} + \frac{dk\pi_{\mathbf{f}_r} m\nu_{\geq 2}}{2^k\beta nd\beta} \right)^{n\pi_{\mathbf{f}_r}} \leq k^{O(n\pi_{\mathbf{f}_r})} (\epsilon + \pi_{\mathbf{f}_r})^{n\pi_{\mathbf{f}_r}}.$$

Again recalling  $H(x) + x \log c \leq \log(1+c) \leq c$  we bound  $\mathbf{c}_{\mathbf{f}}^{\pi, \mathbf{x}, \delta} \mathbf{p}_{\mathbf{f}}^{\pi, \mathbf{x}, \delta} \leq k^{O(n\pi_{\mathbf{f}_r})} e^{n\beta\epsilon}$ . Now assume by symmetry that  $\pi_{\mathbf{f}_r} \leq \pi_{\mathbf{r}_f}$ : then  $2n\epsilon \geq n\pi_{\mathbf{f}_r} + n\pi_{\mathbf{r}_f} = \Delta$  and  $\mathbf{c}_{\mathbf{f}} \mathbf{p}_{\mathbf{f}} \leq k^{O(n\epsilon)}$ . Combining with (40) then gives that

$$\text{if } 10\epsilon \leq \beta \text{ then } \mathbb{E}[Z^2[\pi]; \Omega_B] \leq \frac{\mathbb{E}[Z[\pi^1]; \Omega_B]}{2^{n\epsilon k/3}} \leq \frac{\mathbb{E}[Z[\pi^1]; \Omega_B]}{2^{\Delta k/6}},$$

concluding the proof. □

*Proof of Propn. 4.7.* Follows from Lem. 4.9. □

*Proof of Thm. 4.1.* Follows by combining Propns. 4.3 and 4.7. □

## 5. NEGATIVE-DEFINITENESS OF FREE ENERGY HESSIANS

In this section we prove Thm. 2.

**Proposition 5.1.** *The Hessians  $H\Phi(\ast\mathbf{h})$  and  $H_2\Phi(\ast\mathbf{h})$  are negative-definite.*

**5.1. Derivatives of the Bethe functional.** Let  $\mathbf{h} \in \Delta^\circ$  with  $\dot{\mathbf{h}}$  and  $\hat{\mathbf{h}}$  both symmetric, and let  $\delta$  be any signed measure on  $\text{supp } \varphi$  (not necessarily symmetric) with  $\mathbf{h} + s\delta \in \Delta^\circ$  for sufficiently small  $|s|$ . Then

$$k \partial_s^2 \Phi(\mathbf{h} + s\delta)|_{s=0} = -k \langle (\dot{\delta}/\dot{\mathbf{h}})^2 \rangle_{\dot{\mathbf{h}}} - d \langle (\hat{\delta}/\hat{\mathbf{h}})^2 \rangle_{\hat{\mathbf{h}}} + dk \langle (\bar{\delta}/\bar{\mathbf{h}})^2 \rangle_{\bar{\mathbf{h}}}$$

where  $a/b$  denotes the vector given by coordinate-wise division of  $a$  by  $b$ , and  $\langle \cdot \rangle_h$  denotes integration with respect to measure  $h$ , e.g.  $\langle (\bar{\delta}/\bar{\mathbf{h}})^2 \rangle_{\bar{\mathbf{h}}} = \sum_{\sigma} \delta(\sigma)^2 / \bar{h}(\sigma)$ .

Given fixed marginals  $\bar{\delta}$ ,  $\langle(\delta/\hbar)^2\rangle_{\hbar}$  is minimized by  $\delta(\underline{\sigma}) = \hbar(\underline{\sigma})d^{-1}\sum_{i=1}^d\dot{\chi}_{\sigma_i}$  with  $\dot{\chi}$  chosen to satisfy the margin constraint — which, after a little algebra, becomes the vector equation  $\bar{H}^{-1}\bar{\delta} = d^{-1}[I + (d-1)\hat{M}]\dot{\chi}$  where  $\bar{H} \equiv \text{diag}(\bar{h})$  and  $\hat{M}$  denotes the stochastic matrix

$$\hat{M}_{\sigma,\sigma'} \equiv \bar{h}(\sigma)^{-1} \sum_{\underline{\sigma}} \dot{\hbar}(\underline{\sigma}) \mathbf{1}\{(\sigma_1, \sigma_2) = (\sigma, \sigma')\}. \quad (41)$$

If such  $\dot{\chi}$  exists, then the minimal value of  $\langle(\delta/\hbar)^2\rangle_{\hbar}$  subject to marginals  $\bar{\delta}$  is  $\langle\bar{\delta}, \dot{\chi}\rangle$ . Define analogously the stochastic matrix  $\hat{M}$  corresponding to  $\hat{h}$ : if both  $\hat{L} \equiv I + (d-1)\hat{M}$  and  $\hat{L} \equiv I + (k-1)\hat{M}$  are non-singular, then the maximum of  $k\partial_{\eta}^2\Phi(\mathbf{h} + \eta\delta)|_{\eta=0}$  over all  $\delta$  with marginal  $\bar{\delta}$  is given by

$$-dk\bar{\delta}^t[(\bar{H}\hat{L})^{-1} + (\bar{H}\hat{L})^{-1} - \bar{H}^{-1}]\bar{\delta} = -dk(\bar{H}^{-1/2}\bar{\delta})^t F(\bar{H}^{-1/2}\bar{\delta})$$

where  $F \equiv (\bar{H}^{1/2}\hat{L}\bar{H}^{-1/2})^{-1} + (\bar{H}^{1/2}\hat{L}\bar{H}^{-1/2})^{-1} - I$ . It is clear from (41) that  $\hat{M}$  and  $\hat{M}$  are  $\bar{h}$ -reversible, therefore  $F$  is symmetric. Since  $\sum_{\sigma}\bar{\delta}(\sigma) = 0$  we consider only the action  $F'$  of  $F$  on the space of vectors orthogonal to  $\bar{h}^{1/2}$ . At a global maximizer we know  $F'$  to be negative-semidefinite, so if  $\det F \neq 0$  then it is in fact negative-definite. Thus let  $\hat{M}$ ,  $\hat{M}$ ,  $\hat{M}_2$ ,  $\hat{M}_2$  denote the Markov transition matrices corresponding (via (41)) to  $^*\mathbf{h}$ ,  $^*\hat{\mathbf{h}}$ ,  $^*\hat{\mathbf{h}}$ ,  $^*\hat{\mathbf{h}}$  respectively. In §5.2 we will prove that the matrices

$$\begin{aligned} \hat{L} &\equiv I + (d-1)\hat{M}, & \hat{L}_2 &\equiv I + (d-1)\hat{M}_2, \\ \hat{L} &\equiv I + (k-1)\hat{M}, & \hat{L}_2 &\equiv I + (k-1)\hat{M}_2, \\ L &\equiv I - (d-1)(k-1)\hat{M}\hat{M}, & L_2 &\equiv I - (d-1)(k-1)\hat{M}_2\hat{M}_2 \end{aligned} \quad (42)$$

are all non-singular. Propn. 5.1 then follows by noting that  $F = \bar{H}^{1/2}\hat{L}^{-1}L\hat{L}^{-1}\bar{H}^{-1/2}$ .

**5.2. Calculation of transition matrices.** Recall the notation  $\dot{q} \equiv \dot{q}^*$ ,  $q_{\mathbf{x}} \equiv 1 - \dot{q} = 2^{-k}$ , and  $v_{\mathbf{x}} \equiv 1 - v = 2^{-k}$ . Recalling (24) that  $\bar{h}^*(\sigma)$  is proportional to  $\dot{h}_{\sigma}^*\hat{h}_{\sigma}^*$ , we record here that

$$\bar{h} = [2(1 + q_{\mathbf{x}}v_{\mathbf{x}})]^{-1} \times \begin{pmatrix} \text{of} & \text{oo} & \text{fo} & \text{ff} & \text{of} & \text{oo} & \text{fo} & \text{ff} \\ \dot{q}v & \dot{q}v_{\mathbf{x}} & 2q_{\mathbf{x}}v_{\mathbf{x}} & \dot{q}v & \dot{q}v_{\mathbf{x}} & 2q_{\mathbf{x}}v_{\mathbf{x}} & \dot{q}v & \dot{q}v_{\mathbf{x}} \end{pmatrix} \quad (43)$$

**Lemma 5.2.** *The eigenvalues of  $\hat{M}$  counted with geometric multiplicity are*

$$\text{eigen}(\hat{M}) = (1, 1, 1, \lambda, \lambda, -\lambda, -\lambda) \quad \text{with } \lambda = 2^{-3k/2}.$$

*The matrix  $\hat{M}_2$  is given by  $\hat{M} \otimes \hat{M}$ ; consequently both  $\hat{L}$  and  $\hat{L}_2$  are non-singular.*

*Proof.* The transition matrix  $\hat{M} \in \mathbb{R}^{7 \times 7}$  is block diagonal with blocks  $\hat{m}_{\mathbf{x}}$ ,  $\hat{m}_0$ ,  $\hat{m}_1$  where  $\hat{m}_{\mathbf{x}}$  is the one-dimensional identity matrix (the action of  $\hat{M}$  on  $\{\text{ff}\}$ ), and for  $x = 0, 1$  the matrix  $\hat{m}_x \in \mathbb{R}^{3 \times 3}$  gives action of  $\hat{M}$  on  $\{\text{xf}, \text{xx}, \text{fx}\}$ . Recalling the definition (41), the entries of  $\hat{M}$  are straightforwardly calculated from (22), (23), and (24): for example,

$$\hat{M}_{00,00} = \frac{\hat{g}_{\mathbf{r}\mathbf{r}}(\hat{g}_{\mathbf{r}\mathbf{r}} + \hat{g}_{\mathbf{r}\mathbf{f}})^{d-2}}{(\hat{g}_{\mathbf{r}\mathbf{r}} + \hat{g}_{\mathbf{r}\mathbf{f}})^{d-1} - (\hat{g}_{\mathbf{r}\mathbf{f}})^{d-1}} = \frac{v_{\mathbf{x}}}{1 - v^{d-1}} = v_{\mathbf{x}} \frac{1 + q_{\mathbf{x}}}{1 - q_{\mathbf{x}}}$$

where the last step uses (13). We therefore find  $\hat{m}_0 = \hat{m}_1 = \hat{m}$  where

$$\hat{m} \equiv \begin{pmatrix} \text{rf} & \text{rr} & \text{fr} \\ \text{rf} & 1-a & a(1-b) & ab \\ \text{rr} & 1-a & a & 0 \\ \text{fr} & 1 & 0 & 0 \end{pmatrix} \quad \text{with} \quad \begin{cases} a \equiv v_{\mathbf{x}}(1 + q_{\mathbf{x}})/(1 - q_{\mathbf{x}}) = 2^{-k}, \\ b \equiv (2v_{\mathbf{x}}q_{\mathbf{x}})/(v\dot{q}) = 4^{-k}, \end{cases}$$

which has  $\text{eigen}(\hat{m}) = (1, ab^{1/2}, -ab^{1/2})$ . Thus  $\text{eigen}(\hat{M}) = (1, \text{eigen}(\hat{m}), \text{eigen}(\hat{m}))$  is as stated above. Since  $^*\hat{\mathbf{h}} = ^*\mathbf{h} \otimes ^*\mathbf{h}$ , clearly  $\hat{M}_2 = \hat{M} \otimes \hat{M}$ , so the lemma is proved.  $\square$

**Lemma 5.3.** *There exist (explicit) irreducible  $\bar{h}$ -reversible transition matrices  $\hat{M}^0, \hat{M}^1$  such that  $\hat{M} = [\hat{M}^0 + \hat{M}^1]/2$ ,  $\hat{M}_2 = [\hat{M}^0 \otimes \hat{M}^0 + \hat{M}^1 \otimes \hat{M}^1]/2$ . The eigenvalues of  $\hat{M}^x$  counted with geometric multiplicity are  $\text{eigen}(\hat{M}^x) = (1, \lambda_2, \dots, \lambda_7)$  with  $2^{k/2}|\lambda_i| \lesssim 1$  for all  $i > 2$ ; consequently both  $\hat{L}$  and  $\hat{L}_2$  are non-singular.*



*Proof.* Clearly  $\hat{M}_{\text{rf},\sigma} = \hat{M}_{\text{rf},\sigma} = \mathbf{1}\{\sigma = \text{rf}\}$ , and it is straightforward to calculate that

$$\hat{M}_{\text{ff},\text{ff}} = 1 - \hat{M}_{\text{ff},\text{rf}} = \delta \equiv q_{\text{f}} \frac{1 + v_{\text{r}}}{1 - v_{\text{r}}} \approx 2^{-k}.$$

The remaining entries of  $\hat{M}$  are easily determined by  $\bar{h}$ -reversibility (see (43)): with  $\gamma \equiv v_{\text{r}}/v = 2^{-k}$  and  $\epsilon \equiv 2q_{\text{f}}/\bar{q} \approx 2^{-k}$ , we calculate

$$\hat{M} = \begin{matrix} & \mathcal{M}_0 & \mathcal{M}_1 & \text{ff} \\ \mathcal{M}_0 & \left( \begin{array}{ccc} \hat{m}/2 & \hat{m}/2 & \hat{b} \\ \hat{m}/2 & \hat{m}/2 & \hat{b} \\ \hat{a}^t/2 & \hat{a}^t/2 & \delta \end{array} \right) & & \\ \mathcal{M}_1 & & & \\ \text{ff} & & & \end{matrix}, \quad \text{where } \begin{pmatrix} \hat{m} & \hat{b} \\ \hat{a}^t & \delta \end{pmatrix} \equiv \begin{matrix} & \text{rf} & \text{rr} & \text{fr} & \text{ff} \\ \text{rf} & \left( \begin{array}{cccc} 1 - \delta - \gamma & \gamma & \epsilon\gamma & \delta - \epsilon\gamma \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 - \delta & 0 & 0 & \delta \end{array} \right) \end{matrix}$$

( $\hat{m} \in \mathbb{R}^{3 \times 3}$  while  $\hat{a}, \hat{b} \in \mathbb{R}^{3 \times 1}$ ). Now consider decomposing  $\hat{M} = 2^{-k} \sum_{\underline{L}} \hat{M}(\underline{L})$  where  $\hat{M}(\underline{L})$  is the transition matrix corresponding to  $\hat{h}(\cdot | \underline{L})$ : there are only two possibilities  $\hat{M}^0, \hat{M}^1$  for  $\hat{M}(\underline{L})$ , depending on whether  $\underline{L}_1 \oplus \underline{L}_2 = 0, 1$ . Since  $\hat{h}(\hat{x}) = 2^{-k} \sum_{\underline{L}} \hat{h}(\hat{x}^1 | \underline{L}) \hat{h}(\hat{x}^2 | \underline{L})$ , we conclude

$$\hat{M}_2 = [\hat{M}^0 \otimes \hat{M}^0 + \hat{M}^1 \otimes \hat{M}^1]/2.$$

The entries of each matrix  $\hat{M}^x$  are easily read from  $\hat{M}$  except the ones giving the transition probabilities within  $\{0\text{f}, 1\text{f}\}$ . We calculate these from (15) to find that

$$\hat{M}^0 = \begin{matrix} & 0\text{f} & 00 & 1\text{f} & 11 & 1\text{f} & 1\text{f} \\ 0\text{f} & \left( \begin{array}{cccccc} \frac{1+B}{2} \hat{m}_{\text{rf},\text{rf}} & 0 & 0 & \frac{1-B}{2} \hat{m}_{\text{rf},\text{rf}} & \gamma & \epsilon\gamma & \delta - \epsilon\gamma \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1-B}{2} \hat{m}_{\text{rf},\text{rf}} & \gamma & \epsilon\gamma & \frac{1+B}{2} \hat{m}_{\text{rf},\text{rf}} & 0 & 0 & \delta - \epsilon\gamma \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2}(1-\delta) & 0 & 0 & \frac{1}{2}(1-\delta) & 0 & 0 & \delta \end{array} \right) \end{matrix}$$

where  $B = v_{\text{r}}/(1 - q_{\text{f}} - 2v_{\text{r}} - q_{\text{f}}v_{\text{r}})$ . Then  $\hat{M}^1$  is defined by exchanging the roles of 0 and 1.

