Statistical Limits of Graphical Channel Models and a Semidefinite Programming Approach

by

Chiheon Kim

B.S., Korea Advanced Institute of Science and Technology (2012)

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

Signature redacted

Author ....
Department of Mathematics
August 21, 2018

Signature redacted

Certified by ....
Michel X. Goemans
Professor of Applied Mathematics
Thesis Supervisor

Signature redacted

Accepted by ....
Jonathan A. Kelner
Graduate Chair

Statistical Limits of Graphical Channel Models and a Semidefinite Programming Approach

by

Chiheon Kim

Submitted to the Department of Mathematics
on August 21, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Community recovery is a major challenge in data science and computer science. The goal in community recovery is to find the hidden clusters from given relational data, which is often represented as a labeled hypergraph where nodes correspond to items needing to be labeled and edges correspond to observed relations between the items.

We investigate the problem of exact recovery in the class of statistical models which can be expressed in terms of graphical channels. In a graphical channel model, we observe noisy measurements of the relations between k nodes while the true labeling is unknown to us, and the goal is to recover the labels correctly. This generalizes both the stochastic block models and spiked tensor models for principal component analysis, which has gained much interest over the last decade. We focus on two aspects of exact recovery: statistical limits and efficient algorithms achieving the statistic limit.

For the statistical limits, we show that the achievability of exact recovery is essentially determined by whether we can recover the label of one node given other nodes labels with fairly high probability. This phenomenon was observed by Abbe *et al.* for generic stochastic block models, and called "local-to-global amplification". We confirm that local-to-global amplification indeed holds for generic graphical channel models, under some regularity assumptions. As a corollary, the threshold for exact recovery is explicitly determined.

For algorithmic concerns, we consider two examples of graphical channel models, (i) the spiked tensor model with additive Gaussian noise, and (ii) the generalization of the stochastic block model for k-uniform hypergraphs. We propose a strategy which we call "truncate-and-relax", based on a standard semidefinite relaxation technique. We show that in these two models, the algorithm based on this strategy achieves exact recovery up to a threshold which orderwise matches the statistical threshold. We complement this by showing the limitation of the algorithm.

Thesis Supervisor: Michel X. Goemans
Title: Professor of Applied Mathematics

Acknowledgments

Above anyone else, I would like to thank my advisor, Michel X. Goemans, for the continuous support of my work, for his valuable guidance as a senior mathematician, and for his patience. It would not have been possible for me to successfully complete the program without his support and generosity.

I have had many opportunities to interact with great mathematicians at MIT. I would like to thank Afonso S. Bandeira for his mentoring and for being an amazing collaborator. Special thanks to all of my colleagues in the department, especially to Francisco Unda for being great company for writing this thesis.

I am grateful for MIT PRIMES, RSI, and all students whom I mentored. I would like to thank you for giving me a great opportunity to work brilliant students. I learned so much from every single one of you.

I would like to thank Priyanka and MIT Syncopasian for their constant emotional support and for being a community that I could belong to throughout my six years of graduate study.

Finally, I would like to thank my amazing family for their love.

Contents

Chapter 1

Introduction

Identifying clusters from relational data is one of the fundamental problems in data science. Such a task can be often formulated as a problem of recovering the true labeling (or community assignment) of the items from a given data set which is a collection of noisy measurements of the similarity between two or more items. We are particularly interested in the setting where those measurements are independent of each other. Such models are referred to as graphical models, or conditional random fields in some literature.

As a motivational example, let us consider a community detection problem on a graph. The stochastic block model (SBM) is one of the simplest generative models which can be used to formally address such a problem. Specifically, in the case of the SBM with two symmetric communities, we observe a graph G on the vertex set V, where two vertices are joined with an edge with a probability only depending on whether they belong to the same community. We assume that true community labels are already assigned to the vertices but are unknown to us. The goal of community detection is to recover the correct labels from the observed graph G. We note that one can only succeed with high probability, as there is still a positive (but low) chance for G being adversarial for recovering all the labels.

In this thesis, we only focus on *exact recovery*, *i.e.*, we are only interested in the solution which is correct everywhere. There are other notions of recovery requirements, such as *almost exact recovery* and *partial recovery*, which lead to their own theories of interest. However, we will not take a deeper look into them as it is out of the scope of this thesis.

For exact recovery, it is well-known in the statistics community that the maximum likelihood estimation or Bayesian estimation achieves the minimum error. This allows us to understand exact recovery by analyzing the performance of the optimal estimator. In the SBM with two equal-sized communities, the optimal estimator reduces to the minimum[1] bisection problem: given a graph G with the vertex set V, find $A \subseteq V$ with half of the vertices such that the number of edges between A and $V \setminus A$ is minimized. Hence, exact recovery is possible if and only if the minimum bisection problem returns the correct communities with high probability, where the probability is taken over all randomness of the model.

Such an optimal estimator is usually hard to compute as it requires optimization over the solution space whose size is exponential. In the example above, to compute the optimal estimator we need to solve the minimum bisection which is indeed an NP-hard problem. Nevertheless, we expect that worst-case instances for those problems are nothing similar to a "typical" instance we observe from the model.

Let us name a few advantages of considering the average-case complexity over a statistical model. First of all, it surpasses the curse of worst-case instances which might not be a good representative of data from real-world applications. Second, such models can be used as a testbed for various algorithms, allowing us to compare them with provable performance guarantees. Also, it is very interesting in complexity theory perspective that many NP-hard problems becomes easy when we consider average-case complexity over a model with a planted solution.

This thesis focuses on a particular class of statistical models which are called the *graphical channel models*. Graphical channels are powerful enough for describing important models such as SBMs and its variants, noise models with planted signal which are used to model principal component analysis, random constraint satisfaction problems with planted solutions, and more.

Let us formulate the recovery problem as a problem of inferring the values of

[1] or maximum, depending on the parameters of the model

latent variables from an instance of observables which are mutually independent and depend on only a few number of latent variables each. For example, in the SBM described above, community assignment of a vertex corresponds to a latent variable and the presence (or non-presence) of an edge corresponds to an observable which only depend on the labels of its endpoints.

Precisely, we consider a model which factorizes along a hypergraph $\mathcal{H} = (V, E)$ where $v \in V$ is associated with a latent variable $\mathbf{x}_v$ and $e \in E$ is associated with an observable $\mathbf{y}_e$. We further restrict our focus on the case that the hypergraph $\mathcal{H}$ is k-uniform for some fixed k (i.e., $|e| = k$ for all hyperedge e), hence the dependence graph on $\mathbf{x}_v$'s and $\mathbf{y}_e$'s forms a bipartite graph which is k-regular on the $\mathbf{y}$-side. This is the graphical channel model.

One may notice that the graphical channel model can be expressed as a graph code on a memory-less noise channel. In this context exact recovery corresponds to the problem of decoding the message from a given corrupted codeword. The celebrated Shannon's noisy-channel coding theorem tells us that when we fix a memory-less channel with capacity C, there then exists an encoding scheme with rate R which allows decoding with high probability as long as $R < C$. The converse is also true: if $R > C$, then no encoding scheme with rate R allows a stable decoding.

Inspired by Shannon's theorem, we search for the correct notion of "channel capacity" for graphical channel models. Notice the difference between the two settings: in the setting of graphical channel models, we have an encoding scheme which is given to us and have specific structure, while in the setting of Shannon's theorem, we have freedom to choose an encoding scheme while the channel is fixed. We also remark that the graphical channel model corresponds to a sparse code whose rate decays to zero as the length of message $|V|$ grows. This implies that the channel capacity in a traditional sense would be diverging as $|V|$ grows, so we need an asymptotic notion of the capacity in the limit of $|V| \to \infty$.

It was shown in [6] that the Chernoff-Hellinger divergence serves a role of capacity in generic SBMs. As a consequence, one can find a sharp threshold such that exact recovery can be done successfully with probability asymptotically approaching 1 if and only if the capacity is above the threshold. This can be proven by a local-to-global amplification argument as appearing in [1], which means roughly that we can recover all vertex labels as long as we can *locally* recover the label of a vertex given the labels of all other vertices with low failure probability.

We extend this argument to discrete signal recovery in spiked tensor models in Chapter 2, and to a generalized version of SBMs for k-uniform hypergraphs in Chapter 3. Moreover, in Chapter 4 we prove that local-to-global amplification happens in generic graphical channel models as long as it satisfies some regularity conditions on the channels. As a corollary, we determine the sharp threshold value in terms of a certain type of divergence computed on the channel which matches previous works on a specific model.

On the other hand, we ask whether exact recovery can be *efficiently* done. Recall that in the case of the SBM with two symmetric communities, the optimal estimation scheme corresponds to the minimum bisection problem which is NP-hard to solve in the worst case. We may wonder whether there is a polynomial-time algorithm which solves the minimum bisection problem on most typical instances. Indeed, there are several polynomial-time algorithms known to achieve exact recovery up to the statistical threshold in this case [4, 21, 56, 6].

In general, the optimal estimator in the graphical channel model can be expressed as a polynomial optimization problem, which is the problem of finding the maximum (or the minimum) value of a given polynomial over a set which can be described by polynomial inequalities. Although this problem is in general NP-hard (even when the polynomial is quadratic), the sum-of-squares relaxation scheme provides a systematic way to find an approximate solution. Roughly speaking, this scheme first relaxes the original problem to a convex optimization problem which we can solve efficiently, then finds a feasible solution by rounding the solution of the relaxation. Interestingly, in [21] and [56] it was independently pointed out that in the SBM with two symmetric communities, one does not need the rounding step: The relaxed problem will give the exact solution with high probability as long as it is statistically possible.

Inspired by those results, we propose a strategy which we call "truncate-and-relax":

We first truncate the polynomial to a quadratic polynomial, then solve the standard semidefinite relaxation to find the optimum of the truncated polynomial. We show that in planted bisection models and in generalized SBMs for k-uniform hypergraphs, the truncate-and-relax strategy successfully recovers the community labels up to a threshold which orderwise matches with the statistical threshold (See Section 2.4 and Section 3.4, respectively).

1.1 Organization of the thesis

In the rest of this chapter, we motivate our work by reviewing previous works on two particular models: the stochastic block model with two symmetric communities and the spiked Wigner model. We also give a general formulation of the graphical channel model and several important examples of it, formally describing the recovery requirements that we consider, and discuss briefly various relaxation techniques from the spectral to the Sum-of-Squares methods.

Chapter 2 is devoted to spiked models with additive Gaussian noise. We determine the statistical threshold where the sharp phase transition happens for a generic spiked model. We discuss the guarantee of the truncate-and-relax algorithm on the planted bisection model (which is a specific instance of spiked models) and make a comparison with a naive sum-of-squares technique. In Chapter 3 we consider the generalization of the stochastic block model to k-uniform hypergraphs. We first determine the statistical threshold where the sharp phase transition happens, and discuss the performance guarantee of the truncate-and-relax algorithm on this type of model. Finally, in Chapter 4 we discuss the statistical limit of generic graphical channel models. We describe the local-to-global amplification phenomenon and prove that such amplification holds under a set of mild conditions on the channel.

1.2 Problem description and examples

Consider a hypergraph $\mathcal{H} = (V, E)$ with $|V| = n$ and $|e| = k$ for all $e \in E$ (k-uniform). We assign the vertex variable $\mathbf{x}_v$ to each vertex $v \in V$ and the edge variable $\mathbf{y}_e$ to each (hyper)edge $e \in E$. Each vertex variable has a value in the input alphabet $\mathcal{X}$ and each edge variable has a value in the output alphabet $\mathcal{Y}$. We assume that $\mathcal{X}$ is finite to make the notion of exact recovery clear (but $\mathcal{Y}$ is arbitrary). Let Q be a noisy channel which gets $(x_1, \cdots, x_k) \in \mathcal{X}^k$ as an input and outputs a random value in $\mathcal{Y}$ according to some distribution only depending on the value of the input.

A graphical channel model is defined as a probabilistic model for $\mathbf{y}$ given $\mathbf{x}$, where $\mathbf{y}_e$ is obtained by independently sending $(\mathbf{x}_{v_1}, \cdots, \mathbf{x}_{v_k})$ through the channel Q, for each $e = \{v_1, \cdots, v_k\} \in E$. Let us write this as $\mathbf{x} \underset{\mathcal{H}}{\overset{Q}{\to}} \mathbf{y}$, or simply $\mathbf{x} \overset{Q}{\to} \mathbf{y}$ if $\mathcal{H}$ is clear from the context. Furthermore, we assume that the input $\mathbf{x}$ is drawn from a prior distribution P. Now we formulate exact recovery as the following:

Definition 1.1 (Exact recovery in graphical channel model). *Suppose that we are given the prior distribution P, the channel Q and the base hypergraph $\mathcal{H}$. Exact recovery in the graphical channel model defined by P, Q and $\mathcal{H}$ is a task of recovering $\mathbf{x}$ given an instance of $\mathbf{y}$. For $\epsilon > 0$, we say that a deterministic algorithm D which maps $\mathbf{y} \in \mathcal{Y}^E$ to $D(\mathbf{y}) \in \mathcal{X}^V$* achieves exact recovery with error probability ϵ *if*

$$\underset{\substack{\mathbf{x} \sim P \\ \mathbf{y} \overset{Q}{\leftarrow} \mathbf{x}}}{\mathbb{P}} (D(\mathbf{y}) = \mathbf{x}) \geq 1 - \epsilon.$$

The *statistical threshold* can be characterized by minimizing the error probability over all possible D. Note here that we did not require D to be an *efficient* algorithm, *i.e.*, running in polynomial time with respect to the size of input. If we restrict D even further to be efficient, then we would get the *computational threshold*. We are interested in characterizing those two types of thresholds and ask whether they match (exactly or asymptotically) or not.

Now we describe two main questions of this thesis:

- Can we characterize the (sharp) threshold for exact recovery in terms of an appropriate notion of channel capacity? More specifically, is the global recovery threshold only dependent on a local property such as information capacity of Q as in Shannon's theorem?

- In many statistical models on graphs (e.g. the stochastic block model with two communities), exact recovery can be directly achieved by a simple algorithm based on a standard semidefinite relaxation technique. Would the same technique work when we consider higher-order models such as when the base hypergraph is k-uniform for some $k \geq 3$?

The answer for the first question is yes. We call such phenomenon *local-to-global amplification* and we discuss it in full generality in Chapter 4. For the second question, we consider an algorithm based on the truncate-and-relax strategy and analyze it on generalizations of the SBM and the spiked Wigner model for k-uniform hypergraphs. We prove that the algorithm successfully recovers the ground truth in an orderwise optimal parameter regime, but it cannot achieve exact recovery all the way down to the statistical threshold. Moreover, for a certain generalization of the spiked Wigner model, we consider an alternative algorithm using the sum-of-squares relaxation techniques and prove that this algorithm is orderwise suboptimal in contrast to the truncate-and-relax algorithm.

In the remainder of this section, we provide several examples of graphical channel models which were investigated in the literature.

1.2.1 Stochastic block models

The stochastic block model (SBM) has been one of the most fruitful research topics in community detection and clustering. The SBM can be thought of as a generalization of the Erdős-Renyí (ER) model $\mathcal{G}(n, p)$, in which we observe a graph G on n vertices where each pair of vertices (i, j) is joined with an edge independently with probability p. While the ER model is deeply understood and has a wide range of theories developed for it, often real-world network behaves very differently from a typical random graph. It leads us to consider alternative models for random graphs.

For exposition, let us start with a simple version of the SBM, where there are two equal sized communities. Let n be an integer greater than 1 and let p and q be real numbers in $[0, 1]$.

Definition 1.2. *The model* $\mathsf{SBM}(n, p, q)$ *generates a random graph* $G = (V, E)$ *on* n *vertices in the following way:*

(i) A label $\mathbf{x}_v \in \{0, 1\}$ *is assigned for each vertex* $v \in V$*. We choose a labeling with equal number of the vertices labeled 0 and labeled 1, uniformly at random.*

(ii) Each pair of vertices (u, v) *are joined with an edge independently with probability* p *if* $\mathbf{x}_u = \mathbf{x}_v$ *(i.e., in the same community) or probability* q *if* $\mathbf{x}_u \neq \mathbf{x}_v$*.*

Sometimes the vertex labeling is chosen in a way such that each vertex receives a label drawn independently and uniformly at random, while in our definition the labels of vertices are not independent due to the fact that we restrict the size of two communities to be equal. We remark that this does not create a big difference and any result in this thesis applies to either definition with a slight modification.

We can easily get the stochastic block model with multiple communities by choosing a label $\mathbf{x}_v$ in the finite set $\mathcal{X}$ with $|\mathcal{X}| \geq 3$.

The SBM is believed to provide good insights in the field of community detection. We can take advantage of the fact that there is a "true" community structure when a graph is sampled, and we can theorize community recovery problems in formal way. Likewise for the ER model, it exhibits many sharp phase transition behaviors [77, 6, 4], and it was studied for whether such statistical thresholds can be achieved by an efficient algorithm [33, 8]. Also, the SBM was used as a testbed for various algorithms. To name a few, spectral algorithms [72, 94], semidefinite programming based algorithms [4, 56, 59], belief-propagation [38, 7, 9], and approximate message-passing algorithms [93, 30, 40, 67] were considered. We recommend [1] for a survey of this topic.

In this thesis, we consider a generalization of the SBMs for hypergraphs. It was first introduced in [49] and was studied in [51, 50, 46, 20, 52, 69, 34, 15]. Specifically, we consider a model for k-uniform hypergraphs, which we call *the stochastic block model for k-uniform hypergraphs* (k-HSBM). Chapter 3 is devoted to characterizing the sharp threshold for exact recovery in k-HSBM with two communities. We also analyze the truncate-and-relax algorithm on k-HSBM.

1.2.2 Models for Principal component analysis

Principal component analysis (PCA) is a powerful method which is widely used in signal processing and other applications. When we are given a data matrix Y, PCA provides a way to extract a signal from Y, which often can be written as a low-rank matrix.

Let us consider the simple situation where Y is an observation of the rank-one signal $\mathbf{x}\mathbf{x}^T$ corrupted by additive Gaussian noise. Let n be an integer greater than 1 (dimension of the signal) and let σ be a positive real number (scaling of the noise).

Definition 1.3. *The* spiked Wigner model (with Rademacher prior) *is a generative model which outputs an $n \times n$ random symmetric matrix Y where*

$$Y = \mathbf{x}\mathbf{x}^T + \sigma W,$$

with a vector $\mathbf{x} \in \{\pm 1\}^n$ which is chosen uniformly at random and an $n \times n$ random symmetric matrix W whose entries are independent and standard Gaussian variables. W is also called Wigner matrix.

There are numerous works on the spiked Wigner model in the random matrix point of view. For instance, the maximum eigenvalue of Y was analyzed in [45] and it was shown that it starts to deviate from the maximum eigenvalue of W when $\frac{\sigma}{\sqrt{n}}$ becomes greater than 1. This implies a sharp phase transition for detection[2], which is the problem of testing whether Y has a spike in it. Such phase transitions for different priors and noise were further studied in [85].

It is natural to investigate a higher-order generalization of the spiked Wigner model. Montanari and Richard proposed a statistical model for tensor PCA [75]

[2]To be precise, this result considers the spiked Wigner model with unit ball prior.

and analyzed the signal recovery problem under both statistical and computational points of view. Later, an approximate message passing (AMP) algorithm for detection was considered in [68], and an algorithm based on sum-of-squares (SoS) relaxation for (almost) exact recovery was considered in [58, 26]. Both of the algorithms are somewhat believed to be unimprovable.

We remark that those results consider a unit ball prior, *i.e.*, $\mathbf{x} \in \mathbb{R}^n$ where $\|\mathbf{x}\|_2 = 1$, hence it requires an extra care for defining exact recovery. Instead, in Chapter 2 we consider the Rademacher prior, *i.e.*, when $\mathbf{x}$ is chosen uniformly at random from the n-dimensional hypercube $\{-1, +1\}^n$.

1.2.3 Other models

It is clear that the stochastic block model, the spiked Wigner model, and their generalization to higher-order relations fit in the category of graphical channel models. We describe a few more examples of such models, which might be of independent interest.

Censored Block Model and other variants

In the binary censored block model (CBM), we observe a random graph G which is drawn from the Erdős-Renyí ensemble $\mathcal{G}(n, p)$ and labels $\mathbf{y}_{ij}$ of edge $ij \in E(G)$. Each edge-label $\mathbf{y}_{ij}$ is a noisy measurement of $\mathbf{1}\{\mathbf{x}_i = \mathbf{x}_j\}$: precisely, we have

$$\mathbf{y}_{ij} = \begin{cases} \mathbf{1}_{\{\mathbf{x}_i = \mathbf{x}_j\}} & \text{with probability } 1 - \theta \\ 1 - \mathbf{1}_{\{\mathbf{x}_i = \mathbf{x}_j\}} & \text{with probability } \theta. \end{cases}$$

In other words, $\mathbf{y}_{ij}$ is the result of sending $\mathbf{1}_{\{\mathbf{x}_i = \mathbf{x}_j\}}$ through the binary symmetric channel with error probability θ.

We note that the CBM can be formulated as a graphical channel model by encoding all randomness into $\mathbf{y}$. Precisely, we assign one of the labels in $\{0, 1, *\}$ to each pair $\{i, j\} \subseteq V$ independently with probability

$$\mathbb{P}(\mathbf{y}_{ij} = *) = 1 - p, \ \mathbb{P}(\mathbf{y}_{ij} = \mathbf{1}_{\{\mathbf{x}_i = \mathbf{x}_j\}}) = p(1-\theta), \ \text{and} \ \mathbb{P}(\mathbf{y}_{ij} = 1 - \mathbf{1}_{\{\mathbf{x}_i = \mathbf{x}_j\}}) = p\theta.$$

Phase transition for exact recovery was considered in [2, 3] and the sharp threshold was shown in [57]. A generalization of the binary CBM for uniform hypergraphs was considered in [14] and they characterize the threshold for exact recovery, and in a subsequent publication they propose an efficient algorithm which provably achieves exact recovery up to statistical threshold [15].

Remark the similarity between the CBM and the stochastic block model. Indeed, there are many other variants of block models such as labeled SBMs and SBMs with overlapping communities, and graphical channel models are powerful enough to express those examples.

Random CSPs with a planted solution

Many problems in computer science have a form of constraint satisfaction: we are given a collection of predicates each defined on a few number of boolean variables, and the goal of the constraint satisfaction problem (CSP) is to decide whether there is an assignment which satisfies all predicates. There are many important examples of CSPs such as k-satisfiability (k-SAT), k-colorability of graphs, unique games and many more.

Phase transition phenomena in various types of random CSPs such as k-SATs were studied in the last decade [47, 10] and planted counterpart of random CSPs were also considered in the literature [24, 43]. We remark that those random CSPs with a planted solution can be modeled in a broader context of graphical channel models (see [5] for more information).

Parity-check Codes

As we discussed previously, graphical channel models are a generic formulation of an encoding scheme for memoryless channels. Shannon's noisy-channel coding theorem tells us that that the rate of such encoding scheme is at most the capacity of the channel which is given by the mutual information between an random input and the output. Moreover, Shannon proves that the maximum rate can be achieved by a random encoding. Subsequently, it was proved in [42] that a random linear encoding achieves the maximum rate.

Low-density parity check (LDPC) and low-density generating matrix (LDGM) codes are a type of sparse linear codes which provides an encoding with good rate which is easy to compute. To understand the phase transition behavior of those codes, the decoding scheme based on maximum *a posteriori* estimation was considered in [74, 63]. Phase transition in the concentration of the mutual information between the message and the noisy codeword was studied in [64, 19] when the code has constant sparsity. In a high-level view, their argument can be also thought as a sort of local-to-global amplification in detection, which is sometimes called the *decoupling principle* [55].

1.3 Local-to-global amplification

Recall the example of the stochastic block model with two equal-sized communities, denoted by $\mathsf{SBM}(n, p, q)$.

For exact recovery, right parameter regime to work on is where p and q scale with $\frac{\log n}{n}$. To see this, we are going to argue that if p and q decrease faster than $\frac{\log n}{n}$ then the probability that the graph has an isolated vertex converges to one as n grows. Note that we cannot hope to recover the label of any isolated vertex with an error probability less than one half.

Let us calculate the probability for the graph having an isolated vertex. Instead of the SBM, let us first consider the Erdős-Renyí model $\mathcal{G}(n, p_n)$ for exposition. Under $\mathcal{G}(n, p_n)$, a graph G is sampled in the way that each pair of vertices is connected with probability p independently. It is clear that when $p = q$, the ER model coincides with $\mathsf{SBM}(n, p, q)$. For $v \in V$, let E_v be the event that v is isolated, *i.e.*, there is no edge incident to it. Then, we get $\mathbb{P}(E_v) = (1 - p_n)^{n-1} \approx e^{-np_n}$ by a direct calculation.

Note that only the dependence of E_u and E_v is caused by the presence of edge joining u and v. We may expect that E_v's are very close to being independent, and so

$$\mathbb{P}\left(\bigcup_v E_v\right) \approx 1 - \prod_v (1 - \mathbb{P}(E_v)) \approx 1 - e^{-ne^{-np_n}}.$$

The right-hand side converges to 1 if ne^{-np_n} diverges and to 0 if ne^{-np_n} vanishes. Indeed, this argument can be made rigorous and we get the following sharp threshold for the property of having an isolated vertex.

Proposition 1.1. *Suppose that $p_n = c\frac{\log n}{n}$ for some constant c. Then, the probability for G having an isolated vertex converges to 1 if $c < 1$ and it converges to 0 if $c > 1$.*

In 1969, Erdős and Renyí showed a stronger result: $c = 1$ is the sharp threshold for the connectedness of G, *i.e.*, G is disconnected with high probability if $c < 1$, and G is connected with high probability if $c > 1$. Clearly, connectedness implies that G has no isolated vertices. We would like to emphasize that connectedness is a *global* property and the property of having no isolated vertices is a local-like property, in the sense that it can be decomposed into almost independent local events E_v.

Returning to exact recovery in $\mathsf{SBM}(n, p, q)$, the sharp threshold lies at $(\sqrt{a} - \sqrt{b})^2 = 2$ when $p = \frac{a \log n}{n}$ and $q = \frac{b \log n}{n}$ for some $a, b > 0$. This was proved in [78] and independently in [4]. Moreover, in [4] it was shown that an algorithm which is based on semidefinite programming with an additional refinement step achieves exact recovery all the way down to the statistical threshold. Subsequently, it was proved in [56] and independently in [21] that the extra refinement step is not needed, hence simple semidefinite program readily achieves exact recovery.

For the SBM with multiple communities, the statistical threshold for exact recovery was established in [6], along with an efficient algorithm which provably achieves the same threshold. Including this result for generic SBMs, many works establish that sharp threshold is determined by the error probability for local recovery. Such phenomenon is stated explicitly in [1] and called *local-to-global amplification.*

Definition 1.4 (informal). *For a vertex v, the local recovery at v is the problem of recovering the label of v when we are given the labels of all vertices except v,* *in addition to the observation of edge-variables. Local-to-global amplification is a phenomenon that (global) exact recovery is approximately equivalent to the product of local recoveries.*

In the case of $\mathsf{SBM}(n, p, q)$, the local recovery at v can be done by counting the number of edges from v to each communities. Suppose $p > q$ and suppose that the community membership of all other vertices are told. If v is connected to larger number of vertices in one community than another, then we would expect that v is also in that community since $p > q$. Indeed, local recovery at v only succeeds when it is connected to more vertices which have the same label as v. Thus the probability for failing local recovery at v is equal to

$$\mathbb{P}\left(\sum_{u:\mathbf{x}_u=\mathbf{x}_v} \mathbf{1}\{uv \in E\} - \sum_{u:\mathbf{x}_u\neq\mathbf{x}_v} \mathbf{1}\{uv \in E\} < 0 \right).$$

Note that each sum is a sum of independent, identically distributed Bernoulli variables, and specifically this probability can be rewritten as $\mathbb{P}(X - Y < 0)$ where X and Y are independent binomial variables with distributions $\mathrm{Bin}(\frac{n}{2} - 1, p)$ and $\mathrm{Bin}(\frac{n}{2}, q)$ respectively.

Lemma 1.2 ([4]). *Assume that $a > b > 0$. Let X_n and Y_n be random variables with binomial distribution $\mathrm{Bin}(\frac{n}{2}, \frac{a \log n}{n})$ and $\mathrm{Bin}(\frac{n}{2}, \frac{b \log n}{n})$ respectively. Then,*

$$\lim_{n\to\infty} -\frac{1}{\log n} \log \mathbb{P}(X_n - Y_n < 0) = \frac{1}{2}(\sqrt{a} - \sqrt{b})^2.$$

The lemma implies that

$$\mathbb{P}(\text{local recovery at } v \text{ fails}) \approx n^{-\frac{1}{2}(\sqrt{a}-\sqrt{b})^2}$$

and by local-to-global amplification, we get

$$
\begin{aligned}
\mathbb{P}(\text{global recovery fails}) &\approx 1-\exp\left(-\sum_{v\in V}\mathbb{P}(\text{local recovery at } v \text{ fails})\right)\\
&\approx 1-\exp\left(-n^{1-\frac{1}{2}(\sqrt{a}-\sqrt{b})^2}\right)\\
&= \begin{cases} 0 & \text{if } (\sqrt{a}-\sqrt{b})^2 > 2\\ 1 & \text{if } (\sqrt{a}-\sqrt{b})^2 < 2,\end{cases}
\end{aligned}
$$

hence the threshold is at $(\sqrt{a}-\sqrt{b})^2 = 2$.

In Section 3.3 and 2.3, We make similar argument for the stochastic block model for k-uniform hypergraphs and spiked Tensor models (which includes the case of spiked Wigner model) to characterize the exact recovery threshold. This is further generalized in Chapter 4 to generic graphical channel models, under some regularity assumptions.

1.4 Truncate-and-relax algorithm

The semidefinite relaxation technique allows us to consider a relaxed, convex problem instead of the original highly non-convex optimization. It was extensively used to find an approximate solution of NP-hard problems such as max-cut, sparsest-cut, min-bisection, graph coloring, and many more [53, 60]. Usually, this type of approximation algorithms consist of two steps, (i) first we find a solution from a relaxed problem, and (ii) since this solution might not be feasible in the original problem, we round up to get a feasible solution.

Standard semidefinite relaxation techniques were recently applied to statistical problems. In many models, such semidefinite relaxation algorithms directly achieve exact recovery down to the statistical threshold without a rounding step, *i.e.*, the solution which is statistically best is also the optimum solution for the relaxed problem. This includes many examples such as SBMs with two communities [56, 21], SBMs with multiple communities [57, 83], SBMs with growing number of symmetric communities [11] and spiked Wigner model with Rademacher prior (sometimes called $\mathbb{Z}_2$-synchronization with Gaussian noises) [21].

Let us conisder the spiked Wigner model for exposition. In the model we observe a symmetric $n \times n$ matrix Y where

$$Y = \mathbf{x}_0 \mathbf{x}_0^T + \sigma W$$

such that $\mathbf{x}_0 \in \{\pm 1\}^n$ is chosen uniformly at random, $\sigma > 0$, and W is a symmetric matrix with independent, standard Gaussian entries.

The maximum likelihood estimator corresponds to the optimal solution of the following optimization problem

$$\max_{\mathbf{x} \in \{\pm 1\}^n} \mathbf{x}^T Y \mathbf{x}.$$

Let L be the function which maps $\mathbf{x}$ to $\mathbf{x}^T Y \mathbf{x}$. We remark that $L(\mathbf{x})$ is a quadratic function on $\{\pm 1\}^n$ and in general the problem of optimizing a polynomial on $\{\pm 1\}^n$ is NP-hard. A standard way to relax such a problem is by rewriting the problem in terms of matrix optimization with rank constraints

$$\begin{aligned}
\text{maximize} \quad & \sum_{i,j} X_{ij} Y_{ij} \\
\text{subject to} \quad & X_{ii} = 1 \text{ for all } i, \\
& X = X^T \text{ is positive semidefinite}, \\
& \operatorname{rank}(X) = 1,
\end{aligned}$$

and relaxing the rank constraint. As a result, we obtain the following semidefinite program:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i,j} X_{ij} Y_{ij} \\
\text{subject to} \quad & X_{ii} = 1 \text{ for all } i, \\
& X = X^T \text{ is positive semidefinite}.
\end{aligned}$$

Semidefinite programs are convex optimization problems in the form of

$$\begin{aligned} &\text{maximize} && \operatorname{tr}(CX) \\ &\text{subject to} && \operatorname{tr}(A_i X) = b_i \text{ for } i = 1, \cdots, m \\ & && X = X^T \text{ is positive semidefinite.} \end{aligned} \tag{1.1}$$

for some symmetric matrices $A_1, \cdots, A_m$ and C and real numbers $b_1, \cdots, b_m$. It has nice properties that the dual, defined as

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{m} b_i y_i \\ &\text{subject to} && y_1, \cdots, y_m \in \mathbb{R} \\ & && \sum_{i=1}^{m} y_i A_i - C \text{ is positive semidefinite,} \end{aligned} \tag{1.2}$$

is also a semidefinite program. Moreover, if the primal and the dual satisfy some mild regularity assumption, then the strong duality holds: the optimum solution X^* of the primal and the optimum solution y^* of the dual exists and their values are identical.

Proposition 1.3 (Complementary slackness). *Suppose that the strong duality holds. Let X^* be an optimum solution of (1.1) and y^* be an optimum solution of (1.2). Let $S^* = \sum_{i=1}^{m} y_i^* A_i - C$. Then, $(S^*)^T X^* = 0$. The converse is also true: if X^* and y^* are feasible and $(S^*)^T X^* = 0$, then X^* and y^* are optimum solutions.*

Suppose that the strong duality holds. To show that $X = xx^T$ is an optimum solution of (1.1), we only need to show that there exists a dual feasible solution y^* such that

$$\left(\sum_{i=1}^{m} y_i^* A_i - C \right) x = 0$$

by complementary slackness. In the case of the spiked Wigner model, we have that the relaxed problem solves exact recovery if there exists a diagonal matrix D such that $(D - Y)x = 0$ and $D - Y$ is positive semidefinite. By analyzing the spectrum of random matrix $D - Y$, we get a threshold for this algorithm which turns out to coincide with the statistical threshold in spiked Wigner model [21].

We would like to generalize this algorithm to higher-order models (which corresponds to the case that base hypergraph is k-uniform for $k \geq 3$ in graphical channel models). Let us consider a version of spiked Wigner model for 4-dimensional tensors: we are given a symmetric tensor $Y \in (\mathbb{R}^n)^{\otimes 4}$ such that

$$Y = \mathbf{x}_0^{\otimes 4} + \sigma W$$

where $\mathbf{x}_0 \in \{\pm 1\}^n$ is chosen uniformly at random, $\sigma > 0$, and $W \in (\mathbb{R}^n)^{\otimes 4}$ is a random symmetric tensor with independent, standard Gaussian entries. Here, we call a tensor Y symmetric if it is invariant under any permutation of the indices, for instance in 4-dimensional case we have

$$Y_{ijk\ell} = Y_{ij\ell k} = Y_{ikj\ell} = Y_{ik\ell j} = \cdots = Y_{\ell kji}.$$

The maximum-likelihood estimator of $\mathbf{x}$ in this model would be

$$\underset{\mathbf{x} \in \{\pm 1\}^n}{\operatorname{argmax}} \sum_{i,j,k,\ell \in [n]} Y_{ijkl} \cdot \mathbf{x}_i \mathbf{x}_j \mathbf{x}_k \mathbf{x}_\ell.$$

We cannot use a naive semidefinite relaxation technique since the objective function is no longer quadratic. Sum-of-squares (SoS) relaxation scheme provides a systematic way to obtain a sequence of relaxations for a generic polynomial optimization problem which are successively refined at the cost of the size of the relaxed problem. In particular, when the domain of the original problem is $\{\pm 1\}^n$, this converges to the original problem at nth level of relaxation (however, this relaxation would be a semidefinite program with exponential size).

In [75], the statistical threshold of the k-tensor PCA model was characterized. They also provided an algorithm which uses a spectral method on a flattening of the data matrix, but it achieves exact recovery only in a suboptimal regime. This was strengthened in [58] and [26] to the SoS relaxation scheme: They proved that the spectral method is as good as any constant level relaxation for the SoS approach.

We consider an alternative model in Chapter 2 where the signal $\mathbf{x}_0^{\otimes 4}$ is replaced by a rank-two tensor $\mathbf{x}_0^{\ominus 4} := \left(\frac{1+\mathbf{x}_0}{2}\right)^{\ominus 4} + \left(\frac{1-\mathbf{x}_0}{2}\right)^{\ominus 4}$. Note that

$$(\mathbf{x}_0^{\ominus 4})_{ijk\ell} = \begin{cases} 1 & \text{if } (\mathbf{x}_0)_i = (\mathbf{x}_0)_j = (\mathbf{x}_0)_k = (\mathbf{x}_0)_\ell \\ 0 & \text{otherwise,} \end{cases}$$

which is motivated by a natural generalization of the SBM to 4-uniform hypergraphs: a 4-HSBM model such that the probability for an edge to appear only depends on whether all vertices have the same label or not.

Again, the maximum-likelihood estimator is a solution for an polynomial optimization problem over $\{\pm 1\}^n$. One difference from the previous tensor model is that the objective function here is not homogeneous and it has a fairly large quadratic part in it. We propose a strategy called "truncate-and-relax" which consists of the following two steps:

- We first truncate the high degree part of the objective polynomial and get a quadratic polynomial as an alternative objective function.
- We further relax this new quadratic optimization problem using a standard semidefinite relaxation technique and solve the relaxed problem.

Surprisingly, this simple algorithm achieves exact recovery in a parameter regime which is orderwise optimal. We prove this in Section 2.4 and prove an analogous result for k-HSBM in Section 3.4. On the other hand, in Section 2.5 we prove that the SoS relaxation of degree 4 on this tensor model (for 4-tensor case) does not achieve this orderwise optimal regime, as in the rank-one signal case. It suggests us that a naive SoS relaxation might not be a "right" way to approach higher-order community recovery problems.