We see in particular that  $\hat{M}^0$  and  $\hat{M}^1$  are  $\bar{h}$ -reversible,<sup>7</sup> so, writing  $\bar{H} \equiv \text{diag}(\bar{h})$ , the matrix  $\bar{H}^{1/2} \hat{M}^x \bar{H}^{-1/2}$  is symmetric and hence orthogonally diagonalizable. Let  $u$  be a left eigenvector of  $\hat{M}^x$  with eigenvalue  $\lambda$ , such that  $u$  has norm 1 and is orthogonal to the constant vector  $\underline{1}$ . Suppose  $2^k |\lambda| \gg 1$ : the eigenvalue equations

$$\begin{aligned} \lambda u_{xx} &= \gamma u_{-x\text{f}}, & \lambda u_{\text{f}x} &= \epsilon\gamma u_{-x\text{f}}, \\ \lambda u_{\text{ff}} &= (\delta - \epsilon\gamma)(u_{0\text{f}} + u_{1\text{f}}) + \delta u_{\text{ff}} \end{aligned}$$

imply  $|u_{\sigma}| \lesssim (2^k |\lambda|)^{-1} \ll 1$  for all  $\sigma \notin \text{rf}$ , so using  $\|u\| = 1$ ,  $\langle u, \underline{1} \rangle = 0$  we conclude  $|u_{0\text{f}}| \approx |u_{1\text{f}}| \approx 1$  with  $|u_{0\text{f}} + u_{1\text{f}}| \lesssim (2^k |\lambda|)^{-1}$ . The eigenvalue equation for  $u_{0\text{f}}$  gives  $|\lambda| \approx |\lambda u_{0\text{f}}| \lesssim B + (2^k |\lambda|)^{-1}$ ; rearranging and recalling  $B \lesssim 2^{-k}$  then gives  $2^k |\lambda|^2 \lesssim |\lambda| + 1 \lesssim 2$  which proves the lemma.  $\square$

**Lemma 5.4.** *The matrices  $L$  and  $L_2$  are non-singular.*

*Proof.* For  $A \subseteq \mathcal{M}$  write  $\bar{h}_A \equiv (\bar{h} \mathbf{1}_A)/\bar{h}(A)$ , the stationary distribution  $\bar{h}$  conditioned on  $A$ . The vectors

$$u_1 \equiv \bar{h}, \quad u_2 \equiv (\bar{h}_{\mathcal{M}_0} + \bar{h}_{\mathcal{M}_1})/2 - \bar{h}_{\text{ff}}, \quad u_3 \equiv \bar{h}_{\mathcal{M}_0} - \bar{h}_{\mathcal{M}_1}$$

are left eigenvectors of  $\hat{M}$  with eigenvalue 1 such that  $(w_i \equiv u_i/\bar{h}^{1/2})_{i=1}^3$  forms an orthogonal basis for the 1-eigenspace of the symmetrized matrix  $\hat{S} \equiv \bar{H}^{1/2} \hat{M} \bar{H}^{-1/2}$ . Define likewise  $\hat{S} \equiv \bar{H}^{1/2} \hat{M} \bar{H}^{-1/2}$ ; if  $w$  is orthogonal to this 1-eigenspace, then Lem. 5.2 implies  $\|w^t \hat{S} \hat{S}\| = O(\|w^t \hat{S}\|) = O(2^{-3k/2} \|w\|)$ , so it remains to consider the action of  $\hat{S} \hat{S}$  on the 1-eigenspace of  $\hat{S}$ . Clearly  $u_1^t \hat{M} = u_1^t$ , and  $u_3^t \hat{M} = 0$

<sup>7</sup>To see this without explicit calculation of  $\hat{M}^x$ , simply observe that (by symmetry) each  $\hat{h}(\cdot | \underline{L})$  has marginal  $\bar{h}$ , and so from (41) we have  $\bar{h}(\sigma) \hat{M}_{\sigma,\sigma'}^x = \sum_{\hat{x}} \hat{h}(\hat{x} | L_1 \oplus L_2 = x) = \bar{h}(\sigma') \hat{M}_{\sigma',\sigma}^x$ .

by symmetry. Since  $\bar{h}_{\text{ff}}$  is simply the indicator of  $\text{ff}$ , clearly  $\|w_2\| = \|\bar{h}_{\text{ff}}/\bar{h}^{1/2}\| = \bar{h}(\text{ff})^{-1/2} = 2^{k/2}$ ; and we calculate

$$u_2^t \dot{M} \dot{M} = u_2^t \hat{M} = (1/4 \ 0 \ 0 \ 1/4 \ 0 \ 0 \ 0) + O(2^{-k}),$$

therefore  $\|w_2^t \dot{S} \hat{S}\|/\|w_2\| = \|(u_2^t \dot{M} \hat{M})/\bar{h}^{1/2}\|/\|w_2\| = 2^{-k/2}$ . It follows that  $\dot{M} \hat{M}$  (equivalently  $\dot{S} \hat{S}$ ) can have no eigenvalue with absolute value  $\asymp 1/(dk)$ , hence  $L$  is non-singular.

For  $1 \leq i, j \leq 3$  let  $w_{ij} \equiv w_i \otimes w_j$ , and note that if  $w$  is orthogonal to the span of  $(w_{11}, w_{12}, w_{21})$  then  $\|w^t \dot{S}_2 \hat{S}_2\| = O(2^{-3k/2})\|w\|$ . Next note that

$$w_{12}^t \dot{S}_2 \hat{S}_2 = \frac{(u_1^t \otimes u_2^t)[\dot{M}^0 \otimes \hat{M}^0 + \hat{M}^1 \otimes \hat{M}^1]}{2\bar{h}^{1/2}} = \frac{u_1^t \otimes (u_2^t \hat{M})}{\bar{h}^{1/2}} = w_1^t \otimes (w_2^t \hat{S}),$$

so  $\|w_{12}^t \dot{S}_2 \hat{S}_2\|/\|w_{12}\| = \|w_2^t \hat{S}\|/\|w_2\| = 2^{-k/2}$ . Since  $w_{12} \dot{S}_2 \hat{S}_2$  and  $w_{21} \dot{S}_2 \hat{S}_2$  are orthogonal,

$$\frac{\|(aw_{12} + bw_{21})^t \dot{S}_2 \hat{S}_2\|^2}{\|aw_{12} + bw_{21}\|^2} = \frac{|a|^2 \|w_{12}^t \dot{S}_2 \hat{S}_2\|^2 + |b|^2 \|w_{21}^t \dot{S}_2 \hat{S}_2\|^2}{(|a|^2 + |b|^2) \|w_{12}\|^2} = 2^{-k/2}$$

for any  $a, b \in \mathbb{C}$  (not both zero). It follows that  $\dot{M}_2 \hat{M}_2$  (equivalently  $\dot{S}_2 \hat{S}_2$ ) can have no eigenvalue with absolute value  $\asymp 1/(dk)$ , proving that  $L_2$  is also non-singular.  $\square$

*Proof of Propn. 5.1.* As shown in §5.1 the result follows by verifying that the matrices defined in (42) are non-singular, which is done by the lemmas of §5.2.  $\square$

*Proof of Thm. 2.* Recall (Defn. 3.5) that  $Z$  is the sum of  $Z(\mathbf{h})$  over probability measures  $\mathbf{h} \equiv (\hat{\mathbf{h}}, \check{\mathbf{h}})$  on  $\text{supp } \varphi$  such that  $\mathbf{g} \equiv (\dot{\mathbf{g}}, \hat{\mathbf{g}}) \equiv (n\hat{\mathbf{h}}, m\check{\mathbf{h}})$  is integer-valued, and lies in the kernel of matrix  $H_\Delta \equiv (\dot{H} \ -\hat{H})$ . Let  ${}^*Z$  denote the contribution to  $Z$  from (non-normalized) measures  $\mathbf{g}$  within euclidean distance  $n^{1/2} \log n$  of  ${}^*\mathbf{g}$ . Thm. 3.7 and Propn. 5.1 together give  $\mathbb{E}Z = [1 + o(n^{-1})] \mathbb{E}[{}^*Z]$ . By Lem. 6.6, the integer matrix  $H_\Delta$  defines a surjection

$$L' \equiv \{\delta \in \mathbb{R}^{\text{supp } \varphi} : \langle \delta, 1 \rangle = \langle \bar{\delta}, 1 \rangle = 0\} \quad \text{to} \quad \{\bar{\delta} \in \mathbb{R}^{\mathcal{M}} : \langle \bar{\delta}, 1 \rangle = 0\},$$

so  $L \equiv L' \cap (\ker H_\Delta) \cap Z^{\text{supp } \varphi}$  is an  $(\dot{s} + \check{s} - \bar{s} - 1)$ -dimensional lattice with spacings  $\asymp_k 1$ . The measures  $\mathbf{g}$  contributing to  ${}^*Z$  are given by the intersection of the ball  $\{\|\mathbf{g} - {}^*\mathbf{g}\| \leq n^{1/2} \log n\}$  with an affine translation of  $L$ . The expansion (18) then shows that  $\mathbb{E}[{}^*Z]$  defines a convergent Riemann sum, therefore  $\mathbb{E}[{}^*Z] \asymp_k \exp\{n \Phi({}^*\mathbf{h})\}$  as claimed.

In the pair partition function  $Z^2$ , let  ${}_2^*Z$  denote the contribution from (non-normalized) measures  ${}_2\mathbf{g}$  within euclidean distance  $n^{1/2} \log n$  of the independent-copies local maximizer  ${}_2^*\mathbf{g} = {}^*\mathbf{g} \otimes {}^*\mathbf{g}$ . Recall from the statement of Lem. 4.2 that  ${}_2Z$  denotes the contribution to  $Z^2$  from the near-identical measures  ${}_2\Delta$ . Decompose

$$Z^2 = \left( \begin{array}{c} \text{near-independent} \\ \text{contribution } {}_2^*Z \end{array} \right) + \left( \begin{array}{c} \text{near-identical} \\ \text{contribution } {}_2Z \end{array} \right) + (\text{remainder}), \quad (44)$$

and note that for  $d_{\text{ibd}} \leq d \leq d_{\text{ubd}}$ , Thm. 4.1 and Propn. 5.1 together imply that the expectation of the remainder is a negligible fraction of  $\mathbb{E}[Z^2]$ :

$$\mathbb{E}[Z^2] = [1 + o(n^{-1})] (\mathbb{E}[{}_2^*Z] + \mathbb{E}[{}_2Z]).$$

Repeating the argument above gives  $\mathbb{E}[{}_2^*Z] \asymp_k \exp\{n({}_2\Phi({}_2^*\mathbf{h}))\} \asymp_k (\mathbb{E}Z)^2$ , and combining with Propn. 4.7 gives the conclusion  $\mathbb{E}[Z^2] \lesssim_k (\mathbb{E}Z)^2 + n^{O(1)} (\mathbb{E}Z)$ .  $\square$

## 6. FROM CONSTANT TO HIGH PROBABILITY

In this section we prove Thm. 3. Recall from the proof of Thm. 2 that  ${}^*Z$  denotes the contribution to the auxiliary model partition function on  $(G, \underline{L})$  from configurations whose non-normalized empirical measure  $\mathbf{g} \equiv (n\hat{\mathbf{h}}, m\check{\mathbf{h}})$  lies within euclidean distance  $n^{1/2} \log n$  of  ${}^*\mathbf{g}$ . The main result of this section is the following

**Proposition 6.1.** *For  $k \geq k_0$ , there exists a constant  $C_k$  such that*

$$\limsup_{n \rightarrow \infty} \text{Var} \log(*Z + \epsilon \mathbb{E}Z) \leq C_k$$

for all  $d_{\text{ibd}} \leq d < d_*$  and all  $\epsilon \leq \epsilon_k$ .

The proposition easily implies a strengthened version of the second part of Thm. 3:

**Theorem 6.2.** *For  $k \geq k_0$  and  $d_{\text{ibd}} \leq d < d_*$ ,  $\lim_{\epsilon \downarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}(*Z > \epsilon \mathbb{E}Z) = 1$ .*

*Proof.* Write  $L \equiv L(\epsilon) \equiv \log(*Z + \epsilon \mathbb{E}Z)$ . For  $d_{\text{ibd}} \leq d < d_*$  we have  $\mathbb{E}[(*Z)^2] \asymp_k (\mathbb{E}Z)^2$ , so there exists a constant  $\delta \equiv \delta(k, d) > 0$  for which  $\mathbb{P}(*Z \geq \delta \mathbb{E}Z) \geq \delta$ . Consequently

$$\mathbb{E}L \geq (1 - \delta) \log(\epsilon \mathbb{E}Z) + \delta \log((\epsilon + \delta) \mathbb{E}Z) = \log(\epsilon \mathbb{E}Z) + \delta \log(1 + \delta/\epsilon).$$

On the other hand,  $*Z \leq \epsilon \mathbb{E}Z$  if and only if  $L \leq \log(2\epsilon \mathbb{E}Z)$ , which for small  $\epsilon > 0$  is much less than the lower bound on  $\mathbb{E}L$ . Applying Chebychev's inequality and Propn. 6.1 therefore gives

$$\limsup_{n \rightarrow \infty} \mathbb{P}(*Z \leq \epsilon \mathbb{E}Z) \leq \limsup_{n \rightarrow \infty} \frac{\text{Var} L}{(\mathbb{E}L - \log(2\epsilon \mathbb{E}Z))^2} \leq \frac{C_k}{[\delta \log(1 + \delta/\epsilon) - \log 2]^2}.$$

Taking  $n \rightarrow \infty$  followed by  $\epsilon \downarrow 0$  proves the theorem.  $\square$

We shall prove Propn. 6.1 by controlling the increments of the Doob martingale of the random variable  $L \equiv \log(*Z + \epsilon \mathbb{E}Z)$  with respect to the edge-revealing filtration  $(\mathcal{F}_i)_{1 \leq i \leq m}$  for the graph  $G_n \sim \mathcal{G}_{n,d,k}$ . We will show that the variance of  $L$  has two dominant components: the first is an “independent-copies contribution” coming from pair configurations with empirical measure near  $\frac{1}{2}\mathbf{h} \equiv \frac{1}{2}\mathbf{h} \otimes \frac{1}{2}\mathbf{h}$ , which we will show in this section to be  $\lesssim_k 1$ . The other dominant component in the variance of  $L$  is an “identical-copies contribution” coming from pair configurations with empirical measure near  $\{\frac{1}{2}\mathbf{h}, \frac{1}{2}\mathbf{h}\}$ , which we will easily see to be exponentially small in  $n$  simply by the assumption  $d < d_*$ . In [DSS13] we demonstrate how to control the identical-copies contribution assuming only that the first moment is bounded below by a large constant.