1.5 Frequently used notation

We close this chapter by providing a list of commonly used notations through this thesis.

Notations for the asymptotics of a function

In this thesis, we always consider the sequence of statistical models each of which depends on an integer parameter n. We often ask the behavior of those models in the limit of n growing to infinity. For this reason, we are going to rely on *Bachmann-Landau notations* to describe the asymptotic growth of functions in n.

Let f and g be positive real-valued functions in n. We write

$$\begin{aligned}
f(n) = O(g(n)) &\quad (\text{or } f(n) \lesssim g(n)) \quad \text{if} \quad \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty, \\
f(n) = o(g(n)) &\quad (\text{or } f(n) \ll g(n)) \quad \text{if} \quad \limsup_{n\to\infty} \frac{f(n)}{g(n)} = 0, \\
f(n) = \Omega(g(n)) &\quad (\text{or } f(n) \gtrsim g(n)) \quad \text{if} \quad \liminf_{n\to\infty} \frac{f(n)}{g(n)} > 0, \\
f(n) = \omega(g(n)) &\quad (\text{or } f(n) \gg g(n)) \quad \text{if} \quad \liminf_{n\to\infty} \frac{f(n)}{g(n)} = \infty.
\end{aligned}$$

If $f(n)$ is both $O(g(n))$ and $\Omega(g(n))$, then we write $f(n) = \Theta(g(n))$ or $f(n) \asymp g(n)$. Moreover, we write $f(n) \approx g(n)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$. Note that $f(n) \approx g(n)$ implies that $f(n) \asymp g(n)$ but the converse is not true.

We write $f(n) = \pm o(g(n))$ if $|f(n)| = o(g(n))$. Using this notation, $f(n) \approx g(n)$ is sometimes alternatively denoted $f(n) = (1 \pm o(1))g(n)$.

We often write equations and inequalities with asymptotic notations inside. For instance, $p = n^{1-o(1)}$ means that $p = n^{1-f(n)}$ for some positive function f such that $f(n) = o(1)$, or equivalently,

$$1 - \frac{\log p}{\log n} = o(1).$$

All asymptotic notations without a subscript tacitly means that it is the asymptotics with respect to $n \to \infty$. If necessary, we will specify the variable of our concern as a subscript: For example,

$$f(m,k) = O_m(g(m,k)) \quad \Leftrightarrow \quad \limsup_{m\to\infty} \frac{f(m,k)}{g(m,k)} < \infty \text{ for any fixed } k.$$

Notations for graphs and hypergraphs

Let V be a finite set. We denote the collection of subsets of V of size k by $\binom{V}{k}$, hence $\left|\binom{V}{k}\right| = \binom{|V|}{k}$.

For a positive integer n, we denote $\{1, 2, \cdots, n\}$ by $[n]$. We denote the set of n-tuples of a set S by $S^{[n]}$ or simply S^n. The entries of an n-tuple $s \in S^n$ are denoted by $s_1, \cdots, s_n$. Likewise, we denote the set of tuples of S indexed by the elements in V by S^V, that is,

$$S^V = \{(s_{v_1}, \cdots, s_{v_n}) : s_{v_i} \in S \text{ for } i \in [n]\}$$

where $V = \{v_1, \cdots, v_n\}$.

We denote the symmetric group of degree k by $\mathfrak{S}_k$, which is the group of all permutations on $[k] = \{1, \cdots, k\}$.

A graph is a pair $G = (V(G), E(G))$ which consists of a set $V(G)$ of vertices and a set $E(G)$ of edges, which are elements of $\binom{V}{2}$. An k-uniform hypergraph is a pair $\mathcal{H} = (V(\mathcal{H}), E(\mathcal{H}))$ which consists of a set $V(\mathcal{H})$ of vertices and a set $E(\mathcal{H})$ of *hyperedges*, which are elements of $\binom{V}{k}$. We often use the term "edge" for both an edge in a graph and a hyperedge in a hypergraph, if it is clear from the context.

The adjacency matrix A_G of a graph G is the symmetric $|V(G)| \times |V(G)|$ matrix with entries

$$(A_G)_{uv} = \begin{cases} 1 & \text{if } \{u, v\} \in E(G) \\ 0 & \text{otherwise.} \end{cases}$$

For a k-uniform hypergraph $\mathcal{H}$, we define the incidence vector $A_{\mathcal{H}}$ of $\mathcal{H}$ to be the indicator vector of $E(\mathcal{H}) \subset \binom{V(\mathcal{H})}{k}$, *i.e.*, $A_{\mathcal{H}}$ is an element of $\{0,1\}^{\binom{V(\mathcal{H})}{k}}$ with entries

$$(A_{\mathcal{H}})_e = \begin{cases} 1 & \text{if } e \in E(\mathcal{H}) \\ 0 & \text{otherwise.} \end{cases}$$

Remark that the definition of 2-uniform hypergraphs exactly coincides with the definition of graphs. However, the adjacency matrix and the incidence vector of a graph are not the same object; here we abuse the notation by referring both objects as A_G.

Vectors, matrices, and tensors

We assume readers' familiarity with the basic linear algebra concepts such as vector spaces, bases of a vector space, matrices, eigenvalues and eigenvectors, and so on.

In this thesis, all vector spaces under consideration are finite-dimensional and over the field of real numbers $\mathbb{R}$. In many cases, we consider vectors with real entries indexed by elements in a finite set V. We denote the space of such vectors by $\mathbb{R}^V$ rather than $\mathbb{R}^{|V|}$, to emphasize the indexing. We denote vectors in $\mathbb{R}^V$ by bold-faced lower-case letters such as $\mathbf{x}, \mathbf{y}, \cdots$.

The vector whose entries are all equal to zero is denoted $\mathbf{0}$ or $\mathbf{0}_n$ when we need to specify the dimension. Likewise, the vector whose entries are all equal to one is denoted $\mathbf{1}$ or $\mathbf{1}_n$.

The *restriction* of $\mathbf{x}$ onto a set $S \subseteq V$ of coordinates is

$$\mathbf{x}[S] \in \mathbb{R}^S \quad \text{with entries} \quad \mathbf{x}[S]_v = \mathbf{x}_v \text{ for } v \in S.$$

We denote matrices by plain upper-case letters such as $A, B, \cdots$. Often we regard matrices as elements of $\mathbb{R}^{V_1 \times V_2}$ whose rows are indexed by elements of V_1 and columns are indexed by elements of V_2. We call $A \in \mathbb{R}^{V \times V}$ a $V \times V$ matrix, or a square matrix of size $|V|$. The transpose of a square matrix A is denoted by A^T, and a square matrix A is said to be symmetric if $A = A^T$.

A square matrix A is called a diagonal matrix if its entry A_{uv} is zero whenever $u \neq v$. The identity matrix of size n is denoted Id_n or Id if n is clear from the context. The square matrix of size n whose entries are all equal to zero is denoted by $\mathbf{0}_{n \times n}$ or simply $\mathbf{0}$. The trace of a square matrix A is denoted by $\mathrm{tr}(A)$.

The standard inner product of two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^V$ is

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{v \in V} \mathbf{x}_v \mathbf{y}_v,$$

which is equal to $\mathbf{x}^T\mathbf{y}$ or $\mathbf{y}^T\mathbf{x}$.

The p-norm of a vector $\mathbf{x} \in \mathbb{R}^V$ for $p \in [1, \infty]$ is denoted by $\|\mathbf{x}\|_p$, and defined as

$$\|\mathbf{x}\|_p = \left(\sum_{v \in V} |\mathbf{x}_v|^p\right)^{1/p} \quad \text{for } p \in [1, \infty) \quad \text{and} \quad \|\mathbf{x}\|_\infty = \max_{v \in V} |\mathbf{x}_v|.$$

The 2-norm $\|\cdot\|_2$ is also called the Euclidean norm or just *the norm.* Note that $\|\mathbf{x}\| = \sqrt{\| \langle \mathbf{x}, \mathbf{x} \rangle}$.

The standard inner product of two matrices $A, B \in \mathbb{R}^{V_1 \times V_2}$ is

$$\langle A, B \rangle := \sum_{u \in V_1, v \in V_2} A_{uv} B_{uv}.$$

We note that $\langle A, B \rangle = \operatorname{tr}(A^T B)$. The Frobenius norm of a matrix A, denoted $\|A\|_F$, is defined as $\|A\|_F = \sqrt{\langle A, B \rangle}$.

For a symmetric matrix A, we denote its eigenvalues by

$$\lambda_1(A) \leq \cdots \leq \lambda_n(A).$$

The smallest and the largest eigenvalue are denoted by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ respectively. The spectral norm of a matrix $A \in \mathbb{R}^{V_1 \times V_2}$ is defined as

$$\|A\| := \sqrt{\lambda_{\max}(AA^T)}.$$

If A is a square symmetric matrix of size n, then

$$\|A\| = \max_{i \in n} |\lambda_{\max}(A)|.$$

Let A be a symmetric matrix. A is called positive semidefinite if $\lambda_{min}(A) \geq 0$, or equivalently $\mathbf{x}^T A \mathbf{x} \geq 0$ for any $\mathbf{x}$. A is called positive definite if $\lambda_{min}(A) > 0$, or equivalently $\mathbf{x}^T A \mathbf{x} > 0$ for any nonzero $\mathbf{x}$. We denote $A \succeq 0$ if A is positive semidefinite and $A \succ 0$ if A is positive definite. We denote $A \succeq B$ if $A - B \succeq 0$ and $A \succ B$ if $A - B \succ 0$.

The tensor product of vector spaces $\mathcal{U}_1, \cdots, \mathcal{U}_N$ is denoted by

$$\mathcal{U}_1 \otimes \cdots \otimes \mathcal{U}_N = \bigotimes_{i=1}^{N} \mathcal{U}_i.$$

The tensor product of vectors $u_1 \in \mathcal{U}_1, \cdots, u_N \in \mathcal{U}_N$ is denoted

$$u_1 \otimes \cdots \otimes u_N = \bigotimes_{i=1}^{N} u_i.$$

We call a tensor *pure* if it is a tensor product of vectors.

The k-th tensor power of a vector space $\mathcal{U}$ is defined as the tensor product of k copies of $\mathcal{U}$, and we denoted it by $\mathcal{U}^{\otimes k}$. An element $\mathbf{T}$ of $\mathcal{U}^{\otimes k}$ is called a *k-tensor*. The *k-th tensor power* of a vector $\mathbf{x}$ is the tensor product of k copies of $\mathbf{x}$ and we denote it by $u^{\otimes k}$.

Suppose that $\dim(\mathcal{U}) = n$. Given a basis $\mathcal{B} = \{e^1, \cdots, e^n\}$ of $\mathcal{U}$, the basis of $\mathcal{U}^{\otimes k}$ induced by $\mathcal{B}$ is

$$\mathcal{B}^{\otimes k} := \{e^{v_1} \otimes \cdots \otimes e^{v_k} : (v_1, \cdots, v_k) \in [n]^k\},$$

and each k-tensor $\mathbf{T} \in \mathcal{U}^{\otimes k}$ has the unique expression

$$\mathbf{T} = \sum_{(v_1, \cdots, v_k) \in [n]^k} \mathbf{T}_{v_1, \cdots, v_k} \left(e^{v_1} \otimes \cdots \otimes e^{v_k}\right)$$

with respect to $\mathcal{B}^{\otimes k}$. Note that this gives an isomorphism between $(\mathcal{U})^{\otimes k}$ and $\mathbb{R}^{n^k}$, as $\mathbf{T} \leftrightarrow (\mathbf{T}_{v_1, \cdots, v_k})_{(v_1, \cdots, v_k) \in [n]^k}$.

We often work under the setting that $\mathcal{U} = \mathbb{R}^n$ and $\mathcal{B}$ is the standard basis of $\mathbb{R}^n$. In this case, the corresponding isomorphism between $(\mathbb{R}^n)^{\otimes k}$ and $\mathbb{R}^{[n]^k}$ is trivial, and we refer k-tensor to an element in either of the spaces interchangeably.

We remark that 2-tensors and matrices are equivalent. This equivalence can be seen by associating each pure 2-tensor $u \otimes v$ with a rank-one matrix uv^T. In literature, uv^T (or equivalently $u \otimes v$) is often called the outer product or the exterior product of u and v.

We remark the equivalence between 2-tensors in $(\mathbb{R}^V)^{\otimes 2}$ and matrices in $\mathbb{R}^{V\times V}$. In particular, a pure 2-tensor $\mathbf{x}\otimes\mathbf{y}$ corresponds to the rank-one matrix $\mathbf{x}\mathbf{y}^T$. In general, we call a k-tensor $\mathbf{T}$ rank-one if $\mathbf{T} = \mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_k$ for some vectors $\mathbf{x}_1, \cdots, \mathbf{x}_k \in \mathbb{R}^V$.

Let $\pi \in \mathfrak{S}_k$ be a permutation on $[k]$ and let $\mathbf{T}$ be a k-tensor in $(\mathbb{R}^V)^{\otimes k}$. We define $\mathbf{T}^\pi$ as the k-tensor with entries

$$\mathbf{T}^\pi_{v_1,\cdots,v_k} = \mathbf{T}_{v_{\pi(1)},\cdots,v_{\pi(k)}} \text{ for } (v_1,\cdots,v_k) \in V^k.$$

We say $\mathbf{T}$ is symmetric if $\mathbf{T} = \mathbf{T}^\pi$ for any $\pi \in \mathfrak{S}_k$. Note that his notion agrees with the symmetry of a matrix. The *symmetrization* of a k-tensor $\mathbf{T}$, denoted $\mathsf{Sym}\,\mathbf{T}$, is defined as the average of $\mathbf{T}^\pi$ over $\pi \in \mathfrak{S}_k$, *i.e.*,

$$\mathsf{Sym}\,\mathbf{T} = \frac{1}{k!}\sum_{\pi\in\mathfrak{S}_k} \mathbf{T}^\pi.$$

Let $\mathbf{X}$ be a k_1-tensor in $\mathcal{U}^{\otimes k_1}$ and $\mathbf{Y}$ be a k_2-tensor in $\mathcal{U}^{\otimes k_2}$. The tensor product of $\mathbf{X}$ and $\mathbf{Y}$ is denoted by $\mathbf{X}\otimes\mathbf{Y}$ and defined as the (k_1+k_2)-tensor in $\mathcal{U}^{\otimes(k_1+k_2)}$ with entries

$$(\mathbf{X}\otimes\mathbf{Y})_{u_1,\cdots,u_{k_1},v_1,\cdots,v_{k_2}} = \mathbf{X}_{u_1,\cdots,u_{k-1}}\mathbf{Y}_{v_1,\cdots,v_{k_2}}.$$

The symmetric tensor product of $\mathbf{X}$ and $\mathbf{Y}$, denoted $\mathbf{X}\odot\mathbf{Y}$, is defined as the symmetrization of $\mathbf{X}\otimes\mathbf{Y}$. This notation extends to products with multiple arguments, in particular,

$$\mathbf{x}_1 \odot \cdots \odot \mathbf{x}_k = \mathsf{Sym}\,(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_k).$$

The standard inner product of k-tensors $\mathbf{X}$ and $\mathbf{Y}$ in $(\mathbb{R}^V)^{\otimes k}$ is defined as

$$\langle \mathbf{X}, \mathbf{Y}\rangle := \sum_{(v_1,\cdots,v_k)\in V^k} \mathbf{X}_{v_1,\cdots,v_k}\mathbf{Y}_{v_1,\cdots,v_k}.$$

The *Frobenius norm* of a k-tensor $\mathbf{T}$, denoted $\|\mathbf{T}\|_F$, is defined as $\|\mathbf{T}\|_F = \sqrt{\langle \mathbf{T}, \mathbf{T}\rangle}$.

Probability theory

Let μ be a probability measure on the domain $\mathcal{X}$. We always consider the case that $\mathcal{X}$ is a measurable subset of a real finite-dimensional space; hence the corresponding σ-algebra is implicitly assumed. We write $X \sim \mu$ if X is the random variable such that $\mathbb{P}(X \in A) = \mu(A)$ for any (measurable) $A \subseteq \Omega$. We write

$$\mathbb{P}_{X\sim\mu} \quad (\text{or } \mathop{\mathbb{P}}_{X\sim\mu} \text{ for separate-line formulas})$$

to emphasize that the probability is taken over a random draw of X from the distribution μ. If the variable of consideration is clear from the context, we simply write $\mathbb{P}_\mu$. We denote the corresponding expectation operator by $\mathbb{E}_{X\sim\mu}$ or $\mathbb{E}_\mu$.

We denote the integral of a measurable function $f : \Omega \to \mathbb{R}$ with respect to μ by

$$\int f \, d\mu = \int_\Omega f(\mathbf{x}) \, d\mu(\mathbf{x}).$$

If μ has the density function $p(\mathbf{x})$ with respect to a reference measure λ, then we have

$$\int f \, d\mu = \int f(\mathbf{x}) p(\mathbf{x}) \, d\lambda(\mathbf{x}).$$

When the reference measure is clear from the context, we sometimes write $X \sim p$ to denote $X \sim \mu$. The density function p of μ is also denoted by $\frac{d\mu}{d\lambda}$ which is also known as the Radon-Nikodym derivative of μ with respect to λ.

Chapter 2

Spiked Tensor Models[1]

Principal component analysis (PCA) is a powerful method for recovering a signal from a noisy observation when the signal has special properties such as being a low-rank or a sparse matrix. We often achieve recovery by a simple spectral algorithm, which executes a diagonalization (or more generally, a singular-value decomposition) and estimates the original signal within the space spanned by the eigenvectors of a few largest eigenvalues. One of the simplest models for PCA is the spiked Wigner model in which we observe a random $n \times n$ data matrix Y where

$$Y = \beta \mathbf{v}\mathbf{v}^T + W,$$

such that $\mathbf{v}$ is chosen uniformly at random from the unit sphere in $\mathbb{R}^n$, $\beta > 0$ (signal-to-noise ratio), and W is a random symmetric matrix with independent, standard Gaussian entries.

In many applications, we often observe the data with the elements indicating the interactions between three or more items. Examples include the problems such as image segmentation [54], community recovery in hypergraphs [95, 17], topic modeling [18], hypergraph matching [41] and tensor learning in general [12]. This motivates us to consider the generalization of PCA to higher-order tensors. However, decomposing

[1]The contents of this chapter overlaps in significant amount with [61], which is a collaborative work mainly conducted by the author of this thesis.

a given tensor into rank-one tensors or even finding the best rank-one approximation of the given tensor is hard, in contrast to the case of matrices.

Montanari and Richard [75] propose a simple model for tensor PCA which generalizes the spiked Wigner model. In this model, we observe a random k-tensor $\mathbf{Y} \in (\mathbb{R}^n)^{\otimes k}$ where

$$\mathbf{Y} = \beta \mathbf{v}^{\otimes k} + \mathbf{W},$$

such that $\mathbf{v}$ is a random unit vector in $\mathbb{R}^n$, $\beta > 0$, and $\mathbf{W}$ is a symmetric k-tensor with independent, standard Gaussian entries. The authors prove that there exist constants c_1 and c_2 (which may depend on k but independent of n) such that

(i) if $\beta \geq c_1 \epsilon^{-1} \sqrt{n}$, then the maximum-likelihood estimator $\widehat{\mathbf{v}}_{MLE}$ achieves

$$||\widehat{\mathbf{v}}_{MLE} - \mathbf{v}||_2^2 \leq \epsilon$$

with high probability, and

(ii) if $\beta \leq c_2 \sqrt{n}$, then for any estimator $\widehat{\mathbf{v}}$, the distance from $\widehat{\mathbf{v}}$ to $\mathbf{v}$ (or to $\{\pm \mathbf{v}\}$ when k is even) is bounded away from zero in expectation.

Those thresholds are sharpened further for various recovery requirements and other prior distributions in [84].

On the other hand, in [75] the authors also ask whether efficient recovery is possible. They consider a simple algorithm which performs spectral clustering on the unfolding of the tensor. They prove that this algorithm returns $\widehat{\mathbf{v}}_{alg}$ which is close to $\mathbf{v}$ (or $\pm \mathbf{v}$), as long as $\beta \gtrsim n^{\lceil k/2 \rceil/2}$. It is proved in [58] that the sum-of-squares relaxation of degree $2\lceil k/2 \rceil$ can find a good solution as long as $\beta \gtrsim n^{k/4}$ which is better than $n^{\lceil k/2 \rceil/2}$ when k is odd. Subsequently, it is proved in [26] that sum-of-squares technique with higher-degree would not gain much: Essentially, we need $\beta \gtrsim n^{k/4}$ for the sum-of-squares relaxation of any constant degree to find a good solution. This is somewhat believed to be unimprovable, in analogy to the case of weak recovery where such a gap is present with respect to approximate message passing algorithms (see [68]).

We remark that in those statistical models for PCA or tensor PCA, one can only formulate the recovery requirement in terms of the distance between the ground truth $\mathbf{v}$ and the estimator $\widehat{\mathbf{v}}$, as $\mathbf{v}$ is chosen from a continuous prior.

In this chapter, we consider the variants where the spike is chosen from a discrete set and investigate *exact recovery* problem in those models. For instance, the spiked Wigner model with Rademacher prior is defined as follows.

Definition 2.1. *The* spiked Wigner model with Rademacher prior *is a model in which we observe a random symmetric data matrix Y of size n where*

$$Y = \mathbf{x}_0\mathbf{x}_0^T + \sigma W$$

for a randomly chosen vector $\mathbf{x}_0 \in \{\pm 1\}^n$, the noise parameter $\sigma > 0$, and a random symmetric matrix W with independent, standard Gaussian entries.

We remark that this model is sometimes called $\mathbb{Z}_2$-synchronization model with Gaussian noise [21, 59]. Let us consider the following generalization of the spiked Wigner model to k-tensors.

Definition 2.2. *The* single-spiked k-tensor model with Rademacher prior *(or simply the single-spiked model) is the generative model such that a random k-tensor $\mathbf{Y}$ in $(\mathbb{R}^n)^{\otimes k}$ is generated in the way that*

$$\mathbf{Y} = \mathbf{x}_0^{\otimes k} + \sigma\mathbf{W},$$

where $\mathbf{x}_0 \in \{\pm 1\}^n$ is chosen uniformly at random and $\mathbf{W}$ is the symmetrization of the k-tensor $\mathbf{G}$ whose entries are independent, standard Gaussian variables.

Note that the spike $\mathbf{x}_0^{\otimes k}$ in this model has the entries of the form

$$(\mathbf{x}_0^{\otimes k})_{v_1,\cdots,v_k} = (\mathbf{x}_0)_{v_1}\cdots(\mathbf{x}_0)_{v_k}.$$

In other words, each entry of the spike represents the "parity" of the hyperedge $\{v_1, \cdots, v_k\}$ (assuming that $v_1, \cdots, v_k$ are distinct) when we regard $(\mathbf{x}_0)_v \in \{-1, 1\}$

as the community label of the vertex v.

On the other hand, let us consider an alternative way to label the hyperedge $\{v_1, \cdots, v_k\}$ by whether the vertices lie in the same community or not. Let $\mathbf{x}_0^{\ominus k}$ be the k-tensor with entries

$$(\mathbf{x}_0^{\ominus k})_{v_1, \cdots, v_k} = \begin{cases} 1 & \text{if } (\mathbf{x}_0)_{v_1} = \cdots = (\mathbf{x}_0)_{v_k} \\ 0 & \text{otherwise,} \end{cases}$$

and let us define the corresponding spiked k-tensor model.

Definition 2.3. *The* bisection-spiked k-tensor model with Rademacher prior *(or simply the bisection-spiked model) is the generative model such that a random k-tensor $\mathbf{Y}$ in $(\mathbb{R}^n)^{\otimes k}$ is generated in the way that*

$$\mathbf{Y} = \mathbf{x}_0^{\ominus k} + \sigma \mathbf{W},$$

where $\mathbf{x}_0 \in \{\pm 1\}^n$ is chosen uniformly at random and $\mathbf{W}$ is the symmetrization (as defined in Section 1.5) of the k-tensor $\mathbf{G}$ whose entries are independent, standard Gaussian variables.

We investigate the exact recovery problem in a class of models that includes the single-spiked model and the bisection-spiked model. Specifically, we consider the models where the spike can be expressed as a k-tensor $s(\mathbf{x}_0)$ with entries

$$s(\mathbf{x}_0)_{v_1, \cdots, v_k} = s((\mathbf{x}_0)_{v_1}, \cdots, (\mathbf{x}_0)_{v_k})$$

for some symmetric function $s : \{\pm 1\}^k \to \mathbb{R}$.

The outline of this chapter is as follows. We briefly discuss the Fourier analysis of the functions defined on a hypercube, provide the precise definition of the model, and summarize the main results of this Chapter in Section 2.1. The main results of this Chapter are summarized in Section 2.2. Section 2.3 is devoted to characterizing the statistical threshold for exact recovery. We consider the truncate-and-relax algorithm and analyze its performance in Section 2.4. On the other hand, in Section

2.5 we consider an algorithm based on the sum-of-squares relaxation technique and argue that it is suboptimal in comparison to the truncate-and-relax algorithm in the bisection-spiked model.

2.1 Preliminaries

2.1.1 Notations

For simplicity, let us introduce a few notations which will be used throughout this chapter. We will use $\alpha, \beta, \cdots$ to denote tuples in $[n]^k$. For legibility, we denote the entries of the tuple $\alpha \in [n]^k$ by $\alpha(1), \cdots, \alpha(k)$ instead of $\alpha_1, \cdots, \alpha_k$. For $I \subseteq [k]$, the restriction of α on I is denoted by $\alpha(I)$, *i.e.*,

$$\alpha(I) = (\alpha(i_1), \cdots, \alpha(i_{|I|})) \quad \text{where } I = \{i_1 < \cdots < i_{|I|}\}.$$

We often regard α as a function from $[k]$ to $[n]$. In particular, we use the notation $\alpha^{-1}(S)$ to denote

$$\alpha^{-1}(S) = \{i \in [k] : \alpha(i) \in S\}$$

where $S \subseteq [n]$. When $S = \{v\}$, then we simply write $\alpha^{-1}(v)$ instead of $\alpha^{-1}(\{v\})$.

For a vector $\mathbf{x}$ in $\{\pm 1\}^n$, we use $\mathbf{x}_\alpha$ to denote $\prod_{i \in [k]} \mathbf{x}_{\alpha(i)}$. Moreover, we denote $\prod_{i \in I} \mathbf{x}_{\alpha(i)}$ by $\mathbf{x}_{\alpha(I)}$ for $I \subseteq [k]$.

2.1.2 Fourier analysis on the hypercube

In this subsection, we briefly introduce the Fourier analysis of the real-valued functions on the hypercube $\{\pm 1\}^m$.

Let f be a real-valued function on the m-dimensional hypercube $\{\pm 1\}^m$. One way to represent this function is that we specify all evaluations of f at each point in $\{\pm 1\}^m$, *i.e.*,

$$f = \sum_{x \in \{\pm 1\}^m} f(\mathbf{x}) \mathbf{1}_x,$$

where $\mathbf{1}_x$ is a function which has value 1 at x and 0 elsewhere. This is because $\{\mathbf{1}_x\}_{x\in\{\pm 1\}^n}$ forms a basis of the space of real-valued functions on $\{\pm 1\}^m$, or equivalently $\mathbb{R}^{\{\pm 1\}^n}$.

On the other hand, let us consider the family of functions $\{\chi_S\}_{S\subseteq[m]}$ where

$$\chi_S(y) = \prod_{i\in S} y_i.$$

This family forms an orthonormal basis of the space of functions $\{f : \{\pm 1\}^m \to \mathbb{R}\}$ with respect to the inner product

$$\langle\cdot,\cdot\rangle : (f,g) \mapsto \frac{1}{2^m} \sum_{x\in\{\pm 1\}^m} f(\mathbf{x})g(\mathbf{x}).$$

Hence, any real-valued function f on $\{\pm 1\}^m$ can be uniquely written as

$$f = \sum_{S\subseteq[m]} \widehat{f}(S)\chi_S,$$

where $\widehat{f}(S) := \langle f, \chi_S\rangle$.

The quantities $\widehat{f}(S)$ are called the Fourier coefficients of f. The degree of f is defined as the maximum size of S such that $\widehat{f}(S) \neq 0$, and we denote it by $\deg(f)$. For $d \leq \deg(f)$, we define the truncation of f to degree d as the function $f_{\leq d}$ on $\{\pm 1\}^k$ with the Fourier expansion

$$f_{\leq d} = \sum_{I\subseteq[k]:|I|\leq d} \widehat{f}(I)\chi_I.$$

For $S \subseteq [k]$, we define the restriction of f on the index set S as the function $f|_S(z)$ on $\{\pm 1\}^S$ with the Fourier expansion

$$f|_S(z) = \sum_{I\subseteq S} \widehat{f}(I)\chi_I.$$

If f is symmetric, then $\widehat{f}(S) = \widehat{f}(T)$ for any S and T with the same size. In such cases, we denote $\widehat{f}(S)$ by $\widehat{f}(|S|)$, hence we have

$$f = \sum_{r=0}^{m} \widehat{f}(r) \left(\sum_{S:|S|=r} \chi_S \right).$$

Let $\mathbf{x}_1, \cdots, \mathbf{x}_k \in \{\pm 1\}^V$ and let $f : \{\pm 1\}^k \to \mathbb{R}$. We define $f(\mathbf{x}_1, \cdots, \mathbf{x}_k)$ to be a k-tensor in $(\mathbb{R}^V)^{\otimes k}$ with entries

$$f(\mathbf{x}_1, \cdots, \mathbf{x}_k)_{v_1, \cdots, v_k} = f((\mathbf{x}_1)_{v_1}, \cdots, (\mathbf{x}_k)_{v_k}).$$

If $\mathbf{x}_i = \mathbf{x}$ for all $i \in [k]$, then we simply write $f(\mathbf{x})$ instead of $f(\mathbf{x}, \cdots, \mathbf{x})$.

Since $f \mapsto f(\mathbf{x}_1, \cdots, \mathbf{x}_k)$ is linear, we can express $f(\mathbf{x}_1, \cdots, \mathbf{x}_k)$ as

$$f(\mathbf{x}_1, \cdots, \mathbf{x}_k) = \sum_{I \subseteq [k]} \widehat{f}(I)\, \chi_I\left(\mathbf{x}_1, \cdots, \mathbf{x}_k\right).$$

Note that

$$\chi_I\left(\mathbf{x}_1, \cdots, \mathbf{x}_k\right) = \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k$$

where $\mathbf{v}_i = \mathbf{x}_i$ if $i \in I$ and $\mathbf{v}_i = \mathbf{1}$ otherwise. For brevity, we denote $\chi_I\left(\mathbf{x}_1, \cdots, \mathbf{x}_k\right)$ by $(\mathbf{x}_1, \cdots, \mathbf{x}_k)^I$ and $\chi_I(\mathbf{x}, \cdots, \mathbf{x})$ by $\mathbf{x}^I$.

If f is symmetric, then $f(\mathbf{x})$ can be expressed as

$$\begin{aligned} f(\mathbf{x}) &= \sum_{r=0}^{k} \widehat{f}(r) \sum_{I \subseteq [k]:|I|=r} \mathbf{x}^I \\ &= \sum_{r=0}^{k} \binom{k}{r} \widehat{f}(r) \left(\mathbf{x}^{\otimes r} \odot \mathbf{x}^{\otimes (k-r)}\right). \end{aligned}$$

We recall that $\odot$ denotes the symmetric product (see Section 1.5).

Proposition 2.1. *Let $\mathbf{x}, \mathbf{y} \in \{\pm 1\}^V$ such that $\langle \mathbf{x}, \mathbf{1} \rangle = \langle \mathbf{y}, \mathbf{1} \rangle = 0$. Let s and t be real-valued functions on $\{\pm 1\}^k$. Then,*

$$\langle s(\mathbf{x}), t(\mathbf{y}) \rangle = n^k \sum_{I \subseteq [k]} \widehat{s}(I) \widehat{t}(I) \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{n} \right)^{|I|}.$$

Proof. By definition,

$$\langle s(\mathbf{x}), t(\mathbf{y}) \rangle = \sum_{I,J \subseteq [k]} \widehat{s}(I)\widehat{t}(J) \left\langle \mathbf{x}^I, \mathbf{y}^J \right\rangle .$$

Since $\langle \mathbf{x}, \mathbf{1} \rangle$ and $\langle \mathbf{y}, \mathbf{1} \rangle$ are equal to zero, $\left\langle \mathbf{x}^I, \mathbf{y}^J \right\rangle$ is nonzero only if $I = J$ and $\left\langle \mathbf{x}^I, \mathbf{y}^I \right\rangle = n^{k-|I|} \langle \mathbf{x}, \mathbf{y} \rangle^{|I|}$. Thus we get the desired result. $\square$

In particular, if s and t are symmetric we get

$$\langle s(\mathbf{x}), t(\mathbf{y}) \rangle = n^k \sum_{r=0}^{k} \binom{k}{r} \widehat{s}(r)\widehat{t}(r) \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{n} \right)^r .$$

2.1.3 Description of the model

Let us formally describe the spiked k-tensor model which we consider in this chapter. Let n and k be integers such that $n \geq k \geq 2$. Let s be a symmetric function on $\{\pm 1\}^k$ which is not identical to the zero function and let $\sigma = \sigma(n) > 0$.

Definition 2.4. *The* s-spiked k-tensor model *is a generative random k-tensor model such that $\mathbf{Y} \in (\mathbb{R}^n)^{\otimes k}$ is generative in the way that*

$$\mathbf{Y} = s(\mathbf{x}_0) + \sigma \mathbf{W}$$

where $\mathbf{x}_0$ is randomly chosen from $\{\pm 1\}^n$ and $\mathbf{W}$ is the symmetrization $\mathsf{Sym}\,\mathbf{G}$ *of a random k-tensor $\mathbf{G}$ with independent, standard Gaussian entries.*

Here $\mathbf{G}$ is not a symmetric tensor. We chose to define $\mathbf{W}$ as the symmetrization of $\mathbf{G}$ for a simpler analysis. One may note that the entries of $\mathbf{W}$ do not have the same variance. For instance, we have

$$\mathbf{W}_{1,\cdots,1} = \mathbf{G}_{1,\cdots,1} \sim N(0,1) \quad \text{but} \quad \mathbf{W}_{1,\cdots,k} = \frac{1}{k!} \sum_{\pi \in \mathfrak{S}_k} \mathbf{W}_{\pi(1),\cdots,\pi(k)} \sim N(0, 1/k!).$$

We remark that $1 - o(1)$ fraction of α's in $[n]^k$ has distinct entries $\alpha(1), \cdots, \alpha(k)$. Hence, "typical" entries of $\mathbf{W}$ is distributed as $N(0, 1/k!)$ and the effect of other entries is negligible.

We remark that the single-spiked model and the bisection-spiked model are equivalent to the s_1-spiked model and the s_2-spiked model respectively, where

$$s_1(z) = z_1 \cdots z_k \quad \text{and} \quad s_2(z) = \frac{1}{2^k}\left(\prod_{i=1}^{k}(1+z_i) + \prod_{i=1}^{k}(1-z_i)\right).$$

2.2 Main results

Exact recovery in the s-spiked k-tensor model is the problem of recovering $\mathbf{x}_0$ from an observation of

$$\mathbf{Y} = s(\mathbf{x}_0) + \sigma \mathbf{W}.$$

We note that if $s(z) = s(-z)$, then $s(\mathbf{x}) = s(-\mathbf{x})$ for any $\mathbf{x}$, so we can only hope to recover the ground truth $\mathbf{x}_0$ up to a global sign flip.

Definition 2.5. *We say that exact recovery is achievable if there exists an estimator $\widehat{\mathbf{x}}$ such that*

$$\mathbb{P}(\widehat{\mathbf{x}} = \mathbf{x}_0) = 1 - o(1)$$

if $s(z) \neq s(-z)$ for some $z \in \{\pm 1\}^k$, or

$$\mathbb{P}(\widehat{\mathbf{x}} \in \{\mathbf{x}_0, -\mathbf{x}_0\}) = 1 - o(1)$$

if $s(z) = s(-z)$ for any $z \in \{\pm 1\}^k$. We say that exact recovery is impossible if for any estimator fails to recover $\mathbf{x}$ (or up to a global sign flip when $s(z) = s(-z)$ for all z) with probability $1 - o(1)$.

From now on, we restrict our focus to the s-spiked k-tensor model where **the prior $\mathbf{x}_0$ is chosen uniformly at random among the vectors $\mathbf{x}_0 \in \{\pm 1\}^n$ satisfying** $\mathbf{1}^T \mathbf{x}_0 = 0$, as opposed to the uniform prior on $\{\pm 1\}^n$. However, we remark that the proof technique for the balanced prior easily translates to the case of uniform prior. We also remark that such results can be explicitly obtained by applying the more general result which appears in Chapter 4.

Regarding the statistical threshold of exact recovery, we get the following result.

Theorem 2.2. *Suppose* $\deg(s) \geq 2$. *Let* $\phi_s(t)$ *be*

$$\phi_s(t) = \sum_{r=0}^{k} \binom{k}{r} \widehat{s}(r)^2 t^r,$$

and let σ_s^* *be the positive real number satisfying*

$$(\sigma_s^*)^2 = \phi_s'(1) \cdot \frac{n^{k-1}}{2 \log n}.$$

Then, exact recovery is achievable if $\sigma < (1-\epsilon)\sigma_s^*$ *for some* $\epsilon > 0$, *and exact recovery is impossible if* $\sigma > (1+\epsilon)\sigma_s^*$ *for some* $\epsilon > 0$.

We prove this theorem by analyzing the maximum-likelihood estimator $\widehat{\mathbf{x}}_{ML}$, which can be described as the optimal solution for the following maximization problem:

$$\max_{\mathbf{x} \in \{\pm 1\}^n : \mathbf{1}^T \mathbf{x} = 0} \langle \mathbf{Y}, s(\mathbf{x}) \rangle .$$

We analyze the probability that $\mathbf{x}_0$ is not the unique optimum of the function $\mathbf{x} \mapsto \langle \mathbf{Y}, s(\mathbf{x}) \rangle$, which is equal to

$$p_{ML,fail} := \mathbb{P}\left(\bigcup_{\mathbf{x}: s(\mathbf{x}) \neq s(\mathbf{x}_0)} \{ \langle \mathbf{Y}, s(\mathbf{x}) \rangle \geq \langle \mathbf{Y}, s(\mathbf{x}_0) \rangle \} \right).$$

The proof can be found in Section 2.3.

On the other hand, we consider an algorithm which is based on the truncate-and-relax strategy which we have discussed briefly in Section 1.4.

Let us consider the truncation $s_{\leq d}$ of s to degree d. Let $\widehat{\mathbf{x}}^{(d)}_{trunc}$ be the estimator defined as

$$\widehat{\mathbf{x}}^{(d)}_{trunc} := \operatorname*{argmax}_{\mathbf{x} \in \{\pm 1\}^n : \mathbf{1}^T \mathbf{x} = 0} \langle \mathbf{Y}, s_{\leq d}(\mathbf{x}) \rangle ,$$

where ties are broken arbitrarily. We are particularly interested in $\widehat{\mathbf{x}}^{(2)}_{trunc}$, which is the maximizer of the function

$$\begin{aligned}\langle \mathbf{Y}, s_{\leq 2}(\mathbf{x})\rangle &= \langle s(\mathbf{x}_0) + \sigma \mathbf{W}, s_{\leq 2}(\mathbf{x})\rangle \\ &= \langle s(\mathbf{x}_0), s_{\leq 2}(\mathbf{x})\rangle + \sigma \langle \mathbf{W}, s_{\leq 2}(\mathbf{x})\rangle .\end{aligned}$$

Since the entries in $s_{\leq 2}(\mathbf{x})$ are quadratic polynomials in $\mathbf{x}$, we can write $\langle \mathbf{Y}, s_{\leq 2}(\mathbf{x})\rangle$ as

$$\langle \mathbf{Y}, s_{\leq 2}(\mathbf{x})\rangle = \begin{bmatrix} 1 & \mathbf{x}^T \end{bmatrix} Y \begin{bmatrix} 1 \\ \mathbf{x} \end{bmatrix}$$

for some symmetric matrix Y of size $(n+1)$, whose rows and columns are indexed by $0, 1, \cdots, n$. Here we may think $\mathbb{R}^{[n]}$ as the subspace of $\mathbb{R}^{\{0\}\cup[n]}$ which consists of the vectors whose entry at 0th index is zero.

We consider the standard semidefinite relaxation of the optimization problem

$$\max_{\mathbf{x}\in\{\pm 1\}^n : \mathbf{1}^T\mathbf{x}=0} \begin{bmatrix} 1 & \mathbf{x}^T \end{bmatrix} Y \begin{bmatrix} 1 \\ \mathbf{x} \end{bmatrix},$$

that is,

$$\begin{aligned} \max \quad & \langle Y, X\rangle \\ \text{subject to} \quad & X_{ii} = 1 \text{ for } i \in \{0\} \cup [n] \\ & \langle X, J\rangle = 0 \text{ where } J = \begin{bmatrix} 0 & \mathbf{0}_n^T \\ \mathbf{0}_n & \mathbf{1}_n \mathbf{1}_n^T \end{bmatrix} \\ & X \succeq 0,\ X = X^T \in \mathbb{R}^{(\{0\}\cup[n])\times(\{0\}\cup[n])}. \end{aligned}$$

The truncate-and-relax algorithm solves the relaxation and outputs an optimum solution X^*. We say X^* *recovers* $\mathbf{x}_0$ if

$$X^* = \begin{bmatrix} 1 & \mathbf{x}_0^T \\ \mathbf{x}_0 & \mathbf{x}_0\mathbf{x}_0^T \end{bmatrix},$$

and we say X^* *recovers* $\mathbf{x}_0$ *up to a global sign flip* if

$$X^* = \begin{bmatrix} 1 & * \\ * & \mathbf{x}_0\mathbf{x}_0^T \end{bmatrix}.$$

We require the algorithm to recover $\mathbf{x}_0$ when $s_{\leq 2}$ is not even, or to recover $\mathbf{x}_0$ up to a global sign flip when $s_{\leq 2}$ is even.

Theorem 2.3. *Suppose that $\widehat{s}(2) \neq 0$. Let $\phi_{s,trunc}(t)$ be*

$$\phi_{s,trunc}(t) := \phi_{s_{\leq 2}}(t) = \sum_{r=0}^{2} \binom{k}{r} \widehat{s}(r)^2 t^r,$$

and let $\sigma^*_{s,trunc}$ be

$$\sigma^*_{s,trunc} = \sqrt{\phi'_{s,trunc}(1) \cdot \frac{n^{k-1}}{2\log n}}.$$

The truncate-and-relax algorithm achieves exact recovery if $\sigma > (1+\epsilon)\sigma^*_{s,trunc}$ for some $\epsilon > 0$. Moreover, this analysis is tight: If $\sigma < (1-\epsilon)\sigma^*_{s,trunc}$ for some $\epsilon > 0$, then the truncate-and-relax algorithm fails to recover $\mathbf{x}_0$ with probability $1 - o(1)$.

The proof can be found in Section 2.4. We remark that our analysis is only about the probability for the truncate-and-relax algorithm *directly* recovering $\mathbf{x}_0$. Hence, there is a possibility for algorithms with an additional rounding step achieving the statistical threshold, but it is out of scope of this thesis.

2.2.1 Single-spiked model vs Bisection-spiked model

Recall that the single-spiked model and the bisection-spiked model are instances of s-spiked model, with $s = s_1$ and $s = s_2$ respectively where

$$\begin{aligned} s_1(z) &= z_1 \cdots z_k = z_{[k]} \\ s_2(z) &= \frac{1}{2^k}\left(\prod_{i=1}^{k}(1+z_i) + \prod_{i=1}^{k}(1-z_i)\right) = \frac{1}{2^{k-1}} \sum_{\substack{I \subseteq [k] \\ |I| \text{ even}}} z_I. \end{aligned}$$

The Fourier coefficients of s_1 are

$$\widehat{s_1}(r) = \begin{cases} 1 & \text{if } r = k \\ 0 & \text{otherwise,} \end{cases}$$

and so $\phi_{s_1}(t) = t^k$.

Corollary 2.4 (Single-spiked model). *The statistical threshold for exact recovery in the single-spiked model is*

$$\sigma^*_{s_1} = \sqrt{k} \cdot \frac{n^{\frac{k-1}{2}}}{\sqrt{2\log n}}.$$

Remark that the truncate-and-relax algorithm cannot be used for the single-spiked model since $\widehat{s_1}(2) = 0$. Instead, we consider the *sum-of-squares (SoS) algorithm* (see Section 2.5 for the definition).

Theorem 2.5. *The SoS algorithm achieves exact recovery in the single-spiked* 4*-tensor model if* $\sigma \lesssim \frac{n}{\mathsf{polylog}(n)}$. *On the other hand, the SoS algorithm fails to recover* $\mathbf{x}_0$ *in the single-spiked* 4*-tensor model if* $\sigma \gtrsim n \cdot \mathsf{polylog}(n)$.

The Fourier coefficients of s_2 are

$$\widehat{s_2}(r) = \begin{cases} \frac{1}{2^{k-1}} & \text{if } r \text{ is even} \\ 0 & \text{otherwise.} \end{cases}$$

We have

$$\begin{aligned} \phi_{s_2}(t) &= \frac{1}{2^{2k-1}}\left((1-t)^k + (1+t)^k\right), \text{ and} \\ \phi_{s_2,trunc}(t) &= 1 + \frac{1}{2^{2k-2}}\binom{k}{2} t^2. \end{aligned}$$

Corollary 2.6 (Bisection-spiked model). *The statistical threshold for exact recovery in the bisection-spiked model is*

$$\sigma^*_{s_2} = \sqrt{\frac{k}{2^k}} \cdot \frac{n^{\frac{k-1}{2}}}{\sqrt{2\log n}}.$$

Moreover, the threshold for the truncate-and-relax algorithm achieving exact recovery is

$$\sigma^*_{s_2,trunc} = \sqrt{\frac{k(k-1)}{2^{2k-2}}} \cdot \frac{n^{\frac{k-1}{2}}}{\sqrt{2\log n}}.$$

For comparison with the single-spiked model, we also consider the sum-of-squares algorithm on the bisection-spiked model.

Theorem 2.7. *The sum-of-squares algorithm fails to recover $\mathbf{x}_0$ in the bisection-spiked 4-tensor model if $\sigma \gtrsim n \cdot \mathsf{polylog}(n)$.*

Informally speaking, it suggests us that (at least in the bisection-spiked model) it is better to "forget" higher moments of the data if we use semidefinite programming techniques. We note that Lesieur et al. [68] observed a similar phenomenon for the detection problem and the approximate message passing algorithm.

2.3 Statistical threshold: Proof of Theorem 2.2

Let $\mathbf{x}_0$ be a vector in $\{\pm 1\}^n$ such that $\mathbf{1}^T\mathbf{x}_0 = 0$. For each subset S of $[n]$, let $\mathbf{x}^{(S)}$ be the vector obtained by flipping the sign of $(\mathbf{x}_0)_v$ for the indices $v \in S$, *i.e.*,

$$\mathbf{x}^{(S)}_v = \begin{cases} (\mathbf{x}_0)_v & \text{if } v \notin S \\ -(\mathbf{x}_0)_v & \text{if } v \in S. \end{cases}$$

Here $\mathbf{x}^{(S)}$ is balanced if $\sum_{v\in S}(\mathbf{x}_0)_v = 0$. For simplicity, let us call such S also *balanced* (with respect to $\mathbf{x}_0$).

Note that $p_{ML,fail}$ is the probability that $\mathbf{x}^{(S)}$ outperforms $\mathbf{x}_0$ for some balanced S, *i.e.*,

$$p_{ML,fail} = \mathbb{P}\left(\bigcup_{\substack{S:\ S\neq\emptyset \\ S \text{ balanced}}} E_S \right)$$

where E_S is the event that $\langle \mathbf{Y}, s(\mathbf{x}^{(S)})\rangle \geq \langle \mathbf{Y}, s(\mathbf{x}_0)\rangle$ holds.

Let $t_S = \langle s(\mathbf{x}_0), s(\mathbf{x}_0) - s(\mathbf{x}^{(S)}) \rangle$ and $G_S = \langle \mathbf{W}, s(\mathbf{x}^{(S)}) - s(\mathbf{x}_0) \rangle$. By definition,

$$E_S \quad \Leftrightarrow \quad -t_S + \sigma G_S \geq 0 \quad \Leftrightarrow \quad G_S \geq \frac{t_S}{\sigma}.$$

By proposition 2.1, we have

$$\begin{aligned} \langle s(\mathbf{x}_0), s(\mathbf{x}_0) \rangle &= n^k \phi_s \left(\frac{\mathbf{x}_0^T \mathbf{x}_0}{n} \right) = n^k \phi_s(1) \\ \langle s(\mathbf{x}_0), s(\mathbf{x}^{(S)}) \rangle &= n^k \phi_s \left(\frac{\mathbf{x}_0^T \mathbf{x}^{(S)}}{n} \right) = n^k \phi_s \left(\frac{n - 2|S|}{n} \right), \end{aligned}$$

and so $t_S = n^k \left(\phi_s(1) - \phi_s \left(1 - \frac{2|S|}{n} \right) \right)$.

On the other hand,

$$\begin{aligned} G_S &= \langle \mathbf{W}, s(\mathbf{x}^{(S)}) - s(\mathbf{x}_0) \rangle \\ &= \langle \mathsf{Sym}\, \mathbf{G}, s(\mathbf{x}^{(S)}) - s(\mathbf{x}_0) \rangle = \langle \mathbf{G}, s(\mathbf{x}^{(S)}) - s(\mathbf{x}_0) \rangle . \end{aligned}$$

This follows from the fact that for any symmetric k-tensor $\mathbf{S}$ and a k-tensor $\mathbf{T}$ we have

$$\langle \mathbf{S}, \mathbf{T} \rangle = \frac{1}{k!} \sum_{\pi \in \mathfrak{S}_k} \langle \mathbf{S}^\pi, \mathbf{T}^\pi \rangle = \frac{1}{k!} \sum_{\pi \in \mathfrak{S}_k} \langle \mathbf{S}, \mathbf{T}^\pi \rangle = \langle \mathbf{S}, \mathsf{Sym}\, \mathbf{T} \rangle ,$$

and that $s(\mathbf{x}^{(S)})$ and $s(\mathbf{x}_0)$ are symmetric. Thus, G_S is a centered Gaussian variable with variance

$$\begin{aligned} \|s(\mathbf{x}^{(S)}) - s(\mathbf{x}_0)\|_F^2 &= \|s(\mathbf{x}_0)\|_F^2 + \|s(\mathbf{x}^{(S)})\|_F^2 - 2 \langle s(\mathbf{x}_0), s(\mathbf{x}^{(S)}) \rangle \\ &= 2n^k \left(\phi_s(1) - \phi_s \left(1 - \frac{2|S|}{n} \right) \right) \\ &= 2t_s. \end{aligned}$$

Here the second equality follows again from Proposition 2.1.

As a result, we get

$$\mathbb{P}(E_S) = \mathbb{P}(G_S \geq t_s/\sigma) = \Phi \left(\sqrt{\frac{t_s}{2\sigma^2}} \right),$$

where $\Phi(\mathbf{x}) = \mathbb{P}_{g\sim N(0,1)}(g \geq x)$ is the complementary cumulative distribution function of a standard Gaussian variable.

We provide the complete proof of Theorem 2.2 in the following two subsections.

2.3.1 Proof of the achievability when $\sigma < (1-\epsilon)\sigma_s^*$

For achievability, our goal is to prove that

$$p_{ML,fail} \to 0 \qquad \text{as} \qquad n \to \infty$$

if $\sigma < (1-\epsilon)\sigma_s^*$ for some $\epsilon > 0$.

Suppose first that s is even, *i.e.*, $s(\mathbf{x}) = s(-\mathbf{x})$ for any $\mathbf{x} \in \{\pm 1\}^n$. In this case, we have $\mathbb{P}(E_S) = \mathbb{P}(E_{[n]\setminus S})$ and

$$\begin{aligned} p_{ML,fail} &= \mathbb{P}(\widehat{\mathbf{x}}_{ML} \notin \{\mathbf{x}_0, -\mathbf{x}_0\}) \\ &= \mathbb{P}\left(\bigcup_{\substack{S:\, S\notin\{\emptyset,[n]\} \\ S \text{ balanced}}} E_S\right) \leq \sum_{\substack{S:\, S\notin\{\emptyset,[n]\} \\ S \text{ balanced}}} \mathbb{P}(E_S) = 2 \sum_{\substack{S:\, 1\leq|S|\leq\frac{n}{2} \\ S \text{ balanced}}} \mathbb{P}(E_S). \end{aligned}$$

Hence, we can easily adapt the proof for the case that s is not even to the case that s is even. For this reason, in the rest of this subsection **we assume that s is not even**, or equivalently, there is an odd $r \in \{0, 1, \cdots, k\}$ such that $\widehat{s}(r) \neq 0$.

We argued in the previous part of this section that

$$\mathbb{P}(E_S) = \Phi\left(\sqrt{\frac{t_S}{2\sigma^2}}\right) \quad \text{where} \quad t_S = n^k\left(\phi_s(1) - \phi_s\left(1 - \frac{2|S|}{n}\right)\right).$$

Since t_S only depends on the size of S, let us write t_r for t_S if $|S| = r$. We get

$$p_{ML,fail} \leq \sum_{r=1}^{n-1} \Phi\left(\sqrt{\frac{t_r}{2\sigma^2}}\right) \cdot \#(S : |S| = r, \text{ balanced}).$$

Moreover,

$$\#(S : |S| = r,\ \text{balanced}) = \begin{cases} \binom{n/2}{r/2}^2 & \text{if } r \text{ is even} \\ 0 & \text{otherwise,} \end{cases}$$

since choosing a balanced set S of size r is equivalent to choosing a set S_- of size $r/2$ from the indices labeled with -1 and a set S_+ of size $r/2$ from the indices labeled with $+1$. Thus,

$$p_{ML,fail} \leq \sum_{r=1}^{n-1} \binom{n/2}{r/2}^2 \Phi\left(\sqrt{\frac{t_r}{2\sigma^2}}\right).$$

We have $\Phi(\mathbf{x}) \leq e^{-x^2/2}$ by a Chernoff bound. Hence,

$$\begin{aligned} \Phi\left(\sqrt{\frac{t_r}{2\sigma^2}}\right) &\leq \exp\left(-\frac{t_r}{4\sigma^2}\right) \\ &= \exp\left(-\frac{n^k}{4\sigma^2}\left(\phi_s(1) - \phi_s\left(1 - \frac{2r}{n}\right)\right)\right) \\ &= \exp\left(-r\log n \cdot \left(\frac{\sigma_s^*}{\sigma}\right)^2 \frac{\phi_s(1) - \phi_s\left(1 - \frac{2r}{n}\right)}{\phi'(1) \cdot \frac{2r}{n}}\right), \end{aligned}$$

since $(\sigma_s^*)^2 = \phi_s'(1) \cdot \frac{n^{k-1}}{2\log n}$. When $\sigma < (1-\epsilon)\sigma_s^*$, we have $(\sigma_s^*/\sigma)^2 > (1-\epsilon)^{-2} \geq 1+\epsilon$ and so

$$\Phi\left(\sqrt{\frac{t_r}{2\sigma^2}}\right) \leq \exp\left(-(1+\epsilon)r\log n \cdot \frac{\phi_s(1) - \phi_s\left(1 - \frac{2r}{n}\right)}{\phi_s'(1) \cdot \frac{2r}{n}}\right). \tag{2.1}$$

Claim. *ϕ_s is increasing, strictly convex function on* $(0, 1)$.

Proof of Claim. Recall that

$$\phi_s(t) = \sum_{r=0}^{k} \binom{k}{r} \widehat{s}(r)^2 t^r.$$

By direct calculation, we get $\phi_s'(t) > 0$ and $\phi_s''(t) > 0$ for any $t \in (0, 1)$ since $\widehat{s}(r) > 0$ for some $r \geq 2$. It implies that ϕ_s is increasing and strictly convex on $(0, 1)$. $\square$

This claim implies that there exists a unique point $t_0 \in (0, 1)$ satisfying

$$\phi_s(1) - \phi_s(1 - t_0) = (1 - \epsilon/2)t_0\phi_s'(1),$$

and we have

$$\phi_s(1) - \phi_s(1-t) \geq \begin{cases} (1-\epsilon/2)t\phi_s'(1) & \text{for } t \in [0, t_0] \\ \phi_s(1) - \phi_s(1-t_0) & \text{for } t \in (t_0, 1]. \end{cases}$$

Moreover, when $t \in (1, 2]$, we have

$$\begin{aligned} \phi_s(1) - \phi_s(1-t) &= \sum_{r=0}^{k} \binom{k}{r} \widehat{s}(r)^2 (1 - (1-t)^r) \\ &\geq \sum_{r:\text{odd}} \binom{k}{r} \widehat{s}(r)^2 (1 + (t-1)^r) \\ &\geq \sum_{r:\text{odd}} \binom{k}{r} \widehat{s}(r)^2, \end{aligned}$$

which is strictly positive since we assumed that $\widehat{s}(r) \neq 0$ for some odd r. Thus,

$$\frac{\phi_s(1) - \phi_s\left(1 - \frac{2r}{n}\right)}{\phi_s'(1) \cdot \frac{2r}{n}} \geq \begin{cases} 1 - \epsilon/2 & \text{if } \frac{2r}{n} \leq t_0 \\ C \cdot \frac{n}{r} & \text{if } \frac{2r}{n} > t_0 \end{cases}$$

for some constant $C > 0$ which is independent of n. Plugging it in (2.1), we get

$$\Phi\left(\sqrt{\frac{t_r}{2\sigma^2}}\right) \leq \begin{cases} \exp\left(-(1+\epsilon/2) r \log n\right) & \text{if } r \leq \frac{t_0}{2} n \\ \exp\left(-Cn \log n\right) & \text{if } r > \frac{t_0}{2} n. \end{cases}$$

Thus,

$$\begin{aligned} p_{ML,fail} \leq & \sum_{r:\text{even}} \binom{n/2}{r/2}^2 \Phi\left(\sqrt{\frac{t_r}{2\sigma^2}}\right) \\ \leq & \sum_{r \geq 2} \left(\frac{n}{2}\right)^r (n^{-1-\epsilon/2})^r + 2^n e^{-Cn \log n} \\ \leq & \sum_{r \geq 2} (2n^{\epsilon/2})^{-r} + e^{-Cn \log n + O(n)} \\ \lesssim & \ n^{-\epsilon} + o(1), \end{aligned}$$

which converges to 0 as n grows.

2.3.2 Proof of the impossibility when $\sigma > (1+\epsilon)\sigma_s^*$

Recall that when $|S| = r$, we have

$$\mathbb{P}(E_S) = \Phi\left(\sqrt{\frac{t_r}{2\sigma^2}}\right)$$

where

$$t_r = n^k\left(\phi_s(1) - \phi_s\left(1 - \frac{2r}{n}\right)\right).$$

Since ϕ_s is convex on $(0,1)$, we have

$$t_r \le n^k \phi_s'(1) \cdot \frac{2r}{n} = (\sigma_s^*)^2 \cdot 4r \log n$$

as long as $r < \frac{n}{2}$. In particular, when $S = \{u,v\}$, we have $t_2 \le 8(\sigma_s^*)^2 \log n$ and

$$\mathbb{P}(E_{\{u,v\}}) \ge \Phi\left(2\left(\frac{\sigma_s^*}{\sigma}\right)\sqrt{\log n}\right).$$

By a standard tail estimation of standard normal, we have

$$\Phi(\mathrm{x}) \ge \frac{e^{-x^2/2}}{\sqrt{2\pi}}\left(\frac{1}{x} - \frac{1}{x^3}\right)$$

and it implies that when $\sigma_s^* \asymp \sigma$,

$$\mathbb{P}(E_{\{u,v\}}) \ge \exp\left(-2\left(\frac{\sigma_s^*}{\sigma}\right)^2 \log n - O(\log\log n)\right) = n^{-2(\sigma_s^*/\sigma)^2 - o(1)}.$$

There are $\frac{n^2}{4}$ many pairs of vertices u,v that we can choose. Hence, to have an event $E_{\{u,v\}}$ to happen with high probability, we must have $\mathbb{P}(E_{\{u,v\}}) = \Omega(n^{-2})$; otherwise we will get

$$\mathbb{P}\left(\bigcup E_{\{u,v\}}\right) \le \frac{n^2}{4}\mathbb{P}(E_{\{u,v\}}) = o(1).$$

Indeed, this condition becomes sufficient if the events $E_{\{u,v\}}$ were independent, which is unfortunately not the case here. Nevertheless, we would like to argue that those events are "nearly" independent so that $\mathbb{P}(E_{\{u,v\}}) = \Omega(n^{-2+\epsilon})$ for some $\epsilon > 0$ implies that $p_{ML,fail} = 1 - o(1)$.

Let V_+ and V_- be the partition of $[n]$ where

$$V_+ = \{v \in [n] : (\mathbf{x}_0)_v = 1\} \quad \text{and} \quad V_- = \{v \in [n] : (\mathbf{x}_0)_v = -1\}.$$

For each pair $(u, v) \in (V_+, V_-)$, we will denote $E_{\{u,v\}}$ by E_{uv} and $\mathbf{x}^{(\{u,v\})}$ by $\mathbf{x}^{(uv)}$. Recall the definition of E_{uv}: E_{uv} is the event that G_{uv} is at least t_2/σ, where

$$G_{uv} = \left\langle \mathbf{G}, s(\mathbf{x}^{(uv)}) - s(\mathbf{x}_0) \right\rangle \quad \text{and} \quad t_2 = n^k \left(\phi_s(1) - \phi_s\left(1 - \frac{4}{n}\right)\right).$$

Let U be a subset of $[n]$ such that $|U \cap V_+| = |U \cap V_-|$. We will write $(u, v) \in U$ if $u \in U \cap V_+$ and $v \in U \cap V_-$. Clearly we have

$$p_{ML,fail} \geq \mathbb{P}\left(\bigcup_{(u,v)\in U} E_{uv} \right),$$

and our goal is to prove that the right-hand side converges to 1 if $\mathbb{P}(E_{uv}) \gtrsim n^{-2+c}$ for some $c > 0$, under an appropriate choice of U.

Let us first investigate the set of variables which the event E_{uv} depends on. Expanding G_{uv}, we have

$$G_{uv} = \sum_{\alpha \in [n]^k} \mathbf{G}_\alpha \left(s(\mathbf{x}^{(uv)})_\alpha - s(\mathbf{x}_0)_\alpha \right).$$

Note that $s(\mathbf{x}^{(uv)})_\alpha = s(\mathbf{x}_0)_\alpha$ if $\alpha^{-1}(\{u, v\}) = \emptyset$, hence G_{uv} only depends on the variables $\mathbf{G}_\alpha$ where $\alpha^{-1}(\{u, v\}) \neq \emptyset$.

We partition the set $\{\alpha : \alpha^{-1}(\{u,v\}) \neq \emptyset\}$ into the $\mathcal{A}_u$, $\mathcal{A}_v$ and $\mathcal{A}_{uv}$ where

$$\begin{aligned}
\mathcal{A}_u &:= \{\alpha : |\alpha^{-1}(u)| = 1,\ \alpha^{-1}(U \setminus \{u\}) = \emptyset\} \\
\mathcal{A}_v &:= \{\alpha : |\alpha^{-1}(v)| = 1,\ \alpha^{-1}(U \setminus \{v\}) = \emptyset\} \\
\mathcal{A}_{uv} &:= \{\alpha : |\alpha^{-1}(U)| \geq 2,\ \alpha^{-1}(\{u,v\}) \neq \emptyset\}.
\end{aligned}$$

Let G_u, G_v and G'_{uv} be the variables such that

$$\begin{aligned}
G_u &= \sum_{\alpha \in \mathcal{A}_u} \mathbf{G}_\alpha \left(s(\mathbf{x}^{(u)})_\alpha - s(\mathbf{x}_0)_\alpha\right) \\
G_v &= \sum_{\alpha \in \mathcal{A}_v} \mathbf{G}_\alpha \left(s(\mathbf{x}^{(v)})_\alpha - s(\mathbf{x}_0)_\alpha\right) \\
G'_{uv} &= \sum_{\alpha \in \mathcal{A}_{uv}} \mathbf{G}_\alpha \left(s(\mathbf{x}^{(uv)})_\alpha - s(\mathbf{x}_0)_\alpha\right).
\end{aligned}$$

Here $\mathbf{x}^{(u)}$ is the vector with entries

$$\mathbf{x}^{(u)}_w = \begin{cases} -(\mathbf{x}_0)_u & \text{if } w = u \\ (\mathbf{x}_0)_w & \text{otherwise.} \end{cases}$$

Note that $s(\mathbf{x}^{(uv)})_\alpha = s(\mathbf{x}^{(u)})_\alpha$ if $\alpha \in \mathcal{A}_u$ since in that case we have $\alpha^{-1}(v) = \emptyset$. Similarly, we have $s(\mathbf{x}^{(uv)})_\alpha = s(\mathbf{x}^{(v)})$ if $\alpha \in \mathcal{A}_v$, so $G_{uv} = G_u + G_v + G'_{uv}$.

Informal argument. We would like to argue that the effect of G'_{uv} in G_{uv} is negligible. To see this, note that

$$|\mathcal{A}_u| = |\mathcal{A}_v| = k(n - |U|)^{k-1} \approx kn^{k-1} - k(k-1)|U|n^{k-2}$$

and

$$\begin{aligned}
|\mathcal{A}_{uv}| &\leq \sum_{w \in \{u,v\}} \sum_{i=1}^{k} \#(\alpha : \alpha(i) = w, \alpha([k] \setminus \{i\}) \cap U \neq \emptyset) \\
&= 2k\left(n^{k-1} - (n - |U|)^{k-1}\right) \\
&\approx 2k(k-1)|U|n^{k-2}.
\end{aligned}$$

Hence $|\mathcal{A}_u| = |\mathcal{A}_v| \gg |\mathcal{A}_{uv}|$ when $|U| = o(n)$ as

$$|\mathcal{A}_{uv}| = o(n^{k-1}) \quad \text{and} \quad |\mathcal{A}_u| = |\mathcal{A}_v| = \Omega(n^{k-1}).$$

Thus, we expect

$$\max_{(u,v)\in U} G_u + G_v + G'_{uv} \approx \max_{(u,v)\in U} G_u + G_v = \max_{u\in U\cap V_+} G_u + \max_{v\in U\cap V_-} G_v,$$

and we can bound the last quantity as

$$\begin{aligned}
\mathbb{P}\left(\max_{u\in U\cap V_+} G_u + \max_{v\in U\cap V_-} G_v < \frac{t_2}{\sigma}\right) \leq & \ \mathbb{P}\left(\max_{u\in U\cap V_+} G_u < \frac{t_2}{2\sigma}\right) + \mathbb{P}\left(\max_{v\in U\cap V_-} G_v < \frac{t_2}{2\sigma}\right) \\
= & \prod_{u\in U\cap V_+} \mathbb{P}(G_u < t_2/2\sigma) + \prod_{v\in U\cap V_-} \mathbb{P}(G_v < t_2/2\sigma).
\end{aligned}$$

Together with the approximation $\mathbb{E}\, G_u^2 \approx \frac{1}{2}\,\mathbb{E}\, G_{uv}^2 = t_2$ we get

$$\mathbb{P}(G_u \geq t_2/2\sigma) \approx \Phi\left(\frac{\sqrt{t_2}}{2\sigma}\right) \geq \Phi\left(\frac{\sqrt{8(\sigma_s^*)^2 \log n}}{2\sigma}\right) \geq n^{-(\sigma_s^*/\sigma)^2 - o(1)}.$$

It implies that

$$\begin{aligned}
\prod_{u\in U\cap V_+} \mathbb{P}(G_u < t_2/2\sigma) \lesssim & \ \left(1 - n^{-(\sigma_s^*/\sigma)^2}\right)^{|U|/2} \\
\leq & \ \exp\left(-\frac{1}{2} \cdot |U| n^{-(\sigma_s^*/\sigma)^2}\right)
\end{aligned}$$

which is $o(1)$ if $|U| \gg n^{(\sigma_s^*/\sigma)^2}$. This can be achieved by letting $|U| = n^{1-o(1)}$ as $|U| \gg n^c$ for any $c < 1$. □

Let us make this argument precise. Let $|U| = \frac{n}{\gamma(n)}$, where $\gamma(n)$ will be chosen later in the proof. Note that

$$\max_{(u,v)\in U} G_u + G_v + G'_{uv} \geq \frac{t_2}{\sigma} \quad \Leftrightarrow \quad \bigcup_{(u,v)\in U} E_{uv}.$$

Since $\sigma > (1+\epsilon)\sigma_s^*$, we have

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{(u,v)\in U} E_{uv}\right) &\geq \mathbb{P}\left(\max_{(u,v)\in U} G_u + G_v + G'_{uv} \geq \frac{t_2}{\sigma}\right) \\
&\geq \mathbb{P}\left(\max_{(u,v)\in U} G_u + G_v + G'_{uv} \geq \frac{8(\sigma_s^*)^2 \log n}{(1+\epsilon)\sigma_s^*}\right) \\
&= \mathbb{P}\left(\max_{(u,v)\in U} G_u + G_v + G'_{uv} \geq \frac{8\sigma_s^* \log n}{1+\epsilon}\right)
\end{aligned}$$

Moreover,

$$\begin{aligned}
\max_{(u,v)\in U}(G_u + G_v + G'_{uv}) \geq & \max_{(u,v)\in U}(G_u + G_v) - \max_{(u,v)\in U}(-G'_{uv}) \\
= & \max_{u\in U\cap V_+} G_u + \max_{v\in U\cap V_-} G_v - \max_{(u,v)\in U}(-G'_{uv}).
\end{aligned}$$

Thus,

$$p_{fail,ML} \geq \mathbb{P}\left(\max_{u\in U\cap V_+} G_u + \max_{v\in U\cap V_-} G_v - \max_{(u,v)\in U}(-G'_{uv}) \geq \frac{8\sigma_s^* \log n}{1+\epsilon}\right). \tag{2.2}$$

To estimate this probability, we need to understand the typical values of

$$\max_{u\in U\cap V_+} G_u, \ \max_{v\in U\cap V_-} G_v, \text{ and } \max_{(u,v)\in U}(-G'_{uv}).$$

The following is a folklore result on the tail bound of the maximum of Gaussian variables (for instance, see Problem 3.5 in [92]). For completeness, we include the proof of Lemma 2.8 later in this section.

Lemma 2.8. *Let $g_1, \cdots, g_N$ be centered Gaussian variables, which are not necessarily independent. Suppose that the variances $\mathbb{E}\, g_i^2$ are bounded by some constant $M > 0$ (which may depend on N). Then for any constant $\epsilon > 0$,*

$$\mathbb{P}\left(\max_{i\in[N]} g_i > \sqrt{2(1+\epsilon)M\log N}\right) \leq N^{-\epsilon}.$$

Moreover, if g_i are i.i.d. centered Gaussians with variance M, then for any constant

$\epsilon > 0$,

$$\mathbb{P}\left(\max_{i\in[N]} g_i < \sqrt{2(1-\epsilon)M\log N}\right) \leq N^{-\epsilon+o(1)}.$$

To apply Lemma 2.8 to our setting, let us compute the variances $\mathbb{E}\, G_u^2$, $\mathbb{E}\, G_v^2$ and $\mathbb{E}(G'_{uv})^2$. First remark that G_u's and G_v's where $u \in U \cap V_+$ and $v \in U \cap V_-$ are i.i.d. Gaussians. Moreover, G'_{uv} is independent of G_u and G_v. Hence,

$$\mathbb{E}\, G_{uv}^2 = 2(\mathbb{E}\, G_u^2) + \mathbb{E}(G'_{uv})^2.$$

We have already seen that

$$\mathbb{E}\, G_{uv}^2 = \|s(\mathbf{x}^{(S)}) - s(\mathbf{x}_0)\|_F^2 = 2t_2,$$

hence

$$\mathbb{E}(G'_{uv})^2 = 2t_2 - 2\,\mathbb{E}\, G_u^2.$$

It remains to compute the value of $\mathbb{E}\, G_u^2$ for some $u \in U$.

Proposition 2.9. *For any $u \in U$,*

$$\mathbb{E}\, G_u^2 = 4\phi'(1)(n-|U|)^{k-1} = 8\left(1-\frac{1}{\gamma(n)}\right)^{k-1}(\sigma_s^*)^2\log n.$$

By Proposition 2.9, we get

$$\begin{aligned}
\mathbb{E}(G'_{uv})^2 &= 2t_2 - 8\phi'(1)(n-|U|)^{k-1} \\
&= 2t_2 - 16(1-\gamma(n)^{-1})^{k-1}(\sigma_s^*)^2\log n \\
&\leq 16(\sigma_s^*)^2\log n\left(1-(1-\gamma(n)^{-1})^{k-1}\right) \\
&\leq \frac{16(k-1)}{\gamma(n)}\cdot(\sigma_s^*)^2\log n.
\end{aligned}$$

The second inequality follows from $t_2 \leq 8(\sigma_s^*)^2\log n$ and the last inequality follows from $(1-x)^{k-1} \geq 1-(k-1)x$ for any $x \in [0,1]$.

Let $\eta = \frac{\epsilon}{2}$. By Lemma 2.8, with probability $1 - 2\left(\frac{n}{2\gamma(n)}\right)^{-2\eta}$ we have

$$\max_{u \in U \cap V_+} G_u + \max_{v \in U \cap V_-} G_v \geq 2\left(16(1-\eta)\left(1 - \frac{1}{\gamma(n)}\right)^{k-1} \cdot (\sigma_s^*)^2 \log n \cdot \log \frac{n}{2\gamma(n)}\right)^{1/2}$$

and

$$\max_{(u,v) \in U} (-G'_{uv}) \leq \left(2(1+\eta)\frac{16(k-1)}{\gamma(n)} \cdot (\sigma_s^*)^2 \log n \cdot 2 \log \frac{n}{2\gamma(n)}\right)^{1/2}.$$

Letting $\gamma(n) = \log^3 n$, this reduces to

$$\max_{u \in U \cap V_+} G_u + \max_{v \in U \cap V_-} G_v - \max_{(u,v) \in U} (-G'_{uv}) \geq 8\sigma_s^* \log n \left(\sqrt{1-\eta} - o(1)\right).$$

Note that

$$\frac{1}{1+\epsilon} < \sqrt{1 - \frac{\epsilon}{2}} = \sqrt{1-\eta}$$

for sufficiently small $\epsilon > 0$. Together with (2.2), we get

$$p_{fail,ML} \geq 1 - 2\left(\frac{n}{\log^3 n}\right)^{-\epsilon} = 1 - o(1)$$

as desired.

Let us complete the section by presenting the proof of Lemma 2.8 and the proof of Proposition 2.9 here.

Proof of Lemma 2.8. Since $\Phi(\mathbf{x}) \leq e^{-x^2/2}$,

$$\mathbb{P}(g_i > \sqrt{2cM \log N}) \leq \exp\left(-\frac{2(1+\epsilon)M \log N}{2\,\mathbb{E}\, g_i^2}\right) \leq N^{-(1+\epsilon)}.$$

By union bound, we get

$$\mathbb{P}\left(\max_{i \in [N]} g_i > \sqrt{2(1+\epsilon)M \log N}\right) \leq N \cdot N^{-(1+\epsilon)} = N^{-\epsilon}$$

which converges to 0 as desired.