**6.1. Doob martingale and coupling argument.** We shall prove Propn. 6.1 in a slightly more general setting, for application in [DSS13]. Given a parameter  $w(n)$ , let

$$W = (a_{m-w(n)+1}, \dots, a_m)$$

be the last  $w(n)$  clauses in  $G$ . Let  $G^\dagger \equiv G \setminus W$  be the graph with these clauses removed.

**Proposition 6.3.** *For  $k \geq k_0$  and  $w(n) \lesssim_{k,\epsilon} \log n$ , there exists a constant  $C_k$  such that*

$$\limsup_{n \rightarrow \infty} \text{Var} \log(\mathbb{E}[*Z | G^\dagger] + \epsilon \mathbb{E}Z) \leq C_k$$

for all  $d_{\text{ibd}} \leq d < d_*$  and all  $\epsilon \leq \epsilon_k$ .

**Remark 6.4.** Propn. 6.1 follows trivially from Propn. 6.3 by taking  $w(n) \equiv 0$ . We use Propn. 6.1 only in the present paper. However the methods of Propn. 6.3 will be applied in [DSS13] to prove a sharper bound on the threshold fluctuations.

Write  $L^\dagger \equiv L(G^\dagger, \epsilon) \equiv \mathbb{E}[*Z | G^\dagger] + \epsilon \mathbb{E}Z$ . Consider forming  $G$  step by step as follows. Start with  $n$  isolated vertices, each equipped with  $d$  incident half-edges. For each  $1 \leq i \leq m$ , choose a uniformly random set of  $k$  unmatched half-edges, and join these into the  $i$ -th clause  $a_i$  (the clause is also assigned random literals). Let  $\mathcal{F} \equiv (\mathcal{F}_i)_{0 \leq i \leq m}$  be the filtration generated by this clause-revealing process. The Doob martingale of  $L^\dagger$  with respect to  $\mathcal{F}$  is the process  $(\mathbb{E}[L^\dagger | \mathcal{F}_i])_{1 \leq i \leq m^\dagger}$ , and we have the Doob variance decomposition

$$\text{Var} L^\dagger = \sum_{i=1}^{m^\dagger} \overbrace{\mathbb{E} \left[ (\mathbb{E}[L^\dagger | \mathcal{F}_i] - \mathbb{E}[L^\dagger | \mathcal{F}_{i-1}])^2 \right]}^{\text{denote this Var}_i L^\dagger}.$$

To prove Propn. 6.1 we control the incremental fluctuations  $\text{Var}_i L^\dagger$ , beginning a coupling argument. Let  $G$  be a random  $(d, k)$ -regular graph, and let  $A$  be the clauses with indices between  $\max\{i, m^\dagger - k + 1\}$  and  $m^\dagger$ . Let  $\mathcal{X}$  be the set of variable-incident half-edges involved in the clauses  $A$ . Let  $\hat{G}$  be another random  $(d, k)$ -regular graph which agrees with  $G$ , except that we randomly resample an independent arrangement  $\hat{A}$  of clauses on  $\mathcal{X}$ :

$$\begin{aligned} A &\equiv (a_{\max\{i, m^\dagger - k + 1\}}, \dots, a_{m^\dagger}), \\ \hat{A} &\equiv (\hat{a}_{\max\{i, m^\dagger - k + 1\}}, \dots, \hat{a}_{m^\dagger}). \end{aligned}$$

Write  $\mathcal{S}_j \equiv \sigma(\mathcal{F}_j, a_{m^\dagger}) = \sigma(a_1, \dots, a_j, \mathcal{X}, a_{m^\dagger})$ , and write  $\mathbb{E}_j$  for expectation conditioned on  $\mathcal{S}_j$ . Then the random variable  $\mathbb{E}[L^\dagger | \mathcal{F}_i] - \mathbb{E}[L^\dagger | \mathcal{F}_{i-1}]$  is equidistributed as

$$\mathbb{E}_{i-1} \Delta \quad \text{where } \Delta \equiv L(G^\dagger, \epsilon) - L(\hat{G}^\dagger, \epsilon).$$

The law of  $(G, \hat{G})$  depends on  $i$ , but is the same for all  $i \leq m^\dagger - k + 1$ , so it follows from Jensen's inequality that for all such  $i$  we have

$$\text{Var}_i L^\dagger = \mathbb{E}[(\mathbb{E}_{i-1} \Delta)^2] \leq \mathbb{E}[(\mathbb{E}_{m^\dagger - k} \Delta)^2] \leq \mathbb{E}[\mathbb{E}_{m^\dagger - k} (\Delta^2)].$$

Recall  $\mathbb{E}_{m^\dagger - k}$  means that we condition on the graph  $G^\circ$  with clauses  $(a_1, \dots, a_{m^\dagger - k})$ , as well as on  $(\mathcal{X}, a_{m^\dagger})$ , then average over the possible arrangements of  $(A, \hat{A})$  on a set  $\mathcal{X}$  of  $k^2$  half-edges. For  $m^\dagger - k + 1 < i \leq m^\dagger$ ,  $|\mathcal{X}| = (m - i + 1)k < k^2$ , so  $\mathbb{E}_{i-1}$  averages over a smaller set of possible  $(A, \hat{A})$ . However, since the number of possible  $(A, \hat{A})$  is always  $\lesssim_k 1$ , we can afford to replace the average by a sum: thus, for all  $1 \leq i \leq m^\dagger$ , we have

$$\text{Var}_i L^\dagger = \mathbb{E}[(\mathbb{E}_{i-1} \Delta)^2] \leq \sum_{A, \hat{A}} \mathbb{E} \left[ \left( L(G^\circ \cup A, \epsilon) - L(G^\circ \cup \hat{A}, \epsilon) \right)^2 \right],$$

where the sum is over all arrangements of clauses  $(A, \hat{A})$  on a set of  $k^2$  half-edges. Write

$$X(A) \equiv \mathbb{E}[\star Z(G) | G^\dagger = G \setminus W = G^\circ \cup A].$$

Then summing over  $1 \leq i \leq m^\dagger$  we have

$$\text{Var} L^\dagger \leq m^\dagger \sum_{A, \hat{A}} \mathbb{E}[V(A, \hat{A})^2] \quad \text{where } V(A, \hat{A}) \equiv \log \frac{X(A) + \epsilon \mathbb{E}Z}{X(\hat{A}) + \epsilon \mathbb{E}Z}.$$

From now on we fix  $(A, \hat{A})$  and abbreviate  $X \equiv X(A)$ ,  $\hat{X} \equiv X(\hat{A})$ ,  $V \equiv V(A, \hat{A})$ . The remainder of this section is devoted to proving

$$\limsup_{n \rightarrow \infty} n \mathbb{E}[V^2] \lesssim_k 1, \tag{45}$$

which implies Propn. 6.3.

**6.2. Fourier decomposition.** Consider the graph  $G^\circ$  with clauses  $(a_1, \dots, a_{m^\dagger - k})$ . Its unmatched variable-incident half-edges are partitioned into  $\mathcal{X}$  (the  $k^2$  half-edges that will participate in the clauses  $A$  or  $\hat{A}$ ) and  $\mathcal{W}$  (the remaining half-edges, which will participate in the clauses  $W$ ).

We shall define a certain local neighborhood  $T$  of the half-edges in  $\mathcal{X}$ , such that  $G^\partial \equiv G^\circ \setminus T$  is a graph with unmatched half-edges in disjoint sets  $\mathcal{W}$  (as before) and  $\mathcal{Z}$  (leaves of  $T$  without  $\mathcal{W}$ ); see Fig. 3.<sup>8</sup> We will then define a certain  $T$ -measurable event  $\mathbf{B}$  with probability on the order of  $1/n$  (roughly speaking, the event that  $T$  contains a cycle). On the event  $\mathbf{B}^c$ , which occurs with probability close to one, we bound

$$V^2 \leq \frac{(X - \hat{X})^2}{(\epsilon \mathbb{E}Z)^2}. \tag{46}$$

On the event  $\mathbf{B}^c$  we bound  $V^2$  differently; this will be explained in (47) below.

<sup>8</sup>Each unmatched half-edge is incident to a variable, and does not include a clause.

Let  $\Psi_W(\underline{\sigma}_{\mathcal{W}})$  be the indicator that the clauses  $W$  are satisfied by the configuration  $\underline{\sigma}_{\mathcal{W}}$ . Write  $\mathcal{G} \equiv \mathcal{U} \cup \mathcal{W}$  and let  ${}^*Z^\partial[\underline{\sigma}_{\mathcal{G}}]$  be the partition function on  $G^\partial$  given boundary conditions  $\underline{\sigma}_{\mathcal{G}}$ . Let  $\kappa(\underline{\sigma}_{\mathcal{U}}|\underline{\sigma}_{\mathcal{W}})$  be the partition function on  $T \cup A$  given boundary conditions  $\underline{\sigma}_{\mathcal{G}}$ . Then

$${}^*Z = \sum_{\underline{\sigma}_{\mathcal{W}}} \Psi_W(\underline{\sigma}_{\mathcal{W}}) \sum_{\underline{\sigma}_{\mathcal{U}}} \kappa(\underline{\sigma}_{\mathcal{U}}|\underline{\sigma}_{\mathcal{W}}) {}^*Z^\partial[\underline{\sigma}_{\mathcal{G}}].$$

The same expansion holds for  ${}^*Z$  except that  $\kappa$  is replaced by  $\hat{\kappa}$ , referring to the partition function on  $T \cup A$ . Then

$$X = \mathbb{E}[{}^*Z | G \setminus W] = \sum_{\underline{\sigma}_{\mathcal{W}}} \bar{\varphi}(\underline{\sigma}_{\mathcal{W}}) \sum_{\underline{\sigma}_{\mathcal{U}}} \kappa(\underline{\sigma}_{\mathcal{U}}|\underline{\sigma}_{\mathcal{W}}) {}^*Z^\partial[\underline{\sigma}_{\mathcal{G}}]$$

where  $\bar{\varphi}$  denotes the average of  $\Psi_W$  over the possibilities of  $W$ .

Write  $h \equiv (\hat{h}, \hat{h})$  for the Bethe fixed point corresponding to the first moment maximizer  ${}^*h$ . We will see (Lem. 6.5 below) that in the graph  $G^\partial$  the distribution of the boundary spins  $\underline{\sigma}_{\mathcal{U}}$  is very close to the product measure  $p(\underline{\sigma}_{\mathcal{U}})$  where the  $\sigma_u$  ( $u \in \mathcal{U}$ ) are independent and identically distributed according to  $\hat{h}$ . Take  $(b_1, \dots, b_{|\mathcal{M}|})$  to be an orthonormal basis for  $L^2(\mathcal{M}, \hat{h})$  with  $b_1 \equiv 1$ . Then the product functions

$$b_{\underline{s}}(\underline{\sigma}_{\mathcal{U}}) \equiv \prod_{u \in \mathcal{U}} b_{s(u)}(\sigma_u) \quad (\underline{s} \in [|\mathcal{M}|]^\mathcal{U})$$

form an orthonormal basis for the space  $L^2(\mathcal{M}^\mathcal{U}, p)$ , and for any  $f$  in this space we denote its Fourier coefficients

$$f^\wedge(\underline{s}) \equiv \langle f, b_{\underline{s}} \rangle_p = \sum_{\underline{\sigma}_{\mathcal{U}}} p(\underline{\sigma}_{\mathcal{U}}) f(\underline{\sigma}_{\mathcal{U}}) b_{\underline{s}}(\underline{\sigma}_{\mathcal{U}}).$$

Recall that  $\underline{\tau}$  denotes a pair  $(\sigma^1, \sigma^2)$ . By an abuse of notation we also write  $p$  for the product measure on  $\underline{\tau}$ ,  $p(\underline{\tau}) \equiv p(\sigma^1)p(\sigma^2)$ . The functions  $b_{\underline{s}^1, \underline{s}^2}(\underline{\tau}_{\mathcal{U}}) \equiv b_{\underline{s}^1}(\sigma_{\mathcal{U}}^1) b_{\underline{s}^2}(\sigma_{\mathcal{U}}^2)$  form an orthonormal basis for  $L^2(\mathcal{M}^{2\mathcal{U}}, p)$  and define a Fourier transform on that space. By Plancherel's identity,

$$X = \sum_{\underline{\sigma}_{\mathcal{W}}} \bar{\varphi}(\underline{\sigma}_{\mathcal{W}}) \sum_{\underline{s}} \kappa^\wedge(\underline{s}|\underline{\sigma}_{\mathcal{W}}) F^\wedge[\underline{s}|\underline{\sigma}_{\mathcal{W}}].$$

Let  $X_\emptyset$  be the contribution in the above expansion from  $\underline{s} = \emptyset$  (meaning that  $\underline{s}$  is identically one). On the event  $\mathbf{B}$  we bound

$$\begin{aligned} V^2 &= \left( \log \frac{X + \epsilon \mathbb{E}Z}{X_\emptyset + \epsilon \mathbb{E}Z} + \log \frac{X_\emptyset + \epsilon \mathbb{E}Z}{\hat{X}_\emptyset + \epsilon \mathbb{E}Z} + \log \frac{\hat{X}_\emptyset + \epsilon \mathbb{E}Z}{\hat{X} + \epsilon \mathbb{E}Z} \right)^2 \\ &\leq 3 \frac{(X - X_\emptyset)^2 + (\hat{X} - \hat{X}_\emptyset)^2}{(\epsilon \mathbb{E}Z)^2} + 3 \frac{(X_\emptyset - \hat{X}_\emptyset)^2}{\min\{X_\emptyset, \hat{X}_\emptyset\}^2}. \end{aligned} \quad (47)$$

We next expand the squared terms in (46) and (47) and separate out the ‘‘independent-copies’’ and ‘‘identical-copies’’ contributions. Abbreviate  $D \equiv X - \hat{X}$ . Then (46) expands as

$$\begin{aligned} D^2 &= \sum_{\underline{\tau}_{\mathcal{W}}} \underbrace{\bar{\varphi}(\sigma_{\mathcal{W}}^1) \bar{\varphi}(\sigma_{\mathcal{W}}^2)}_{\bar{\varphi}^\otimes(\underline{\tau}_{\mathcal{W}})} \sum_{\underline{\tau}_{\mathcal{U}}} \underbrace{(\kappa - \hat{\kappa})(\sigma_{\mathcal{U}}^1|\sigma_{\mathcal{W}}^1)(\kappa - \hat{\kappa})(\sigma_{\mathcal{U}}^2|\sigma_{\mathcal{W}}^2)}_{\varpi^\otimes(\underline{\tau}_{\mathcal{U}}|\underline{\tau}_{\mathcal{W}})} {}_2Z^\partial[\underline{\tau}_{\mathcal{G}}]. \\ &= \sum_{\underline{\tau}_{\mathcal{W}}} \bar{\varphi}^\otimes(\underline{\tau}_{\mathcal{W}}) \sum_{\underline{\tau}_{\mathcal{U}}} \varpi^\otimes(\underline{\tau}_{\mathcal{U}}|\underline{\tau}_{\mathcal{W}}) {}_2Z^\partial[\underline{\tau}_{\mathcal{G}}] \end{aligned} \quad (48)$$

where  ${}_2Z^\partial$  denotes the *pair* partition function on  $G^\partial$ . In the manner of (44) we decompose  ${}_2Z^\partial$  into a near-independent contribution  ${}_2Z^\partial$ , a near-identical contribution  $\ddot{{}_2Z^\partial}$ , and a remainder term which has expectation  $o(n^{-1}) \mathbb{E}[Z^2]$ . Substituting into the above expansion gives the corresponding decomposition of  $D^2$ :

$$D^2 = \left( \text{near-independent contribution } \ddot{{}_2D} \right) + \left( \text{near-identical contribution } \ddot{{}_2D} \right) + (\text{remainder}). \quad (49)$$

An analogous expansion holds for each squared term in (47). In each case the remainder term has expectation  $o(n^{-1}) \mathbb{E}[Z^2]$  uniformly over  $d_{\text{ibd}} \leq d \leq d_{\text{ubd}}$ , and so can be ignored. The near-identical contribution has expectation  $o(n^{-1}) \mathbb{E}[Z^2]$  uniformly over the integers  $d_{\text{ibd}} \leq d < d_{\text{ubd}}$ , and so can be ignored for the main result of this paper. In related work [DSS13] we will show how to control these terms closer to the threshold, yielding a sharper bound on the threshold fluctuations.