On the other hand, if g_i are i.i.d. $N(0, M)$, then

$$\begin{aligned}
\mathbb{P}\left(\max_{i\in[N]} g_i \le \sqrt{2(1-\epsilon)M\log N}\right) &= \prod_{i=1}^{N} \mathbb{P}\left(g_i \le \sqrt{2(1-\epsilon)M\log N}\right) \\
&= \left[1-\Phi\left(\sqrt{2(1-\epsilon)\log N}\right)\right]^N.
\end{aligned}$$

Since $\Phi(\mathbf{x}) \ge e^{-x^2/2}(x^{-1}-x^{-3})$,

$$\left[1-\Phi\left(\sqrt{2(1-\epsilon)\log N}\right)\right]^N \le (1-N^{-(1-\epsilon)+o(1)})^N \le e^{-N^{\epsilon+o(1)}},$$

which is less than $N^{-\epsilon+o(1)}$ as desired. □

Proof of Proposition 2.9. Note that

$$\begin{aligned}
\mathbb{E}\, G_u^2 &= \sum_{\alpha\in\mathcal{A}_u} \left(s(\mathbf{x}^{(u)})_\alpha - s(\mathbf{x}_0)_\alpha\right)^2 \\
&= \sum_{i\in[k]} \sum_{\substack{\alpha:\alpha(i)=u,\\ \alpha^{-1}(U)=\{i\}}} \left(\sum_{I\subseteq[k]} \widehat{s}(I)\left((\mathbf{x}^{(u)})_{\alpha(I)} - (\mathbf{x}_0)_{\alpha(I)}\right)\right)^2
\end{aligned}$$

since $\mathbf{G}_\alpha$ are i.i.d. standard Gaussians.

Note that

$$(\mathbf{x}^{(u)})_{\alpha(I)} = \begin{cases} (\mathbf{x}_0)_{\alpha(I)} & \text{if } i \notin I \\ -(\mathbf{x}_0)_{\alpha(I)} & \text{if } i \in I. \end{cases}$$

It implies that

$$\begin{aligned}
\mathbb{E}\, G_u^2 &= \sum_{\substack{\alpha:\alpha(i)=u,\\ \alpha^{-1}(U)=\{i\}}} \left(-2\sum_{I\subseteq[k]:I\ni i} \widehat{s}(I)(\mathbf{x}_0)_{\alpha(I)}\right)^2 \\
&= 4\sum_{\substack{\alpha:\alpha(i)=u,\\ \alpha^{-1}(U)=\{i\}}} \sum_{I,J:I\cap J\ni i} \widehat{s}(I)\widehat{s}(J)(\mathbf{x}_0)_{\alpha(I)}(\mathbf{x}_0)_{\alpha(J)} \\
&= 4\sum_{I,J:I\cap J\ni i} \widehat{s}(I)\widehat{s}(J)\left(\sum_{\substack{\alpha:\alpha(i)=u,\\ \alpha^{-1}(U)=\{i\}}} (\mathbf{x}_0)_{\alpha(I)}(\mathbf{x}_0)_{\alpha(J)}\right).
\end{aligned}$$

Note that

$$\begin{aligned}
\sum_{\substack{\alpha:\alpha(i)=u,\\ \alpha^{-1}(U)=\{i\}}} (\mathbf{x}_0)_{\alpha(I)}(\mathbf{x}_0)_{\alpha(J)} = &\ (n-|U|)^{k-1-|I\setminus J|-|J\setminus I|} \prod_{j\in(I\setminus J)\cup(J\setminus I)} \left(\sum_{\alpha(j)\in[n]\setminus U} (\mathbf{x}_0)_{\alpha(j)}\right) \\
= &\ \begin{cases} (n-|U|)^{k-1} & \text{if } I=J \\ 0 & \text{otherwise,} \end{cases}
\end{aligned}$$

since $\sum_{v\in[n]\setminus U}(\mathbf{x}_0)_v = 0$. Hence we get

$$\begin{aligned}
\mathbb{E}\, G_u^2 &= 4(n-|U|)^{k-1}\sum_{i=1}^{k}\sum_{I\subseteq[k]:I\ni i}\widehat{s}(I)^2 \\
&= 4(n-|U|)^{k-1}\sum_{I\subseteq[k]}\widehat{s}(I)^2\cdot|I| \\
&= 4(n-|U|)^{k-1}\sum_{r=0}^{k}\binom{k}{r}\widehat{s}(r)^2\cdot r = 4\phi_s'(1)(n-|U|)^{k-1},
\end{aligned}$$

which is equal to $8\left(1-\frac{1}{\gamma(n)}\right)^{k-1}(\sigma_s^*)^2\log n$. □

2.4 Truncate-and-Relax algorithm

Let us recall a few definitions we need in this section. The truncation of s to degree 2 is

$$s_{\leq 2} = \sum_{I\subseteq[k]:|I|\leq 2}\widehat{s}(I)\chi_I$$

and the estimator $\widehat{\mathbf{x}}^{(2)}_{trunc}$ is an optimum solution to

$$\max_{\mathbf{x}\in\{\pm1\}^n:\mathbf{1}^T\mathbf{x}=0} \langle \mathbf{Y}, s_{\leq 2}(\mathbf{x})\rangle\,. \tag{2.3}$$

The truncate-and-relax algorithm simply solves the *standard semidefinite relaxation* of the optimization problem (2.3), which we describe in Section 2.4.1.

2.4.1 Binary quadratic optimization

We give a quick overview on the binary quadratic optimization problems and semidefinite relaxation techniques applied to this class of problems.

Binary Quadratic Programs (BQPs) are a class of combinatorial optimization problems with binary variables, quadratic objective function and linear or quadratic constraints. Note that (2.3) is an example of BQP since the variables $\mathbf{x}_v$ take value in $\{\pm 1\}$, the objective function $\langle \mathbf{Y}, s_{\leq 2}(\mathbf{x})\rangle$ is a quadratic function in $\mathbf{x}$, and the constraint $\mathbf{1}^T\mathbf{x} = 0$ is linear. It implies that the techniques to solve BQP readily apply for computing $\widehat{\mathbf{x}}^{(2)}_{trunc}$. However, in general BQP is NP-hard to solve even when it is unconstrained, as it includes well-known hard problems such as the problem of finding the maximum cut or the minimum bisection of a graph.

Alternatively, convex relaxation approaches can be used to find an approximate solution for a given BQP. In particular, we consider Semidefinite Programs (SDPs) which are a broad class of convex optimization problems where the variables forms a symmetric positive semidefinite matrix, the objective function is given as the linear function in the variables and constraints are given by linear matrix inequalities. We remark that SDP can be solved up to precision ϵ in $\mathsf{poly}(n, \log(1/\epsilon))$ time.

Let us consider the following unconstrained BQP:

$$\max_{\mathbf{x}\in\{\pm 1\}^n} \mathbf{x}^T Q\mathbf{x} + 2\ell^T\mathbf{x}. \tag{2.4}$$

Let $\widetilde{X} = \begin{bmatrix} 1 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{x}\mathbf{x}^T \end{bmatrix}$ and $\widetilde{Q} = \begin{bmatrix} 0 & \ell^T \\ \ell & Q \end{bmatrix}$. By definition, the objective function $\mathbf{x}^T Q\mathbf{x} + 2\ell^T\mathbf{x}$ can be written as $\left\langle \widetilde{Q}, \widetilde{X} \right\rangle$. Moreover, the matrix $\widetilde{X}$ is positive semidefinite and has rank one, and its diagonal entries are equal to one as $\widetilde{X}_{00} = 1$ and $\widetilde{X}_{ii} = \mathbf{x}_i^2 = 1$. Conversely, if $\widetilde{X}$ satisfies those conditions then $\mathbf{x}$ is a vector in $\{\pm 1\}^n$.

Hence, the original BQP (2.4) is equivalent to

$$\begin{aligned} &\text{maximize} && \left\langle \widetilde{Q}, \widetilde{X} \right\rangle \\ &\text{subject to} && \widetilde{X}_{ii} = 1 && \text{for } i = 0, \cdots, n \\ &&& \widetilde{X} \succeq 0, \ \operatorname{rank}(\widetilde{X}) = 1. \end{aligned}$$

By removing the rank constraint ($\operatorname{rank}(\widetilde{X}) = 1$), we get the following SDP relaxation:

$$\begin{aligned} &\text{maximize} && \left\langle \widetilde{Q}, \widetilde{X} \right\rangle \\ &\text{subject to} && \widetilde{X}_{ii} = 1 \quad \text{for } i = 0, \cdots, n \\ &&& \widetilde{X} \succeq 0. \end{aligned} \tag{2.5}$$

The dual of (2.5) corresponds to the following semidefinite program:

$$\begin{aligned} &\text{minimize} && \operatorname{tr}(D) \\ &\text{subject to} && D - \widetilde{Q} \succeq 0, \\ &&& D \text{ is a diagonal matrix of size } (n+1). \end{aligned} \tag{2.6}$$

It is easy to see the *weak duality*, that is, for any primal feasible $\widetilde{X}$ and dual feasible D,

$$\left\langle D - \widetilde{Q}, \widetilde{X} \right\rangle \geq 0$$

and so

$$\left\langle \widetilde{Q}, \widetilde{X} \right\rangle \leq \left\langle D, \widetilde{X} \right\rangle = \operatorname{tr}(D)$$

since $\widetilde{X}_{ii} = 1$ for any $i = 0, \cdots, n$. Hence, the optimum value of the primal (2.5) is always bounded above by the optimum value of the dual (2.6). For the *strong duality*, we may need some constraint qualification to hold, such as Slater's condition.

Proposition 2.10 (Slater's condition). *If there is a strictly feasible $\widetilde{X}$ for the primal (i.e., $\widetilde{X} \succ 0$), then the optimum values of the primal and dual are equal, and the optimum is attained in the dual. Similarly, if there is a strictly feasible D for the dual (i.e., $D - \widetilde{Q} \succ 0$), then again the optimum values are equal and the optimum is* *attained in the primal.*

We remark that in the primal-dual pair (2.5) and (2.6), the Slater's condition is satisfied since

$$\widetilde{X} = \mathsf{Id}_{n+1} \quad \text{and} \quad D = 2\|\widetilde{Q}\| \cdot \mathsf{Id}_{n+1}$$

are strictly feasible. Once we have the strong duality, we can classify the primal-dual solution pairs in terms of complementary slackness.

Proposition 2.11 (Complementary slackness). *Suppose that Slater's conditions are satisfied. Then, $\widetilde{X}$ and D are primal and dual optimal solutions if and only if the Karush-Kuhn-Tucker (KKT) conditions holds: (i) $\widetilde{X}$ is primal feasible, (ii) D is dual feasible, and (iii) $\left\langle \widetilde{X}, D - \widetilde{Q} \right\rangle = 0$.*

The proof of Proposition 2.10 and Proposition 2.11 is omitted here, as they can be found in many textbooks on convex optimizations (we refer the interested readers to a textbook by Boyd and Vandenberghe [28]).

2.4.2 Matrix expression

To formulate the truncate-and-relax algorithm, we would like to first express $\langle \mathbf{Y}, s_{\leq 2}(\mathbf{x}) \rangle$ in the form of $\left\langle \widetilde{Y}, \widetilde{X} \right\rangle$ where $\widetilde{X} = \begin{bmatrix} 1 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{x}\mathbf{x}^T \end{bmatrix}$.

Note that

$$\begin{aligned} \langle \mathbf{Y}, s_{\leq 2}(\mathbf{x}) \rangle &= \langle s(\mathbf{x}_0) + \sigma \mathbf{W}, s_{\leq 2}(\mathbf{x}) \rangle \\ &= \langle s(\mathbf{x}_0), s_{\leq 2}(\mathbf{x}) \rangle + \sigma \langle \mathbf{W}, s_{\leq 2}(\mathbf{x}) \rangle . \end{aligned}$$

By Proposition 2.1, we have

$$\begin{aligned} \langle s(\mathbf{x}_0), s_{\leq 2}(\mathbf{x}) \rangle &= n^k \sum_{r=0}^{k} \binom{k}{r} \widehat{s}(r) \widehat{s_{\leq 2}}(r) \left(\frac{\mathbf{x}_0^T \mathbf{x}}{n} \right)^r \\ &= n^k \sum_{r=0}^{2} \binom{k}{r} \widehat{s}(r)^2 \left(\frac{\mathbf{x}_0^T \mathbf{x}}{n} \right)^r . \end{aligned}$$

Recall the definition

$$\phi_{s,trunc}(t) = \phi_{s_{\leq 2}}(t) = \sum_{r=0}^{2} \binom{k}{r} \widehat{s}(r)^2 t^r.$$

Thus, we have

$$\langle s(\mathbf{x}_0), s_{\leq 2}(\mathbf{x}) \rangle = n^k \phi_{s,trunc}\left(\frac{\mathbf{x}_0^T \mathbf{x}}{n}\right). \tag{2.7}$$

Moreover, we can write

$$\phi_{s,trunc}\left(\frac{\mathbf{x}_0^T \mathbf{x}}{n}\right) = \left\langle \widetilde{S}, \widetilde{X} \right\rangle \quad \text{and} \quad \langle s(\mathbf{x}_0), s_{\leq 2}(\mathbf{x}) \rangle = n^k \left\langle \widetilde{S}, \widetilde{X} \right\rangle,$$

where

$$\widetilde{S} = \left[\begin{array}{c|c} \widehat{s}(0)^2 & \dfrac{k\widehat{s}(1)^2}{2} \cdot \dfrac{\mathbf{x}_0^T}{n} \\ \hline \dfrac{k\widehat{s}(1)^2}{2} \cdot \dfrac{\mathbf{x}_0}{n} & \dbinom{k}{2} \widehat{s}(2)^2 \cdot \dfrac{\mathbf{x}_0 \mathbf{x}_0^T}{n^2} \end{array}\right]. \tag{2.8}$$

On the other hand, since $s_{\leq 2}(\mathbf{x})$ is symmetric we have $\langle \mathbf{W}, s_{\leq 2}(\mathbf{x}) \rangle = \langle \mathbf{G}, s_{\leq 2}(\mathbf{x}) \rangle$ and

$$\langle \mathbf{G}, s_{\leq 2}(\mathbf{x}) \rangle = \left\langle \mathbf{G}, \sum_{I \subseteq [k] : |I| \leq 2} \widehat{s}(|I|) \mathbf{x}^I \right\rangle = \sum_{I \subseteq [k] : |I| \leq 2} \widehat{s}(|I|) \left\langle \mathbf{G}, \mathbf{x}^I \right\rangle.$$

As a quick reminder, $\mathbf{x}^I$ is defined as the k-tensor $\chi_I(\mathbf{x})$ (See Section 2.1.2).

Let $\mathsf{Proj}_I(\cdot)$ be the linear operator which maps a k-tensor $\mathbf{T}$ to an $|I|$-tensor $\mathsf{Proj}_I(\mathbf{T})$ with entries

$$[\mathsf{Proj}_I(\mathbf{T})]_\beta = \sum_{\substack{\alpha \in [n]^k : \\ \alpha(I) = \beta}} \mathbf{T}_\alpha \quad \text{for } \beta \in [n]^d,$$

for $I = \{i_1 < \cdots < i_d\} \subseteq [k]$. By definition, we get $\langle \mathbf{T}, \mathbf{x}^I \rangle = \langle \mathsf{Proj}_I(\mathbf{T}), \mathbf{x}^{\otimes |I|} \rangle$ and

$$
\begin{aligned}
\langle \mathbf{G}, s_{\leq 2}(\mathbf{x}) \rangle &= \sum_{r=0}^{2} \widehat{s}(r) \left\langle \sum_{I \subseteq [k] : |I| = r} \mathsf{Proj}_I(\mathbf{G}), \mathbf{x}^{\otimes r} \right\rangle \\
&= \left\langle \left[\begin{array}{c|c} \widehat{s}(0)\, \mathsf{Proj}_\emptyset(\mathbf{G}) & \frac{\widehat{s}(1)}{2} \sum_{i=1}^{k} \mathsf{Proj}_{\{i\}}(\mathbf{G})^T \\ \hline \frac{\widehat{s}(1)}{2} \sum_{i=1}^{k} \mathsf{Proj}_{\{i\}}(\mathbf{G}) & \widehat{s}(2) \sum_{1 \leq i < j \leq k} \mathsf{Sym}\,\mathsf{Proj}_{\{i,j\}}(\mathbf{G}) \end{array} \right], \begin{bmatrix} 1 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{x}\mathbf{x}^T \end{bmatrix} \right\rangle .
\end{aligned}
$$

Letting

$$
\begin{aligned}
g_0 &= \mathsf{Proj}_\emptyset(\mathbf{G}) = \sum_\alpha \mathbf{G}_\alpha, \qquad \mathbf{g} = \sum_{i=1}^{k} \mathsf{Proj}_{\{i\}}(\mathbf{G}), \quad \text{and} \\
G &= \sum_{1 \leq i < j \leq k} \frac{1}{2} \left(\mathsf{Proj}_{\{i,j\}}(\mathbf{G}) + \mathsf{Proj}_{\{i,j\}}(\mathbf{G})^T \right),
\end{aligned}
\tag{2.9}
$$

and

$$
\widetilde{W} = \begin{bmatrix} \widehat{s}(0)\, g_0 & \frac{\widehat{s}(1)}{2} \mathbf{g}^T \\ \frac{\widehat{s}(1)}{2} \mathbf{g} & \widehat{s}(2) G \end{bmatrix}, \tag{2.10}
$$

we get $\langle \mathbf{G}, s_{\leq 2}(\mathbf{x}) \rangle = \left\langle \widetilde{W}, \widetilde{X} \right\rangle$.

In summary, by putting (2.7), (2.8), (2.9) and (2.10) together, we get

$$
\langle \mathbf{Y}, s_{\leq 2}(\mathbf{x}) \rangle = \left\langle n^k \widetilde{S} + \sigma \widetilde{W}, \widetilde{X} \right\rangle .
$$

2.4.3 Standard semidefinite relaxation

Let

$$
\widetilde{Y} = n^k \widetilde{S} + \sigma \widetilde{W}
$$

where $\widetilde{S}$ and $\widetilde{W}$ are defined as in (2.7), (2.8), (2.9) and (2.10). We recall that $\widehat{\mathbf{x}}^{(2)}_{trunc}$ is the optimum solution of the following BQP, which is equivalent to (2.3):

$$
\max_{\mathbf{x} \in \{\pm 1\}^n : \mathbf{x}^T \mathbf{1} = 0} \left\langle \widetilde{Y}, \widetilde{X} \right\rangle . \tag{BQP}
$$

Let us introduce a few notations for the ease of reading. Given a vector $\mathbf{x} \in \mathbb{R}^n$, we denote the extended vector $\begin{bmatrix} 1 & \mathbf{x}^T \end{bmatrix}^T$ in $\mathbb{R}^{n+1}$ by $\widetilde{\mathbf{x}}$. We use the convention that the indices of the vectors in $\mathbb{R}^{n+1}$ are labeled $0, 1, \cdots, n$ to emphasize that the zeroth coordinate is augmented. Likewise, the rows and the columns of the matrices in $\mathbb{R}^{(n+1)\times(n+1)}$ are labeled $0, 1, \cdots, n$.

Let $\widetilde{M}$ be a symmetric matrix in $\mathbb{R}^{(n+1)\times(n+1)}$. We often decompose $\widetilde{M}$ into blocks as

$$\widetilde{M} = \begin{bmatrix} a & \mathbf{b}^T \\ \mathbf{b} & C \end{bmatrix},$$

where $a \in \mathbb{R}$, $\mathbf{b} \in \mathbb{R}^n$ and $C \in \mathbb{R}^{n\times n}$. We call $\widetilde{M}$ *block-diagonal* if $\mathbf{b} = \mathbf{0}$.

We consider the following SDP relaxation of (BQP) as we discussed in Section 2.4.1:

$$\begin{aligned} \text{maximize} \quad & \left\langle \widetilde{Y}, \widetilde{X} \right\rangle \\ \text{subject to} \quad & \widetilde{X}_{ii} = 1 \text{ for } i = 0, \cdots, n, \\ & \widetilde{X} \succeq 0. \end{aligned} \tag{SDP}$$

We remark that here we drop the balance constraint $\mathbf{x}^T\mathbf{1} = 0$ in (SDP) for a simpler analysis, although it is straightforward to obtain a relaxation including the balance constraint, for example, by adding

$$\left\langle \begin{bmatrix} 0 & \mathbf{0}_n^T \\ \mathbf{0}_n & \mathbf{1}_n\mathbf{1}_n^T \end{bmatrix}, \widetilde{X} \right\rangle = 0$$

to (SDP).

We denote the optimal solution of (SDP) by $\widehat{X}_{sdp}$. The dual program of (SDP) is

$$\begin{aligned} \text{minimize} \quad & \operatorname{tr}(D) \\ \text{subject to} \quad & D - \widetilde{Y} \succeq 0, \\ & D \text{ is a diagonal matrix in } \mathbb{R}^{(n+1)\times(n+1)}. \end{aligned} \tag{SDP*}$$

By Slater's condition (Proposition 2.10) and KKT conditions (Proposition 2.11), a primal-dual pair $(\widehat{X}, D)$ is optimal if $\widehat{X}$ is primal feasible, D is dual feasible and

$$\left\langle D - \widetilde{Y}, \widetilde{X} \right\rangle = 0, \quad \text{i.e.,} \quad \left\langle D, \widetilde{X} \right\rangle = \left\langle \widetilde{Y}, \widetilde{X} \right\rangle.$$

We would like say that (SDP) succeeds if $\widetilde{\mathbf{x}_0}\widetilde{\mathbf{x}_0}^T$ is the "unique" solution of (SDP) with probability $1 - o(1)$.

Note that when $s_{\leq 2}$ is even, then $\mathbf{x}_0$ and $-\mathbf{x}_0$ are indistinguishable in the SDP (SDP). Hence, in this case we say that (SDP) achieves exact recovery if all the optimum solutions are of the form

$$\widetilde{X} = \mu \begin{bmatrix} 1 & \mathbf{x}_0^T \\ \mathbf{x}_0 & \mathbf{x}_0\mathbf{x}_0^T \end{bmatrix} + (1-\mu) \begin{bmatrix} 1 & -\mathbf{x}_0^T \\ -\mathbf{x}_0 & \mathbf{x}_0\mathbf{x}_0^T \end{bmatrix}.$$

The following proposition tells us that if $\widetilde{Y}$ is block-diagonal (which is equivalent to that $s_{\leq 2}$ is even), then there must be an optimum solution which is also block-diagonal.

Proposition 2.12. *If $\widehat{s}(1) = 0$, then there exists an optimum solution $\widetilde{X}^*$ of* (SDP) *which is block diagonal,* i.e., $\widetilde{X}^* = \begin{bmatrix} * & \mathbf{0}_n^T \\ \mathbf{0}_n & * \end{bmatrix}$.

Proof. Let $\widetilde{X}$ be a feasible solution of (SDP) which decomposes as $\widetilde{X} = \begin{bmatrix} x_0 & \mathbf{x}^T \\ \mathbf{x} & X \end{bmatrix}$.

Let $\widetilde{X}' = \begin{bmatrix} x_0 & -\mathbf{x}^T \\ -\mathbf{x} & X \end{bmatrix}$.

We claim that $\widetilde{X}'$ is feasible. First of all, $\widetilde{X}'_{ii} = 1$ is satisfied for all $i = 0, \cdots, n$ by definition. Moreover, note that

$$\widetilde{X}' = D^{-1}\widetilde{X}D \quad \text{where} \quad D = \operatorname{diag}([1 \ \ -\mathbf{1}_n^T]^T) = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{bmatrix}.$$

Thus, $\widetilde{X}' \succeq 0$ if and only if $\widetilde{X} \succeq 0$.

If $\widetilde{Y}$ is block diagonal, we get

$$\left\langle \widetilde{Y}, \widetilde{X} \right\rangle = \left\langle \widetilde{Y}, \widetilde{X}' \right\rangle = \left\langle \widetilde{Y}, \frac{\widetilde{X} + \widetilde{X}'}{2} \right\rangle.$$

Note that $\frac{\widetilde{X}+\widetilde{X}'}{2}$ is also feasible by the convexity of SDPs, and

$$\frac{\widetilde{X} + \widetilde{X}'}{2} = \begin{bmatrix} x_0 & \mathbf{0}^T \\ \mathbf{0} & X \end{bmatrix}$$

so it is block-diagonal. As a result, we obtain a block-diagonal optimal solution $\frac{\widetilde{X}+\widetilde{X}'}{2}$ of (SDP) as desired. $\square$

To be precise, we say the *truncate-and-relax algorithm achieves exact recovery* if any optimum solution of (SDP) lies in

- the span of
$$\widetilde{X}_0 = \widetilde{\mathbf{x}}_0 \widetilde{\mathbf{x}}_0^T = \begin{bmatrix} 1 & \mathbf{x}_0^T \\ \mathbf{x}_0 & \mathbf{x}_0\mathbf{x}_0^T \end{bmatrix},$$
if $\widehat{s}(1) \neq 0$ (i.e., s is not even), or

- the span of
$$\widetilde{X}_0 = \begin{bmatrix} 1 & \mathbf{x}_0^T \\ \mathbf{x}_0 & \mathbf{x}_0\mathbf{x}_0^T \end{bmatrix} \quad \text{and} \quad \widetilde{X}_0' = \begin{bmatrix} 1 & -\mathbf{x}_0^T \\ -\mathbf{x}_0 & \mathbf{x}_0\mathbf{x}_0^T \end{bmatrix},$$
if $\widehat{s}(1) = 0$.

By KKT conditions, $\widetilde{X}_0$ is optimal if and only if

$$\left\langle D - \widetilde{Y}, \widetilde{X}_0 \right\rangle = 0 \quad \text{or equivalently} \quad \left(D - \widetilde{Y}\right)\widetilde{\mathbf{x}}_0 = \mathbf{0}$$

since $\widetilde{X}_0 = \widetilde{\mathbf{x}}_0\widetilde{\mathbf{x}}_0^T$. For each $v \in [n]$, we have

$$\begin{aligned}\left(\left(D-\widetilde{Y}\right)\widetilde{\mathbf{x}}_0\right)_v = 0 \quad &\Leftrightarrow \quad D_{vv}\left(\widetilde{\mathbf{x}}_0\right)_v = \sum_{w\in[n]} \widetilde{Y}_{vw}\left(\widetilde{\mathbf{x}}_0\right)_w \\ &\Leftrightarrow \quad D_{vv} = \sum_{w\in[n]} \widetilde{Y}_{vw}\left(\widetilde{\mathbf{x}}_0\right)_v\left(\widetilde{\mathbf{x}}_0\right)_w,\end{aligned} \tag{2.11}$$

as $\left(\widetilde{\mathbf{x}}_0\right)_v^2 = 1$ for all $v \in [n]$.

Let $D_0 = \operatorname{diag}(\widetilde{\mathbf{x}}_0)$ and $\Gamma = D_0\widetilde{Y}D_0$. Then, (2.11) can be rewritten as $D = \operatorname{diag}(\Gamma\mathbf{1}_{n+1})$ since

$$\left(\Gamma\mathbf{1}_{n+1}\right)_v = \sum_{w\in[n]}\Gamma_{vw} = \sum_{w\in[n]}\widetilde{Y}_{vw}\left(\widetilde{\mathbf{x}}_0\right)_v\left(\widetilde{\mathbf{x}}_0\right)_w = D_{vv}.$$

We define the *Laplacian* of a symmetric matrix M as the symmetric matrix

$$L_M = \operatorname{diag}\left(M\mathbf{1}\right) - M$$

of the same size. In particular, the Laplacian L_Γ can be written as

$$\begin{aligned}L_\Gamma = \operatorname{diag}(\Gamma\mathbf{1}_{n+1}) - \Gamma = \ & D - D_0\widetilde{Y}D_0 \\ = \ & D_0(D-\widetilde{Y})D_0,\end{aligned}$$

hence $D - \widetilde{Y} = D_0 L_\Gamma D_0$.

Proposition 2.13. *The truncate-and-relax algorithm achieves exact recovery if* $L_\Gamma \succeq 0$ *and*

- $\lambda_2(L_\Gamma) > 0$ *with probability* $1-o(1)$, *if* $\widehat{s}(1) \neq 0$, *or*
- $\lambda_3(L_\Gamma) > 0$ *with probability* $1-o(1)$, *if* $\widehat{s}(1) = 0$.

Proof. Note that the KKT conditions are satisfied as long as $L_\Gamma \succeq 0$. We claim that the extra condition on λ_2 (or λ_3 if s is even) implies the uniqueness of the solution.

We first note that $L_\Gamma\mathbf{1}_{n+1} = \mathbf{0}_{n+1}$ (immediate from the definition), and if $\widehat{s}(1) = 0$

then additionally we have

$$L_\Gamma \begin{bmatrix} 1 \\ -\mathbf{1}_n \end{bmatrix} = \mathbf{0}_{n+1}.$$

The condition $\lambda_2(L_\Gamma) > 0$ forces that the null-space of L_Γ is spanned by $\mathbf{1}_{n+1}$, and $\lambda_3(L_\Gamma) > 0$ in the case of $\widehat{s}(1) = 0$ forces that the null-space of L_Γ is spanned by $\mathbf{1}_{n+1}$ and $\begin{bmatrix} 1 \\ -\mathbf{1}_n \end{bmatrix}$.

This implies that if $\widetilde{X}_1$ is an optimal solution of (SDP) then $D_0\widetilde{X}_1D_0$ must lie in either the space spanned by $\mathbf{1}_{n+1}\mathbf{1}_{n+1}^T$, or the space spanned by $\mathbf{1}_{n+1}\mathbf{1}_{n+1}^T$ and $\begin{bmatrix} 1 & -\mathbf{1}_n^T \\ -\mathbf{1}_n & \mathbf{1}_n\mathbf{1}_n^T \end{bmatrix}$. This implies the achievability of the truncate-and-relax algorithm. □

2.4.4 Proof of Theorem 2.3

Now we are ready to prove Theorem 2.3. Recall that

$$\begin{aligned} \phi_{s,trunc}(t) = \phi_{s_{\leq 2}}(t) &= \sum_{r=0}^{2} \binom{k}{r} \widehat{s}(r)^2 t^r, \text{ and} \\ \sigma^*_{s,trunc} &= \sqrt{\phi'_{s,trunc}(1)\frac{n^{k-1}}{2\log n}}. \end{aligned}$$

We first prove that the truncate-and-relax algorithm fails if $\sigma > (1+\epsilon)\sigma^*_{s,trunc}$ using the result from Section 2.3.

Theorem 2.14. *If $\sigma > (1+\epsilon)\sigma^*_{s,trunc}$ for some $\epsilon > 0$, then*

- *If $\widehat{s}(1) \neq 0$, then*

$$\mathbb{P}(\widehat{\mathbf{x}}^{(2)}_{trunc} \neq \mathbf{x}_0) = 1 - o(1).$$

- *If $\widehat{s}(1) = 0$, then*

$$\mathbb{P}(\widehat{\mathbf{x}}^{(2)}_{trunc} \notin \{\mathbf{x}_0, -\mathbf{x}_0\}) = 1 - o(1).$$

Moreover, the truncate-and-relax algorithm fails with probability $1 - o(1)$.

Proof. Since $\langle s(\mathbf{x}_0), s_{\leq 2}(\mathbf{x}_0)\rangle = \langle s_{\leq 2}(\mathbf{x}_0), s_{\leq 2}(\mathbf{x}_0)\rangle$, we have

$$\widehat{\mathbf{x}}^{(2)}_{trunc} = \underset{\mathbf{x}\in\{\pm 1\}^n : \mathbf{1}^T\mathbf{x}=0}{\operatorname{argmax}} \left\langle s_{\leq 2}(\mathbf{x}_0) + \sigma \mathbf{W}, s_{\leq 2}(\mathbf{x})\right\rangle ,$$

which is exactly the ML estimator in the $s_{\leq 2}$-spiked k-tensor model. Hence, the statistical threshold is at

$$\sigma^*_{s_{\leq 2}} = \sqrt{\phi'_{s_{\leq 2}}(1)\frac{n^{k-1}}{2\log n}} = \sigma^*_{s,trunc}.$$

In particular, if $\sigma > (1+\epsilon)\sigma^*_{s,trunc}$ for some $\epsilon > 0$, then

- $\mathbb{P}(\widehat{\mathbf{x}}^{(2)}_{trunc} \neq \mathbf{x}_0) = 1 - o(1)$ when $s_{\leq 2}$ is not even, and
- $\mathbb{P}(\widehat{\mathbf{x}}^{(2)}_{trunc} \notin \{\mathbf{x}_0, -\mathbf{x}_0\}) = 1 - o(1)$ when $s_{\leq 2}$ is even,

as desired.

Suppose that $s_{\leq 2}$ is not even. Suppose that $\widehat{\mathbf{x}}^{(2)}_{trunc} = \mathbf{x}$ for some $\mathbf{x} \in \{\pm 1\}^n$ such that $\mathbf{1}^T\mathbf{x} = 0$ and $\mathbf{x} \neq \mathbf{x}_0$. Then, $\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}^T$ is also a feasible solution for (SDP) and

$$\left\langle \widetilde{Y}, \widetilde{\mathbf{x}}\widetilde{\mathbf{x}}^T\right\rangle > \left\langle \widetilde{Y}, \widetilde{\mathbf{x}}_0\widetilde{\mathbf{x}}_0{}^T\right\rangle .$$

Hence, $\widehat{X}_{sdp} \neq \widetilde{\mathbf{x}}_0\widetilde{\mathbf{x}}_0{}^T$. By the same argument,

$$\widehat{X}_{sdp} \notin \operatorname{span}\left(\left\{\begin{bmatrix} 1 & \mathbf{x}_0^T \\ \mathbf{x}_0 & \mathbf{x}_0\mathbf{x}_0^T\end{bmatrix}, \begin{bmatrix} 1 & -\mathbf{x}_0^T \\ -\mathbf{x}_0 & \mathbf{x}_0\mathbf{x}_0^T\end{bmatrix}\right\}\right)$$

if $s_{\leq 2}$ is even.

This implies that if $\widehat{\mathbf{x}}^{(2)}_{trunc}$ fails to recover $\mathbf{x}_0$ with probability $1 - o(1)$, then the truncate-and-relax algorithm also fails with probability $1 - o(1)$, as desired. □

On the other hand, we prove the achievability of the truncate-and-relax algorithm by arguing that dual solution given by the complementary slackness is with high probability feasible.

Theorem 2.15. *If $\sigma < (1-\epsilon)\sigma^*_{s,trunc}$ for some $\epsilon > 0$, then the truncate-and-relax algorithm achieves exact recovery.*

Proof. By Proposition 2.13, it is sufficient to prove that

- L_Γ is positive semidefinite, and
- its second or third smallest eigenvalue (depending on the parity of s) is strictly positive.

Without loss of generality, we may assume that $\widehat{s}(0) = 0$ since it has no effect on L_Γ (in fact, any change in the diagonal of Γ has no effect on L_Γ). Note that by linearity of the Laplacian operator $L_{(\cdot)}$, we have

$$L_\Gamma = n^k L_{D_0 \widetilde{S} D_0} + \sigma L_{D_0 \widetilde{W} D_0}.$$

First of all, let us take a look at $L_{D_0\widetilde{S}D_0}$. We get

$$D_0 \widetilde{S} D_0 = \left[\begin{array}{c|c} 0 & \dfrac{k\widehat{s}(1)^2}{2} \cdot \dfrac{\mathbf{1}_n^T}{n} \\ \hline \dfrac{k\widehat{s}(1)^2}{2} \cdot \dfrac{\mathbf{1}_n}{n} & \dbinom{k}{2} \widehat{s}(2)^2 \cdot \dfrac{\mathbf{1}_n \mathbf{1}_n^T}{n^2} \end{array}\right]$$

and so

$$\begin{aligned} L_{D_0\widetilde{S}D_0} &= \left[\begin{array}{c|c} \dfrac{k\widehat{s}(1)^2}{2} & -\dfrac{k\widehat{s}(1)^2}{2} \cdot \dfrac{\mathbf{1}_n^T}{n} \\ \hline -\dfrac{k\widehat{s}(1)^2}{2} \cdot \dfrac{\mathbf{1}_n}{n} & \left(\dfrac{k\widehat{s}(1)^2}{2} + \dbinom{k}{2}\widehat{s}(2)^2\right) \cdot \dfrac{\mathsf{Id}_n}{n} - \dbinom{k}{2}\widehat{s}(2)^2 \cdot \dfrac{\mathbf{1}_n\mathbf{1}_n^T}{n^2} \end{array}\right] \\ &= \frac{k\widehat{s}(1)^2}{2} \begin{bmatrix} 1 & -\dfrac{\mathbf{1}_n^T}{n} \\ \dfrac{\mathbf{1}_n}{n} & \dfrac{\mathbf{1}_n\mathbf{1}_n^T}{n^2} \end{bmatrix} + \frac{1}{n}\left(\frac{k\widehat{s}(1)^2}{2} + \binom{k}{2}\widehat{s}(2)^2\right) \begin{bmatrix} 0 & \mathbf{0}_n^T \\ \mathbf{0}_n & \mathsf{Id}_n - \dfrac{\mathbf{1}_n\mathbf{1}_n^T}{n} \end{bmatrix}. \end{aligned} \tag{2.12}$$

On the other hand,

$$D_0 \widetilde{W} D_0 = \begin{bmatrix} 0 & \dfrac{\widehat{s}(1)}{2}(D_0\mathbf{g})^T \\ \dfrac{\widehat{s}(1)}{2} D_0 \mathbf{g} & \widehat{s}(2) D_0 G D_0 \end{bmatrix}$$

and

$$L_{D_0\widetilde{W}D_0} = \frac{\widehat{s}(1)}{2}\begin{bmatrix} \mathbf{x}_0^T\mathbf{g} & -(D_0\mathbf{g})^T \\ -D_0\mathbf{g} & \mathrm{diag}(D_0\mathbf{g}) \end{bmatrix} + \widehat{s}(2)\begin{bmatrix} 0 & \mathbf{0}_n^T \\ \mathbf{0}_n & L_{D_0GD_0} \end{bmatrix}. \tag{2.13}$$

Let

$$\widetilde{\mathbf{u}} := \frac{1}{\sqrt{n+1}}\begin{bmatrix} 1 \\ \mathbf{1}_n^T \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{v}} := \sqrt{\frac{n}{n+1}}\begin{bmatrix} 1 \\ -\frac{\mathbf{1}_n^T}{n} \end{bmatrix}.$$

Then,

$$Id_{n+1} - \widetilde{\mathbf{u}}\widetilde{\mathbf{u}}^T - \widetilde{\mathbf{v}}\widetilde{\mathbf{v}}^T = \begin{bmatrix} 0 & \mathbf{0}_n^T \\ \mathbf{0}_n & \mathrm{Id}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T. \end{bmatrix}$$

Hence, any vector $\mathbf{z}$ in $\mathbb{R}^{n+1}$ can be written as

$$\mathbf{z} = a\widetilde{\mathbf{u}} + b\widetilde{\mathbf{v}} + \begin{bmatrix} 0 \\ \mathbf{w} \end{bmatrix},$$

where $a = \widetilde{\mathbf{u}}^T\mathbf{z}$, $b = \widetilde{\mathbf{v}}^T\mathbf{z}$, and for some $\mathbf{w} \in \mathbb{R}^n$ satisfying $\mathbf{1}_n^T\mathbf{w} = 0$.