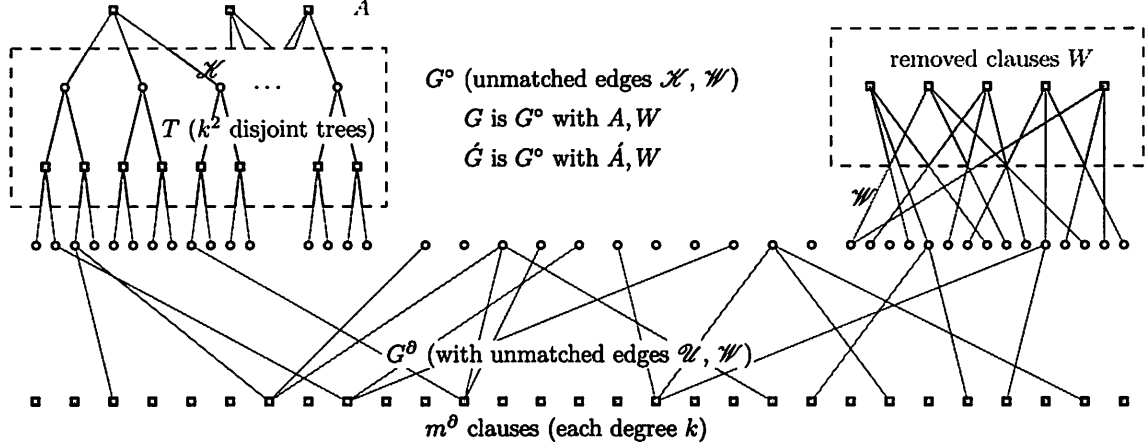


FIGURE 3.  $G^o \equiv$  graph with clauses  $a_1, \dots, a_{i-1}$ ;  $G^\delta \equiv G^o \setminus T$  (gray)

Let  $B_\ell^o(\mathcal{X})$  denote the ball of graph distance  $\ell$  about  $\mathcal{X}$  in the graph  $G^o$ ; the leaves of  $B_\ell^o(\mathcal{X})$  are half-edges incident to variables ( $\ell$  odd) or clauses ( $\ell$  even). We shall fix a maximum depth  $2t$  (where  $t$  will eventually be a large constant depending on  $\epsilon$ ) and set  $T = B_{2t}^o(\mathcal{X})$  where

$$t' \equiv \min \left\{ \{t\} \cup \left\{ B_{2\ell}^o \text{ does not consist of } |\mathcal{X}| \text{ disjoint trees} \right\} \right\}. \quad (50)$$

With this definition,  $\mathcal{W}$  can only intersect  $T$  in its leaves, which are variable-incident half-edges at distance  $2t'$  from  $\mathcal{X}$ . Recall  $\mathcal{Q}$  denotes the leaves of  $T$  without  $\mathcal{W}$ . Let  $\mathbb{E}_T$  denote expectation conditioned on  $T$ . Applying this to (48) and (49), the near-independent contribution to  $\mathbb{E}_T[D^2]$  can be expressed as

$$\begin{aligned} \mathbb{E}_T[D^2] &= \sum_{\mathcal{I}_{\mathcal{W}}} \bar{\varphi}^{\otimes}(\mathcal{I}_{\mathcal{W}}) \sum_{\mathcal{I}_{\mathcal{Q}}} \varpi^{\otimes}(\mathcal{I}_{\mathcal{Q}} | \mathcal{I}_{\mathcal{W}}) \frac{\mathbb{F}(\mathcal{I}_{\mathcal{Q}} | \mathcal{I}_{\mathcal{W}})}{p(\mathcal{I}_{\mathcal{Q}})} p(\mathcal{I}_{\mathcal{Q}}), \\ &= \sum_{\mathcal{I}_{\mathcal{W}}} \bar{\varphi}^{\otimes}(\mathcal{I}_{\mathcal{W}}) \sum_{\underline{s}^1, \underline{s}^2} \varpi^\wedge(\underline{s}^1 | \underline{\sigma}_{\mathcal{W}}^1) \varpi^\wedge(\underline{s}^2 | \underline{\sigma}_{\mathcal{W}}^2) \mathbb{F}^\wedge(\underline{s}^1, \underline{s}^2 | \mathcal{I}_{\mathcal{W}}). \end{aligned} \quad (51)$$

We emphasize that  $\varpi = \kappa - \hat{\kappa}$  and  $\mathbb{F}$  depend on  $T$  although for convenience we have suppressed it from the notation.

**6.3. Expansion of partition function.** We now estimate the Fourier coefficients of the function  $\mathbb{F}$  appearing in (51): recalling that  $\mathcal{I}_{\mathcal{G}^\delta} \equiv (\mathcal{I}_{\mathcal{Q}}, \mathcal{I}_{\mathcal{W}})$ ,

$$\mathbb{F}^\wedge(\underline{s}^1, \underline{s}^2 | \mathcal{I}_{\mathcal{W}}) = \sum_{\mathcal{I}_{\mathcal{Q}}} \mathbb{E}_T[Z^\delta[\mathcal{I}_{\mathcal{G}^\delta}]] b_{\underline{s}^1, \underline{s}^2}(\mathcal{I}_{\mathcal{Q}}).$$

We begin by estimating  $\mathbb{E}_T[Z^\delta[\mathcal{I}_{\mathcal{G}^\delta}]]$ . Note this depends on  $T$  only through the numbers  $n^\delta, m^\delta$  of clauses in  $G^\delta$ , which are determined by  $T$ . Denote  $\nu \equiv n - n^\delta$  (the number of variables in  $T$ ),  $\mu \equiv m - m^\delta$  (the number of clauses in  $T \cup W$ ) and note  $\nu + \mu \lesssim_{k, \epsilon} \log n$ .



It is most convenient here to work with the *non-normalized* empirical measures  $g \equiv (\dot{g}, \hat{g})$ . The associated non-normalized marginal edge counts are given by

$$\begin{aligned} \dot{H}\dot{g} & \text{ empirical marginal of variable-incident half-edges,} \\ \hat{H}\hat{g} & \text{ empirical marginal of clause-incident half-edges,} \end{aligned}$$

where  $\dot{H}, \hat{H}$  are the marginalization matrices corresponding to  $\dot{\varphi}, \hat{\varphi}$  as defined in §3.3. The empirical measure  $g$  can contribute to  $\mathbb{E}_T[Z^\partial[\underline{\sigma}_{\mathcal{G}}]]$  only if

$$\langle \dot{g}, 1 \rangle = n^\partial, \langle \hat{g}, 1 \rangle = m^\partial, \text{ and } \dot{H}\dot{g} - \hat{H}\hat{g} = \bar{H}^{\underline{\sigma}_{\mathcal{G}}}$$

where  $\langle \dot{g}, 1 \rangle$  indicates the total mass of  $\dot{g}$ , and  $\bar{H}^{\underline{\sigma}_{\mathcal{G}}}$  denotes the non-normalized edge empirical measure associated to  $\underline{\sigma}_{\mathcal{G}}$ . The contribution to  $\mathbb{E}_T[Z^\partial[\underline{\sigma}_{\mathcal{G}}]]$  from such  $g$  is given by

$$\Xi(g) \equiv \dot{\varphi}^{\dot{g}} \hat{\varphi}^{\hat{g}} \binom{\langle \dot{g}, 1 \rangle}{\dot{g}} \binom{\langle \hat{g}, 1 \rangle}{\hat{g}} / \binom{\langle \dot{H}\dot{g}, 1 \rangle}{\dot{H}\dot{g}}$$

(with multi-index notation). In the leading exponential term  $\Xi(g)$  behaves exactly as  $\mathbb{E}Z(h)$ ,

$$\Xi(g) = \exp\{n^\partial \Phi(g/n^\partial) + O_k(\mu + \nu)\}, \quad (52)$$

so we see immediately that the dominant contribution to  $\mathbb{E}[Z^\partial[\underline{\sigma}_{\mathcal{G}}]]$  comes from near  ${}^*g$ . The following lemma gives a more precise analysis:

**Lemma 6.5.** *Fix  $\underline{\sigma}_{\mathcal{G}}$  and assume there exists a signed measure  $\delta \equiv (\dot{\delta}, \hat{\delta})$  such that*

$$\langle \dot{\delta}, 1 \rangle = \nu, \langle \hat{\delta}, 1 \rangle = \mu, \text{ and } \hat{H}\hat{\delta} - \dot{H}\dot{\delta} = \bar{H}^{\underline{\sigma}_{\mathcal{G}}}$$

(where  $\delta$  depends on  $\underline{\sigma}_{\mathcal{G}}$ , and is non-zero only on the support of  $\varphi$ ). Then

$$\frac{\mathbb{E}_T[Z^\partial[\underline{\sigma}_{\mathcal{G}}]]}{p(\underline{\sigma}_{\mathcal{G}}) \mathbb{E}_T[Z^\partial]} = 1 + O_k\left(\frac{\|\delta\|_1 \log n}{n^{1/2}}\right) \quad (53)$$

*Proof.* Let  $g$  run over empirical measures contributing to the expected partition function  $\mathbb{E}Z$  on a  $(d, k)$ -regular graph with  $n$  variables and  $m$  clauses. Away from the simplex boundary, the empirical measures contributing to  $\mathbb{E}_T[Z^\partial[\underline{\sigma}_{\mathcal{G}}]]$  are parametrized by  $g - \delta$ . Then, writing  $(a)_b$  for the falling factorial  $a!/(a-b)!$ , some algebra gives

$$\frac{\Xi(g - \delta)}{\Xi(g)} = \frac{\overbrace{\bar{z}^{\nu d}}^{1/c}}{\dot{z}^\nu \hat{z}^\mu} \underbrace{\left[ \frac{n^\nu}{(n)_\nu} \frac{m^\mu}{(m)_\mu} \frac{(nd)_{\nu d}}{(nd)^{\nu d}} \right]}_s \underbrace{\left[ \frac{(\dot{g})_{\dot{\delta}}}{({}^*\dot{g})_{\dot{\delta}}} \frac{(\hat{g})_{\hat{\delta}}}{({}^*\hat{g})_{\hat{\delta}}} \frac{(\dot{H}\dot{g})_{\dot{H}\dot{\delta}}}{(\dot{H}\dot{g})_{\dot{H}\dot{\delta}}} \right]}_{e(g, \delta)} \underbrace{\left[ \frac{\overbrace{\dot{z}^\nu \hat{z}^\mu}^c}{\bar{z}^{\nu d}} \frac{({}^*\dot{h}/\dot{\varphi})_{\dot{\delta}} ({}^*\hat{h}/\hat{\varphi})_{\hat{\delta}}}{({}^*\bar{h})_{\hat{H}\hat{\delta}}} \right]}_{r(\delta)}.$$

(The above uses multi-index notation, so for example  $(\dot{g})_{\dot{\delta}}$  indicates the product of  $(\dot{g}(\underline{\alpha}))_{\dot{\delta}(\underline{\alpha})}$  over  $\underline{\alpha} \in \text{supp } \dot{\varphi}$ .) Using the relation (34) between  ${}^*h$  and the Bethe fixed point  $h \equiv (\dot{h}, \hat{h})$  we have

$$r(\delta) = p(\underline{\sigma}_{\mathcal{G}}), \quad c = [\dot{z}\hat{z}^{d/k}/\bar{z}^d]^\nu \cdot \hat{z}^{|\mathcal{G}|/k} = \exp\{\nu {}^*\Phi\} \cdot \hat{z}^{|\mathcal{G}|/k}. \quad (54)$$

The factor  $s$  does not depend on  $(g, \delta)$ , and we note

$$|s - 1| \lesssim_k \frac{(\nu + \mu)^2}{n} \leq \frac{\|\delta\|_1^2}{n}, \quad |e(g, \delta) - 1| \lesssim_k \frac{\|\delta\|_1 (\|g - {}^*g\|_1 + \|\delta\|_1)}{n}.$$

Let  $B$  denote the subset of measures  $g$  within euclidean distance  $n^{1/2} \log n$  of  ${}^*g$ : since we established that  $\Phi$  has negative-definite Hessian around  ${}^*g$ ,

$$\frac{\mathbb{E}_T[Z^\partial[\underline{\sigma}_{\mathcal{G}}]]}{e^{o(1/n^2)} \mathbb{E}Z} = \frac{\sum_{g \in B} \Xi(g - \delta)}{\sum_{g \in B} \Xi(g)}.$$

Then, by the preceding calculations, the right-hand side above equals

$$\frac{s \cdot r(\delta) \sum_{g \in B} \Xi(g) e(g, \delta)}{c \sum_{g \in B} \Xi(g)} = \frac{r(\delta)}{c} \left( 1 + O\left(\frac{\|\delta\|_1 \log n}{n^{1/2}}\right) \right)$$

from which (53) easily follows.  $\square$

Lem. 6.5 applies for any factor model with free energy attaining a local maximum at  $^*g$  with negative-definite Hessian. In particular, it applies to both the first- and second-moment versions of the auxiliary model. We now show how to construct the required measures  $\delta$  for the pair auxiliary model (the construction for the first-moment version being similar but simpler). The construction is based on the following

**Lemma 6.6.** *For any  $\tau, \tau' \in \mathcal{M}^2$ ,*

(a) *There exists a signed measure  $\dot{\delta} \equiv \dot{\delta}_{\tau-\tau'}$  such that*

$$\text{supp } \dot{\delta} \subseteq \text{supp } \dot{\varphi}, \quad \langle \dot{\delta}, 1 \rangle = 0, \quad \text{and} \quad \dot{H} \dot{\delta} = \mathbf{1}_\tau - \mathbf{1}_{\tau'}.$$

(b) *There exists a signed, integer-valued measure  $\hat{\delta} \equiv \hat{\delta}_{\tau-\tau'}$  such that*

$$\text{supp } \hat{\delta} \subseteq \text{supp } \dot{\varphi}, \quad \langle \hat{\delta}, 1 \rangle = 0, \quad \text{and} \quad \hat{H} \hat{\delta} = \mathbf{1}_\tau - \mathbf{1}_{\tau'}.$$

*Proof.* (a) Let  $\dot{H}_1$  be the first-moment version of  $\dot{H}$ , so that  $\dot{H} = \dot{H}_1 \otimes \dot{H}_1$ . Define a graph on  $\mathcal{M}$  by putting an edge between  $\sigma$  and  $\sigma'$  if there exists a signed (but not necessarily integer-valued) measure  $\dot{\delta} = \dot{\delta}_{\sigma-\sigma'}$  such that

$$\text{supp } \dot{\delta} \subseteq \text{supp } \dot{\varphi}, \quad \langle \dot{\delta}, 1 \rangle = 0, \quad \text{and} \quad \dot{H}_1 \dot{\delta} = \mathbf{1}_\sigma - \mathbf{1}_{\sigma'}.$$

We will show that the graph is connected, hence complete. Since  $\underline{\sigma} = (\mathbf{f}\mathbf{f}^d)$  and  $\underline{\sigma}' = (\mathbf{0}\mathbf{0}^d)$  are both valid variable configurations, we see that  $\mathbf{f}\mathbf{f}$  is connected to  $\mathbf{0}\mathbf{0}$ , and likewise to  $\mathbf{1}\mathbf{1}$ . Next, since

$$\underline{\sigma} = (\mathbf{0}\mathbf{0}^2, \mathbf{0}\mathbf{f}^{d-2}) \quad \text{and} \quad \underline{\sigma}' = (\mathbf{0}\mathbf{0}^3, \mathbf{0}\mathbf{f}^{d-3})$$

are both valid,  $\mathbf{0}\mathbf{0}$  is connected to  $\mathbf{0}\mathbf{f}$ , and likewise  $\mathbf{1}\mathbf{1}$  to  $\mathbf{1}\mathbf{f}$ . Finally note that if

$$\dot{\delta} = \mathbf{1}\{(\mathbf{f}\mathbf{0}, \mathbf{0}\mathbf{f}^{d-1})\} - 2\mathbf{1}\{(\mathbf{0}\mathbf{0}^2, \mathbf{0}\mathbf{f}^{d-2})\} + \mathbf{1}\{(\mathbf{0}\mathbf{0}^3, \mathbf{0}\mathbf{f}^{d-3})\}$$

then  $\dot{H}_1 \dot{\delta} = \mathbf{1}_{\mathbf{f}\mathbf{0}} - \mathbf{1}_{\mathbf{0}\mathbf{0}}$ , so  $\mathbf{0}\mathbf{0}$  is connected to  $\mathbf{f}\mathbf{0}$ , and likewise  $\mathbf{1}\mathbf{1}$  to  $\mathbf{f}\mathbf{1}$ , which proves that the graph on  $\mathcal{M}$  is connected. The claim for  $\dot{H} = \dot{H}_1 \otimes \dot{H}_1$  follows straightforwardly since we can go from  $\tau$  to  $\tau'$  first by changing  $\sigma^1$  to  $(\sigma^1)'$ , then changing  $\sigma^2$  to  $(\sigma^2)'$ , using the above manipulations.

(b) Define a graph on  $\mathcal{M}^2$  by putting an edge between  $\tau$  and  $\tau'$  if the required  $\hat{\delta}_{\tau-\tau'}$  exists, where we emphasize that here  $\hat{\delta}_{\tau-\tau'}$  is required to be integer-valued. We will show that the graph is connected, hence complete. In this proof we will write  $\tau \equiv \mathbf{i}\mathbf{o} \equiv (\mathbf{i}, \mathbf{o})$  where  $\mathbf{i}$  is the (pair) variable-to-clause message, and  $\mathbf{o}$  is the (pair) clause-to-variable message.

Suppose  $\tau = \mathbf{i}\mathbf{o}$  and  $\tau' = \mathbf{i}'\mathbf{o} = (\mathbf{i}', \mathbf{f}\mathbf{f})$ . Let  $\tilde{\tau} = (\mathbf{f}\mathbf{f}, \mathbf{f}\mathbf{f})$ . Then both  $(\tilde{\tau}^{k-1}, \tau)$  and  $(\tilde{\tau}^{k-1}, \tau')$  are valid clause configurations, so  $\tau, \tau'$  are connected.

Next suppose  $\tau = \mathbf{i}\mathbf{o} = (\mathbf{i}, \mathbf{0}\mathbf{f})$  and  $\tau'' = \mathbf{i}\mathbf{o}'' = (\mathbf{i}, \mathbf{f}\mathbf{f})$ . Let  $\tilde{\tau} = (\mathbf{1}\mathbf{f}, \mathbf{f}\mathbf{f})$ . Then  $(\tilde{\tau}^{k-1}, \tau)$  is a valid clause configuration if we take the clause to have all the same literals. On the other hand,  $(\tilde{\tau}^{k-1}, \tau'')$  is a valid clause configuration if we take the clause to have all the same literals except in the first coordinate. It follows that  $\tau, \tau''$  are connected.

By the same reasoning, any  $\tau = \mathbf{i}\mathbf{o}$  is connected to  $\tau'' = (\mathbf{i}, \mathbf{f}\mathbf{f})$ . Then  $\tau''$  is in turn connected to  $\tau''' = (\mathbf{i}', \mathbf{f}\mathbf{f})$ , which is connected to  $\tau' = (\mathbf{i}', \mathbf{o}')$ . This proves our claim that that the graph on  $\mathcal{M}^2$  is complete.  $\square$

Fix a reference spin  $\bar{\tau}$ , say  $\bar{\tau} = (\bar{\sigma}^1, \bar{\sigma}^2) = (\mathbf{f}\mathbf{f}, \mathbf{f}\mathbf{f})$ . Lem. 6.6b implies we can find signed integer measures  $\delta^{\text{var}}$  and  $\delta^{\text{cl}}$  such that

$$\begin{aligned} \langle \dot{\delta}^{\text{var}}, 1 \rangle &= 1, & \langle \hat{\delta}^{\text{var}}, 1 \rangle &= 0, & \dot{H} \dot{\delta}^{\text{var}} - \hat{H} \hat{\delta}^{\text{var}} &= d\mathbf{1}_{\bar{\tau}} \\ \langle \dot{\delta}^{\text{cl}}, 1 \rangle &= 0, & \langle \hat{\delta}^{\text{cl}}, 1 \rangle &= 1, & \dot{H} \dot{\delta}^{\text{cl}} - \hat{H} \hat{\delta}^{\text{cl}} &= -k\mathbf{1}_{\bar{\tau}} \end{aligned} \tag{55}$$

(Note in the pair auxiliary model for NAE-SAT, we may trivially take  $\hat{\delta}^{\text{var}}(\hat{r})$  to be the indicator of  $\hat{r} = (\hat{r}^1, \hat{r}^2) = (\mathbf{ff}^d, \mathbf{ff}^d)$  with  $\hat{\delta}^{\text{var}}$  identically zero; and likewise take  $\hat{\delta}^{\text{cl}}(\hat{r})$  to be the indicator of  $\hat{r} = (\hat{r}^1, \hat{r}^2) = (\mathbf{ff}^k, \mathbf{ff}^k)$  with  $\hat{\delta}^{\text{cl}}$  identically zero. In a more general setting, the statement of Lem. 6.6b suffices to construct measures satisfying (55).) One can then check that

$$\delta = \nu \delta^{\text{var}} + \mu \delta^{\text{cl}} - \sum_{y \in \mathcal{Y}} \delta_{\tau_y - \bar{\tau}} \quad (56)$$

satisfies the conditions of Lem. 6.5 (using that  $-\nu d + \mu k = n^\partial d - m^\partial k = |\mathcal{Y}|$ ). It follows that in the pair auxiliary model,

$$\mathbb{E}_T[\mathbb{Z}^\partial[\mathcal{I}_{\mathcal{Y}}]] = \left(1 + O_k\left(\frac{(|T| + |\mathcal{Y}|) \log n}{n^{1/2}}\right)\right) \frac{p(\mathcal{I}_{\mathcal{Y}})}{c^2} \mathbb{E}[\mathbb{Z}^\partial],$$

and since  $\mathbb{E}[\mathbb{Z}^\partial] \asymp_k (\mathbb{E}Z)^\partial$  we see that  $\mathbb{E}_T[\mathbb{Z}^\partial[\mathcal{I}_{\mathcal{Y}}]] \asymp_k \mathbb{E}_T[\mathbb{Z}^\partial[\sigma_{\mathcal{Y}}^1]] \mathbb{E}_T[\mathbb{Z}^\partial[\sigma_{\mathcal{Y}}^2]]$ .

Now write  $Z^\partial[\cdot, \sigma_{\mathcal{Y}}]$  for the sum of  $Z^\partial[\sigma_{\mathcal{Y}}, \sigma_{\mathcal{Y}}]$  over all  $\sigma_{\mathcal{Y}}$  with  $\sigma_{\mathcal{Y}}$  fixed. We can find  $\delta$  with  $\|\delta\|_1 \lesssim_k |T|$  such that  $\hat{H}\delta - \hat{H}\hat{\delta} = \bar{H}^{\sigma_{\mathcal{Y}}}$ . If  $g$  contributes to  $Z^\partial[\cdot, \sigma_{\mathcal{Y}}]$ , then  $g - \delta$  contributes to  $Z^\partial[\sigma_{\mathcal{Y}}, \sigma_{\mathcal{Y}}]$ , so we conclude

$$\mathbb{E}_T[Z^\partial[\sigma_{\mathcal{Y}}, \sigma_{\mathcal{Y}}]] = \left(1 + O_k\left(\frac{|T| \log n}{n^{1/2}}\right)\right) p(\sigma_{\mathcal{Y}}) \mathbb{E}_T[Z^\partial[\cdot, \sigma_{\mathcal{Y}}]]. \quad (57)$$

The analogous expansion holds for the second-moment version  $\mathbb{E}_T[\mathbb{Z}^\partial[\mathcal{I}_{\mathcal{Y}}, \mathcal{I}_{\mathcal{Y}}]]$ , and so

$$\frac{\mathbb{F}^\wedge(\underline{s}|\mathcal{I}_{\mathcal{Y}})}{\mathbb{F}^\wedge(\emptyset|\mathcal{I}_{\mathcal{Y}})} = \frac{\sum_{\mathcal{I}_{\mathcal{Y}}} \mathbb{E}_T[\mathbb{Z}^\partial[\mathcal{I}_{\mathcal{Y}}, \mathcal{I}_{\mathcal{Y}}]] \mathbf{b}_{\underline{s}}(\mathcal{I}_{\mathcal{Y}})}{\mathbb{E}_T[\mathbb{Z}^\partial[\cdot, \mathcal{I}_{\mathcal{Y}}]]} = \langle 1, \mathbf{b}_{\underline{s}} \rangle_p + O_k\left(\frac{|T| \log n}{n^{1/2}}\right). \quad (58)$$

Recall that  $\langle 1, \mathbf{b}_{\underline{s}} \rangle_p$  is simply the indicator that  $\underline{s} = \emptyset$ , so we see that the higher-order Fourier coefficients of  $\mathbb{F}$  are small. The following lemma yields a more precise analysis of these coefficients under the assumption that  $k$  divides  $|\mathcal{Y}|$ , which includes the typical case where  $T$  consists of  $k$  disjoint trees not involving  $\mathcal{Y}$ .

**Lemma 6.7.** *If  $k$  divides  $|\mathcal{Y}|$ , then for a constant  $C_k$  there exist coefficients  $\xi_j \equiv (\xi_j(\sigma))_\sigma$ , indexed by  $0 \leq j \leq C_k$ , such that  $\|\xi_j\|_\infty \lesssim_k n^{-1/2}$  and*

$$\frac{\mathbb{E}_T[Z^\partial[\sigma_{\mathcal{Y}}]] \cdot c}{p(\sigma_{\mathcal{Y}}) \mathbb{E}_T[Z^\partial[\cdot, \sigma_{\mathcal{Y}}]]} = 1 + \langle \bar{H}^{\sigma_{\mathcal{Y}}}, \xi_0 \rangle + \sum_{j=0}^{C_k} \langle \bar{H}^{\sigma_{\mathcal{Y}}}, \xi_j \rangle^2 + O_k\left(\frac{(\log n)^6}{n^{3/2}}\right) \quad (59)$$

where  $c = 1 + o_n(1)$  and does not depend on  $\sigma_{\mathcal{Y}}$ .