Let $\mathbf{z} = a\widetilde{\mathbf{u}} + b\widetilde{\mathbf{v}} + \begin{bmatrix} 0 \\ \mathbf{w} \end{bmatrix}$ where $\mathbf{w} \in \mathbb{R}^n$ with $\mathbf{1}_n^T\mathbf{w} = 0$. Then,

$$\mathbf{z}^T L_{D_0\widetilde{S}D_0}\mathbf{z} = \frac{k\widehat{s}(1)^2}{2}\left(\frac{n+1}{n}\right)\cdot b^2 + \frac{1}{n}\left(\frac{k\widehat{s}(1)^2}{2} + \binom{k}{2}\widehat{s}(2)^2\right)\|\mathbf{w}\|_2^2.$$

Moreover,

$$\begin{aligned} \mathbf{z}^T L_{D_0\widetilde{W}D_0}\mathbf{z} = \;& b^2\left(\frac{\widehat{s}(1)}{2}\left(\frac{n+1}{n}\cdot\mathbf{x}_0^T\mathbf{g}\right)\right) + \widehat{s}(1)\sqrt{\frac{n+1}{n}}(\mathbf{g}^T D_0\mathbf{w})b \\ & + \mathbf{w}^T\left(\frac{\widehat{s}(1)}{2}\mathrm{diag}(D_0\mathbf{g}) + \widehat{s}(2)L_{D_0GD_0}\right)\mathbf{w} \\ \geq \;& -b^2\left(\frac{n+1}{n}\cdot\frac{|\widehat{s}(1)|}{2}|\mathbf{x}_0^T\mathbf{g}|\right) - |b|\|\mathbf{w}\|\left(\sqrt{\frac{n+1}{n}}\|D_0\mathbf{g}\|\right) \\ & - \|\mathbf{w}\|_2^2\left\|\frac{\widehat{s}(1)}{2}\mathrm{diag}(D_0\mathbf{g}) + \widehat{s}(2)L_{D_0GD_0}\right\|. \end{aligned}$$

Lemma 2.16. *L_Γ is positive semidefinite with probability $1 - o(1)$ if the followings holds:*

(i) There exists a constant $c > 0$ such that

$$n^{k-1}\left(\frac{k\widehat{s}(1)^2}{2} + \binom{k}{2}\widehat{s}(2)^2\right) - \sigma\left\|\frac{\widehat{s}(1)}{2}\mathrm{diag}(D_0\mathbf{g}) + \widehat{s}(2)L_{D_0GD_0}\right\| \geq cn^{k-1}$$

with probability $1 - o(1)$,

(ii) $|\mathbf{x}_0^T\mathbf{g}| \ll n^k/\sigma$ and $\|D_0\mathbf{g}\| \ll n^{k-1/2}/\sigma$ with probability $1 - o(1)$.

Indeed, the condition (i) and (ii) implies that $\lambda_2(L_\gamma) > 0$ when $\widehat{s}(1) \neq 0$ and $\lambda_3(L_\Gamma) > 0$ when $\widehat{s}(1) = 0$, with probability $1 - o(1)$.

Proof. We have argued that for $\mathbf{z} = a\widetilde{\mathbf{u}} + b\widetilde{\mathbf{v}} + \begin{bmatrix} 0 \\ \mathbf{w} \end{bmatrix}$ where $\mathbf{w} \in \mathbb{R}^n$ with $\mathbf{1}_n^T\mathbf{w} = 0$,

$$\mathbf{z}^T L_\Gamma \mathbf{z} \geq \eta_1 b^2 + 2\eta_2 |b| \|\mathbf{w}\|_2 + \eta_3\|\mathbf{w}\|_2^2$$

where

$$\begin{aligned}
\eta_1 =& \; n^k\left(\frac{n+1}{n}\cdot\frac{k\widehat{s}(1)^2}{2}\right) - \sigma\left(\frac{n+1}{n}\cdot\frac{|\widehat{s}(1)|}{2}|\mathbf{x}_0^T\mathbf{g}|\right) \\
\eta_2 =& \; -\sigma\frac{1}{2}\left(|\widehat{s}(1)|\sqrt{\frac{n+1}{n}}\|D_0\mathbf{g}\|\right) \\
\eta_3 =& \; n^{k-1}\left(\frac{k\widehat{s}(1)^2}{2} + \binom{k}{2}\widehat{s}(2)^2\right) - \sigma\left\|\frac{\widehat{s}(1)}{2}\mathrm{diag}(D_0\mathbf{g}) + \widehat{s}(2)L_{D_0GD_0}\right\|
\end{aligned}$$

Note that if $\eta_2^2 - \eta_1\eta_3 \leq 0$ and $\eta_3 \geq 0$ then $\mathbf{z}^T L_\Gamma \mathbf{z}$ is always nonnegative. Thus, $L_\Gamma \succeq 0$.

If condition (i) holds, then $\eta_3 \gtrsim n^{k-1}$. If condition (ii) holds, then $\eta_1 \asymp n^k$ and $|\eta_2| \ll n^{k-1/2}$. It implies that

$$\eta_2^2 \ll n^{2k-1} \lesssim \eta_1\eta_3,$$

for sufficiently large n and we have $L_\Gamma \succeq 0$.

We remark that if $\eta_2^2 - \eta_1\eta_3$ is strictly negative, then $\lambda_2(L_\Gamma) > 0$ and this holds if $\widehat{s}(1) \neq 0$ and the condition (i) and (ii) hold. On the other hand, if $\widehat{s}(1) = 0$ then we have $\eta_1 = \eta_2 = 0$ so $\eta_2^2 - \eta_1\eta_3 = 0$. In this case, the condition (ii) implies that $\eta_3 > 0$ and we have $\lambda_3(L_\Gamma) > 0$. □

It is straightforward to see that with probability $1 - o(1)$,

$$|\mathbf{x}_0^T \mathbf{g}| = O(n^{\frac{k}{2}}) \quad \text{and} \quad \|D_0\mathbf{g}\| = \|\mathbf{g}\| = O(n^{\frac{k}{2}})$$

since $\mathbf{g}$ is a sum of k Gaussian vectors with i.i.d. entries with variance n^{k-1}. When $\sigma \asymp n^{\frac{k-1}{2}}/\sqrt{\log n}$, we get

$$n^{k-1/2}/\sigma \asymp n^{k/2}\sqrt{\log n}$$

so the condition (ii) is satisfied.

It remains to show that if $\sigma < (1-\epsilon)\sigma^*_{s,trunc}$ for some $\epsilon > 0$, then the condition (i) holds, *i.e.*, there exists $c > 0$ such that

$$\sigma \left\| \frac{\widehat{s}(1)}{2}\mathrm{diag}(D_0\mathbf{g}) + \widehat{s}(2)L_{D_0GD_0} \right\| \le (1-c)n^{k-1}\left(\frac{k\widehat{s}(1)^2}{2} + \binom{k}{2}\widehat{s}(2)^2\right) \tag{2.14}$$

with probability $1 - o(1)$.

We first note that

$$\left\| \frac{\widehat{s}(1)}{2}\mathrm{diag}(D_0\mathbf{g}) + \widehat{s}(2)L_{D_0GD_0} \right\| \le |\widehat{s}(2)| \|D_0GD_0\| + \max_{v\in[n]} \left| \frac{\widehat{s}(1)}{2}(D_0\mathbf{g})_v + \widehat{s}(2)(L_{D_0GD_0})_{vv} \right|,$$

by the triangle inequality. Recall the definition of G:

$$G = \sum_{1\le i<j\le k} \mathsf{Sym}\,\mathsf{Proj}_{\{i,j\}}(\mathbf{G}).$$

Here, observe that $\mathsf{Proj}_{\{i,j\}}(\mathbf{G})$ is a matrix with i.i.d. Gaussian entries with variance n^{k-2}. Hence, $\|\mathsf{Proj}_{\{i,j\}}(\mathbf{G})\| \lesssim \sqrt{n^{k-2}} \cdot \sqrt{n}$ and

$$\|D_0GD_0\| = \|G\| \le \sum_{1\le i<j\le k} \|\,\mathsf{Sym}\ \mathsf{Proj}_{\{i,j\}}(\mathbf{G})\| \lesssim n^{\frac{k-1}{2}}. \tag{2.15}$$

On the other hand, we bound

$$\max_{v\in[n]}\left|\frac{\widehat{s}(1)}{2}(D_0\mathbf{g})_v+\widehat{s}(2)(L_{D_0GD_0})_{vv}\right|$$

by using the Lemma 2.8 on the maximum of Gaussian variables. Expanding the v-th term, we get

$$\begin{aligned}\frac{\widehat{s}(1)}{2}(D_0\mathbf{g})_v+\widehat{s}(2)(L_{D_0GD_0})_{vv} &= \frac{\widehat{s}(1)}{2}(\mathbf{x}_0)_v\mathbf{g}_v+\widehat{s}(2)\sum_{u\in[n]}G_{uv}(\mathbf{x}_0)_u(\mathbf{x}_0)_v\\ &= (\mathbf{x}_0)_v\cdot\left\langle\mathbf{e}^v,\frac{\widehat{s}(1)}{2}\mathbf{g}+\widehat{s}(2)G\mathbf{x}_0\right\rangle,\end{aligned}$$

where $\mathbf{e}^v$ is the vth standard unit vector of $\mathbb{R}^n$. Since

$$\mathbf{g}=\sum_{i\in[k]}\mathsf{Proj}_{\{i\}}(\mathbf{G})\quad\text{and}\quad G=\sum_{1\le i<j\le[k]}\mathsf{Sym\,Proj}_{\{i,j\}}(\mathbf{G}),$$

we get

$$\langle\mathbf{e}^v,\mathbf{g}\rangle=k\left\langle\mathbf{G},\mathbf{1}^{\otimes(k-1)}\odot\mathbf{e}^v\right\rangle$$

and

$$\langle\mathbf{e}^v,G\mathbf{x}_0\rangle=\binom{k}{2}\left\langle\mathbf{G},\mathbf{1}^{\otimes(k-2)}\odot\mathbf{e}^v\odot\mathbf{x}_0\right\rangle.$$

Thus,

$$\begin{aligned}\frac{\widehat{s}(1)}{2}(D_0\mathbf{g})_v+\widehat{s}(2)(L_{D_0GD_0})_{vv} = &\Big\langle\mathbf{G},\frac{k\widehat{s}(1)}{2}(\mathbf{1}^{\otimes(k-1)}\odot\mathbf{e}^v)\\ &+\binom{k}{2}\widehat{s}(2)(\mathbf{1}^{\otimes(k-2)}\odot\mathbf{e}^v\odot\mathbf{x}_0)\Big\rangle\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\left((D_0\mathbf{g})_v+(L_{D_0GD_0})_{vv}\right)^2\\ &=\left\|\frac{k\widehat{s}(1)}{2}(\mathbf{1}^{\otimes(k-1)}\odot\mathbf{e}^v)+\binom{k}{2}\widehat{s}(2)(\mathbf{1}^{\otimes(k-2)}\odot\mathbf{e}^v\odot\mathbf{x}_0)\right\|_F^2\\ &=\left(\frac{k\widehat{s}(1)^2}{4}+\binom{k}{2}\frac{\widehat{s}(2)^2}{2}\right)n^{k-1}+O(n^{k-2})\\ &=\frac{\phi'_{s,trunc}(1)}{4}n^{k-1}+O(n^{k-2}).\end{aligned}$$

Here the last equality follows from that

$$\phi'_{s,trunc}(1) = \sum_{r=0}^{2} \binom{k}{r} r \cdot \widehat{s}(r)^2 = k\widehat{s}(1)^2 + 2\binom{k}{2}\widehat{s}(2)^2.$$

As a result, with probability $1 - o(1)$ we get

$$\max_{v\in[n]} \left| \frac{\widehat{s}(1)}{2}(D_0\mathbf{g})_v + \widehat{s}(2)(L_{D_0GD_0})_{vv} \right| \leq \frac{(1+\eta)}{2}\sqrt{\phi'_{s,trunc}(1) \cdot 2n^{k-1}\log n}, \tag{2.16}$$

for any constant $\eta > 0$.

Putting (2.15) and (2.16) together, we have

$$\begin{aligned}
&\sigma\left|\left|\frac{\widehat{s}(1)}{2}\mathrm{diag}(D_0\mathbf{g}) + \widehat{s}(2)L_{D_0GD_0}\right|\right| \\
&\leq O\left(\sigma n^{\frac{k-1}{2}}\right) + \frac{(1+\eta)}{2}\sqrt{\sigma^2\phi'_{s,trunc}(1) \cdot 2n^{k-1}\log n} \\
&< o\left(n^{k-1}\right) + \frac{(1+\eta)}{2} \cdot (1-\epsilon)^2\phi'_{s,trunc}(1) \cdot n^{k-1},
\end{aligned}$$

since $\sigma < (1-\epsilon)\sigma^*_{s,trunc} = (1-\epsilon)\sqrt{\phi'_{s,trunc}(1)\frac{n^{k-1}}{2\log n}}$. Letting $\eta = \epsilon$, we get

$$\begin{aligned}
\sigma\left|\left|\frac{\widehat{s}(1)}{2}\mathrm{diag}(D_0\mathbf{g}) + \widehat{s}(2)L_{D_0GD_0}\right|\right| &\leq (1-\epsilon)n^{k-1} \cdot \frac{\phi'_{s,trunc}(1)}{2} \\
&= (1-\epsilon)n^{k-1}\left(\frac{k\widehat{s}(1)^2}{2} + \binom{k}{2}\widehat{s}(2)^2\right),
\end{aligned}$$

which confirms that the condition (i) holds. □

2.5 Sum-of-Squares relaxation

Let us first briefly discuss Sum-of-Squares based relaxation algorithms. Given a polynomial $p \in \mathbb{R}[\mathbf{x}]$, consider the problem of finding the maximum of $p(\mathbf{x})$ over $\mathbf{x} \in \mathbb{R}^n$ satisfying polynomial equalities $q_1(\mathbf{x}) = 0, \cdots, q_m(\mathbf{x}) = 0$. Most hard combinatorial optimization problems can be reduced into this form, including max-cut, k-colorability, and general constraint satisfaction problems. The *Sum-of-Squares hierarchy* (SoS) is a systematic way to relax a polynomial optimization problem to a sequence of increasingly strong convex programs, each leading to a larger semidefinite program. See [23] for a good exposition of the topic.

There are many different ways to formulate the SoS hierarchy [89, 82, 79, 65]. Here we choose to follow the description based on *pseudo-expectation functionals* [23].

Suppose that we are interested in maximizing an n-variate polynomial $g(\mathbf{x})$ over the n-dimensional hypercube $\{\pm 1\}^n$. For instance, consider $g(\mathbf{x}) = \langle \mathbf{Y}, \mathbf{x}^{\otimes 4} \rangle$ where $\mathbf{Y} = \mathbf{x}_0^{\otimes 4} + \sigma \mathbf{W}$ is generated from the single-spiked 4-tensor model. Here maximizing $g(\mathbf{x})$ corresponds to the problem of computing the maximum-likelihood estimator under this single-spiked model.

Observe that

$$\max_{\mathbf{x} \in \{\pm 1\}^n} g(\mathbf{x}) = \max_{\mu} \mathop{\mathbb{E}}_{\mathbf{x} \sim \mu} g(\mathbf{x}), \tag{2.17}$$

where μ ranges over all probability distributions over all balanced vectors in $\{\pm 1\}^n$. We remark that the space of all probability distributions is *convex*, but possibly has an exponentially large dimension. In particular, in this case we have

$$\left\{ \mu \in \mathbb{R}^{\{\pm 1\}^n} : \sum_{\mathbf{x} \in \{\pm 1\}^n} \mu(\mathbf{x}) = 1,\ \mu(\mathbf{x}) \geq 0 \text{ for all } \mathbf{x} \in \{\pm 1\}^n \right\},$$

which is $2^n - 1$ dimensional.

A linear functional $\widetilde{\mathbb{E}}$ on $\mathbb{R}[\mathbf{x}]$ is called *pseudo-expectation functional (p.e.f.) of degree* $d = 2\ell$ if it satisfies

- $\widetilde{\mathbb{E}}\, 1 = 1$, and
- $\widetilde{\mathbb{E}}\, q(\mathbf{x})^2 \geq 0$ for any $q \in \mathbb{R}[\mathbf{x}]$ of degree at most ℓ.

Clearly, any "true" expectation functional is also a p.e.f. of any degree, but the converse is not true. We relax (2.17) as

$$\begin{aligned} \text{maximize} \quad & \widetilde{\mathbb{E}}\, g(\mathbf{x}) \\ \text{subject to} \quad & \widetilde{\mathbb{E}} \text{ is a p.e.f. of degree} \leq 2\ell. \\ & \widetilde{\mathbb{E}}[q(\mathbf{x})(\mathbf{x}_v^2 - 1)] = 0 \text{ for any } v \in [v] \text{ and } q \text{ of degree} \leq 2\ell - d. \end{aligned}$$

The space of p.e.f.s of degree at most 2ℓ can be described as an affine section of the semidefinite cone of dimension $O(n^{2\ell})$. As ℓ increases, the space gets smaller and it coincides with the space of true expectations when $\ell = n$ for the polynomials on the hypercube $\{\pm 1\}^n$.

2.5.1 Pseudo-expectation Functionals and Moment matrices

We continue to use the Fourier-theoretic notations introduced in Section 2.1.2.

Let V be a finite set. We denote the collection of subsets of V of size at most ℓ by $\binom{V}{\leq \ell}$, *i.e.*, $\binom{V}{\leq \ell} = \bigcup_{i=0}^{\ell} \binom{V}{i}$. Moreover, we denote the size of $\binom{[n]}{\leq \ell}$ by $\binom{n}{\leq \ell}$. Thus we have $\binom{n}{\leq \ell} = \sum_{i=0}^{\ell} \binom{n}{i}$.

We denote the space of multilinear polynomials of $\mathbf{x}$ by $\mathcal{V}$, and the subspace of $\mathcal{V}$ consisting of multilinear polynomials of degree at most d by $\mathcal{V}_d$. Recall that any real-valued function f on the hypercube $\{\pm 1\}^n$ has the unique Fourier expansion

$$f(\mathbf{x}) = \sum_{S \subseteq [n]} \widehat{f}(S) \mathbf{x}_S,$$

which is a multilinear polynomial in $\mathbf{x}$. To avoid confusion, here we choose to use $\mathbf{x}_S = \prod_{i \in S} \mathbf{x}_i$ directly, instead of the character $\chi_S(\cdot) : \mathbf{x} \mapsto \mathbf{x}_S$.

Let M be a symmetric matrix of size $\binom{n}{\leq \ell}$. We use the convention that the rows and the columns of M are indexed by the subsets of $[n]$ of size at most ℓ. To avoid confusion, we use $M[S, T]$ instead of $M_{S,T}$ to denote the entries of M in the remainder of this section.

Pseudo-expectation functionals

For our purpose, we consider pseudo-expectation functionals defined on the hypercube $\{\pm 1\}^n$. See [23] for general definition.

Let ℓ be a positive integer and $d = 2\ell$. A *pseudo-expectation functional (p.e.f.) of degree d on* $\{\pm 1\}^n$ is a linear functional $\widetilde{\mathbb{E}}$ on the space $\mathcal{V}_d$ of multilinear polynomials of degree at most d such that

(i) $\widetilde{\mathbb{E}}[1] = 1$, and

(ii) $\widetilde{\mathbb{E}}[q^2] \geq 0$ for any q with $\deg(q) \leq \ell = d/2$.

We say $\widetilde{\mathbb{E}}$ *satisfies* the system of equations $\{p_i(\mathbf{x}) = 0\}_{i=1}^m$ if

$$\widetilde{\mathbb{E}}[p_1 q_1 + \cdots + p_m q_m] = 0$$

for any $q_1, \cdots, q_m \in \mathcal{V}$ such that $\deg(q_i) \leq d - \deg(p_i)$ for all $i \in [m]$.

We remark that the followings holds:

- If $\widetilde{\mathbb{E}}$ is a p.e.f. of degree d, then $\widetilde{\mathbb{E}}$ is also a p.e.f. of degree d' for any $d' \leq d$.

- If $\widetilde{\mathbb{E}}$ is a p.e.f. of degree $2n$, then $\widetilde{\mathbb{E}}$ is a true expectation functional of a probability distribution which is supported on the variety

$$P := \{\mathbf{x} \in \{\pm 1\}^n : p_i(\mathbf{x}) = 0 \text{ for all } i \in [m]\}.$$

The second fact implies that

$$\max_{\mathbf{x} \in P} p_0(\mathbf{x}) \quad \Leftrightarrow \quad \max_{\substack{\widetilde{\mathbb{E}}\text{:p.e.f. of degree } 2n \\ \text{satisfying } \{p_i(\mathbf{x}) = 0\}_{i \in [m]}}} \widetilde{\mathbb{E}}[p_0]$$

Let d be an even integer such that $d > \max(\deg(p_0), \deg(p_1), \cdots, \deg(p_m))$. We define the *Sum-of-Squares (SoS) relaxation of degree d* as the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & \widetilde{\mathbb{E}}[p_0] \\ \text{subject to} \quad & \widetilde{\mathbb{E}} \text{ is a p.e.f. on } \{\pm 1\}^n \text{ of degree } d, \\ & \widetilde{\mathbb{E}} \text{ satisfies } \{p_i(\mathbf{x}) = 0\}_{i=1}^m. \end{aligned} \tag{SoS_d}$$

We note that the value of (SoS_d) decreases as d grows, and it reaches the optimum value $\max_{\mathbf{x} \in P} p_0(\mathbf{x})$ of the original problem at $d = 2n$.

Moment matrices and SoS-symmetry

We are going to show that (SoS_d) can be written as a semidefinite program of size $\binom{n}{\leq d/2}$. For this, we introduce a matrix interpretation of the SoS relaxation.

We say M is *SoS-symmetric* if $M[S,T] = M[S',T']$ whenever $S \oplus T = S' \oplus T'$, where $S \oplus T$ denotes the symmetric difference $(S \setminus T) \cup (T \setminus S)$. Note that when $\mathbf{x} \in \{\pm 1\}^n$ we always have

$$\mathbf{x}_S \mathbf{x}_T = \left(\prod_{i \in S} \mathbf{x}_i\right)\left(\prod_{j \in T} \mathbf{x}_j\right) = \prod_{i \in S \oplus T} \mathbf{x}_i = \mathbf{x}_{S \oplus T},$$

hence the matrix X with entries $X[S,T] = \mathbf{x}_{S \oplus T}$ is SoS-symmetric.

Given $f \in \mathcal{V}_{2\ell}$, we say M *represents* f if

$$\widehat{f}(U) = \sum_{\substack{S,T \in \binom{[n]}{\leq \ell}: \\ S \oplus T = U}} M[S,T].$$

By definition, if M represents f, then $\langle M, X \rangle = f(\mathbf{x})$ where X is the SoS-symmetric matrix with entries $X[S,T] = \mathbf{x}_{S \oplus T}$. We denote the *unique* SoS-symmetric matrix representing f by M_f, *i.e.*,

$$M_f[S,T] = \widehat{f}(S \oplus T) \cdot \left(\#\left((S',T') : S' \oplus T' = S \oplus T\right)\right)^{-1}.$$

Let $\mathcal{L}$ be a linear functional on $\mathcal{V}$. By linearity, $\mathcal{L}$ is determined by the values $(\mathcal{L}[\mathbf{x}_S] : S \subseteq [n])$ and we often write $\mathcal{L}_S$ to denote $\mathcal{L}[\mathbf{x}_S]$. Let $X_{\mathcal{L}}$ be the SoS-symmetric matrix of size $\binom{n}{\leq \ell}$ with entries $X[S,T] = \mathcal{L}[\mathbf{x}_{S \oplus T}]$. The matrix $X_{\mathcal{L}}$ is called the *moment matrix of $\mathcal{L}$ of degree 2ℓ*. We remark that the matrix X with entries $X[S,T] = \mathbf{x}_{S \oplus T}$ corresponds to $X_{\delta_{\mathbf{x}}}$ where

$$\delta_{\mathbf{x}}(\cdot) : p \mapsto p(\mathbf{x})$$

is the evaluation functional at $\mathbf{x}$, or equivalently the expectation functional of the one-point distribution on $\{\mathbf{x}\}$.

By definition, we have

$$\begin{aligned}
\mathcal{L}[f] &= \sum_{U \in \binom{[n]}{\leq 2\ell}} \widehat{f}(U)\mathcal{L}[\mathbf{x}_U] \\
&= \sum_{S,T \in \binom{[n]}{\leq \ell}} M_f[S,T] X_{\mathcal{L}}[S,T] = \langle X_{\mathcal{L}}, M_f \rangle
\end{aligned}$$

for any f with $\deg(f) \leq 2\ell$.

Let $\widetilde{\mathbb{E}}$ be a p.e.f. of degree $d = 2\ell$ for some integer $\ell \geq 1$, satisfying the system $\{p_i(\mathbf{x}) = 0\}_{i=1}^m$. Let $X_{\widetilde{\mathbb{E}}}$ be the moment matrix of $\widetilde{\mathbb{E}}$ of degree 2ℓ (hence $X_{\widetilde{\mathbb{E}}}$ is a SoS-symmetric matrix of size $\binom{n}{\leq \ell}$).

We remark that $X_{\widetilde{\mathbb{E}}}$ satisfies the followings:

(i) $X[\emptyset, \emptyset] = 1$.

(ii) X is positive semidefinite and SoS-Symmetric.

(iii) $\langle M_f, X_{\widetilde{\mathbb{E}}} \rangle = 0$ for any $f \in \{\sum_{i=1}^m p_i q_i : \deg(q_i) \leq d - \deg(p_i)\ \forall i \in [m]\}$.

The converse is also true: If X satisfies (i), (ii) and (iii) then $X = X_{\widetilde{\mathbb{E}}}$ for some p.e.f. $\widetilde{\mathbb{E}}$ of degree d satisfying $\{p_i(\mathbf{x}) = 0\}_{i \in [m]}$.

Hence, (SoS_d) can be written as the following SDP:

$$\begin{aligned}
\text{maximize} \quad & \langle M_{p_0}, X \rangle \\
\text{subject to} \quad & X_{\emptyset,\emptyset} = 1, \\
& \langle M_q, X \rangle = 0 \text{ for all } q \in \mathcal{B}, \\
& X \succeq 0, \text{ and } X \text{ is SoS-symmetric,}
\end{aligned} \qquad (\mathsf{SDP}_d)$$

where $\mathcal{B} = \{\mathbf{x}_S p_i : i \in [m],\ |S| \leq d - \deg(p_i)\}$.

2.6 Proof of Theorem 2.5

Let us recall the setting of the single-spiked 4-tensor model: We observe a 4-tensor

$$\mathbf{Y}_{\otimes} = \mathbf{x}_0^{\otimes 4} + \sigma \mathbf{W}$$

where $\mathbf{x}_0$ is a vector in $\{\pm 1\}^n$ satisfying $\mathbf{1}^T \mathbf{x}_0 = 0$, $\sigma > 0$ is a noise-scaling factor, and $\mathbf{W} = \mathsf{Sym}\, \mathbf{G}$ where $\mathbf{G}$ is a 4-tensor with i.i.d. $N(0,1)$ entries.

Let $f_{\otimes}(\mathbf{x}) = \langle \mathbf{Y}, \mathbf{x}^{\otimes 4} \rangle$. The maximum-likelihood estimator $\widehat{\mathbf{x}}_{ML}$ is the optimum solution of

$$\max_{\mathbf{x} \in \{\pm 1\}^n : \mathbf{1}^T \mathbf{x} = 0} f_{\otimes}(\mathbf{x}).$$

Let us consider the corresponding SoS relaxation of degree 4:

$$\begin{aligned} \text{maximize} \quad & \widetilde{\mathbb{E}}[f_{\otimes}] \\ \text{subject to} \quad & \widetilde{\mathbb{E}} \text{ is a degree 4 p.e.f. on } \{\pm 1\}^n \\ & \text{satisfying } \sum_{i=1}^{n} \mathbf{x}_i = 0. \end{aligned}$$

Let $\mathbb{E}_{U(\pm \mathbf{x}_0)}$ be the expectation operator of the uniform distribution on $\{\pm \mathbf{x}_0\}$. Note that $\mathbb{E}_{U(\pm \mathbf{x}_0)} = \frac{1}{2}(\delta_{\mathbf{x}_0} + \delta_{-\mathbf{x}_0})$ so

$$\mathbb{E}_{U(\pm \mathbf{x}_0)}[\mathbf{x}_S] = \begin{cases} (\mathbf{x}_0)_S & \text{if } |S| \text{ is even} \\ 0 & \text{if } |S| \text{ is odd.} \end{cases}$$

If $\mathbb{E}_{U(\pm \mathbf{x}_0)}$ is the optimal solution of the relaxation, then we can recover $\mathbf{x}_0$ up to a global sign flip from its quadratic moments. First we give an upper bound on σ for the SoS algorithm to achieve exact recovery, in the case of the single-spiked model.

Theorem 2.17 (Theorem 2.5, Achievablility). *If $\sigma \lesssim \frac{n}{\sqrt{\log n}}$, then $\widetilde{\mathbb{E}} = \mathbb{E}_{U(\{\pm \mathbf{x}_0\})}$ is the optimum solution for (2.6) with probability $1 - o(1)$.*

We can reduce Theorem 2.17 to the matrix version of the problem using a tensor flattening, as in [75]. Given a 4-tensor $\mathbf{Y}$, the canonical flattening of $\mathbf{Y}$ is defined as

$n^2 \times n^2$ matrix Y with entries $Y_{(i,j),(k,\ell)} = \mathbf{Y}_{ijk\ell}$. Note that

$$Y = \mathrm{vec}\left(\mathbf{x}_0^{\otimes 2}\right) \mathrm{vec}\left(\mathbf{x}_0^{\otimes 2}\right)^T + \sigma W$$

where $\mathrm{vec}\left(\mathbf{x}_0^{\otimes 2}\right)$ is the vectorization of $\mathbf{x}_0^{\otimes 2}$, and W is the flattening of $\mathbf{W}$. Note that this is an instance of $\mathbb{Z}_2$-synchronization model with Gaussian noises. It follows that with high probability the exact recovery is possible when $\sigma \lesssim \frac{n}{\sqrt{\log n}}$ (see Proposition 2.3 in [21]).

We complement this result by proving that this bound on σ is tight up to a multiple of poly-logarithmic factor.

Theorem 2.18 (Theorem 2.5, Impossibility). *Let $c > 0$ be a small constant. If $\sigma \geq n(\log n)^{1/2+c}$, then there exists a pseudo-expectation $\widetilde{\mathbb{E}}$ of degree 4 on the hypercube $\{\pm 1\}^n$ satisfying $\sum_{i=1}^n \mathbf{x}_i = 0$ such that $\widetilde{\mathbb{E}}[f_\otimes] > f_\otimes(\mathbf{x}_0)$ with probability $1 - o(1)$.*

We ask the performance of the SoS relaxation of degree 4 on the bisection-spiked 4-tensor model. Let us recall the setting of the bisection-spiked model: We observe a 4-tensor $\mathbf{Y}_\ominus$ such that

$$\mathbf{Y}_\ominus = \mathbf{x}_0^{\ominus 4} + \sigma \mathbf{W}$$

where $\mathbf{x}_0$ is a vector in $\{\pm 1\}^n$ with $\mathbf{1}^T \mathbf{x}_0 = 0$, $\sigma > 0$, and $\mathbf{W} = \mathsf{Sym}\,\mathbf{G}$ where $\mathbf{G}$ is a random 4-tensor with i.i.d. $N(0,1)$ entries. Let

$$f_\ominus(\mathbf{x}) = \left\langle \mathbf{Y}_\ominus, \mathbf{x}^{\ominus 4} \right\rangle,$$

and let us consider the corresponding SoS relaxation of degree 4:

$$\begin{aligned} &\text{maximize} && \widetilde{\mathbb{E}}[f_\ominus] \\ &\text{subject to} && \widetilde{\mathbb{E}} \text{ is a degree 4 p.e.f. on } \{\pm 1\}^n \\ & && \text{satisfying } \sum_{i=1}^n \mathbf{x}_i = 0. \end{aligned} \tag{2.18}$$

Theorem 2.19 (Theorem 2.7, rephrased). *Let $c > 0$ be a small constant. If $\sigma \geq$*

$n(\log n)^{1/2+c}$, *then there exists a pseudo-expectation* $\widetilde{\mathbb{E}}$ *of degree 4 on the hypercube* $\{\pm 1\}^n$ *satisfying* $\sum_{i=1}^n \mathbf{x}_i = 0$ *such that* $\widetilde{\mathbb{E}}[f_{\ominus}] > f_{\ominus}(\mathbf{x}_0)$ *with probability* $1 - o(1)$.

2.6.1 Proof of Theorem 2.18

Let us define $g_{\otimes}$ and $g_{\ominus}$ be

$$g_{\otimes}(\mathbf{x}) = \langle \mathbf{G}, \mathbf{x}^{\otimes 4} \rangle \quad \text{and} \quad g_{\ominus}(\mathbf{x}) = \langle \mathbf{G}, \mathbf{x}^{\ominus 4} \rangle .$$

Then,

$$f_{\otimes}(\mathbf{x}) = \langle \mathbf{x}_0^{\otimes 4}, \mathbf{x}^{\otimes 4} \rangle + \sigma g_{\otimes}(\mathbf{x}) \quad \text{and} \quad f_{\ominus}(\mathbf{x}) = \langle \mathbf{x}_0^{\ominus 4}, \mathbf{x}^{\ominus 4} \rangle + \sigma g_{\ominus}(\mathbf{x}).$$

Let $\widetilde{\mathbb{E}}$ be a p.e.f. of degree 4 on $\{\pm 1\}^n$ satisfying $\sum_{i=1}^n \mathbf{x}_i = 0$. Note that

$$\langle \mathbf{x}_0^{\otimes 4}, \mathbf{x}^{\otimes 4} \rangle = n^k (\frac{\mathbf{x}_0^T \mathbf{x}}{n})^4$$

and

$$\langle \mathbf{x}_0^{\ominus 4}, \mathbf{x}^{\ominus 4} \rangle = \frac{n^k}{2^{2(k-1)}} \left(1 + 6 \left(\frac{\mathbf{x}_0^T \mathbf{x}}{n} \right)^2 + \left(\frac{\mathbf{x}_0^T \mathbf{x}}{n} \right)^4 \right),$$

hence $\widetilde{\mathbb{E}}[f] \geq \sigma \widetilde{\mathbb{E}}[g]$ in either cases. We would like to construct a p.e.f. $\widetilde{\mathbb{E}}$ depending on $\mathbf{G}$ so that $\sigma \widetilde{\mathbb{E}}[g]$ exceeds $f(\mathbf{x}_0)$ with high probability.

Lemma 2.20. *There exists a p.e.f.* $\widetilde{\mathbb{E}}$ *of degree 4 on* $\{\pm 1\}^n$ *satisfying* $\sum_{i=1}^n \mathbf{x}_i = 0$ *which only depend on* $\mathbf{G}$ *such that*

$$\widetilde{\mathbb{E}}[g] \gtrsim \frac{n^3}{(\log n)^{1/2+o(1)}}$$

for either $g = g_{\ominus}$ *or* $g = g_{\otimes}$.

Proof of Theorem 2.18. Note that in either cases $g(\mathbf{x}_0)$ is a Gaussian random variable with variance $O(n^4)$. So, $|g(\mathbf{x}_0)| \lesssim n^2 \log n$ with probability $1 - o(1)$. Let $\widetilde{\mathbb{E}}$ be the pseudo-expectation satisfying the conditions in Lemma 2.20. Then, with probability

$1 - o(1)$ we have

$$\widetilde{\mathbb{E}}[f] - f(\mathbf{x}_0) \geq \; -\Theta(n^4) + \sigma\left(\widetilde{\mathbb{E}}[g] - |g(\mathbf{x}_0)|\right)$$
$$\geq \; -\Theta(n^4) + \sigma\left(\frac{n^3}{(\log n)^{1/2+o(1)}} - O(n^2 \log n)\right).$$

When $\sigma > n(\log n)^{1/2+c}$ for some constant $c > 0$, we have

$$\sigma\left(\frac{n^3}{(\log n)^{1/2+o(1)}}\right) \gg n^4,$$

hence $\widetilde{\mathbb{E}}[f] - f(\mathbf{x}_0) > 0$. □

In the remainder of the section, we prove Lemma 2.20.