*Proof.* In this case a particularly simple choice of  $\delta$  satisfies  $\hat{H}\delta - \hat{H}\hat{\delta} = \bar{H}^{\sigma_{\mathcal{Y}}}$ : take

$$\delta \equiv \delta^{\sigma_{\mathcal{Y}}} \equiv (|\mathcal{Y}|/k) \delta^{\text{cl}} - \sum_{u \in \mathcal{Y}} \delta_{\sigma_u - \bar{\sigma}},$$

which is a signed integer measure by the assumption that  $k$  divides  $|\mathcal{Y}|$ . Notice that  $\delta^{\sigma_{\mathcal{Y}}}$  can be expressed as a linear function of  $\bar{H}^{\sigma_{\mathcal{Y}}}$ :

$$\delta^{\sigma_{\mathcal{Y}}} = \sum_{\sigma} \bar{H}^{\sigma_{\mathcal{Y}}}(\sigma) \delta^\sigma, \quad \text{where } \delta^\sigma \equiv k^{-1} \delta^{\text{cl}} - \delta^{\sigma - \bar{\sigma}}. \quad (60)$$

Let  $G'$  be a graph with  $n^\partial = n - \nu$  variables and  $w(n)$  clauses missing, so that there is a set  $\mathcal{Y}$  of unmatched variable-incident half-edges of size  $w(n)k$ . Let  $Z'[\cdot, \sigma_{\mathcal{Y}}]$  denote the partition function on  $G'$  subject to  $\sigma_{\mathcal{Y}}$ . Write  $B'$  for the set of non-normalized empirical measures  $g$  that contribute to  $Z'[\cdot, \sigma_{\mathcal{Y}}]$  and lie within distance  $n^{1/2} \log n$  of  $\ast g$ . Then, with  $\delta = \delta^{\sigma_{\mathcal{Y}}}$ , a similar calculation as in Lem. 6.5 gives

$$\frac{\mathbb{E}_T[Z^\partial[\sigma_{\mathcal{Y}}, \sigma_{\mathcal{Y}}]]}{e^{o(1/n^2)} \mathbb{E}[Z'[\cdot, \sigma_{\mathcal{Y}}]]} = \frac{\sum_{g \in B'} \Xi(g - \delta)}{\sum_{g' \in B'} \Xi(g')} = \frac{s \cdot p(\sigma_{\mathcal{Y}}) \sum_{g \in B'} \Xi(g) e(g, \delta)}{c \sum_{g' \in B'} \Xi(g')}. \quad (61)$$

We now make a more precise expansion of the term  $e(g, \delta)$ : it is a product of ratios  $(a)_b / (*a)^b$  where  $*a \asymp_k n$ ,  $|a - *a| \leq n^{1/2} \log n$ , and  $b \lesssim_k |\mathcal{U}|$ . For each ratio we estimate

$$\begin{aligned} (a)_b / a^b &= \prod_{i=0}^{b-1} (1 - i/a) = 1 - \frac{b(b-1)}{2(*a)} + O_k\left(\frac{b^2 \log n}{n^{3/2}}\right), \\ a^b / (*a)^b &= 1 + \frac{b(a - *a)}{*a} + \frac{b^2(a - *a)^2}{2(*a)^2} + O_k\left(\frac{b^3(\log n)^3}{n^{3/2}}\right), \\ \text{thus } (a)_b / (*a)^b &= 1 + b\left[\frac{a - *a + 1/2}{*a}\right] + \frac{b^2}{2}\left[\left(\frac{a - *a}{*a}\right)^2 - \frac{1}{*a}\right] + O_k\left(\frac{b^3(\log n)^3}{n^{3/2}}\right). \end{aligned}$$

Let  $A(g) \equiv (\dot{A}(\hat{g}), \hat{A}(\hat{g}))$  and  $B(g) \equiv (\dot{B}(\hat{g}), \hat{B}(\hat{g}))$  be defined by

$$A(g) \equiv \frac{g - *g + 1/2}{*g}, \quad B(g) \equiv \frac{1}{2}\left[\left(\frac{g - *g}{*g}\right)^2 - \frac{1}{*g}\right].$$

Define also the averaged versions

$$A^{\text{avg}} \equiv \frac{\sum_{g \in B} \Xi(g) A(g)}{\sum_{g' \in B} \Xi(g')}, \quad B^{\text{avg}} \equiv \frac{\sum_{g \in B} \Xi(g) B(g)}{\sum_{g' \in B} \Xi(g')}.$$

In view of (52), since  $\Phi$  has negative-definite Hessian at  $*h$ , we have  $[\|A^{\text{avg}}\|_\infty]^2 + \|B^{\text{avg}}\|_\infty \lesssim_k 1/n$ . Substituting the above into (61) gives

$$\frac{\mathbb{E}_T[Z^\partial[\underline{\sigma}_{\mathcal{U}}, \underline{\sigma}_{\mathcal{W}}]]}{p(\underline{\sigma}_{\mathcal{U}})} = \left(\mathbb{E}[Z'[\cdot, \underline{\sigma}_{\mathcal{W}}]]\right) \frac{\underline{s}}{c} \left[1 + \langle \delta^{\underline{\sigma}_{\mathcal{U}}}, A^{\text{avg}} \rangle + \langle (\delta^{\underline{\sigma}_{\mathcal{U}}})^2, B^{\text{avg}} \rangle + o\left(\frac{(\log n)^4}{n^{3/2}}\right)\right]$$

Recalling (60), we can express

$$\langle \delta, A^{\text{avg}} \rangle = \langle \bar{H}^{\underline{\sigma}_{\mathcal{U}}}, \xi_0 \rangle \quad \text{where } \xi_0(\sigma) = \langle \delta^\sigma, A^{\text{avg}} \rangle.$$

Meanwhile  $\langle \delta^2, B^{\text{avg}} \rangle$  is a sum of terms  $\delta^2(\underline{\hat{\sigma}}) \dot{B}^{\text{avg}}(\underline{\hat{\sigma}})$  and  $\delta^2(\underline{\hat{\sigma}}) \hat{B}^{\text{avg}}(\underline{\hat{\sigma}})$ , which can be written as

$$\delta^2(\underline{\hat{\sigma}}) \dot{B}^{\text{avg}}(\underline{\hat{\sigma}}) = \langle \bar{H}^{\underline{\sigma}_{\mathcal{U}}}, \xi_{\underline{\hat{\sigma}}} \rangle^2 \quad \text{where } \xi_{\underline{\hat{\sigma}}}(\sigma) \equiv \delta^\sigma(\underline{\hat{\sigma}}) [\dot{B}^{\text{avg}}(\underline{\hat{\sigma}})]^{1/2},$$

and similarly  $\delta^2(\underline{\hat{\sigma}}) \hat{B}^{\text{avg}}(\underline{\hat{\sigma}}) = \langle \bar{H}^{\underline{\sigma}_{\mathcal{U}}}, \xi_{\underline{\hat{\sigma}}} \rangle^2$ . The result follows.  $\square$

An analogous expansion as (59) holds for the second-moment partition function  $\mathbb{E}_T[Z^\partial[\underline{\mathcal{I}}_{\mathcal{U}}, \underline{\mathcal{I}}_{\mathcal{W}}]]$  with  $\underline{\mathcal{I}}_{\mathcal{W}}$  fixed. Lem. 6.7 easily implies bounds on the Fourier coefficients of the function  $\mathbb{F}(\underline{\mathcal{I}}_{\mathcal{U}} | \underline{\mathcal{I}}_{\mathcal{W}})$  of (51): for  $\underline{s} \in [|\mathcal{M}|]^\mathcal{U}$  write  $\{\underline{s}\} \equiv \{u \in \mathcal{U} : s_u \neq 1\}$ . For  $\underline{s}^1, \underline{s}^2 \in [|\mathcal{M}|]^\mathcal{U}$  write  $\{\underline{s}^1 \underline{s}^2\} \equiv \{\underline{s}^1\} \cup \{\underline{s}^2\}$ . If  $\{\underline{s}^1 \underline{s}^2\} = \{u, u'\}$  for some  $u \neq u'$  in  $\mathcal{U}$ , the Fourier coefficient captures the quadratic term in (59):

$$\begin{aligned} \frac{\mathbb{F}^\wedge(\underline{s} | \underline{\mathcal{I}}_{\mathcal{W}}) \cdot c}{\mathbb{F}^\wedge(\emptyset | \underline{\mathcal{I}}_{\mathcal{W}})} &= \sum_{j=1}^{C_k} \sum_{\underline{\mathcal{I}}_{\mathcal{U}}} p(\underline{\mathcal{I}}_{\mathcal{U}}) b_{\underline{s}}(\underline{\mathcal{I}}_{\mathcal{U}}) \xi_j(\tau_u) \xi_j(\tau_{u'}) + O_k\left(\frac{(\log n)^6}{n^{3/2}}\right) \\ &= \sum_{j=1}^{C_k} \langle \xi_j, b_{s(u)} \rangle_p \langle \xi_j, b_{s(u')} \rangle_p = O_k(1/n). \end{aligned}$$

If  $|\{\underline{s}^1 \underline{s}^2\}| \geq 3$  then the Fourier coefficient no longer captures the quadratic term. Altogether, in the case  $k$  divides  $|\mathcal{U}|$  we obtain the following in addition to (58):

$$\frac{\mathbb{F}^\wedge(\underline{s} | \underline{\mathcal{I}}_{\mathcal{W}})}{\mathbb{F}^\wedge(\emptyset | \underline{\mathcal{I}}_{\mathcal{W}})} \text{ is } \begin{cases} \lesssim_k n^{-1} & \text{for } |\{\underline{s}^1 \underline{s}^2\}| \geq 2; \\ \lesssim_k n^{-3/2} (\log n)^6 & \text{for } |\{\underline{s}^1 \underline{s}^2\}| \geq 3. \end{cases} \quad (62)$$

**6.4. Local neighborhood Fourier coefficients.** Let us recall again the definition (50) of the local neighborhood  $T$  of the unmatched  $\mathcal{K}$  in  $G^\circ$  (equipped with random literals), with leaves joining  $T$  to the graph  $G^\circ$  considered in §6.3. Recall also that  $A, \hat{A}$  are arbitrary choices of clauses (with literals) to place on  $\mathcal{K}$ . For now we suppress the dependence on  $\underline{\sigma}_{\mathcal{W}}$ , and abbreviate  $\kappa(\underline{\sigma}_{\mathcal{U}}) \equiv \kappa(\underline{\sigma}_{\mathcal{U}}|\underline{\sigma}_{\mathcal{W}})$  for the partition function on  $T \cup A$  given boundary configuration  $\underline{\sigma}_{\mathcal{U}} \equiv (\underline{\sigma}_{\mathcal{U}}, \underline{\sigma}_{\mathcal{W}})$ . Similarly we write  $\hat{\kappa}(\underline{\sigma}_{\mathcal{U}}) \equiv \hat{\kappa}(\underline{\sigma}_{\mathcal{U}}|\underline{\sigma}_{\mathcal{W}})$ . In this subsection we control the Fourier coefficients  $\varpi^\wedge = (\kappa - \hat{\kappa})^\wedge$  appearing in (51).

Let  $\mathbf{T}$  denote the event that  $T$  consists of  $|\mathcal{K}|$  tree components with  $T \cap \mathcal{W} = \emptyset$ . Let  $\mathbf{C}^\circ$  denote the event that  $T$  either contains a single cycle or has a single intersection with  $\mathcal{W}$  (but not both), but still consists of  $|\mathcal{K}|$  components. Lastly let  $\mathbf{C}'$  be the event that  $B_{2\nu-2}^\circ(\mathcal{K})$  has  $|\mathcal{K}|$  components, but  $T = B_{2\nu}^\circ(\mathcal{K})$  has  $|\mathcal{K}| - 1$  components (cf. (50)) and is disjoint from  $\mathcal{W}$ .

**Lemma 6.8.** *For  $T \in \mathbf{T}$ ,  $\varpi^\wedge(\underline{s}) = 0$  for all  $|\{\underline{s}\}| \leq 1$ ; and  $\kappa^\wedge(\emptyset)|_{\mathbf{T}}$  takes a constant value  $\bar{\kappa}^\wedge(\emptyset)$ , which does not depend on  $\underline{\sigma}_{\mathcal{W}}$ , the literals on  $T$ , or the clauses  $A$ . For  $T \in \mathbf{C}^\circ$ ,  $\varpi^\wedge(\emptyset) = 0$ .*

*Proof.* On the event  $\mathbf{T} \cup \mathbf{C}^\circ$ , the graphs  $T \cup A$  and  $T \cup \hat{A}$  are isomorphic ignoring the literals. If  $|\{\underline{s}\}| \leq 1$  then  $\mathbf{b}_{\underline{s}}$  depends at most on the spin of a single edge  $e \in \mathcal{U}$ . For  $T \in \mathbf{T}$ , using the symmetry of NAE-SAT one can produce an involution  $\iota: \underline{\sigma}_{\mathcal{U}} \mapsto \hat{\underline{\sigma}}_{\mathcal{U}}$  on  $\mathcal{M}^{\mathcal{U}}$  which keeps  $\sigma_e$  fixed, is measure-preserving with respect to  $\mathbf{p}$ , and satisfies  $\kappa(\underline{\sigma}_{\mathcal{U}}) = \hat{\kappa}(\hat{\underline{\sigma}}_{\mathcal{U}})$ : set  $\hat{\sigma}_u$  to be  $\sigma_u$  or  $-\sigma_u$  depending on whether the sum of literals along the unique path joining  $e$  to  $u$  in  $T \cup A$  differs in parity from the corresponding sum in  $T \cup \hat{A}$ . Then

$$\kappa^\wedge(\underline{s}) = \sum_{\underline{\sigma}_{\mathcal{U}}} \mathbf{p}(\iota \underline{\sigma}_{\mathcal{U}}) \mathbf{b}_{\underline{s}}(\iota \underline{\sigma}_{\mathcal{U}}) \kappa(\iota \underline{\sigma}_{\mathcal{U}}) = \sum_{\underline{\sigma}_{\mathcal{U}}} \mathbf{p}(\underline{\sigma}_{\mathcal{U}}) \mathbf{b}_{\underline{s}}(\underline{\sigma}_{\mathcal{U}}) \hat{\kappa}(\underline{\sigma}_{\mathcal{U}}) = \hat{\kappa}^\wedge(\underline{s}).$$

It is also clear that this value does not depend on  $\underline{\sigma}_{\mathcal{W}}$ , proving our claim on  $\mathbf{T}$ . A similar argument proves  $\varpi^\wedge(\emptyset) = 0$  on  $\mathbf{C}^\circ$ .  $\square$

**Lemma 6.9.** *For  $T \in \mathbf{T}$ ,  $\varpi^\wedge(\underline{s}) \lesssim_k \bar{\kappa}^\wedge(\emptyset)/[4^{(k-4)t}]$  for all  $|\{\underline{s}\}| = 2$ .*

*Proof.* Since  $|\{\underline{s}\}| = 2$  we may write  $\mathbf{b}_{\underline{s}}(\underline{\sigma}_{\mathcal{U}}) = f(\sigma_u)g(\sigma_w)$  for  $f \equiv \mathbf{b}_{s(u)}$  and  $g \equiv \mathbf{b}_{s(w)}$ . If  $u, w$  belong in the same connected component of  $T$ , arguing as in the proof of Lem. 6.8 gives  $\varpi^\wedge(\underline{s}) = 0$ , so assume they belong to different components. Since  $T \in \mathbf{T}$  we may define the random measure  $\mu_T(\underline{\sigma}_{\mathcal{U}}) \equiv \mathbf{p}(\underline{\sigma}_{\mathcal{U}}) \kappa(\underline{\sigma}_{\mathcal{U}}) / \bar{\kappa}^\wedge(\emptyset)$ , and similarly  $\hat{\mu}_T$ . Then

$$\varpi^\wedge(\underline{s}) = \bar{\kappa}^\wedge(\emptyset) \langle \mu_T - \hat{\mu}_T, fg \rangle = \bar{\kappa}^\wedge(\emptyset) [\text{Cov}_\mu(f, g) - \text{Cov}_{\hat{\mu}}(f, g)],$$

since Lem. 6.8 implies  $f$  has the same expectation with respect to  $\mu$  or  $\hat{\mu}$  (and likewise  $g$ ). Let  $\gamma_u$  denote the (unique) path joining  $u$  to  $\mathcal{K}$  in  $T$ , and likewise  $\gamma_w$ . Let  $N$  denote the event that on the path  $\gamma \equiv \gamma_u \cup \gamma_w$  there exists a clause  $a$  such that, among the  $k - 2$  variables in  $(\partial a) \setminus \gamma$ , there exist two variables  $v', v''$  with

$$\sigma_{v' \rightarrow a} = \sigma_{v'' \rightarrow a} = \mathbf{f} \quad \text{or} \quad L_{av'} \oplus \sigma_{v' \rightarrow a} = -L_{av''} \oplus \sigma_{v'' \rightarrow a}.$$