Outline

We note that our method shares a similar idea which appears in [58] and [26].

Let us consider the case that $g = g_\otimes$, *i.e.*, $g(\mathbf{x}) = \langle \mathbf{G}, \mathbf{x}^{\otimes 4}\rangle$ where $\mathbf{G}$ has independent standard Gaussian entries. We would like to construct $\widetilde{\mathbb{E}} = \widetilde{\mathbb{E}}_{\mathbf{G}}$ which has large correlation with $\mathbf{G}$. If we simply let

$$\widetilde{\mathbb{E}}[\mathbf{x}_{i_1}\mathbf{x}_{i_2}\mathbf{x}_{i_3}\mathbf{x}_{i_4}] = \frac{1}{24}\sum_{\pi \in \mathfrak{S}_4} \mathbf{G}_{i_{\pi(1)}, i_{\pi(2)}, i_{\pi(3)}, i_{\pi(4)}}$$

for $\{i_1 < i_2 < i_3 < i_4\} \subseteq [n]$ and $\widetilde{\mathbb{E}}[\mathbf{x}_T]$ be zero if $|T| \leq 3$, then

$$\widetilde{\mathbb{E}}[g] = \frac{1}{24}\sum_{1 \leq i_1 < i_2 < i_3 < i_4 \leq n}\left(\sum_{\pi \in \mathfrak{S}_4} \mathbf{G}_{i_{\pi(1)}, i_{\pi(2)}, i_{\pi(3)}, i_{\pi(4)}}\right)^2$$

so the expectation of $\widetilde{\mathbb{E}}[g]$ over $\mathbf{G}$ would be equal to $\binom{n}{4} \approx \frac{n^4}{24}$. However, in this case $\widetilde{\mathbb{E}}$ does not satisfies the equality $\sum_{i=1}^n \mathbf{x}_i = 0$ nor the conditions for pseudo-expectations.

To overcome this, we first project the $\widetilde{\mathbb{E}}$ constructed above to the space of linear functionals which satisfy the equality constraints ($\mathbf{x}_i^2 = 1$ and $\mathbf{1}^T\mathbf{x} = 0$). Then, we take a convex combination of the projection and a pseudo-expectation to control the spectrum of the functional. More specifically, we take the following steps to construct a pseudoexpectation functional $\widetilde{\mathbb{E}}$ of degree 4.

(1) **Removing degeneracy.** We first establish the one-to-one correspondence between the collection of linear functionals on n-variate even multilinear polynomials of degree at most 4 and the collection of linear functionals on $(n-1)$-variate multilinear polynomials of degree at most 4 by posing $\mathbf{x}_n = 1$. We observe that this correspondence preserves positivity.

(2) **Description of the linear constraint $\mathbf{1}^T\mathbf{x} = 0$.** Let ψ be a linear functional on $(n-1)$-variate multilinear polynomials of degree at most 4. We may think ψ as a vector in $\mathbb{R}^{\binom{n-1}{\leq 4}}$. Then, we can write the condition that the functional ψ satisfies $\sum_{i=1}^{n-1} \mathbf{x}_i + 1 = 0$ as $A\psi = 0$ for some matrix A.

(3) **Projection.** Let $w \in \mathbb{R}^{\binom{n-1}{\leq 4}}$ be the coefficient vector of $g(\mathbf{x})$. Let Π be the projection matrix to the space $\{x : Ax = 0\}$. In other words,

$$\Pi = \mathsf{Id}_{\binom{n-1}{\leq 4}} - A^T(AA^T)^\dagger A$$

where $(\cdot)^\dagger$ denotes the pseudo-inverse. Let e be the first column of Π and $\psi_1 = \frac{\Pi w}{e^T w}$. Then $(\psi_1)_\emptyset = 1$ and $A\psi_1 = 0$ by definition.

(4) **Convex combination.** Let $\psi_0 = \frac{e}{e^T e}$. We note that ψ_0 corresponds to the expectation operator of uniform distribution on $\{x \in \{\pm 1\}^n : \mathbf{1}^T x = 0\}$.

We will construct ψ by

$$\psi = (1-\epsilon)\psi_0 + \epsilon\psi_1$$

with an appropriate constant ϵ. Equivalently,

$$\psi = \psi_0 + \frac{\epsilon}{e^T w} \cdot \left(\Pi - \frac{ee^T}{e^T e}\right) w.$$

(5) **Spectrum analysis.** We bound the spectrum of the functional $\left(\Pi - \frac{ee^T}{e^T e}\right) w$ to decide the size of ϵ for ψ being positive semidefinite.

We describe each step in detail in the remainder of this section.

Step 1: Removing degeneracy.

We remark that either $g =$ Observe that g is even, *i.e.*, $g(\mathbf{x}) = g(-x)$ for any $\mathbf{x} \in \{\pm 1\}^n$. To maximize such an even function, we claim that we may only consider the pseudo-expectations such that whose odd moments are zero.

Proposition 2.21. *Let $\widetilde{\mathbb{E}}$ be a pseudo-expectation of degree 4 on hypercube satisfying $\sum_{i=1}^n x_i = 0$. Let p be a degree 4 multilinear polynomial which is even. Then, there exists a pseudo-expectation $\widetilde{\mathbb{E}}'$ of degree 4 such that $\widetilde{\mathbb{E}}[p] = \widetilde{\mathbb{E}}'[p]$ and $\widetilde{\mathbb{E}}'[x_S] = 0$ for any $S \subseteq [n]$ of odd size.*

Proof. Let $\widetilde{\mathbb{E}}$ be a pseudo-expectation of degree 4 on hypercube satisfying $\sum_{i=1}^n \mathbf{x}_i = 0$. Let us define a linear functional $\widetilde{\mathbb{E}}_0$ on the space of multilinear polynomials of degree at most 4 so that $\widetilde{\mathbb{E}}_0[\mathbf{x}_S] = (-1)^{|S|}\,\widetilde{\mathbb{E}}[x_S]$ for any $S \in \binom{[n]}{\leq 4}$. Then, for any multilinear polynomial q of degree at most 2, we have

$$\widetilde{\mathbb{E}}_0[q(\mathbf{x})^2] = \widetilde{\mathbb{E}}[q(-x)^2] \geq 0.$$

Also, $\widetilde{\mathbb{E}}_0$ satisfies $\widetilde{\mathbb{E}}_0[1] = 1$ and

$$\widetilde{\mathbb{E}}_0\left[\left(\sum_{i=1}^n \mathbf{x}_i\right) q(\mathbf{x})\right] = -\,\widetilde{\mathbb{E}}\left[\left(\sum_{i=1}^n \mathbf{x}_i\right) q(-\mathbf{x})\right] = 0$$

for any q of degree 3. Thus, $\widetilde{\mathbb{E}}_0$ is a valid pseudo-expectation of degree 4 satisfying $\sum_{i=1}^n x_i = 0$.

Let $\widetilde{\mathbb{E}}' = \frac{1}{2}(\widetilde{\mathbb{E}} + \widetilde{\mathbb{E}}_0)$. This is again a valid pseudo-expectation, since the space of pseudo-expectations is convex. We have $\widetilde{\mathbb{E}}'[p(\mathbf{x})] = \widetilde{\mathbb{E}}[p(\mathbf{x})] = \widetilde{\mathbb{E}}_0[p(\mathbf{x})]$ since p is even, and $\widetilde{\mathbb{E}}'[\mathbf{x}_S] = (1 + (-1)^{|S|})\,\widetilde{\mathbb{E}}[\mathbf{x}_S] = 0$ for any S of odd size. $\square$

Let $\mathcal{E}$ be the space of all pseudo-expectations of degree 4 on n-dimensional hypercube with zero odd moments. Let $\mathcal{E}'$ be the space of all pseudo-expectations of degree

4 on $(n-1)$-dimensional hypercube. We claim that there is a bijection between two spaces.

Proposition 2.22. *Let $\psi \in \mathcal{E}$. Let us define a linear functional ψ' on the space of $(n-1)$-variate multilinear polynomials of degree at most 4 so that for any $T \subseteq [n-1]$ with $|T| \leq 4$*

$$\psi'[\mathbf{x}_T] = \begin{cases} \psi[\mathbf{x}_{T\cup\{n\}}] & \text{if } |T| \text{ is odd} \\ \psi[\mathbf{x}_T] & \text{otherwise.} \end{cases}$$

Then, $\psi \mapsto \psi'$ is a bijective mapping from $\mathcal{E}$ to $\mathcal{E}'$.

Proof. We say linear functional ψ on the space of polynomials of degree at most 2ℓ is *positive semidefinite* if $\psi[q^2] \geq 0$ for any q of degree ℓ.

Note that the mapping $\psi' \mapsto \psi$ where $\psi[\mathbf{x}_S] = \psi'[\mathbf{x}_{S\setminus\{n\}}]$ for any $S \subseteq [n]$ of even size is the inverse of $\psi \mapsto \psi'$. Hence, it is sufficient to prove that ψ is positive semidefinite if and only if ψ' is positive semidefinite.

($\Rightarrow$) Let q be an n-variate polynomial of degree at most 2. Let q_0 and q_1 be polynomials in $x_1, \cdots, x_{n-1}$ such that

$$q(\mathbf{x}_1, \cdots, \mathbf{x}_n) = q_0(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}) + x_n q_1(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}).$$

We get $\psi'[q^2] = \psi'[(q_0 + \mathbf{x}_n q_1)^2] = \psi'[(q_0^2 + q_1^2) + 2\mathbf{x}_n q_0 q_1]$. For $i = 1, 2$, let q_{i0} and q_{i1} be the even part and the odd part of q_i, respectively. Then we have

$$\begin{aligned} \psi'[q^2] &= \psi'[(q_{00}^2 + q_{01}^2 + q_{10}^2 + q_{11}^2) + 2x_n(q_{00}q_{11} + q_{01}q_{10})] \\ &= \psi[(q_{00}^2 + q_{01}^2 + q_{10}^2 + q_{11}^2) + 2(q_{00}q_{11} + q_{01}q_{10})] \\ &= \psi[(q_{00} + q_{11})^2 + (q_{10} + q_{01})^2] \geq 0. \end{aligned}$$

The first equality follows from that $\psi'[q] = 0$ for odd q. Hence, ψ' is positive semidefinite.

($\Leftarrow$) Let q be an $(n-1)$-variate polynomial of degree at most 2. Let q_0 and q_1 be the even part and the odd part of q, respectively. Then,

$$\psi[q^2] = \psi[(q_0^2 + q_1^2) + 2q_0q_1].$$

Note that $q_0^2 + q_1^2$ is even and q_0q_1 is odd. So,

$$\psi[q^2] = \psi'[(q_0^2 + q_1^2) + 2\mathbf{x}_n q_0 q_1] = \psi'[(q_0 + \mathbf{x}_n q_1)^2] \geq 0.$$

Hence ψ is positive semidefinite. $\square$

In addition to the proposition, we note that ψ satisfies $\sum_{i=1}^{n} \mathbf{x}_i = 0$ if and only if ψ' satisfies $1 + \sum_{i=1}^{n-1} \mathbf{x}_i = 0$. Hence, maximizing $\widetilde{\mathbb{E}}[g]$ over $\widetilde{\mathbb{E}} \in \mathcal{E}$ satisfying $\sum_{i=1}^{n} \mathbf{x}_i = 0$ is equivalent to

$$\max_{\psi' \in \mathcal{E}'} \psi'[g'] \quad \text{subject to} \quad \psi' \text{ satisfies } 1 + \sum_{i=1}^{n-1} \mathbf{x}_i = 0,$$

where $g'(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}) = g(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}, 1)$.

Step 2: A matrix description of the constraint $\sum_{i=1}^{n-1} \mathbf{x}_i + 1 = 0$.

Let $\mathcal{F}$ be the set of linear functional on the space of $(n-1)$-variate multilinear polynomials of degree at most 4. We often regard a functional $\psi \in \mathcal{F}$ as a $\binom{n-1}{\leq 4}$ dimensional vector with entries $\psi_S = \psi[\mathbf{x}_S]$ where S is a subset of $[n-1]$ of size at most 4. The space $\mathcal{E}'$ of pseudo-expectations of degree 4 (on $(n-1)$-variate multilinear polynomials) is a convex subset of $\mathcal{F}$.

Observe that $\psi \in \mathcal{F}$ satisfies $1 + \sum_{i=1}^{n-1} \mathbf{x}_i = 0$ if and only if

$$\psi\left[\left(1 + \sum_{i=1}^{n-1} \mathbf{x}_i\right) \mathbf{x}_S\right] = 0$$

for any $S \subseteq [n-1]$ with $|S| \leq 3$.

Let s, t and u be integers such that $0 \leq s, t \leq 4$ and $0 \leq u \leq \min(s,t)$. Let $M_{s,t}^{u}$

be the matrix of size $\binom{n-1}{\leq 4}$ such that

$$(M^u_{s,t})_{S,T} = \begin{cases} 1 & \text{if } |S| = s,\ |T| = t, \text{ and } |S \cap T| = u \\ 0 & \text{otherwise} \end{cases}$$

for $S, T \in \binom{[n-1]}{\leq 4}$. Then, the condition that $\psi \in \mathcal{F}$ satisfying $1 + \sum_{i=1}^{n-1} \mathbf{x}_i = 0$ can be written as $A\psi = 0$ where

$$A = M^0_{0,0} + M^0_{0,1} + \sum_{s=1}^{3} (M^{s-1}_{s,s-1} + M^s_{s,s} + M^s_{s,s+1}).$$

Before we move onto the projection step, let us take a brief look into the algebra generated by the matrices $M^u_{s,t}$ for $0 \leq s, t \leq 4$ and $0 \leq u \leq \min(s,t)$.

Step 2.5: Algebra generated by $M^u_{s,t}$

Let m be a positive integer greater than 8. For nonnegative integers s, t, u, let $M^u_{s,t}$ be the $\binom{m}{\leq 4} \times \binom{m}{\leq 4}$ matrix with

$$(M^u_{s,t})_{S,T} = \begin{cases} 1 & \text{if } |S| = s,\ |T| = t, \text{ and } |S \cap T| = u \\ 0 & \text{otherwise,} \end{cases}$$

for $S, T \subseteq [m]$ with $|S|, |T| \leq 4$. Let $\mathcal{A}$ be the algebra of matrices

$$\sum_{0 \leq s,t \leq 4} \sum_{u=0}^{s \wedge t} x^u_{s,t} M^u_{s,t}$$

with complex numbers $x^u_{s,t}$. This algebra $\mathcal{A}$ is a C^*-algebra: it is a complex algebra which is closed under taking complex conjugate. $\mathcal{A}$ is a subalgebra of the Terwilliger algebra of the Hamming cube $H(m, 2)$ [90], [87].

Note that $\mathcal{A}$ has dimension 55 which is the number of triples (s, t, u) with $0 \leq s, t \leq 4$ and $0 \leq u \leq s \wedge t$.

Define

$$\beta^u_{s,t,r} := \sum_{p=0}^{s\wedge t}(-1)^{p-t}\binom{p}{u}\binom{m-2r}{p-r}\binom{m-r-p}{s-p}\binom{m-r-p}{t-p}$$

for $0 \le s,t \le 4$ and $0 \le r,u \le s\wedge t$. The following theorem says that matrices in the algebra $\mathcal{A}$ can be written in a block-diagonal form with small sized blocks.

Theorem 2.23 ([87]). *There exists an orthogonal $\binom{m}{\le 4} \times \binom{m}{\le 4}$ matrix U such that for $M \in \mathcal{A}$ with*

$$M = \sum_{s,t=0}^{4}\sum_{u=0}^{s\wedge t} x^u_{s,t} M^u_{s,t},$$

the matrix $U^T M U$ is equal to the matrix

$$\begin{pmatrix} C_0 & 0 & 0 & 0 & 0 \\ 0 & C_1 & 0 & 0 & 0 \\ 0 & 0 & C_2 & 0 & 0 \\ 0 & 0 & 0 & C_3 & 0 \\ 0 & 0 & 0 & 0 & C_4 \end{pmatrix}$$

where each C_r is a block diagonal matrix with $\binom{m}{r} - \binom{m}{r-1}$ repeated, identical blocks of order $5-r$:

$$C_r = \begin{pmatrix} B_r & 0 & \cdots & 0 \\ 0 & B_r & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_r \end{pmatrix},$$

and

$$B_r = \left(\sum_u \binom{m-2r}{s-r}^{-1/2}\binom{m-2r}{t-r}^{-1/2}\beta^u_{s,t,r}x^u_{s,t}\right)^4_{s,t=r}.$$

For brevity, let us denote this block-diagonalization of M by the tuple of matrices $(B_0, B_1, B_2, B_3, B_4)$ where $B_r \in \mathbb{R}^{(5-r)\times(5-r)}$.

Step 3: Projection.

Let $c_\otimes$ be the coefficient vector of $g'_\otimes$ where

$$g'_\otimes(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}) = g_\otimes(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}, 1).$$

Then, we get

$$(c_\otimes)_S = \frac{1}{N_{|S|}} \sum_{\substack{(i_1,\cdots,i_4)\in[n]^4: \\ S=\{i_1\}\oplus\cdots\oplus\{i_k\}}} \mathbf{G}_{i_1,\cdots,i_4}$$

for $S \in \binom{[n-1]}{0} \cup \binom{[n-1]}{2} \cup \binom{[n-1]}{4}$, and

$$(c_\otimes)_S = \frac{1}{N_{|S|}} \sum_{\substack{(i_1,\cdots,i_4)\in[n]^4: \\ S\cup\{n\}=\{i_1\}\oplus\cdots\oplus\{i_k\}}} \mathbf{G}_{i_1,\cdots,i_4}$$

for $S \in \binom{[n-1]}{1} \cup \binom{[n-1]}{3}$, where $N_0 = 3n^2 - 2n$, $N_1 = N_2 = 12n - 16$, and $N_3 = N_4 = 24$. Note that $c_\otimes$ has independent entries, since

$$\{(i_1, \cdots, i_4) \in [n]^4 : S = \{i_1\} \oplus \cdots \oplus \{i_4\}\}$$

for $|S| = 0, 2, 4$ and

$$\{(i_1, \cdots, i_4) \in [n]^4 : S \cup [n] = \{i_1\} \oplus \cdots \oplus \{i_4\}\}$$

for $|S| = 1, 3$ forms a partition on $[n]^4$ and the entries of $\mathbf{G}$ are independent. Hence, the covariance matrix of $c_\otimes$ is diagonal and invertible.

Let Σ be the inverse of the covariance matrix $\mathbb{E}[c_\otimes c_\otimes^T]$ of $c_\otimes$. We note that $\Sigma_{S,S} = N_{|S|}$ since $(c_\otimes)_S$ is the average of $N_{|S|}$ independent standard Gaussians, so its variance is equal to $N_{|S|}^{-1}$. Let $w = \Sigma^{1/2} c_\otimes$. Clearly,

$$\mathbb{E}\, ww^T = \mathbb{E}\, \Sigma^{1/2} c_\otimes c_\otimes^T \Sigma^{1/2} = \Sigma^{1/2} \Sigma^{-1} \Sigma^{1/2} = \mathsf{Id}.$$

Hence, w has independent standard Gaussian entries.

Let $\Pi = \mathsf{Id} - A^T(AA^T)^\dagger A$ where $(AA^T)^\dagger$ is the Moore-Penrose pseudoinverse of AA^T and Id is the identity matrix of size $\binom{n-1}{\leq 4}$. Then, Π is the orthogonal projection matrix onto the nullspace of A. Since A, A^T and Id are all in the algebra $\mathcal{A}$, the projection matrix Π is also in $\mathcal{A}$.

Let e be the first column of Π and

$$\psi_0 := \frac{e}{e^T e} \quad \text{and} \quad \psi_1 := \frac{\Pi w}{e^T w}.$$

We have $A\psi_0 = A\psi_1 = 0$ by definition of Π, and $(\psi_0)_\emptyset = (\psi_1)_\emptyset = 1$ since $(\Pi w)_\emptyset = e^T w$.

Let ϵ be a real number with $0 < \epsilon < 1$ and $\psi = (1-\epsilon)\psi_0 + \epsilon\psi_1$. This functional still satisfies $A\psi = 0$ and $\psi_\emptyset = 1$, regardless of the choice of ϵ. We would like to choose ϵ such that ψ is positive semidefinite with high probability.

Step 4: Analysis on the spectrum of ψ.

Consider the functional $\psi_0 = \frac{e}{e^T e}$. It has entries

$$(\psi_0)_S = \begin{cases} 1 & \text{if } S = \emptyset \\ -\frac{1}{n-1} & \text{if } |S| = 1 \text{ or } 2 \\ \frac{3}{(n-1)(n-3)} & \text{if } |S| = 3 \text{ or } 4, \end{cases}$$

for $S \subseteq [n-1]$ of size at most 4. We note that this functional corresponds to the degree 4 or less moments of the uniform distribution on the set of vectors $\mathbf{x} \in \{\pm 1\}^{n-1}$ satisfying $\sum_{i=1}^{n-1} \mathbf{x}_i + 1 = 0$.

Proposition 2.24. *Let ψ be a vector in $\mathbb{R}^{\binom{n-1}{\leq 4}}$ such that $A\psi = 0$ and p be an $(n-1)$-variate multilinear polynomial of degree at most 2. Suppose that $\psi_0[p^2] = 0$. Then, $\psi[p^2] = 0$.*

Proof. Let $\mathcal{U} = \{\mathbf{x} \in \{\pm 1\}^{n-1} : \sum_{i=1}^{n-1} \mathbf{x}_i + 1 = 0\}$. Note that ψ_0 is the expectation functional of the uniform distribution on $\mathcal{U}$ as we seen above. Hence, $\psi_0[p^2] = 0$ if and only if $p(\mathbf{x})^2 = 0$ for any $x \in \mathcal{U}$.

On the other hand, the functional ψ is a linear combination of functionals $\{p \mapsto p(\mathbf{x}) : x \in \mathcal{U}\}$ since $A\psi = 0$. Hence, if $\psi_0[p^2] = 0$ then $\psi[p^2] = 0$ as $p(\mathbf{x})^2 = 0$ for any $x \in \mathcal{U}$. $\square$

Recall that $\psi = (1-\epsilon)\psi_0 + \epsilon\psi_1$ where $\psi_0 = \frac{e}{e^Te}$ and $\psi_1 = \frac{\Pi w}{e^Tw}$. Let $\psi_1' = e^Tw \cdot (\psi_1 - \psi_0)$. Then,

$$\begin{aligned}\psi_1' &= \Pi w - \frac{e^Tw}{e^Te}e \\ &= \left(\Pi - \frac{ee^T}{e^Te}\right)w\end{aligned}$$

and $\psi = \psi_0 + \frac{\epsilon}{e^Tw}\psi_1'$. We note that $A\psi_1' = 0$ since ψ_1' is a linear combination of ψ_0 and ψ_1.

Let X_{ψ_0} and $X_{\psi_1'}$ be the moment matrix of ψ_0 and ψ_1' respectively. Let X_ψ be the moment matrix of ψ. Clearly,

$$X_\psi = X_{\psi_0} + \frac{\epsilon}{e^Tw}X_{\psi_1'}.$$

Moreover, for any $p \in \mathbb{R}^{\binom{n-1}{\leq 2}}$ satisfying $X_{\psi_0}p = 0$, we have $X_{\psi_1'}p = 0$ by the proposition. Hence, $X_\psi \succeq 0$ if

$$\frac{\epsilon}{|e^Tw|}\|X_{\psi_1'}\| \leq \lambda_{\min,\neq 0}(X_{\psi_0})$$

where $\lambda_{\min,\neq 0}$ denotes the minimum nonzero eigenvalue.

We note that e^Tw and $\|X_{\psi_1'}\|$ are independent random variables. It follows from that w is a gaussian vector with i.i.d. standard entries, and that e and $\left(\Pi - \frac{ee^T}{e^Te}\right)$ are orthogonal. Hence, we can safely bound e^Tw and $\|X_{\psi_1'}\|$ separately.

To bound $\|X_{\psi_1'}\|$ we need the following theorem.

Theorem 2.25 (Matrix Gaussian ([91])). *Let $\{A_k\}$ be a finite sequence of fixed, symmetric matrices with dimension d, and let $\{\xi_k\}$ be a finite sequence of independent standard normal random variables. Then, for any $t \geq 0$,*

$$\mathbb{P}\left[\left\|\sum_k \xi_k A_k\right\| \geq t\right] \leq d \cdot e^{-t^2/2\sigma^2} \quad \textit{where} \quad \sigma^2 := \left\|\sum_k A_k^2\right\|.$$

For each $U \subseteq [n-1]$ with size at most 4, let Y_U be the $\binom{n-1}{\leq 2} \times \binom{n-1}{\leq 2}$ matrix with entries

$$(Y_U)_{S,T} = \begin{cases} 1 & \text{if } S \oplus T = U \\ 0 & \text{otherwise.} \end{cases}$$

We can write $X_{\psi_1'}$ as

$$X_{\psi_1'} = \sum_{\substack{U \subseteq [n-1] \\ |U| \leq 4}} (\psi_1')_U Y_U.$$

Since $\psi_1' = \left(\Pi - \frac{ee^T}{e^T e}\right) w$, we have

$$\begin{aligned} X_{\psi_1'} &= \sum_{\substack{U \subseteq [n-1] \\ |U| \leq 4}} \sum_{\substack{V \subseteq [n-1] \\ |V| \leq 4}} \left(\Pi - \frac{ee^T}{e^T e}\right)_{U,V} w_V Y_U \\ &= \sum_V w_V \left(\sum_U \left(\Pi - \frac{ee^T}{e^T e}\right)_{U,V} Y_U \right). \end{aligned}$$

By Theorem 2.25, $\|X_{\psi_1'}\|$ is roughly bounded by $(\|\Sigma_X\| \log n)^{1/2}$ where

$$\Sigma_X := \sum_V \left(\sum_U \left(\Pi - \frac{ee^T}{e^T e}\right)_{U,V} Y_U \right)^2.$$

Proposition 2.26. *For each $I, J \in \binom{[n-1]}{\leq 2}$, the (I, J) entry of Σ_X only depends on $|I|$, $|J|$ and $|I \cap J|$,* i.e., *Σ_X is in the algebra $\mathcal{A}$.*

Proof. Note that

$$\begin{aligned} \Sigma_X &= \sum_V \sum_{U_1, U_2} \left(\Pi - \frac{ee^T}{e^T e}\right)_{U_1,V} \left(\Pi - \frac{ee^T}{e^T e}\right)_{V,U_2} Y_{U_1} Y_{U_2} \\ &= \sum_{U_1, U_2} \left(\left(\Pi - \frac{ee^T}{e^T e}\right)^2 \right)_{U_1,U_2} Y_{U_1} Y_{U_2} \\ &= \sum_{U_1, U_2} \left(\Pi - \frac{ee^T}{e^T e}\right)_{U_1,U_2} Y_{U_1} Y_{U_2}. \end{aligned}$$

Hence,

$$(\Sigma_X)_{I,J} = \sum_{K \in \binom{[n-1]}{\leq 2}} \left(\Pi - \frac{ee^T}{e^T e} \right)_{I \oplus K, J \oplus K},$$

which is invariant under any permutation π on $[n-1]$ as $\Pi - \frac{ee^T}{e^Te}$ is. It implies that $\Sigma_X \in \mathcal{A}$. $\square$

Since $\left\| \Pi - \frac{ee^T}{e^Te} \right\| = 1$, we get $\|\Sigma_X\| \leq \binom{n-1}{2} = \left(\frac{1}{2} + o(1)\right) n^2$.

Proposition 2.27. *If $\epsilon < o_n\left(\frac{1}{n\sqrt{\log n}}\right)$, then with probability $1 - o(1)$ the moment matrix X_ψ is positive-semidefinite.*

Proof. By theorem 2.25, we have

$$\mathbb{P}\left(\|X_{\psi_1'}\| \geq t\right) \leq \binom{n-1}{\leq 2} \cdot e^{-t^2/2\|\Sigma_X\|}.$$

Let $t = 3n\sqrt{\log n}$. Since $\|\Sigma_X\| \leq (1/2 + o(1))n^2$, we have $\|X_{\psi_1'}\| \leq 3n\sqrt{\log n}$ with probability $1 - n^{-\Omega(1)}$. On the other hand, note that

$$\mathbb{P}\left(|e^T w| \leq t\right) \leq \frac{t}{\sqrt{2\pi}}.$$

It implies that $|e^T w| > g(n)$ with probability $1 - o(1)$ for any $g(n) = o_1(1)$. Thus,

$$\frac{\|X_{\psi_1'}\|}{|e^T w|} \lesssim \frac{n\sqrt{\log n}}{g(n)}$$

almost asymptotically surely. Together with the fact that $\lambda_{\min, \neq 0}(X_{\psi_0}) = 1 - o(1)$, we have $X_\psi \succeq 0$ whenever $\epsilon < \frac{g(n)}{n\sqrt{\log n}}$ for some $g(n) = o(1)$. $\square$

Step ∞: Putting it all together.

We have constructed a linear functional ψ on the space of $(n-1)$-variate multilinear polynomials of degree at most 4, which satisfies (i) $\psi[1] = 1$, (ii) ψ satisfies $\sum_{i=1}^{n-1} \mathbf{x}_i + 1 = 0$, and (iii) $\psi[p^2] \geq 0$ for any p of degree 2. It implies that ψ is a valid pseudo-expectation of degree 4.

Now, let us compute $\psi[g'_{\otimes}]$ and $\psi[g'_{\ominus}]$ where

$$g'_{\otimes}(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}) = g_{\otimes}(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}, 1)$$

and

$$g'_{\ominus}(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}) = g_{\ominus}(\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}, 1).$$

Recall that $c_{\otimes}$ is the coefficient vector of $g'_{\otimes}$. We also define $c_{\ominus}$ to be the coefficient vector of $g'_{\ominus}$.

Hence,

$$\begin{aligned}
\psi[g'_{\otimes}] = c_{\otimes}^T \psi &= w^T \Sigma^{1/2} \left(\frac{e}{e^T e} + \frac{\epsilon}{e^T w} \left(\Pi - \frac{ee^T}{e^T e} \right) w \right) \\
&= \frac{e^T \Sigma^{1/2} w}{e^T e} + \epsilon \cdot \frac{w^T \Sigma^{1/2} \left(\Pi - \frac{ee^T}{e^T e} \right) w}{e^T w}.
\end{aligned}$$

Note that

$$\begin{aligned}
\mathbb{E}\left[w^T \Sigma^{1/2} \left(\Pi - \frac{ee^T}{e^T e} \right) w \right] &= \left\langle \Sigma^{1/2} \left(\Pi - \frac{ee^T}{e^T e} \right), \mathbb{E}[ww^T] \right\rangle \\
&= \operatorname{tr}\left(\Sigma^{1/2} \left(\Pi - \frac{ee^T}{e^T e} \right) \right),
\end{aligned}$$

which is at least $(\sqrt{6}/12 - o(1))n^4$. Also, $|e^T w| = O(1)$ and $|e^T \Sigma^{1/2} w| = O(n)$ with high probability. Hence, with probability $1 - o(1)$, we have

$$\psi[g'_{\otimes}] \gtrsim O(n) + \frac{n^4}{n(\log n)^{1/2+o(1)}} \gtrsim \frac{n^3}{(\log n)^{1/2+o(1)}}.$$

We can compute $\psi[g'_{\ominus}]$ in the same fashion, as $c_{\ominus}$ can be expressed as Bw for some matrix $B \in \mathcal{A}$. As a result, we get

$$\psi[g'_{\ominus}] \gtrsim \frac{n^3}{(\log n)^{1/2+o(1)}}$$

as well as $\psi[g'_{\otimes}]$.

Chapter 3

Stochastic Block Model for k-uniform Hypergraphs[1]

Identifying clusters from relational data is one of fundamental problems in computer science. It has many applications such as analyzing social networks [80], detecting protein-protein interactions [71, 32], finding clusters in Hi-C genomic data [29], image segmentation [88], recommendation systems [70, 86] and many others. The goal is to find a community structure from relational measurements between data points.

Although many clustering problems are known to be NP-hard, typical data we encounter in applications are very different from the worst-case instances. This motivates us to study probabilistic models and average-case complexity for them. The stochastic block model (SBM) is one such model that has received much attention in the past few decades. In the SBM, we observe a random graph on the finite set of nodes where each pair of nodes is independently joined by an edge with probability only depending on the community membership of the endpoints.

It is natural to consider the community detection problem for higher-order relations. A number of authors have already considered problems of learning from complex relational data [13, 54, 12] and it has several applications such as folksonomy [48, 96], computer vision [54], and network alignment problems for protein-protein

[1]The contents of this chapter overlaps in significant amount with [62], which is a collaborative work mainly conducted by the author of this thesis.

interactions [73].

We consider a version of SBM for higher-order relations, which we call *the stochastic block model for k-uniform hypergraph* (k-HSBM): we observe a random k-uniform hypergraph such that each set of nodes of size k appears independently as an (hyper-)edge with probability only depending on the community labels of nodes in it. k-HSBM was first introduced in [49] and investigated for its statistical limit in terms of detection [68], the minimax misclassification ratio [69, 34], and as a testbed for algorithms including naive spectral method [50, 51, 52], spectral method along with local refinements [1, 34, 15] and approximate-message passing algorithms [20, 68].

We focus on exact recovery, where our goal is to fully recover the community labels of the nodes from a random k-uniform hypergraph drawn from the model. For exact recovery, the maximum a posteriori (MAP) estimator always outperforms any other estimators in the sense that it has the highest probability of correctly recovering the solution. We prove that for the k-HSBM with two equal-sized and symmetric communities, exact recovery shows a sharp phase transition behavior, and moreover, the threshold can be characterized by the success of a certain type of local refinement. This type of phenomenon was mentioned as "local-to-global amplification" in [1], and was proved in [4] for the usual SBM with two symmetric communities (corresponds to 2-HSBM) and more generally in [6] for SBMs with fixed number of communities. Our result can be regarded as a direct generalization of [6] to k-uniform hypergraphs.

Furthermore, we analyze the truncate-and-relax algorithm on k-HSBM and prove that it achieves exact recovery in a parameter regime which is orderwise optimal.

We remark that in [1] it was suggested that the local refinement methods together with an efficient partial recovery algorithm would imply the efficient exact recovery up to the information-theoretic threshold. An explicit algorithm exploiting this idea appears in [34, 15] with a provable threshold for their algorithm to be successful. We note that the threshold of the algorithm of [34] matches with the statistical threshold we derive, hence there is no gap between statistical and computational thresholds. On the other hand, we prove that the truncate-and-relax algorithm does not achieve the exact statistical threshold when $k \geq 4$.

3.1 Introduction

3.1.1 The Stochastic Block Model for graphs: An overview

Before we discuss the main topic of this chapter, let us summarize the previous works regarding the stochastic block model for graphs.

The stochastic block model (SBM) has been one of the most fruitful research topics in community detection and clustering. One benefit of it is that, being a generative model we can formally study the probability of inferring the ground truth. While data from the real-world can behave differently, the SBM is believed to provide good insights in the field of community detection and has been studied for its sharp phase transition behavior [77, 6, 4], computational vs. information-theoretic gaps [33, 8], and as a test bed for various algorithms such as spectral methods [72, 94], semidefinite programs [4, 56, 59], belief-propagation methods [38, 7, 9], and approximate message-passing algorithms [93, 30, 40, 67]. We recommend [1] for a survey of this topic.

For the sake of exposition, let us consider the symmetric SBM with two equal-sized clusters, also known as the planted bisection model. Let n be a positive integer, and let p and q be real numbers in $[0, 1]$. The planted bisection model with parameter n, p and q is a generative model which outputs a random graph G on n vertices such that (i) the bipartition (A, B) of V defining two equal-sized clusters is chosen uniformly at random, and (ii) each pair $\{u, v\}$ in V is connected independently with probability p if u and v are in the same cluster, or probability q otherwise. Note that this model coincides with Erdős-Rényi random graph model $\mathcal{G}(n, p)$, when p and q are equal.

The goal is to find the ground truth (A, B) either approximately or exactly, given a sampled graph G. We may ask the following questions regarding the quality of the solution.

- (Exact recovery) When can we find (A, B) exactly (up to symmetry) with high probability?
- (Almost exact recovery) Can we find a bipartition such that the vanishing portion of the vertices are mislabeled?

- (Detection) Can we find a bipartition such that the portion of mislabeled vertices is less than $\frac{1}{2} - \epsilon$ for some positive constant ϵ?

There are a number of works regarding these questions in the algorithmic point of view or in the sense of statistical achievability. The following is a short list of the states-of-the-art works regarding the model:

- Suppose that $p = \frac{a \log n}{n}$ and $q = \frac{b \log n}{n}$ where a and b are positive constants not depending on n. Then, exact recovery is possible if and only if $(\sqrt{a} - \sqrt{b})^2 > 2$. Moreover, there are efficient algorithms which achieves the information-theoretic threshold [4, 56].

- Suppose that $p = \frac{a}{n}$ and $q = \frac{b}{n}$ where a and b are positive constants not depending on n. Then, the detection is possible if and only if $\frac{(a-b)^2}{2(a+b)} > 1$. Moreover, there are efficient algorithms achieving the information-theoretic threshold [76, 77, 72].

We further note that those sharp phase transition behaviours and algorithms achieving the threshold are found for general stochastic block models [8, 1], [68, 67]. This paper focuses on exact recovery.

3.1.2 The Stochastic Block Model for hypergraphs

The stochastic block model for hypergraphs (HSBM) is a natural generalization of the SBM for graphs which was first introduced in [49]. Informally, the HSBM can be thought as a generative model which returns a hypergraph with unknown clusters, and each hyperedge appears in the hypergraph independently with the probability depending on the community labels of the vertices involved in the hyperedge.

In [49], the authors consider the HSBM under the setting that the hypergraph generated by the model is k-uniform and dense. They consider a spectral algorithm on a version of hypergraph Laplacian, and prove that the algorithm exactly recovers the partition for $k > 3$ with probability $1 - o(1)$. Subsequently, the same authors extended their results to sparse, non-uniform (but bounded order) setting, studying partial recovery [51, 50, 52].

We note that sparsity is an important factor to address in recovery problems of different types: exact recovery, almost exact recovery, and detection. In the case of the SBM for graphs, we recall that the average degree must be $\Omega(\log n)$ to assure exact recovery, and the average degree must be $\Omega(1)$ to assure detection. Conversely, the point of the sharp phase transition lies exactly in those regimes. We may expect similar behaviour for the k-uniform HSBM. For exact recovery, it was confirmed that the phase transition occurs in the regime of logarithmic average degree, by analyzing the optimal minimax risk of k-uniform HSBM [69, 34]. For detection, the phase transition occurs in the regime of constant average degree [20]. The authors of [20] proposed a conjecture specifying the exact threshold point, based on the performance of belief-propagation algorithm. Also, such results for the weighted HSBM were independently proved in [15] and a exact threshold of the censored block model for uniform hypergraphs was classified in [14].

In this paper, we consider a specific k-uniform HSBM with two equal-sized clusters. Let us remark that in the SBM case, we had two parameters p and q where the probability that an edge $\{i, j\}$ appears in the graph is p or q depending on whether i and j are in the same cluster or not. For an hyperedge of size greater than 2, there are different ways to generalize this notion, but we will focus on a simple model that the probability that a set e of size k appears as a hyperedge depends on whether e is completely contained in a cluster or not.

Let n be a positive even number and let $V = [n]$ be the set of vertices of the hypergraph $\mathcal{H}$. Let $k \geq 2$ be an integer. Let p and q be numbers between 0 and 1, possibly depending on n. We denote the collection of size k subsets of V by $\binom{V}{k}$. The k-HSBM with parameters k, n, p and q, denoted $\mathsf{HSBM}(n, p, q; k)$, is a model which samples a k-uniform hypergraph $\mathcal{H}$ on the vertex set V according to following rules.

- $\mathbf{x}_0$ is a vector in $\{\pm 1\}^V$ chosen uniformly at random, among those with the equal number of -1's and 1's. We may think -1 and 1 as community labels.

- Each $e = \{e_1, \cdots, e_k\}$ in $\binom{V}{k}$ appears independently as an hyperedge with probability

$$\mathbb{P}(e \in E(\mathcal{H})) = \begin{cases} p & \text{if } (\mathbf{x}_0)_{e_1} = (\mathbf{x}_0)_{e_2} = \cdots = (\mathbf{x}_0)_{e_k} \\ q & \text{otherwise.} \end{cases}$$

We say e is *in-cluster* with respect to $\mathbf{x}_0$ for the first case, and *cross-cluster* w.r.t. $\mathbf{x}_0$ for the other case.

Our goal is to find the clusters from a given hypergraph $\mathcal{H}$ generated from the model. We specially focus on exact recovery, formally defined as follows.

Definition 3.1. *We say exact recovery in* $\mathsf{HSBM}(n, p, q; k)$ *is possible if there exists an estimator* $\widehat{\mathbf{x}_0}$ *which only fails to recover* $\mathbf{x}_0$ *up to a global sign flip with vanishing probability,* i.e.,

$$\mathop{\mathbb{P}}_{(\mathbf{x}_0, \mathcal{H}) \sim \mathsf{HSBM}(n,p,q;k)} (\widehat{\mathbf{x}_0}(\mathcal{H}) \notin \{\mathbf{x}_0, -\mathbf{x}_0\}) = o(1).$$

On the other hand, we say exact recovery in $\mathsf{HSBM}(n, p, q; k)$ *is impossible if any estimator* $\widehat{\mathbf{x}_0}$ *fails to recover* $\mathbf{x}_0$ *up to a global sign flip with probability* $1 - o(1)$, i.e.,

$$\mathop{\mathbb{P}}_{(\mathbf{x}_0, \mathcal{H}) \sim \mathsf{HSBM}(n,p,q;k)} (\widehat{\mathbf{x}_0}(\mathcal{H}) \notin \{\mathbf{x}_0, -\mathbf{x}_0\}) = 1 - o(1) \text{ for any } \widehat{\mathbf{x}_0}.$$

We remark that $\mathcal{H}$ must be connected for exact recovery to be successful. In Erdős-Rényi (ER) model for random hypergraphs, it is known that a random hypergraph from the ER model is connected with high probability only if the expected average degree is at least $\frac{c(k-1)\log n}{\binom{n-1}{k-1}}$ for some $c > 1$[2]. Together with the works in [15] and [34], this motivates us to work on the parameter regime where

$$p = \frac{\alpha \log n}{\binom{n-1}{k-1}} \quad \text{and} \quad q = \frac{\beta \log n}{\binom{n-1}{k-1}}$$

[2]The proof for this result is a direct adaptation of the proof in [27] for $k = 2$, *i.e.*, random graph model. See [35, 25, 36] for phase transitions regarding giant components, which justifies the regime for partial recovery and detection.

for some constant α and β.

3.1.3 Main results

We first establish a sharp phase transition behaviour for exact recovery in the stochastic block model for k-uniform hypergraphs. We will assume that the parameter k is a fixed positive integer not depending on n, and edge probabilities decay as

$$p = \frac{\alpha \log n}{\binom{n-1}{k-1}} \quad \text{and} \quad q = \frac{\beta \log n}{\binom{n-1}{k-1}}$$

where α and β are fixed positive constants. Asymptotics in this paper are based on n growing to infinity, unless noted otherwise.

Theorem 3.1. *Exact recovery in* $\mathsf{HSBM}(n, p, q; k)$ *is possible if* $I(\alpha, \beta) > 1$, *and impossible if* $I(\alpha, \beta) < 1$ *where* $I(\alpha, \beta) = \frac{1}{2^{k-1}}(\sqrt{\alpha} - \sqrt{\beta})^2$.

In case of exact recovery, the maximum a posteriori (MAP) estimator achieves the minimum error probability. The MAP estimator corresponds to the maximum-likelihood (ML) estimator in this model since the partition is chosen from a uniform distribution. Hence, it is sufficent to analyze the performance of the ML estimator to prove Theorem 3.1.

On the other hand, we ask whether there exists an efficient algorithm which recover the hidden partition $\mathbf{x}_0$ achieving the information-theoretic threshold. Note that the ML estimator (which achieves the minimum error probability) is given by

$$\widehat{\mathbf{x}}_{MLE}(H) = \underset{\mathbf{x} \in \{\pm 1\}^V : \mathbf{1}^T \mathbf{x} = 0}{\operatorname{argmax}} \; \underset{(\mathbf{x}_0, \mathcal{H}) \sim \mathsf{HSBM}(n,p,q;k)}{\mathbb{P}} (\mathcal{H} = H | \mathbf{x}_0 = \mathbf{x}).$$

This is in general hard to compute. For example, when $k = 2$ and $p > q$, it reduces to find a balanced bipartition with the minimum number of edges crossing given a graph G, also known as MIN-BISECTION problem which is NP-hard. However, there is a simple and efficient algorithm which works up to the threshold of the ML estimator in case of $k = 2$. This algorithm is based on a standard semidefinite relaxation of MIN-BISECTION [53].

For general k-HSBM, we consider the truncate-and-relax algorithm which we briefly discussed in Section 1.4. Given a k-uniform hypergraph H on the vertex set V, let us define a weighted graph (G_H, w) on the same vertex set where the weights are given by

$$w_{ij} = \#(e \in E(H) : \{i, j\} \subseteq e)$$

for each $\{i, j\} \in \binom{V}{2}$. Let $\widehat{\mathbf{x}}_{\text{trunc}}$ be an optimal solution of

$$\begin{aligned} \text{maximize} \quad & \sum_{ij \in \binom{V}{2}} w_{ij} \mathbf{x}_i \mathbf{x}_j \\ \text{subject to} \quad & \mathbf{x} \in \{\pm 1\}^V, \ \mathbf{1}^T \mathbf{x} = 0, \end{aligned}$$

which is equivalent to finding the min-bisection of the weighted graph (G_H, w). Now, consider the following semidefinite program:

$$\begin{aligned} \text{maximize} \quad & \sum_{ij \in \binom{V}{2}} w_{ij} X_{ij} \\ \text{subject to} \quad & \sum_{i,j \in V} X_{ij} = 0, \\ & X_{ii} = 1 \text{ for all } i \in V, \\ & X = X^T \succeq 0. \end{aligned} \tag{3.1}$$

This program is a relaxation of the min-bisection problem above, since for any feasible x in the original problem corresponds to a feasible solution $X = xx^T$ in the relaxed problem.

The ML estimator attempts to maximize the function

$$f_H(\mathbf{x}) = \log \underset{(\mathbf{x}_0, \mathcal{H}) \sim \mathsf{HSBM}(n,p,q;k)}{\mathbb{P}} (\mathcal{H} = H | \mathbf{x}_0 = \mathbf{x})$$

over the vectors in the hypercube $\{\pm 1\}^V$ with equal number of -1's and 1's. This function $f_H(\mathbf{x})$ can be written as a multilinear polynomial in $\mathbf{x}$ (See Section 2.1.2). Let $f_H^{(2)}(\mathbf{x})$ be the quadratic part of $f_H(\mathbf{x})$. Then, maximizing $f_H^{(2)}(\mathbf{x})$ is equivalent to find the min-bisection of (G_H, w). We note that this algorithm is equivalent to the truncate-and-relax algorithm for the spiked tensor models as in Section 2.4.

Now, let $\widehat{\Sigma}(H)$ be the solution of (3.1). We prove that this estimator correctly recovers the hidden partition with high probability up to a threshold which is order-wise optimal.

Theorem 3.2. *Suppose $\alpha > \beta$. Then $\widehat{\Sigma}(\mathcal{H})$ is equal to $\mathbf{x}_0\mathbf{x}_0^T$ with probability $1 - o(1)$ if $I_{sdp}(\alpha, \beta) > 1$ where*

$$I_{sdp}(\alpha, \beta) = \frac{k-1}{2^{2k}} \cdot \frac{(\alpha - \beta)^2}{\left(\frac{k}{2^k}\alpha + \left(1 - \frac{k}{2^k}\right)\beta\right)}.$$

It is natural to ask whether this analysis is tight. The proof proceeds by constructing a dual solution which certifies that $\mathbf{x}_0\mathbf{x}_0^T$ is the unique optimum of (3.1) with high probability. Following [21], the dual solution (if exists) is completely determined by (G_H, w) which has the form of a "Laplacian" matrix. Precisely, the major part of the proof is devoted to prove that the matrix L of size $V \times V$ with entries

$$L_{ij} = \begin{cases} -w(ij)(\mathbf{x}_0)_i(\mathbf{x}_0)_j & \text{if } i \neq j \\ \sum_{i' \in V \setminus \{i\}} w(ii')(\mathbf{x}_0)_i(\mathbf{x}_0)_{i'} & \text{if } i = j, \end{cases}$$

is positive-semidefinite with high probability. We use the Matrix Bernstein inequality to prove that the fluctuation $\|L - \mathbb{E}\, L\|$ is smaller compared to the minimum eigenvalue of $\mathbb{E}\, L$ w.h.p., under the assumption $I_{sdp}(\alpha, \beta) > 1$. However, we believe that it can be improved by a direct analysis of $\|L - \mathbb{E}\, L\|$. Numerical simulations and discussions which supports our belief can be found in Section 3.5.

Finally, we complement Theorem 3.2 by providing a lower bound of the truncate-and-relax algorithm. Recall that the algorithm tries to find a solution in the relaxed problem (3.1). It implies that if the min-bisection of (G_H, w) is not the correct partition $\mathbf{x}_0$, then the truncate-and-relax algorithm will also return a solution which is not equal to $\mathbf{x}_0\mathbf{x}_0^T$. Hence, we have

$$\mathbb{P}(\widehat{\Sigma}(\mathcal{H}) \neq \mathbf{x}_0\mathbf{x}_0^T) \geq \mathbb{P}(\widehat{\mathbf{x}}_{\text{trunc}}(\mathcal{H}) \notin \{\mathbf{x}_0, -\mathbf{x}_0\}).$$

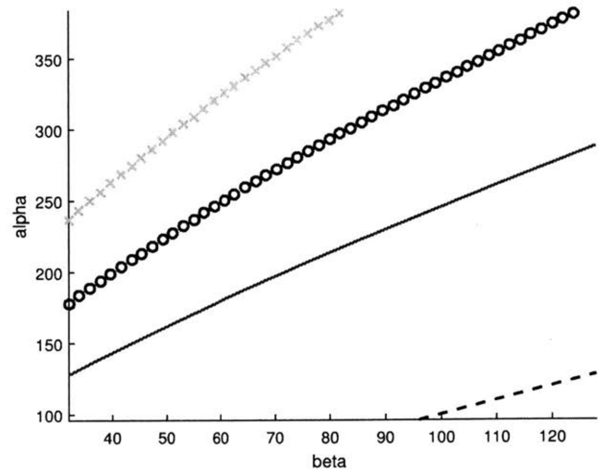

Figure 3-1: Visualization of I, I_2, I_{sdp} when $k = 6$: (a) the solid line represents $I(\alpha, \beta) = 1$, (b) the circled line represents $I_2(\alpha, \beta) = 1$, and (c) the x-marked line represents $I_{sdp}(\alpha, \beta) = 1$. The dashed black line is the graph of $\alpha = \beta$.

We find a sharp threshold for the estimator $\widehat{\mathbf{x}}_{\text{trunc}}(\mathcal{H})$ recovering $\mathbf{x}_0$ or $-\mathbf{x}_0$ successfully.

Theorem 3.3. *Suppose $\alpha > \beta$. Let $I_2(\alpha, \beta)$ be defined as following:*

$$I_2(\alpha, \beta) = \max_{t \geq 0} \frac{1}{2^{k-1}} \left[\alpha(1 - e^{-(k-1)t}) + \sum_{a=1}^{k-1} \beta \binom{k-1}{a} \left(1 - e^{-(k-1-2a)t}\right) \right]$$

If $I_2(\alpha, \beta) < 1$, then $\widehat{\mathbf{x}}_{\text{trunc}}(\mathcal{H})$ is not equal to neither $\mathbf{x}_0$ nor $-\mathbf{x}_0$ with probability $1 - o(1)$. On the other hand, if $I_2(\alpha, \beta) > 1$, then $\widehat{\mathbf{x}}_{\text{trunc}}(\mathcal{H})$ is either of $\mathbf{x}_0$ or $-\mathbf{x}_0$ with probability $1 - o(1)$.

We note that

$$I(\alpha, \beta) = \max_{t \geq 0} \frac{1}{2^{k-1}} \left(\alpha(1 - e^{-(k-1)t}) + \beta(1 - e^{(k-1)t}) \right),$$

hence $I(\alpha, \beta) < I_2(\alpha, \beta)$ for any $\alpha > \beta > 0$. Figure 3-1 shows the relations between I, I_2 and I_{sdp} for $k = 6$.

Theorem 3.3 and the discussion above implies that the truncate-and-relax algorithm fails with probability $1 - o(1)$ if $I_2(\alpha, \beta) > 1$. We conjecture that this is the correct threshold of the performance of the algorithm. In future work, we will attempt to prove this conjecture by improving the matrix concentration bound as discussed above.

Conjecture 3.1.1. *If* $I_2(\alpha, \beta) > 1$, *then* $\widehat{\Sigma}(\mathcal{H}) = \mathbf{x}_0\mathbf{x}_0^T$ *with probability* $1 - o(1)$.

3.2 Maximum-likelihood estimator

Recall that $\widehat{\mathbf{x}}_{MLE}(H)$ is a maximizer of the likelihood probability $\mathbb{P}(\mathcal{H} = H|\mathbf{x}_0 = \mathbf{x})$ (ties are broken arbitrarily). Let $f_H(\mathbf{x}) = \log \mathbb{P}(\mathcal{H} = H|\mathbf{x}_0 = \mathbf{x})$ for $\mathbf{x} \in \{\pm 1\}^V$.

For brevity, let us first introduce a few notations. For a vector $\mathbf{x} \in \{\pm 1\}^V$, let $\mathbf{x}^{\ominus k}$ be a vector in $\{0,1\}^{\binom{V}{k}}$ where

$$(\mathbf{x}^{\ominus k})_e = \begin{cases} 1 & \text{if } x_{e_1} = x_{e_2} = \cdots = x_{e_k} \\ 0 & \text{otherwise,} \end{cases}$$

for each $e = \{e_1, \cdots, e_k\} \subseteq V$. **Here we abuse the notation $\mathbf{x}^{\ominus k}$ which was originally defined as a k-tensor in Chapter 2**; one may think this new definition as the restriction of the k-tensor version onto k-tuples with distinct indices.

Let H be a k-uniform hypergraph on the vertex set V with the edge set $E(H)$. Let A_H be the vector in $\{0,1\}^{\binom{V}{k}}$ such that

$$(A_H)_e = \begin{cases} 1 & \text{if } e \in E(H) \\ 0 & \text{otherwise} \end{cases}$$

for each $e \in \binom{V}{k}$. Note that

$$\begin{aligned} \langle A_H, \mathbf{x}^{\ominus k} \rangle &= \sum_{e \in \binom{V}{k}} (A_H)_e (\mathbf{x}^{\ominus k})_e \\ &= \sum_{e \in E(H)} \mathbf{1}\{e \text{ is in-cluster with respect to } x\}. \end{aligned}$$

Hence, $\langle A_H, \mathbf{x}^{\ominus k} \rangle$ is equal to the number of in-cluster edges in H with respect to the partition x.

The ML estimator tries to find the "best" partition $\mathbf{x} \in \{\pm 1\}^V$ with equal number of 1's and -1's. Intuitively, if $p > q$, *i.e.*, in-cluster edges appears more likely than cross-cluster edges (*assortative*), then the best partition will correspond to $\mathbf{x}$ such that the number of in-cluster edges w.r.t. $\mathbf{x}$ is maximized. On the other hand, if $p < q$ (*disassortative*) then the best partition will corresponds to the minimizer, respectively. The following proposition confirms this intuition. We defer the proof to Section 3.7.1.

Proposition 3.4. *The ML estimator* $\widehat{\mathbf{x}}_{MLE}(H)$ *is the maximizer (minimizer, respectively) of* $\langle A_H, \mathbf{x}^{\ominus k} \rangle$ *if* $p > q$ *(if* $p < q$*, respectively) over all* $\mathbf{x} \in \{\pm 1\}^V$ *such that* $\mathbf{1}^T \mathbf{x} = 0$.

3.3 Threshold for exact recovery in k-HSBM

We prove Theorem 3.1 throughout this section. The techniques we use can be seen as a hypergraph extension of the techniques used in [4].

Informally, we are going to argue that the event for the ground truth $\mathbf{x}_0$ being the best guess (i.e. $\mathbf{x}_0$ is the global optimum of the likelihood function) can be approximately decomposed into the events E_v for $v \in V$, where E_v is the vent that the likelihood function does not increase if we flip the sign of $(\mathbf{x}_0)_v$. In [1], Abbe call this type of phenomenon *local-to-global amplification.* In Chapter 4, we argue that local-to-global amplification is universal: It holds in a broader class of graphical models including the spiked tensor model and the HSBM.

Let p_{fail} be the probability that the ML estimator fails to recover the hidden partition, *i.e.*,

$$p_{fail} = \mathop{\mathbb{P}}_{(\mathbf{x}_0, \mathcal{H}) \sim \mathsf{HSBM}(n,p,q;k)} \left(\widehat{\mathbf{x}}_{MLE}(\mathcal{H}) \notin \{\mathbf{x}_0, -\mathbf{x}_0\} \right).$$

As we have seen in the previous section, the ML estimator $\widehat{\mathbf{x}}_{MLE}(H)$ is a maximizer of $\langle A_H, \mathbf{x}^{\ominus k}\rangle$ over the choices of $\mathbf{x} \in \{\pm 1\}^V$ such that $\mathbf{1}^T\mathbf{x} = 0$. Thus,

$$p_{fail} = \mathbb{P}\left(\exists \mathbf{x} \notin \{\pm \mathbf{x}_0\} \text{ s.t. } \mathbf{1}^T\mathbf{x} = 0 \text{ and } \langle A_{\mathcal{H}}, \mathbf{x}^{\ominus k}\rangle \geq \left\langle A_{\mathcal{H}}, \mathbf{x}_0^{\ominus k}\right\rangle\right).$$

3.3.1 Lower bound: Impossibility

We first prove the impossibility part of Theorem 3.1. For concreteness, we focus on the assortative case, *i.e.*, $p > q$ but the proof can be easily adapted for the disassortative case.

Before we prove the lower bound, let us consider the usual stochastic block model for graphs which corresponds to $k = 2$ in order to explain the intuition of the proof. Given a sample G, partition $\mathbf{x}_0$ and a vertex $v \in V$, let us define the *in-degree* of v as

$$\text{indeg}_{G,\mathbf{x}_0}(v) = \#(vw \in E(G) : (\mathbf{x}_0)_v = (\mathbf{x}_0)_w),$$

and the *out-degree* of v as

$$\text{outdeg}_{G,\mathbf{x}_0}(v) = \#(vw \in E(G) : (\mathbf{x}_0)_v \neq (\mathbf{x}_0)_w).$$

We will omit the subscript $G, \mathbf{x}_0$ if the context is clear.

Suppose that there are vertices v and w from different clusters such that the in-degree of each vertex is smaller than the out-degree of each vertex. In this case, swapping the label of v and w will yield a new balanced partition with greater number of in-cluster edges, hence the ML estimator will fail to recover $\mathbf{x}_0$. Now, suppose that

$$\mathbb{P}(\text{indeg}(v) < \text{outdeg}(v)) = \omega(n^{-1})$$

for all v. If those events were independent, we would get

$$\begin{aligned}\mathbb{P}\left(\text{indeg}(v) \geq \text{outdeg}(v)\ \forall v : (\mathbf{x}_0)_v = 1\right) &= \prod_{v:(\mathbf{x}_0)_v=1} \mathbb{P}(\text{indeg}(v) \geq \text{outdeg}(v)) \\ &\leq (1 - \omega(n^{-1}))^{n/2} \leq e^{-\omega(1)},\end{aligned}$$

and similar for w. It would imply that there is a "bad" pair (v, w) with probability $1 - o(1)$ hence the ML estimator fails with probability $1 - o(1)$. We remark that this argument is not mathematically because the in-degrees of vertex v and w (as well as out-degrees of them) are *not* independent as they share a variable indicating whether $\{v, w\}$ is an edge or not. However, we can overcome it by conditioning on highly probable event which makes those events independent, as in [4].

We extend the definitions of in-degree and out-degree for the k-HSBM as

$$\begin{aligned}
\mathrm{indeg}_{H,\mathbf{x}_0}(v) = &\quad \#(e \in E(H) : v \in e, \quad e \text{ is in-cluster w.r.t. } \mathbf{x}_0), \\
\mathrm{outdeg}_{H,\mathbf{x}_0}(v) = &\quad \#(e \in E(H) : v \in e, \quad e \text{ is cross-cluster but} \\
& \qquad\qquad\qquad\qquad\qquad e \setminus \{v\} \text{ is in-cluster w.r.t. } \mathbf{x}_0).
\end{aligned}$$

Observe that they coincide with the corresponding definition for the usual SBM ($k = 2$). We note that the sum of in-degree and out-degree is not equal to the degree of v, the number of hyperedges in H containing v when $k \geq 3$. We extended those definitions in this way because any edge e which is neither in-cluster nor cross-cluster but $e \setminus \{v\}$ is in-cluster does not contribute on $\langle A_{\mathcal{H}}, \mathbf{x}^{\ominus k} \rangle$ when we flip the sign of the label of v.

Now, note that the in-degree and the out-degree of v are independent binomial random variables with different parameters. To estimate the probability

$$\mathbb{P}\left(\mathrm{indeg}(v) - \mathrm{outdeg}(v) < 0\right),$$

we provide a tight estimate for the tail probability of a weighted sum of independent binomial variables in Section 3.6. Precisely we prove that

$$\mathbb{P}(\mathrm{indeg}(v) - \mathrm{outdeg}(v) < -\delta \log n) = n^{-I(\alpha,\beta)+o(1)}$$

as long as $\delta = \delta(n)$ vanishes as n grows, where

$$I(\alpha, \beta) = \frac{1}{2^{k-1}} \left(\sqrt{\alpha} - \sqrt{\beta}\right)^2.$$

As we discussed, if $I(\alpha,\beta) < 1$ then the tail probability is of order $\omega(n^{-1})$ and it implies that the ML estimator fails with probability $1 - o(1)$.

Theorem 3.5. *Let* $I(\alpha,\beta) = \frac{1}{2^{k-1}}(\sqrt{\alpha} - \sqrt{\beta})^2$. *If* $I(\alpha,\beta) < 1$, *then* $p_{fail} = 1 - o(1)$.

Proof. Let $A = \{v \in V : (\mathbf{x}_0)_v = +1\}$ and $B = V \setminus A$. For $a \in A$ and $b \in B$, let us define $\mathbf{x}_0^{(ab)}$ to be the vector obtained by flipping the signs of $(\mathbf{x}_0)_a$ and $(\mathbf{x}_0)_b$. By definition, $\mathbf{x}_0^{(ab)}$ is balanced. We are going to prove that with high probability there exist $a \in A$ and $b \in B$ such that $\left\langle A_{\mathcal{H}}, \mathbf{x}_0^{\ominus k} \right\rangle \le \left\langle A_{\mathcal{H}}, (\mathbf{x}_0^{(ab)})^{\ominus k} \right\rangle$. For simplicity, let $\Sigma = \mathbf{x}_0^{\ominus k}$ and $\Sigma^{(ab)} = (\mathbf{x}_0^{(ab)})^{\ominus k}$.

Note that

$$\begin{aligned}\langle A_{\mathcal{H}}, \Sigma \rangle - \left\langle A_{\mathcal{H}}, \Sigma^{(ab)} \right\rangle &= \left(\mathrm{indeg}_{\mathcal{H},\mathbf{x}_0}(a) - \mathrm{outdeg}_{\mathcal{H},\mathbf{x}_0}(a)\right) \\ &\quad + \left(\mathrm{indeg}_{\mathcal{H},\mathbf{x}_0}(b) - \mathrm{outdeg}_{\mathcal{H},\mathbf{x}_0}(b)\right).\end{aligned}$$

For $v \in V$, let E_v be the event such that

$$\mathrm{indeg}_{\mathcal{H},\mathbf{x}_0}(v) - \mathrm{outdeg}_{\mathcal{H},\mathbf{x}_0}(v) \le 0$$

holds. Then, $E_a \cap E_b$ implies that $\langle A_{\mathcal{H}}, \Sigma \rangle - \left\langle A_{\mathcal{H}}, \Sigma^{(ab)} \right\rangle \le 0$. Hence

$$p_{fail} = \mathbb{P}\left(\exists a \in A, b \in B : \langle A_{\mathcal{H}}, \Sigma \rangle - \left\langle A_{\mathcal{H}}, \Sigma^{(ab)} \right\rangle\right) \ge \mathbb{P}\left(\bigcup_{a \in A} E_a \cap \bigcup_{b \in B} E_b\right).$$

Informal overview. We recall that if E_v for $v \in V$ were mutually independent, we can exactly express the right-hand side as

$$\left(1 - \prod_{a \in A} \mathbb{P}(\neg E_a)\right)\left(1 - \prod_{b \in B} \mathbb{P}(\neg E_b)\right)$$

but unfortunately it is not the case. To see this, let us fix $a \in A$ and $a' \in A$. Then, we have

$$\begin{aligned}
\operatorname{indeg}(a) - \operatorname{outdeg}(a) &= \sum_{e \ni a : e \subseteq A} (A_{\mathcal{H}})_e - \sum_{e \ni a : e \cap A = \{a\}} (A_{\mathcal{H}})_e, \text{ and} \\
\operatorname{indeg}(a') - \operatorname{outdeg}(a') &= \sum_{e \ni a' : e \subseteq A} (A_{\mathcal{H}})_e - \sum_{e \ni a' : e \cap A = \{a'\}} (A_{\mathcal{H}})_e.
\end{aligned}$$

They share variables $(A_{\mathcal{H}})_e$ for e satisfying $\{a, a'\} \subseteq e \subseteq A$. The expected contribution of those variables is $p\binom{|A|-2}{k-2} = o(1)$, so we may expect

$$\mathbb{P}(E_a \cup E_{a'}) \approx 1 - \mathbb{P}(\neg E_a)\,\mathbb{P}(\neg E_{a'}).$$

In the similar spirit, we are going to prove that for an appropriate choice of $U \subseteq V$, the events $\{E_a\}_{a \in U \cap A}$ and $\{E_b\}_{b \in U \cap B}$ are approximately independent, so

$$\begin{aligned}
p_{fail} &\geq \mathbb{P}\left(\left(\bigcup_{a \in A \cap U} E_a \right) \cap \left(\bigcup_{b \in B \cap U} E_b \right) \right) \\
&\approx \left(1 - \prod_{a \in A \cap U} \mathbb{P}(\neg E_a) \right) \left(1 - \prod_{b \in B \cap U} \mathbb{P}(\neg E_b) \right).
\end{aligned}$$

Together with the tight estimate on $\mathbb{P}(E_v)$, it would give us a good lower bound on p_{fail}. □

Let $U \subseteq V$ be a set of size γn where $|U \cap A| = |U \cap B|$. We will choose $\gamma = \gamma(n)$ later to be poly-logarithmically decaying function in n. Let $\mathcal{S}$ be the set of $e \in \binom{V}{k}$ such that e contains at least two vertices in U. We would like to condition on the values of $\{(A_{\mathcal{H}})_e\}_{e \in \mathcal{S}}$, which captures all dependency occurring among E_v's for $v \in U$.

Let $\delta = \delta(n)$ be a positive number depending on n which we will choose later, and let F be the event that the inequality

$$\max_{v \in U} \sum_{e \in \mathcal{S} : e \ni v} (A_{\mathcal{H}})_e \leq \delta \log n$$

holds. For each $a \in A \cap U$, let E'_a be the event that the inequality

$$\sum_{e \subseteq A : e \cap U = \{a\}} (A_{\mathcal{H}})_e - \sum_{\substack{e : e \cap U = \{a\} \\ e \setminus \{a\} \subseteq B}} (A_{\mathcal{H}})_e \leq -\delta \log n$$

is satisfied. We claim that $E'_a \cap F \subseteq E_a$. It follows from the direct calculation, as if we assume $E'_a \cap F$, then

$$\begin{aligned}
\text{indeg}(v) - \text{outdeg}(v) &= \sum_{e \ni a : e \subseteq A} (A_{\mathcal{H}})_e - \sum_{e \ni a : e \cap A = \{a\}} (A_{\mathcal{H}})_e \\
&\leq \sum_{e \subseteq A : e \cap U = \{a\}} (A_{\mathcal{H}})_e - \sum_{\substack{e : e \cap U = \{a\} \\ e \setminus \{a\} \subseteq B}} (A_{\mathcal{H}})_e + \sum_{e \in \mathcal{S} : e \ni a} (A_{\mathcal{H}})_e \\
&\leq \sum_{e \subseteq A : e \cap U = \{a\}} (A_{\mathcal{H}})_e - \sum_{\substack{e : e \cap U = \{a\} \\ e \setminus \{a\} \subseteq B}} (A_{\mathcal{H}})_e + \delta \log n \\
&\leq 0.
\end{aligned}$$

We get

$$p_{fail} \geq \mathbb{P}\left(\bigcup_{a \in A \cap U} E_a \cap \bigcup_{b \in B \cap U} E_b \right) \geq \mathbb{P}\left(\bigcup_{a \in A \cap U} E'_a \cap \bigcup_{b \in B \cap U} E'_b \middle| F \right) \mathbb{P}(F).$$

Note that E'_v only depends on the set of variables $\{(A_{\mathcal{H}})_e : e \cap U = \{v\}\}$, which are mutually disjoint for $v \in U$. Also, $\{(A_{\mathcal{H}})_e : e \in \mathcal{S}\}$ is disjoint with any of those sets of variables. Hence, events F and $\{E'_v\}_{v \in U}$ are mutually independent, and we get

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{a \in A \cap U} E'_a \cap \bigcup_{b \in B \cap U} E'_b \middle| F \right) &= \mathbb{P}\left(\bigcup_{a \in A \cap U} E'_a \right) \mathbb{P}\left(\bigcup_{b \in B \cap U} E'_b \right) \\
&= \left(1 - \prod_{a \in A \cap U} \mathbb{P}(\neg E'_a) \right) \left(1 - \prod_{b \in B \cap U} \mathbb{P}(\neg E'_b) \right)
\end{aligned}$$

We claim that

$$\prod_{a \in A \cap U} \mathbb{P}(\neg E'_a) = o(1), \quad \prod_{b \in B \cap U} \mathbb{P}(\neg E'_b) = o(1) \quad \text{and} \quad \mathbb{P}(\neg F) = o(1),$$

for appropriate choice of γ and δ. This immediately implies that $p_{fail} = 1 - o(1)$ as desired.

Let us first prove $\mathbb{P}(\neg F) = o(1)$. Let X_v be the random variable defined as

$$X_v := \sum_{e \in \mathcal{S}: e \ni v} (A_{\mathcal{H}})_e,$$

for $v \in U$. We have

$$\mathbb{P}(\neg F) = \mathbb{P}(\exists v \in U : X_v > \delta \log n) \leq \sum_{v \in U} \mathbb{P}(X_v > \delta \log n)$$

by a union bound. Note that

$$\begin{aligned} \mathbb{E} X_v &\leq \max(p, q) \#(e : e \ni v, |e \cap U| \geq 2) \\ &= \frac{\max(\alpha, \beta) \log n}{\binom{n-1}{k-1}} \left(\binom{n-1}{k-1} - \binom{n-|U|}{k-1} \right) \\ &\lesssim \left(1 - (1-\gamma)^{k-1}\right) \log n = \Theta(\gamma \log n). \end{aligned}$$

Using a standard Chernoff bound, we get the following lemma.

Lemma 3.6. *Let X be a sum of independent Bernoulli variables such that $\mathbb{E} X = \Theta(\gamma \log n)$ where $\gamma = o_n(\log^{-1} n)$. Let δ be a positive number which decays to 0 as n grows, with $\delta = \omega_n(\log^{-1} n)$. Then,*

$$\mathbb{P}\left(X > \delta \log n\right) \leq n^{-\delta \log \frac{\delta}{\gamma} + o(1)}.$$

Proof. See Section 3.6.1. $\square$

Letting $\gamma = \log^{-3} n$ and $\delta = (\log \log n)^{-1}$, we get

$$\delta \log \frac{\delta}{\gamma} = \frac{3 \log \log n - \log \log \log n}{\log \log n} = 3 - o(1)$$

and so $\mathbb{P}(\neg F) = n^{-3+o(1)} = o(1)$.

Now, we would like to prove that

$$\prod_{a \in A \cap U} \mathbb{P}(\neg E'_a) = o(1) \quad \text{and} \quad \prod_{b \in B \cap U} \mathbb{P}(\neg E'_b) = o(1)$$

by showing that

$$\mathbb{P}(E'_v) \geq n^{-I(\alpha,\beta)+o(1)} \quad \text{where } I(\alpha, \beta) = \frac{1}{2^{k-1}}(\sqrt{\alpha} - \sqrt{\beta})^2$$

for any $v \in U$. This implies that

$$\begin{aligned} \prod_{a \in A \cap U} \mathbb{P}(\neg E'_a) &\leq \left(1 - n^{-I(\alpha,\beta)+o(1)}\right)^{\gamma n/2} \\ &\leq \exp\left(-\frac{\gamma}{2} n^{1-I(\alpha,\beta)+o(1)}\right), \end{aligned}$$

and since we assumed that $I(\alpha, \beta) < 1$ and $\gamma = \log^{-3} n$, we get

$$\prod_{a \in A \cap U} \mathbb{P}(\neg E'_a) \leq e^{-\frac{n^{1-I(\alpha,\beta)}}{\log^3 n}} = o(1),$$

and similarly $\prod_{b \in B \cap U} \mathbb{P}(\neg E'_b) = o(1)$ as desired.

To estimate the probability that E'_a happens, let Y_a and Z_a be random variables defined as

$$Y_a = \sum_{e \subseteq A : e \cap U = \{a\}} (A_{\mathcal{H}})_e \quad \text{and} \quad Z_a = \sum_{\substack{e : e \cap U = \{a\} \\ e \setminus \{a\} \subseteq B}} (A_{\mathcal{H}})_e.$$

Recall that E'_a is the event that $Y_a - Z_a \leq -\delta \log n$ holds.