Note that  $N$  is  $(T, \underline{\sigma}_{\mathcal{U} \setminus \{u, w\}})$ -measurable, and on this event  $\text{Cov}_\mu(f, g) = \text{Cov}_{\hat{\mu}}(f, g) = 0$ . For any fixed realization of literals on  $T$  the probability (with respect to the law  $\mathbf{p}$  of the spins on  $\mathcal{U} \setminus \{u, w\}$ ) that  $N$  fails is  $\lesssim 4^{-(k-4)t}$  simply by the 0/1 symmetry in  $\mathbf{p}$ , so we find

$$|\varpi^\wedge(\underline{s})| \lesssim 4^{-(k-4)t} \bar{\kappa}^\wedge(\emptyset) \max_{s \in [|\mathcal{K}|]} \|\mathbf{b}_s\|_\infty^2 \lesssim_k 4^{-(k-4)t} \bar{\kappa}^\wedge(\emptyset),$$

since the relation  $\|\mathbf{b}_s\|_q^2 = 1$  implies  $\|\mathbf{b}_s\|_\infty^2 \leq (\min_{\sigma \in \mathcal{A}} \dot{q}_\sigma)^{-1} \lesssim_k 1$ .  $\square$

**Lemma 6.10.** *If  $T \in \mathbf{C}'$  then  $\kappa^\wedge(\emptyset) = \hat{\kappa}^\wedge(\emptyset)$  and*

$$|\varpi^\wedge(\emptyset)| \lesssim_k \frac{[\kappa^\wedge(\emptyset) \wedge \hat{\kappa}^\wedge(\emptyset)]}{(4^k/k^{O(1)})^t}.$$

*Proof.* Let  $\kappa^\circ(\underline{\sigma}_{\mathcal{A}}, \underline{\sigma}_{\mathcal{X}})$  denote the partition function on  $T$  given boundary conditions  $(\underline{\sigma}_{\mathcal{A}}, \underline{\sigma}_{\mathcal{X}}, \underline{\sigma}_{\mathcal{H}})$ , and define the random measure

$$\mu^\circ(\underline{\sigma}_{\mathcal{X}}) \equiv [(\kappa^\circ)^\wedge(\emptyset)]^{-1} \sum_{\underline{\sigma}_{\mathcal{A}}} \mathbf{p}(\underline{\sigma}_{\mathcal{A}}) \kappa^\circ(\underline{\sigma}_{\mathcal{A}}, \underline{\sigma}_{\mathcal{X}}).$$

If  $\iota(\underline{\sigma}_{\mathcal{X}})$  is the indicator that  $\underline{\sigma}_{\mathcal{X}}$  is valid for clauses  $A$  (likewise  $\iota$  for clauses  $A'$ ) then

$$\varpi^\wedge(\emptyset) = \sum_{\underline{\sigma}_{\mathcal{A}}, \underline{\sigma}_{\mathcal{X}}} \mathbf{p}(\underline{\sigma}_{\mathcal{A}}) \kappa^\circ(\underline{\sigma}_{\mathcal{A}}, \underline{\sigma}_{\mathcal{X}}) [(\iota - \hat{\iota})(\underline{\sigma}_{\mathcal{X}})] = (\kappa^\circ)^\wedge(\emptyset) \sum_{\underline{\sigma}_{\mathcal{X}}} \mu^\circ(\underline{\sigma}_{\mathcal{X}}) [(\iota - \hat{\iota})(\underline{\sigma}_{\mathcal{X}})].$$

By definition of the event  $\mathbf{C}_{\nu'}$ ,

$$\mu^\circ(\underline{\sigma}_{\mathcal{X}}) = \mu^\circ(\sigma_{e'}, \sigma_{e''}) \prod_{e \in \mathcal{X} \setminus \{e', e''\}} \mu^\circ(\sigma_e)$$

where  $\{e', e''\}$  is the unique pair of edges in  $\mathcal{X}$  such that  $B_{2\nu'}^\circ(e')$  and  $B_{2\nu'}^\circ(e'')$  intersect. The graph  $T$  (without  $A$  or  $A'$ ) contains no cycles, so it follows from the symmetry argument of Lem. 6.8 that the marginal of  $\mu^\circ$  on each  $e \in \mathcal{X}$  does not depend on the literals on  $T$ ; further each marginal must simply be  $\mathbf{p}$  from the Bethe recursions. It remains to note (arguing as in the proof of Lem. 6.9) that  $|\mu^\circ(\sigma_{e'}, \sigma_{e''}) - \mu^\circ(\sigma_{e'}) \mu^\circ(\sigma_{e''})| \lesssim 4^{-(k-4)\nu'}$ , from which we conclude

$$[1 + O_k(4^{-(k-4)\nu'})] \kappa^\wedge(\emptyset) = (\kappa^\circ)^\wedge(\emptyset) \sum_{\underline{\sigma}_{\mathcal{X}}} \mathbf{p}(\underline{\sigma}_{\mathcal{X}}) \iota(\underline{\sigma}_{\mathcal{X}}) = [1 + O_k(4^{-(k-4)\nu'})] \hat{\kappa}^\wedge(\emptyset)$$

(where  $\langle \mathbf{p}, \iota \rangle = \langle \mathbf{p}, \hat{\iota} \rangle$  again by symmetry).  $\square$

**6.5. Conclusion.** We now combine the estimates of §6.3 and §6.4 to conclude the proof:

*Proof of Propn. 6.1.* Recalling (46) and (47), we will set  $\mathbf{B}$  to be the union of the events  $\mathbf{C}_{\nu'}$ . First consider (51) for  $T \in \mathbf{T}$ : by Lem. 6.8 there is no contribution from terms  $\underline{\mathbf{z}} = (\underline{\mathbf{z}}^1, \underline{\mathbf{z}}^2)$  with either  $|\{\underline{\mathbf{z}}^i\}| \leq 1$ , so we need only consider  $\underline{\mathbf{z}}$  with  $|\{\underline{\mathbf{z}}^1, \underline{\mathbf{z}}^2\}| \geq 2$ . The number of choices of  $\underline{\mathbf{z}}$  with  $|\{\underline{\mathbf{z}}\}| = 2$  is  $(k^{O(1)}4^k)^t$ . Combining (62) and Lem. 6.9 then gives

$$\frac{\mathbb{E}[\mathbf{z}^* \mathbf{D}; \mathbf{T}]}{(\mathbb{E} \mathbf{Z})^2} = \sum_{\mathcal{I}_{\mathcal{H}}} \bar{\varphi}^{\otimes}(\mathcal{I}_{\mathcal{H}}) \left\{ \prod_{i=1,2} \frac{\kappa^\wedge(\emptyset | \underline{\sigma}_{\mathcal{H}}^i)}{(4^k/k^{O(1)})^t} \right\} \frac{\mathbb{F}^\wedge(\emptyset | \mathcal{I}_{\mathcal{H}})}{(\mathbb{E} \mathbf{Z})^2} \left[ \frac{(k^{O(1)}4^k)^t}{n} + \frac{(\log n)^{O(1)}}{n^{3/2}} \right] \leq \frac{(k^{O(1)}/4^k)^t}{n}.$$

Similarly, for  $T \in \mathbf{C}^\circ$ , by Lem. 6.8 there is no contribution from  $\underline{\mathbf{z}} = \emptyset$ , so (58) gives

$$\frac{\mathbb{E}[\mathbf{z}^* \mathbf{D}; \mathbf{C}^\circ]}{(\mathbb{E} \mathbf{Z})^2} \leq \mathbb{P}(\mathbf{C}^\circ) \cdot O_k\left(\frac{|T| \log n}{n^{1/2}}\right) \leq \frac{(\log n)^{O(1)}}{n^{3/2}}.$$

Note  $\mathbf{B}^c = \mathbf{T} \cup \mathbf{C}^\circ \cup \Omega$  where  $\mathbb{P}(\Omega) \leq (\log n)^{O(1)}/n^2$ , so  $\mathbb{E}[\mathbf{z}^* \mathbf{D}; \Omega]/(\mathbb{E} \mathbf{Z})^2 \leq (\log n)^{O(1)}/n^2$ . This concludes our analysis of (46). Turning to (47), let  $\mathbf{z}^*(\mathbf{X} - \mathbf{X}_\emptyset)$  be the near-independent contribution to  $(\mathbf{X} - \mathbf{X}_\emptyset)^2$ : again applying (58),

$$\frac{\mathbb{E}[\mathbf{z}^*(\mathbf{X} - \mathbf{X}_\emptyset); \mathbf{B}]}{(\mathbb{E} \mathbf{Z})^2} \leq \mathbb{P}(\mathbf{B}) \frac{(\log n)^{O(1)}}{n^{1/2}} \leq \frac{(\log n)^{O(1)}}{n^{3/2}}.$$

As for the second term in (58), applying (6.10) gives

$$\mathbb{E} \left[ \frac{\mathbf{z}^*(\mathbf{X} - \mathbf{X}_\emptyset)}{\min\{\mathbf{X}_\emptyset, \hat{\mathbf{X}}_\emptyset\}^2}; \mathbf{B} \right] \leq \sum_{\nu'} \mathbb{P}(\mathbf{C}_{\nu'}) \left[ \frac{1}{(4^k/k^{O(1)})^{2\nu'}} + \frac{(\log n)^{O(1)}}{n^{1/2}} \right] \leq \frac{1}{n} \sum_{\nu'} \frac{1}{(4^k/k^{O(1)})^{\nu'}} \lesssim_k \frac{1}{n}.$$

Altogether this proves

$$\mathbb{E}[\mathbf{V}^2] \lesssim_k \frac{1}{n} + \frac{(k^{O(1)}/4^k)^t}{n\epsilon^2} + \frac{(\log n)^{O(1)}}{n^{3/2}}$$

where  $(\log n)^{O(1)}$  hides a prefactor that can depend on  $\epsilon$ . By taking  $t$  large this is enough to prove (45), and the result follows as explained in §6.1.  $\square$

## 7. FROM CLUSTERS TO ASSIGNMENTS

In this final section we prove the main theorem.

*Proof of Thm. 1 for  $d \neq d_*$ .* By the results of Thm. 3 and Thm. 6.2 it suffices to show that for  $k \geq k_0$  and  $d_{\text{bd}} \leq d < d_*$ ,  $\lim_{n \rightarrow \infty} \mathbb{P}(Z > 0 \mid Z(\mathbf{h}) > 0) = 1$  holds uniformly over all empirical measures  $\mathbf{h} \in \Delta$  satisfying  $\|(n\hat{\mathbf{h}}, m\hat{\mathbf{h}}) - \mathbf{g}\| \leq n^{1/2} \log n$ .

Given an auxiliary model configuration  $\underline{\sigma}$  on the edges of  $(G, \underline{L})$ , our aim is to complete  $\underline{\sigma}$  to an NAE-SAT solution  $\underline{x}$  on  $(G, \underline{L})$  (meaning that whenever  $x_v$  agrees with  $\eta_v \equiv \dot{m}_d(\underline{\sigma}_v)$  whenever  $\eta_v \neq \mathbf{f}$ ). Clearly, the potential issue is that setting a free variable may cause a chain of forcings resulting in an invalid assignment. We therefore let  $F^\# \equiv F^\#(G, \underline{L}, \underline{\sigma})$  denote the subset of clauses  $a \in F$  such that at least two variables in  $\partial a$  are free, and all rigid variables  $v \in \partial a$  have the same evaluation  $L_{av} \oplus \eta_v \equiv \xi_a$ .<sup>9</sup>

Let  $G^\# \equiv G^\#(G, \underline{L}, \underline{\sigma})$  denote the subgraph of  $G$  induced by the free variables together with the clauses  $F^\#$ . We claim that  $\underline{\sigma}$  has a valid completion to an NAE-SAT solution provided each connected component of  $G^\#$  contains at most one cycle. Indeed, in a tree component of  $G^\#$  one may choose an arbitrary root vertex and assign it an arbitrary value — this may cause a chain of forcings, but no conflict results since there is no cycle. In a unicyclic component  $C$  with cycle  $v_0, a_0, v_1, \dots, a_{n-1}, v_n$  (with indices taken modulo  $n$  so  $v_0 = v_n$ ), setting  $x_{v_i} = -L_{a_i v_i} \oplus \xi_{a_i}$  ensures that all clauses along the cycle are satisfied. Then, by the preceding argument for tree components, there exists a valid completion of  $\underline{x}$  to the remainder of  $C$ , proving our claim.

Conditioned on  $Z(\mathbf{h}) > 0$  we may generate  $(G, \underline{L}, \underline{\sigma})$  — where  $G$  has the law of  $\mathcal{G}_{n,d,k}$ ,  $\underline{L}$  is uniformly random, and  $\underline{\sigma}$  has empirical measure  $\hat{\mathbf{h}}$  — as follows: start with a set  $V$  of  $n$  variables each incident to  $d$  half-edges and a set  $F$  of  $m$  clauses each incident to  $k$  half-edges, and place spins on half-edges according to  $\hat{\mathbf{h}}$  and  $\hat{\mathbf{h}}$ . Then construct the graph by randomly matching clause and variable half-edges in breadth-first search manner started from an initial variable  $o$ , and respecting the given spins  $\sigma$ . It is clear from this construction that up to the time that the process has explored say  $n^{1/3}$  vertices, the evolution of the spins  $\underline{\sigma}$  on the leaves of the exploration tree is very close to the Markovian evolution of the Gibbs measure  $\nu$  described in §3.5. In particular, starting from any free variable  $v$ , the exploration of its connected component  $T_v$  in  $G^\#$  is dominated by a Galton–Watson branching process with offspring numbers distributed as a random variable  $0 \leq Y \leq dk$  with

$$\mathbb{E}Y \lesssim dk \mathbf{h}(F^\# \mid \sigma_1 = \mathbf{f}\mathbf{f}) = \frac{dk \sum_{j=2}^k \binom{k-1}{j-1} \hat{h}_{\mathbf{f}\mathbf{f}}^j [(\hat{h}_{0\mathbf{f}})^{k-j} + (\hat{h}_{1\mathbf{f}})^{k-j}]}{\hat{z}\hat{\mathbf{h}}(\mathbf{f}\mathbf{f})} \lesssim k^3/2^k.$$

By a standard argument the total size of the Galton–Watson tree has an exponential tail,<sup>10</sup> so we may take  $C \equiv C(k)$  such that  $\mathbb{P}(|T_v| \geq C \log n) \leq n^{-10}$ . The probability of seeing more than one cycle in  $T_v$  is then crudely  $\leq n^{-3/2}$ . Taking a union bound over all free variables shows that with high probability no component of  $G^\#$  contains more than a single cycle, so  $\underline{\sigma}$  corresponds to a true NAE-SAT solution as claimed.  $\square$

The above analysis completes the analysis of the SAT–UNSAT transition in the case that the critical threshold  $d_*$  (see Propn. 3.11) is non-integer. We conclude by showing that if  $d_* \in \mathbb{Z}$ , then at  $d = d_*$  the probability that a random NAE-SAT instance  $(\mathcal{G}_{n,d,k}, \underline{L})$  is solvable is asymptotically bounded away from zero and one.

<sup>9</sup>If  $\eta_{\partial a} = (\mathbf{f}^k)$  we also include  $a \in F^\#$ , and arbitrarily define  $\xi_a = 0$ .

<sup>10</sup>Suppose  $Y$  is a non-negative integer random variable with  $\Lambda(\lambda) \equiv \log \mathbb{E}e^{\lambda Y} < \infty$  for some  $\lambda > 0$ , and  $\Lambda'(0) = \mathbb{E}Y < 1$ . Let  $(Y_j)_{j \geq 1}$  be a sequence of i.i.d. random variables distributed as  $Y$ , and  $Z_n \equiv 1 + \sum_{j=1}^n (Y_j - 1)$ . Then the total size of a Galton–Watson tree with offspring distribution  $Y$  has the same law as  $\tau \equiv \inf\{n : Z_n = 0\}$ , and it is clear that the distribution of  $\tau$  has exponential decay:  $\mathbb{P}(\tau > j) \leq \mathbb{P}(Z_j \geq 1) \leq e^{-t} \mathbb{E}e^{tZ_j} = e^{j[\Lambda(t) - t]}$ , and since  $\Lambda'(0) = \mathbb{E}Y < 1$ , by considering  $t > 0$  sufficiently small we can find a constant  $c > 0$  such that  $\mathbb{P}(\tau > j) \leq e^{-cj}$ .



*Proof of Thm. 1 for  $d = d_*$  integer.* The probability of solvability is bounded away from zero by Thm. 2 together with the preceding argument. To see that the probability is bounded away from one, it suffices to show that  $\mathbb{E}[Z | \Omega_C] < 1$  for an event  $\Omega_C$  of asymptotically positive probability. We shall take  $\Omega_C$  to be the event that there is a large (but constant) number of disjoint triangles in the graph. We show below that each additional triangle decreases the expected partition function by a constant factor, so that  $\mathbb{E}[Z | \Omega_C] < 1$  for a sufficiently large (but constant) number of cycles. It is well-known that the number of triangles is asymptotically a non-degenerate Poisson random variable, so  $\Omega_C$  has asymptotically positive probability as required.

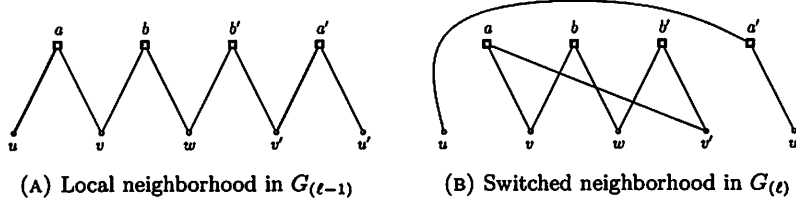


FIGURE 4. Switching argument for case of integer-valued  $d_*$ .

We define recursively a sequence of graphs  $(G_{(\ell)})_{\ell \geq 0}$  by the so-called “switching method.” Start from  $G_{(0)} \equiv \mathcal{G}_{n,d,k}$  ( $d = d_*$ ). For  $\ell \geq 1$ , let  $v, v'$  be a random pair of vertices at distance two in the hypergraph  $G_{(\ell-1)}$ , with common neighbor  $w$ . Say  $v$  is joined to  $w$  by clause  $b$ , and let  $u \neq w$  be another neighbor of  $v$  via a different clause  $a \neq b$ . Likewise say  $v'$  is joined to  $w'$  by clause  $b'$ , and let  $u' \neq w'$  be another neighbor of  $v'$  via a different clause  $a' \neq b'$  (Fig. 4A). Let  $G_{(\ell)}$  be defined by making the switching shown in Fig. 4B. The result will follow by showing that for  $\ell$  bounded by a large constant, this switching decreases the expected partition function by a constant factor.

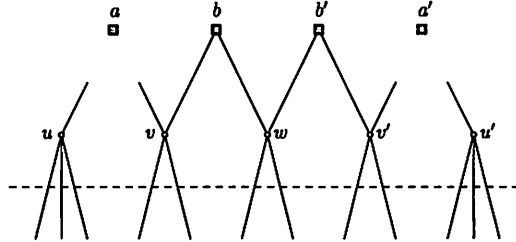


FIGURE 5. Local neighborhood in graph  $G_{(\ell)}^*$  with  $a, a'$  removed

Note that with high probability all previous switchings occur at distance at least say  $(\log n)^{1/2}$  away, so it suffices to prove the claim with  $\ell = 1$ . Consider the graph  $G^*$  with the clauses  $a$  and  $a'$  removed, leaving unmatched half-edges incident to variables (Fig. 5). Write  $\mathbb{P}^*$  for the marginal law, with respect to the auxiliary model on  $G^*$ , for the spins  $\sigma_u, \sigma_v, \sigma_w, \sigma_{v'}$  on the unmatched half-edges incident to  $u, v, u', v'$ ; and write each  $\sigma$  as  $i\sigma$  where  $i$  is the clause-to-variable message while  $\sigma$  is the variable-to-clause message. (For example,  $\sigma_u$  will correspond to  $\sigma_{u \rightarrow a}$  in the original graph versus  $\sigma_{u \rightarrow a'}$  in the switched graph.) We shall compare the probability for  $a$  and  $a'$  to be satisfied within the original graph versus the switched graph: with  $\mathbf{r} \equiv \{0, 1\}$  as above, we claim

$$\begin{aligned} \mathbb{P}^*(\sigma_v \in \mathbf{r}, \sigma_{v'} \in \mathbf{r}) [1 + o(1)] &\geq \mathbb{P}^*(\sigma_v \in \mathbf{r}, \sigma_u \in \mathbf{r}) + 2^{-5k}, & \text{(I)} \\ \mathbb{P}^*(\sigma_u \in \mathbf{r}, \sigma_{u'} \in \mathbf{r}) [1 + o(1)] &\geq \mathbb{P}^*(\sigma_{v'} \in \mathbf{r}, \sigma_{u'} \in \mathbf{r}) & \text{(II)} \end{aligned}$$

(In the above, the first display concerns clause  $a$  while the second concerns  $a'$ ; in both displays the left-hand side is relevant to the switched graph while the right is relevant to the original graph.) Recalling Lem. 6.5,  $\mathbb{P}^*$  is the same up to  $1 + o(1)$  factors as the measure  $\mathbf{P}$  induced on the local

neighborhood by taking boundary conditions given by  $p$  (on the edges cut by the dashed line in Fig. 5, without regard to the structure of  $G^*$ ). Under  $\mathbf{P}$ , clearly  $\sigma_u$ ,  $\sigma_{u'}$ , and  $\sigma_{v'}$  are mutually independent. Since  $v'$  has only  $d - 2$  neighbors coming from the rest of the graph, it is slightly biased towards  $\mathbf{f}$ , which proves (II).

To prove (I) we need to take two effects into account: first,  $\sigma_v$  and  $\sigma_{v'}$  are correlated while  $\sigma_v$  and  $\sigma_u$  are independent; and secondly, as noted in the proof of (II), marginally  $\sigma_{v'}$  is slightly more likely than  $\sigma_u$  to be  $\mathbf{f}$  due to the different structure of the local neighborhood. The correlation goes in our favor while the marginal bias goes against, and we argue that the former dominates. Indeed, as we have seen in the proof of Lem. 6.9, there is an event of probability  $\gamma = 2^{-2k}$  such that on this event  $\sigma_v$  and  $\sigma_{v'}$  must be both rigid (with probability  $z$ ) or both free (with probability  $z$ ), but given the complementary event they are conditionally independent with probability  $x$  for  $\sigma_v$  to be rigid and probability  $y$  for  $\sigma_{v'}$  to be rigid. Thus  $\mathbf{P}(\sigma_v \in \mathbf{r}) = \gamma z + (1 - \gamma)x$  which implies  $x_{\mathbf{f}} \equiv 1 - x = 2^{-k}$ ; likewise and  $\mathbf{P}(\sigma_{v'} \in \mathbf{r}) = \gamma z + (1 - \gamma)y$  which implies  $y_{\mathbf{f}} \equiv 1 - y = 2^{-k}$ . Combining,  $\mathbf{P}(\sigma_v, \sigma_{v'} \in \mathbf{r}) - \mathbf{P}(\sigma_v \in \mathbf{r})\mathbf{P}(\sigma_{v'} \in \mathbf{r})$  is quadratic in  $z$ , and it is straightforward to compute the derivative and see that it is  $\asymp \gamma$  (hence positive) for  $0 \leq z \leq 1$ . Evaluating at  $z = 1$  gives

$$\mathbf{P}(\sigma_v, \sigma_{v'} \in \mathbf{r}) - \mathbf{P}(\sigma_v \in \mathbf{r})\mathbf{P}(\sigma_{v'} \in \mathbf{r}) \geq \gamma(1 - \gamma)(1 - x)(1 - y) = 2^{-4k}.$$

As for the marginal bias, note that the increased chance for  $\sigma_{v'}$  to be free compared with  $\sigma_u$  comes from the fact that  $v$  receives only  $d - 2$  incoming messages from the rest of the graph: thus  $\sigma_{v \rightarrow b}$  is slightly biased towards  $\mathbf{f}$ , and this effect can percolate through the chain  $\sigma_{b \rightarrow w}, \sigma_{w \rightarrow b'}, \sigma_{b' \rightarrow v'}$  to affect  $\sigma_{v'}$ . However the initial bias on  $\sigma_{v \rightarrow b}$  is  $\lesssim 4^{-k}$ , and the effect decreases by a factor  $2^k$  passing through each step in the chain, so the overall bias is  $\lesssim 2^{-6k}$ . Combining the estimates proves (I).

The result follows from (I) and (II) by noting that in a clause with  $k$  random incoming messages which are mutually independent except for possible correlation among the first two, the probability for the clause to be satisfied decreases if the probability for the first two messages to be both rigid increases.  $\square$

## REFERENCES

- [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*, SODA '01, pages 721–722, Philadelphia, PA, USA, 2001. Society for Industrial and Applied Mathematics.
- [AGK04] E. Aurell, U. Gordon, and S. Kirkpatrick. Comparing beliefs, surveys, and random walks. In *NIPS*, 2004.
- [AL07] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12(54):1454–1508, 2007.
- [AM06] D. Achlioptas and C. Moore. Random  $k$ -SAT: two moments suffice to cross a sharp threshold. *SIAM J. Comput.*, 36(3):740–762, 2006.
- [AP04] D. Achlioptas and Y. Peres. The threshold for random  $k$ -SAT is  $2^k \log 2 - O(k)$ . *J. Amer. Math. Soc.*, 17(4):947–973, 2004.
- [BBC<sup>+</sup>01] B. Bollobás, C. Borgs, J. T. Chayes, J. H. Kim, and D. B. Wilson. The scaling window of the 2-SAT transition. *Random Struct. Algor.*, 18(3):201–256, 2001.
- [BMZ05] A. Braunstein, M. Mézard, and R. Zecchina. Survey propagation: an algorithm for satisfiability. *Random Struct. Algor.*, 27(2):201–226, 2005.
- [BS01] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6(23):1–13, 2001.
- [BZ04] A. Braunstein and R. Zecchina. Survey propagation as local equilibrium equations. *J. Stat. Mech. Theory E.*, 2004(06):P06007, 2004.
- [CNRZ03] T. Castellani, V. Napolano, F. Ricci-Tersenghi, and R. Zecchina. Bicolouring random hypergraphs. *J. Phys. A*, 36(43):11037–11053, 2003.
- [Coj13] A. Coja-Oghlan. Random regular  $k$ -SAT. Preprint, arXiv:1310.2728v1, 2013.
- [Coj14] A. Coja-Oghlan. The asymptotic  $k$ -SAT threshold. In *Proc. 46th STOC*, 2014.
- [CP12] A. Coja-Oghlan and K. Panagiotou. Catching the  $k$ -NAESAT threshold. In *Proc. 44th 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.

- [CR92] V. Chvátal and B. Reed. Mick gets some (the odds are on his side) [satisfiability]. In *Proc. 33rd FOCS*, pages 620–627, 1992.
- [CZ12] A. Coja-Oghlan and L. Zdeborová. The condensation transition in random hypergraph 2-coloring. In *Proc. 23rd SODA*, pages 241–250. SIAM, 2012.
- [DM95] B. Davis and D. McDonald. An elementary proof of the local central limit theorem. *J. Theoret. Probab.*, 8(3):693–701, 1995.
- [DRZ08] L. Dall’Asta, A. Ramezanpour, and R. Zecchina. Entropy landscape and non-Gibbs solutions in constraint satisfaction problems. *Phys. Rev. E (3)*, 77(3):031118, 16, 2008.
- [DSS13] J. Ding, A. Sly, and N. Sun. Maximum independent sets on random regular graphs. Preprint, arXiv:1310.4787, 2013.
- [DSS14] J. Ding, A. Sly, and N. Sun. Satisfiability threshold for random regular NAE-SAT. In *Proc. 46th STOC*, New York, NY, USA, 2014. ACM.
- [FdIV01] W. Fernandez de la Vega. Random 2-SAT: results and problems. *Theoretical Computer Science*, 265(1):131–146, 2001.
- [Goe96] A. Goerdt. A threshold for unsatisfiability. *J. Comput. System Sci.*, 53(3):469–486, 1996.
- [GŠV14] A. Galanis, D. Štefankovič, and E. Vigoda. Inapproximability for antiferromagnetic spin systems in the tree non-uniqueness region. In *Proc. 46th STOC*, 2014.
- [JLR00] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
- [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<sup>+</sup>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. *P. Natl. Acad. Sci.*, 104(25):10318–10323, 2007.
- [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.
- [MP85] M. Mézard and G. Parisi. Replicas and optimization. *Journal de Physique Lettres*, 46(17):771–778, 1985.
- [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.
- [MRTZ03] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina. Two solutions to diluted  $p$ -spin models and XORSAT problems. *J. Stat. Phys.*, 111(3-4):505–533, 2003.
- [MS08] E. Maneva and A. Sinclair. On the satisfiability threshold and clustering of solutions of random 3-SAT formulas. *Theoret. Comput. Sci.*, 407(1-3):359–369, 2008.
- [MZK<sup>+</sup>99] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky. Determining computational complexity from characteristic “phase transitions”. *Nature*, 400(6740):133–137, 1999.
- [Par02] G. Parisi. On local equilibrium equations for clustering states. Preprint, arXiv:cs/0212047v2, 2002.
- [PS12] B. Pittel and G. B. Sorkin. The satisfiability threshold for  $k$ -XORSAT. Preprint, arXiv:1212.1905v2, 2012.
- [RW92] R. W. Robinson and N. C. Wormald. Almost all cubic graphs are Hamiltonian. *Random Struct. Algor.*, 3(2):117–125, 1992.
- [RW94] R. W. Robinson and N. C. Wormald. Almost all regular graphs are Hamiltonian. *Random Struct. Algor.*, 5(2):363–374, 1994.
- [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.

\*DEPARTMENT OF STATISTICS, UNIVERSITY OF CHICAGO  
ECKHART HALL, 5734 SOUTH UNIVERSITY AVENUE, CHICAGO, IL 60637A

†DEPARTMENT OF STATISTICS, UNIVERSITY OF CALIFORNIA, BERKELEY  
EVANS HALL, BERKELEY, CA 94720  
— AND —

MATHEMATICAL SCIENCES INSTITUTE, AUSTRALIAN NATIONAL UNIVERSITY  
JOHN DEDMAN BUILDING, UNION LANE, CANBERRA, ACT 0200

‡DEPARTMENT OF STATISTICS, STANFORD UNIVERSITY  
SEQUOIA HALL, 390 SERRA MALL, STANFORD, CA 94305