Lemma 3.7. *Let Y be a binomial random variable from* $\mathrm{Bin}(N,p)$ *and Z be a binomial random variable from* $\mathrm{Bin}(N,q)$ *where $N = \left(\frac{1}{2^{k-1}} \pm o(1)\right)\binom{n-1}{k-1}$, $p = \frac{\alpha \log n}{\binom{n-1}{k-1}}$ and $q = \frac{\beta \log n}{\binom{n-1}{k-1}}$. Let $I(\alpha, \beta) = \frac{1}{2^{k-1}}(\sqrt{\alpha} - \sqrt{\beta})^2$ and let δ be a positive number vanishing as n grows. Then,*

$$\mathbb{P}(Y - Z \leq -\delta \log n) = n^{-I(\alpha,\beta)+o(1)}.$$

In fact, we derive a generic tail bound for weighted sum of binomial random variables (Theorem 3.17) in Section 3.6 of the appendix. Lemma 3.7 is a direct corollary of Theorem 3.17 and we defer the proof to Section 3.6.2. □

3.3.2 Upper bound: Achievability

We are going to use a union bound to prove the upper bound. Let x and $\mathbf{x}_0$ be vectors in $\{-1,+1\}^V$. The *Hamming distance* between $\mathbf{x}$ and $\mathbf{x}_0$ (denoted $d(\mathbf{x},\mathbf{x}_0)$) is defined as the number of $v \in V$ such that $\mathbf{x}_v \neq (\mathbf{x}_0)_v$. Note that if $\mathbf{x}$ and $\mathbf{x}_0$ are balanced, then

$$\begin{aligned} d(\mathbf{x},\mathbf{x}_0) = \;& \#(v \in V : (\mathbf{x}_v,(\mathbf{x}_0)_v) = (1,-1)) + \#(v \in V : (\mathbf{x}_v,(\mathbf{x}_0)_v) = (-1,1)) \\ = \;& \#(v \in V : \mathbf{x}_v = 1) + \#(v \in V : (\mathbf{x}_0)_v = 1) - 2\#(v \in V : \mathbf{x}_v = (\mathbf{x}_0)_v = 1) \\ = \;& n - 2\#(v \in V : \mathbf{x}_v = (\mathbf{x}_0)_v = 1), \end{aligned}$$

hence $d(\mathbf{x},\mathbf{x}_0)$ is even.

Let us fix $\mathbf{x}_0$ and let $\mathcal{H}$ be a k-uniform random hypergraph generated by the model under the ground truth $\mathbf{x}_0$. We note that the distribution of the random variable $\left\langle A_{\mathcal{H}}, \mathbf{x}^{\ominus k} - \mathbf{x}_0^{\ominus k} \right\rangle$ is invariant under the permutation of V preserving $\mathbf{x}_0$, hence it only depends on $d(\mathbf{x},\mathbf{x}_0)$. Hence, there is a quantity $p_{fail}^{(d)}$ which satisfies

$$p_{fail}^{(d)} = \mathbb{P}(\left\langle A_{\mathcal{H}}, \mathbf{x}^{\ominus k} - \mathbf{x}_0^{\ominus k} \right\rangle \geq 0)$$

for any x with $d(\mathbf{x},\mathbf{x}_0) = d$. Moreover, $p_{fail}^{(d)} = p_{fail}^{(n-d)}$ since our model is invariant under a global sign flip.

Recall that the ML estimator fails to recover $\mathbf{x}_0$ if and only if

$$\left\langle A_{\mathcal{H}}, \mathbf{x}^{\ominus k} \right\rangle \geq \left\langle A_{\mathcal{H}}, \mathbf{x}_0^{\ominus k} \right\rangle$$

for some balanced $\mathbf{x} \in \{\pm 1\}^V$ which is neither $\mathbf{x}_0$ nor $-\mathbf{x}_0$. We remark that we count the equality as a failure, which will only make p_{fail} larger. By union bound, we have

$$\begin{aligned} p_{fail} &\leq \sum_{\substack{\mathbf{x}\in\{\pm 1\}^V:\mathbf{1}^T\mathbf{x}=0,\\ 1\leq d(\mathbf{x},\mathbf{x}_0)\leq n-1}} \mathbb{P}(\left\langle A_{\mathcal{H}}, \mathbf{x}^{\ominus k} - \mathbf{x}_0^{\ominus k}\right\rangle \geq 0) \\ &\leq 2 \sum_{\substack{d:1\leq d\leq \frac{n}{2}\\ d \text{ is even}}} p_{fail}^{(d)} \cdot \#(x \in \{\pm 1\}^V : \mathbf{1}^T\mathbf{x} = 0,\ d(\mathbf{x}, \mathbf{x}_0) = d). \end{aligned}$$

We note that there is a one-to-one correspondence between a balanced x and a pair of sets (U_+, U_-) where

$$\begin{aligned} U_+ &= \{v : x_v = -1, (\mathbf{x}_0)_v = 1\} \subseteq \{v : (\mathbf{x}_0)_v = +1\} \\ U_- &= \{v : x_v = 1, (\mathbf{x}_0)_v = -1\} \subseteq \{v : (\mathbf{x}_0)_v = -1\}, \end{aligned}$$

and we must have $d(\mathbf{x}, \mathbf{x}_0) = 2|U_+| = 2|U_-|$ since x is balanced. Hence, the number of balanced x's with $d(\mathbf{x}, \mathbf{x}_0) = d$ is equal to $\binom{n/2}{d/2}^2$. We have

$$p_{fail} \leq 2 \sum_{\substack{d:1\leq d\leq \frac{n}{2}\\ d \text{ is even}}} \binom{n/2}{d/2}^2 p_{fail}^{(d)}.$$

Now, let us formally state the main result of this section.

Theorem 3.8. *Suppose that $I(\alpha, \beta) > 1$. Then,*

$$p_{fail} \leq n^{-\frac{I(\alpha,\beta)-1}{2}+o(1)}$$

and it implies that $p_{fail} = o(1)$.

Proof. Let d be even number in between 1 and $\frac{n}{2}$. Choose any balanced $\mathbf{x}$ with $d(\mathbf{x}_0, \mathbf{x}) = d$, and let X_d be

$$X_d := \left\langle A_{\mathcal{H}}, \mathbf{x}^{\ominus k} - \mathbf{x}_0^{\ominus k}\right\rangle = \sum_{e\in\binom{V}{k}} (A_{\mathcal{H}})_e \left(\mathbf{x}^{\ominus k} - \mathbf{x}_0^{\ominus k}\right)_e.$$

Let $A = \{v : (\mathbf{x}_0)_v = 1\}$ and $A' = \{v : \mathbf{x}_v = 1\}$. We say e crosses A if $e \cap A$ and $e \setminus A$

are both non-empty (and respectively for A'). Then,

$$\left(\mathbf{x}^{\ominus k} - \mathbf{x}_0^{\ominus k}\right)_e = \begin{cases} -1 & \text{if } e \text{ doesn't cross } A' \text{ but crosses } A \\ 1 & \text{if } e \text{ crosses } A' \text{ but doesn't cross } A \\ 0 & \text{otherwise.} \end{cases}$$

Hence $X_d = Y_d - Z_d$ where Y_d and Z_d are binomial variables with $Y_d \sim \mathrm{Bin}(N_1, p)$ and $Z_d \sim \mathrm{Bin}(N_2, q)$ where

$$\begin{aligned} N_1 &= \#(e \in \binom{V}{k} : e \text{ doesn't cross } A' \text{ but crosses } A) \\ N_2 &= \#(e \in \binom{V}{k} : e \text{ crosses } A' \text{ but doesn't cross } A). \end{aligned}$$

A simple combinatorial argument shows that

$$\begin{aligned} N_1 &= \#(e : e \subseteq A' \text{ and } e \text{ crosses } A) + \#(e : e \subseteq V \setminus A' \text{ and } e \text{ crosses } A) \\ &= \left(\binom{|A'|}{k} - \binom{|A \cap A'|}{k} - \binom{|A' \setminus A|}{k}\right) + \left(\binom{n-|A'|}{k} - \binom{|A \setminus A'|}{k} - \binom{n-|A \cup A'|}{k}\right) \\ &= 2\left(\binom{n/2}{k} - \binom{d/2}{k} - \binom{(n-d)/2}{k}\right), \end{aligned}$$

and $N_2 = N_1$ by the symmetry. Hence,

$$p_{fail}^{(d)} = \mathop{\mathbb{P}}_{\substack{Y_d \sim \mathrm{Bin}(N,q) \\ Z_d \sim \mathrm{Bin}(N,p)}} (Y_d - Z_d \geq 0)$$

where $N = 2\left(\binom{n/2}{k} - \binom{d/2}{k} - \binom{(n-d)/2}{k}\right)$, $p = \frac{\alpha \log n}{\binom{n-1}{k-1}}$ and $q = \frac{\beta \log n}{\binom{n-1}{k-1}}$.

We claim that

$$\binom{n/2}{d/2}^2 p_{fail}^{(d)} = \begin{cases} C_1 \cdot n^{-d(I(\alpha,\beta)-1)/2} & \text{if } d < \frac{n}{\log \log n} \\ n^{-C_2 \cdot \frac{n}{\log \log n}} & \text{if } d \geq \frac{n}{\log \log n}. \end{cases}$$

for some positive constants C_1 and C_2 which does not depend on n or d. Assuming that the claim is true, we get

$$\sum_{1 \le d < \frac{n}{\log\log n}} \binom{n/2}{d/2}^2 p_{fail}^{(d)} \le C_1 \sum_{d \ge 1} (n^{-(I(\alpha,\beta)-1)/2})^d \lesssim n^{-\epsilon/2},$$

and

$$\sum_{\frac{n}{\log\log n} \le d \le \frac{n}{2}} \binom{n/2}{d/2}^2 p_{fail}^{(d)} \le \frac{n}{2} \cdot n^{-C_2 \cdot \frac{n}{\log\log n}} \le n^{-\omega_n(1)},$$

hence $p_{fail}^{(d)} = O(n^{-\epsilon/2})$ as desired.

To complete the proof, we are going to use the tail bound derived in Section 3.6.

Let us first focus on the case that $d \ge \frac{n}{\log\log n}$. We have

$$\begin{aligned} N_d &\ge \frac{2}{2^k \cdot k!}\left((n-2k+2)^k - d^k - (n-d)^k\right) \\ &\ge \frac{n^k}{2^{k-1} \cdot k!}\left(\left(1 - \frac{2k-2}{n}\right)^k - \frac{1}{(\log\log n)^k} - \left(1 - \frac{1}{\log\log n}\right)^k\right) \\ &= (1+o(1))\frac{1}{2^{k-1}} \cdot \frac{n}{\log\log n} \cdot \binom{n-1}{k-1}. \end{aligned}$$

We get

$$p_{fail}^{(d)} \le \exp\left(-\Omega\left(\frac{n\log n}{\log\log n}\right)\right)$$

which follows from Theorem 3.17. Since $\binom{n/2}{d/2}^2 \le 2^n$, we get

$$\binom{n/2}{d/2}^2 p_{fail}^{(d)} \le \exp\left(-\Omega\left(\frac{n\log n}{\log\log n}\right) + O(n)\right),$$

which is still $e^{-\Omega(\frac{n\log n}{\log\log n})}$.

If $d < \frac{n}{\log\log n}$, then $N_d = (1+o(1))\frac{d}{2^{k-1}} \cdot \binom{n-1}{k-1}$. Using Theorem 3.17 with $h(n) = d$, $c_1 = 1$, $c_2 = -1$, $\alpha_1 = \alpha$, $\alpha_2 = \beta$, $\rho_1 = \frac{1}{2^{k-1}}$ and $\rho_2 = \frac{1}{2^{k-1}}$, we get

$$p_{fail}^{(d)} \le \exp\left(-(1-o(1))I(\alpha,\beta) \cdot d\log n\right).$$

Since $\binom{n/2}{d/2}^2 \le n^d$ and $I(\alpha, \beta) > 1$, we have

$$\binom{n/2}{d/2}^2 p_{fail}^{(d)} \le \exp\left(-(1-o(1))(I(\alpha,\beta)-1)\cdot d\log n\right) \le C_1 n^{-d\epsilon/2}$$

for some constant $C_1 > 0$ which does not depend on n and it concludes the proof.

□

3.4 Truncate-and-relax algorithm

In this section, we propose an algorithm based on the standard semidefinite relaxation of maximization problem of quadratic function on the hypercube $\{\pm 1\}^V$. We also prove that this algorithm achieves the optimal threshold up to a constant multiplicative factor. We will only focus on the assortative case (i.e. $p > q$) but the algorithm could be adapted for disassortative cases with a different threshold which only depends on α and β, which we will not derive in this paper.

Let $H = (V(H), E(H))$ be a k-uniform hypergraph, and recall that we defined the weighted graph (G_H, w) on the vertex set $V(H)$ where weights are given by

$$w(ij) = \#(e \in E(H) : \{i,j\} \subseteq e).$$

We may think G_H be a multigraph realization of H, by replacing each hyperedge e in H by the k-clique on e. For brevity, let us define the adjacency matrix W of (G_H, w) as the symmetric n by n matrix such that its diagonal entries are zero, and $W_{ij} = w(ij)$ for each pair $\{i,j\} \subseteq V$. We defined the estimator $\widehat{\mathbf{x}}_{\text{trunc}}$ as

$$\widehat{\mathbf{x}}_{\text{trunc}} := \operatorname*{argmax}_{\mathbf{x} \in \{\pm 1\}^V : \mathbf{1}^T \mathbf{x} = 0} \sum_{1 \le i < j \le n} W_{ij} \mathbf{x}_i \mathbf{x}_j.$$

On the other hand, recall that the ML estimator $\widehat{\mathbf{x}}_{MLE}$ is the maximizer of

$$\langle A_H, \mathbf{x}^{\ominus k} \rangle = \sum_{e \in \binom{V}{k}} (A_H)_e (\mathbf{x}^{\ominus k})_e$$

over balanced x's. Note that

$$(\mathbf{x}^{\ominus k})_e = \frac{1}{2^{k-1}} \sum_{I \subseteq e : |I| \text{ even}} \mathbf{x}_I,$$

as we have seen in Chapter 2.

If we collate the terms by its degrees, then we have

$$\langle A_H, \mathbf{x}^{\ominus k} \rangle = \langle A_H, \mathbf{1} \rangle + \frac{1}{2^{k-1}} \sum_e \left((A_H)_e \cdot \frac{1}{2^{k-1}} \sum_{\{i,j\} \subseteq e} \mathbf{x}_i \mathbf{x}_j \right) + (\text{higher order terms}).$$

Let $g_H(\mathbf{x})$ be the homogeneous multilinear polynomial of degree 2 defined as

$$g_H(\mathbf{x}) = \sum_e (A_H)_e \sum_{\{i,j\} \subseteq e} \mathbf{x}_i \mathbf{x}_j,$$

which is a constant multiple of the degree 2 part of $\langle A_H, \mathbf{x}^{\ominus k} \rangle$. Then,

$$g_H(\mathbf{x}) = \sum_e \sum_{\{i,j\} \subseteq e} (A_H)_e \, \mathbf{x}_i \mathbf{x}_j = \sum_{1 \le i < j \le n} W_{ij} \, \mathbf{x}_i \mathbf{x}_j.$$

The truncate-and-relax algorithm tries to solve the standard SDP relaxation of $\max g_H(\mathbf{x})$, as described in Section 2.4. We prove that the optimum of the relaxation is achieved by $\mathbf{x}_0 \mathbf{x}_0^T$ with probability $1 - o(1)$ if α and β satisfies $I_{sdp}(\alpha, \beta) > 1$ where I_{sdp} is defined as in Section 3.1.3.

3.4.1 Laplacian of the adjacency matrix

Before we delve into the semidefinite relaxation that our algorithm uses, let us take a detour with a spectral algorithm which can also be thought as a relaxation of $\max g_H(\mathbf{x})$.

Recall that W is the adjacency matrix of a weighted graph. For $x \in \{\pm 1\}^n$ with the corresponding bisection (A, B) where $A = \{v : x_v = 1\}$ and $B = [n] \setminus A$, we have

$$g_H(\mathbf{x}) = \sum_{1 \le i < j \le n} W_{ij}\mathbf{x}_i\mathbf{x}_j = \sum_{1 \le i < j \le n} W_{ij} - 2 \sum_{i \in A, j \in B} W_{ij}$$

so maximizing $g_H(\mathbf{x})$ is equivalent to the minimum bisection problem ($\mathsf{MIN\text{-}BISECTION}$):

$$\mathsf{MIN\text{-}BISECTION}(W) : \min_{x \in \{\pm 1\}^n : \mathbf{1}^T\mathbf{x} = 0} \sum_{i \in A, j \in B} W_{ij}.$$

The *Laplacian* of W is a matrix L_W defined as $L_W = D - W$ where D is the diagonal matrix with entries

$$D_{vv} = \sum_{u \in [n]} W_{uv}.$$

Equivalently,

$$L_W = \sum_{1 \le i < j \le n} W_{ij}(e_i - e_j)(e_i - e_j)^T$$

where e_i is the vector with the entry $(e_i)_i = 1$ and zero elsewhere. It implies that

$$x^T L_W x = \sum_{i,j=1}^{n} W_{ij}(x_i - x_j)^2 = 4 \sum_{i \in A, j \in B} W_{ij},$$

hence $\mathsf{MIN\text{-}BISECTION}(W)$ is equivalent to minimizing $x^T L_W x$ over balanced x in $\{\pm 1\}^n$.

By relaxing the condition $x \in \{\pm 1\}^n$ to $\|x\|_2^2 = n$, we get

$$\min_{x : \|x\|_2^2 = n, \mathbf{1}^T\mathbf{x} = 0} x^T L_W x.$$

Note that L_W is positive semidefinite and the minimum eigenvalue of L_W is zero, since it is diagonally dominant and $\mathbf{1}^T L_W \mathbf{1} = 0$. Hence, the optimal solution of the relaxed problem corresponds to an eigenvector of the second smallest eigenvalue of L_W, scaled to have norm $\sqrt{n}$.

It motivates us to suggest a spectral algorithm with the following two stages:

- (Relaxation) We compute a unit eigenvector ξ of the second smallest eigenvalue of L_W.

- (Rounding) We round $\sqrt{n} \cdot \xi$ to the closest point on $\{\pm 1\}^n$, which corresponds to x with $x_v = \operatorname{sgn}(\xi_v)$.

We remark that in [49, 51, 52], the authors generalize this idea to the case when we have three or more clusters. Their algorithm computes eigenvectors of k smallest eigenvalues to associate each vertex with a vector in $\mathbb{R}^k$, and uses k-means clustering on them to find the community label of each vertex. Their algorithm has a few advantages such as that it applies to weak recovery and detection as well as exact recovery, that it generalizes to non-uniform models, and that it runs in nearly-linear time in n. However, it only works in a order-wise suboptimal parameter regime, requiring p and q be at least $\Omega\left(\frac{\log^2 n}{n^{k-1}}\right)$ for exact recovery.

Subsequently, in [15] and independently in [34], spectral algorithms with an additional local refinement step were proposed, and it was proved that both algorithms achieve exact recovery in the regime where p and q are $\Omega\left(\frac{\log n}{n^{k-1}}\right)$, which matches the statistical limit in terms of order in n. Also, we note that it was mentioned in [1] that the local refinement technique used for the SBM can be extended to the hypergraph case, together with a partial recovery algorithm in [20]. Finally, we remark that it was proved in [34] that their algorithm achieves the statistical limit shown in this paper. In other words, there is an efficient algorithm which recovers the ground truth almost asymptotically surely whenever $I(\alpha, \beta) > 1$.

3.4.2 Semidefinite relaxation and its dual

Let $X = xx^T$. Then, the condition that $\mathbf{x} \in \{\pm 1\}^V$ and $\mathbf{1}^T\mathbf{x} = 0$ is equivalent to that X is a symmetric n by n positive semidefinite rank-one matrix such that $X_{ii} = 1$ for all $i \in V$ and $\mathbf{1}^T X \mathbf{1} = 0$. If we relax the rank condition, then we get the following semidefinite program equivalent to (3.1) as argued in the introduction:

$$\begin{aligned} \text{maximize} \quad & \langle W, X \rangle \\ \text{subject to} \quad & X_{ii} = 1 \text{ for all } i \in [n] \\ & \langle X, \mathbf{1}\mathbf{1}^T \rangle = 0 \\ & X \succeq 0. \end{aligned} \tag{3.2}$$

The dual of (3.2) is

$$\begin{aligned} \text{minimize} \quad & tr(D) \\ \text{subject to} \quad & D \text{ is } n \times n \text{ diagonal matrix, } \lambda \in \mathbb{R} \\ & D + \lambda \mathbf{1}\mathbf{1}^T - W \succeq 0. \end{aligned} \tag{3.3}$$

We recall that $\widehat{\Sigma}$ was defined as the optimum solution of the primal program (3.2), and we say $\widehat{\Sigma}$ recovers the ground truth $\mathbf{x}_0$ if $\widehat{\Sigma} = \mathbf{x}_0\mathbf{x}_0^T$. In the case of $k = 2$ (the usual SBM), it is known that the relaxed SDP achieves exact recovery up to the statistical threshold even without the local refinement step [56]. We prove that for any $k \geq 2$ our algorithm successfully recover the ground truth, as long as $I_{sdp}(\alpha, \beta) > 1$ which is slightly weaker than the statistical threshold $I(\alpha, \beta) > 1$. On the other hand, we show that for $k \geq 4$ our algorithm fails with probability $1 - o(1)$ for some (α, β) even when exact recovery is statistically posible (see the next section).

Let X be an optimum solution of the primal and let (D, λ) be an optimum solution for dual. Then by complementary slackness we get $\langle X, D + \lambda \mathbf{1}\mathbf{1}^T - W \rangle = 0$. Conversely, if X is a feasible solution for the primal and (D, λ) is a feasible solution for the dual, then X and (D, λ) are optimal if $\langle X, D + \lambda \mathbf{1}\mathbf{1}^T - W \rangle = 0$. It implies that $X = \mathbf{x}_0\mathbf{x}_0^T$ is optimal if there exists dual feasible solution (D, λ) such that

$$\langle \mathbf{x}_0\mathbf{x}_0^T, D + \lambda \mathbf{1}\mathbf{1}^T - W \rangle = 0,$$

which is equivalent to

$$D_{ii} = \sum_{j \in V} W_{ij}(\mathbf{x}_0)_i(\mathbf{x}_0)_j$$

since $D + \lambda \mathbf{1}\mathbf{1}^T - W$ is positive semidefinite. Note that D is completely determined by W and $\mathbf{x}_0$.

Let $\Gamma = \operatorname{diag}(\mathbf{x}_0) W \operatorname{diag}(\mathbf{x}_0)$ and $D_\Gamma = \operatorname{diag}(\Gamma \mathbf{1})$. Note that D_Γ is equal to D defined above. Let $L_\Gamma = D_\Gamma - \Gamma$. Then,

$$L_\Gamma = \operatorname{diag}(\mathbf{x}_0)(D_\Gamma - W)\operatorname{diag}(\mathbf{x}_0)$$

by definition.

Proposition 3.9. *Let Π be the projector matrix onto the orthogonal complement of the span of $\{\mathbf{1}, \mathbf{x}_0\}$,* i.e.,

$$\Pi = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T - \frac{1}{n}\mathbf{x}_0\mathbf{x}_0^T.$$

If $\Pi L_\Gamma \Pi$ is positive semidefinite, then $\mathbf{x}_0\mathbf{x}_0^T$ is an optimal solution for (3.2). Moreover, if the third smallest eigenvalue of $\Pi L_\Gamma \Pi$ is positive, then $\mathbf{x}_0\mathbf{x}_0^T$ is the unique optimum.

Proof. First note that (D_Γ, λ) is a feasible solution for the dual if there exists λ such that $D_\Gamma - W + \lambda \mathbf{1}\mathbf{1}^T$ is positive semidefinite. By multiplying $\operatorname{diag}(\mathbf{x}_0)$ on the both side, it is equivalent to that $L_\Gamma + \lambda \mathbf{x}_0\mathbf{x}_0^T$ is positive semidefinite for some λ. This condition is satisfied if and only if $\Pi L_\Gamma \Pi \succeq 0$ and hence $\mathbf{x}_0\mathbf{x}_0^T$ is an optimal solution for the primal.

Moreover, if $\lambda_3(\Pi L_\Gamma \Pi) > 0$ then there exists λ such that $L_\Gamma + \lambda \mathbf{x}_0\mathbf{x}_0^T$ is positive definite on the orthogonal complement of $\mathbf{1}$. It immediately implies that $\mathbf{x}_0\mathbf{x}_0^T$ is the unique optimal solution for the primal. $\square$

In the remainder of this section, we present and prove a sufficient condition for $\lambda_3(\Pi L_\Gamma \Pi) > 0$. We also present and prove a necessary condition for $\widehat{\mathbf{x}}_{\text{trunc}}$ being $\mathbf{x}_0$ up to global sign flip with high probability.

3.4.3 Performance of the algorithm

We first present the main result of this section.

Theorem 3.10. *Suppose that α and β satisfies*

$$\frac{k-1}{2^{k-1}}(\alpha-\beta) > \sqrt{4(k-1)\left(\frac{k}{2^k}\alpha + \left(1-\frac{k}{2^k}\right)\beta\right)}.$$

Then, $\lambda_3(\Pi L_\Gamma \Pi) > 0$ with probability $1-o(1)$.

We remark that Theorem 3.10 implies Theorem 3.2. To prove 3.10, we are going to use standard concentration result for the norm of the sum of random matrices. We first note that

$$\begin{aligned}\Pi L_\Gamma \Pi &= \mathbb{E}\,\Pi L_\Gamma \Pi + (\Pi L_\Gamma \Pi - \mathbb{E}\,\Pi L_\Gamma \Pi)\\ &= \Pi(\mathbb{E}\,L_\Gamma)\Pi + \Pi(L_\Gamma - \mathbb{E}\,L_\Gamma)\Pi.\end{aligned}$$

We would like to prove that if α and β satisfies the condition in 3.10, then with probability $1-o(1)$,

$$\lambda_3\left(\Pi(\mathbb{E}\,L_\Gamma)\Pi\right) > \|\Pi(L_\Gamma - \mathbb{E}\,L_\Gamma)\Pi\|.$$

Let $\mathbf{1}_e$ be the vector in $\mathbb{R}^V$ where $(\mathbf{1}_e)_i = 1$ when $i \in e$ and $(\mathbf{1}_e)_i = 0$ otherwise. Let $(\mathbf{x}_0)_e = \mathrm{diag}(\mathbf{x}_0)\mathbf{1}_e$. We note that

$$Q_{\mathcal{H}} = \sum_e (A_{\mathcal{H}})_e \left(\mathbf{1}_e\mathbf{1}_e^T - \mathrm{diag}(\mathbf{1}_e)\right),$$

and so

$$\Gamma = \sum_e (A_{\mathcal{H}})_e \left((\mathbf{x}_0)_e(\mathbf{x}_0)_e^T - \mathrm{diag}(\mathbf{1}_e)\right).$$

It implies that

$$\begin{aligned}
L_\Gamma &= \sum_e (A_\mathcal{H})_e L_{(\mathbf{x}_0)_e(\mathbf{x}_0)_e^T - \mathrm{diag}(\mathbf{1}_e)} \\
&= \sum_e (A_\mathcal{H})_e \left((\mathbf{1}_e^T(\mathbf{x}_0)_e)\mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right).
\end{aligned}$$

Proposition 3.11.

$$\Pi(\mathbb{E}\, L_\Gamma)\Pi = \frac{p-q}{2} n \binom{n/2-2}{k-2} \cdot \Pi,$$

hence $\lambda_3(\Pi(\mathbb{E}\, L_\Gamma)\Pi) = \frac{k-1}{2^{k-1}}(\alpha-\beta)\log n + o(\log n)$.

Proof. Note that $\mathbb{E}\, L_\Gamma$ is invariant under the permutation of V which preserves $\mathbf{x}_0$. Hence, we can write $\mathbb{E}\, L_\Gamma$ as

$$\mathbb{E}\, L_\Gamma = \frac{\langle \mathbb{E}\, L_\Gamma, \Pi\rangle}{n-2}\Pi + \frac{\langle \mathbb{E}\, L_\Gamma, \mathbf{1}\mathbf{1}^T\rangle}{n^2}\mathbf{1}\mathbf{1}^T + \frac{\langle \mathbb{E}\, L_\Gamma, \mathbf{x}_0\mathbf{x}_0^T\rangle}{n^2}\mathbf{x}_0\mathbf{x}_0^T.$$

We have $\langle \mathbb{E}\, L_\Gamma, \mathbf{1}\mathbf{1}^T\rangle = \mathbf{1}^T(\mathbb{E}\, L_\Gamma)\mathbf{1} = 0$ by definition of L_Γ. Also,

$$\begin{aligned}
\langle \mathbb{E}\, L_\Gamma, \mathbf{x}_0\mathbf{x}_0^T\rangle &= \sum_e \mathbb{E}(A_\mathcal{H})_e \cdot \langle \mathbf{1}_e^T(\mathbf{x}_0)_e \cdot \mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T, \mathbf{x}_0\mathbf{x}_0^T\rangle \\
&= \sum_e \mathbb{E}(A_\mathcal{H})_e((\mathbf{1}_e^T(\mathbf{x}_0)_e)^2 - k^2) \\
&= q\sum_{r=1}^{k-1}((k-2r)^2 - k^2)\binom{n/2}{r}\binom{n/2}{k-r},
\end{aligned}$$

which is equal to $-qn^2\binom{n-2}{k-2}$. On the other hand,

$$\begin{aligned}
tr(\mathbb{E}\, L_\Gamma) &= \sum_e \mathbb{E}(A_\mathcal{H})_e \cdot tr(\mathbf{1}_e^T(\mathbf{x}_0)_e \cdot \mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T) \\
&= \sum_e \mathbb{E}(A_\mathcal{H})_e \cdot ((\mathbf{1}_e^T(\mathbf{x}_0)_e)^2 - k) \\
&= (k^2-k)\left(q\binom{n}{k} + 2(p-q)\binom{n/2}{k}\right) - qn^2\binom{n-2}{k-2} \\
&= 2(p-q)(k^2-k)\binom{n/2}{k} - qn\binom{n-2}{k-2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\langle \mathbb{E}\, L_\Gamma, \Pi \rangle &= tr(\mathbb{E}\, L_\Gamma) - \frac{1}{n}\left\langle \mathbb{E}\, L_\Gamma, \mathbf{x}_0\mathbf{x}_0^T \right\rangle \\
&= 2(p-q)(k^2-k)\binom{n/2}{k} - qn\binom{n-2}{k-2} + qn\binom{n-2}{k-2} \\
&= \frac{p-q}{2}\cdot n(n-2)\binom{n/2-2}{k-2}.
\end{aligned}$$

We get

$$\Pi\, \mathbb{E}\, L_\Gamma \Pi = \frac{\langle \mathbb{E}\, L_\Gamma, \Pi\rangle}{n-2}\Pi = \frac{p-q}{2} n \binom{n/2-2}{k-2} \cdot \Pi.$$

□

Now, let us bound the operator norm of $\Pi(L_\Gamma - \mathbb{E}\, L_\Gamma)\Pi$. We need the following version of Matrix Bernstein inequality [91].

Theorem 3.12 (Matrix Bernstein inequality). *Let $\{X_k\}$ be a finite sequence of independent, symmetric random matrices of dimension N. Suppose that $\mathbb{E}\, X_k = 0$ and $\|X_k\| \leq M$ almost surely for all k. Then for all $t \geq 0$,*

$$\mathbb{P}\left(\left\|\sum_k X_k\right\| \geq t\right) \leq N \cdot \exp\left(-\frac{t^2/2}{\sigma^2 + Mt/3}\right) \quad \textit{where } \sigma^2 = \left\|\sum_k \mathbb{E}\, X_k^2\right\|.$$

Recall that

$$L_\Gamma = \sum_e (A_{\mathcal{H}})_e \left((\mathbf{1}_e^T(\mathbf{x}_0)_e)\mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right).$$

Hence,

$$\Pi(L_\Gamma - \mathbb{E}\, L_\Gamma)\Pi = \sum_e \left((A_{\mathcal{H}})_e - \mathbb{E}(A_{\mathcal{H}})_e\right) \cdot \Pi \left((\mathbf{1}_e^T(\mathbf{x}_0)_e)\mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)\Pi.$$

We note that

$$\left\|\Pi\left(\mathbf{1}_e^T(\mathbf{x}_0)_e \cdot \mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)\Pi\right\| \leq |\mathbf{1}_e^T(\mathbf{x}_0)_e| + \|(\mathbf{x}_0)_e\|^2 \leq 2k$$

for any $e \in \binom{V}{k}$. By Matrix Bernstein inequality, we have

$$\mathbb{P}\left(\|\Pi(L_\Gamma - \mathbb{E}\, L_\Gamma)\Pi\| \geq t\right) \leq n \cdot \exp\left(-\frac{t^2/2}{\sigma^2 + 2kt/3}\right)$$

where

$$\sigma^2 = \left\|\sum_e \mathbb{E}\left((A_{\mathcal{H}})_e - \mathbb{E}(A_{\mathcal{H}})_e\right)^2 \cdot \left(\Pi\left((\mathbf{1}_e^T(\mathbf{x}_0)_e)\mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)\Pi\right)^2\right\|.$$

If $\sigma = \omega_n(1)$, then we have

$$\begin{aligned} \mathbb{P}\left(\|\Pi(L_\Gamma - \mathbb{E}\, L_\Gamma)\Pi\| \geq (1+\epsilon)\sigma\sqrt{2\log n}\right) &\leq n \cdot \exp\left(-(1+\epsilon)^2\log n + o(\log n)\right) \\ &\leq n^{-2\epsilon + o(1)}, \end{aligned}$$

by letting $t = (1+\epsilon)\sigma\sqrt{2\log n}$.

Proposition 3.13.

$$\sigma^2 \leq 2(k-1)\left(\frac{k}{2^k}\alpha + \left(1 - \frac{k}{2^k}\right)\beta\right)\log n + O\left(\frac{\log n}{n}\right).$$

Proof. Let Y_e be

$$Y_e := \left(\mathbf{1}_e^T(\mathbf{x}_0)_e \cdot \mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)\Pi\left(\mathbf{1}_e^T(\mathbf{x}_0)_e \cdot \mathrm{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right),$$

and let $\Sigma = \sum_e \mathbb{E}(A_{\mathcal{H}})_e Y_e$. We have

$$\begin{aligned} \sigma^2 &= \left\|\sum_e \mathbb{E}\left((A_{\mathcal{H}})_e - \mathbb{E}(A_{\mathcal{H}})_e\right)^2 \cdot \Pi Y_e \Pi\right\| \\ &\leq \left\|\sum_e \mathbb{E}(A_{\mathcal{H}})_e \cdot \Pi Y_e \Pi\right\| = \|\Pi\Sigma\Pi\|, \end{aligned}$$

since Y_e is positive semidefinite and $\mathbb{E}\left((A_{\mathcal{H}})_e - \mathbb{E}(A_{\mathcal{H}})_e\right)^2 \leq \mathbb{E}(A_{\mathcal{H}})_e$. Moreover, we obtain the exact expression of Σ in the following lemma.

Lemma 3.14. $\Sigma = c_1 \cdot \frac{1}{n}\mathbf{x}_0\mathbf{x}_0^T + c_2 \cdot \Pi$ *where*

$$c_1 = (k-1)\beta \log n + O\left(\frac{\log n}{n}\right)$$

and

$$c_2 = 2(k-1)\left(\frac{k}{2^k}\alpha + \left(1 - \frac{k}{2^k}\right)\beta\right)\log n + O\left(\frac{\log n}{n}\right)$$

The proof of Lemma 3.14 is deferred to Section 3.7.2. As a result, we get $\Pi\Sigma\Pi = c_2\Pi$ and $\sigma^2 \le \|\Pi\Sigma\Pi\| = c_2$ as desired. $\square$

We are now ready to prove Theorem 3.10.

Proof of Theorem 3.10. Let ϵ be an arbitrary positive real number. By Matrix Bernstein Inequality and Proposition 3.13, with probability $1 - O(n^{-2\epsilon+o(1)})$ we have

$$\|\Pi(L_\Gamma - \mathbb{E}\, L_\Gamma)\Pi\| < (1+\epsilon)\sqrt{2c_2 \log n}$$

where

$$c_2 = (k-1)\left(k(1-2^{-k+1})\alpha + (2^{-k+1} - k + 2)\beta\right)\log n + o(\log n).$$

It implies that $\lambda_3(\Pi L_\Gamma \Pi) > 0$ with probability $1 - O(n^{-2\epsilon+o(1)})$ if

$$\frac{k-1}{2^{k-1}}(\alpha - \beta)\log n + o(\log n) > (1+\epsilon)\sqrt{2c_2 \log n}.$$

It holds as long as

$$\frac{k-1}{2^{k-1}}(\alpha - \beta) > \sqrt{4(k-1)\left(\frac{k}{2^k}\alpha + \left(1 - \frac{k}{2^k}\right)\beta\right)},$$

as desired. $\square$

3.4.4 Limitation of the algorithm

In the previous section, we proved that the truncate-and-relax algorithm successfully recovers $\mathbf{x}_0$ with high probability if α and β satisfies

$$\frac{k-1}{2^{k-1}}(\alpha-\beta) > \sqrt{4(k-1)\left(\frac{k}{2^k}\alpha + \left(1-\frac{k}{2^k}\right)\beta\right)}.$$

It is natural to ask whether this bound is improvable or not. Recall that $\widehat{\mathbf{x}}_{\text{trunc}}$ is the optimum solution for $\max x^T W x$ over balanced x's in $\{\pm 1\}^n$. Since our algorithm is the relaxed version of it, we have

$$\mathbb{P}\left(\widehat{\Sigma} \neq \mathbf{x}_0\mathbf{x}_0^T\right) \geq \mathbb{P}\left(\widehat{\mathbf{x}}_{\text{trunc}}(\mathcal{H}) \notin \{\mathbf{x}_0, -\mathbf{x}_0\}\right).$$

The following theorem gives a condition on α and β such that that the probability that $\widehat{\mathbf{x}}_{\text{trunc}}(\mathcal{H})$ fails to recover $\mathbf{x}_0$ is $1-o(1)$.

Theorem 3.15. *Let*

$$I_2(\alpha,\beta) = \max_{t\geq 0} \frac{1}{2^{k-1}}\left(\alpha(1-e^{-(k-1)t}) + \beta\sum_{r=1}^{k-1}\binom{k-1}{r}(1-e^{-(k-1-2r)t})\right).$$

If $I_2(\alpha,\beta) < 1$, then $\mathbb{P}\left(\widehat{\mathbf{x}}_{\text{trunc}}(\mathcal{H}) \notin \{-\mathbf{x}_0, \mathbf{x}_0\}\right) = 1-o(1)$. In particular, the truncate-and-relax algorithm fails to recover $\mathbf{x}_0$ with probability $1-o(1)$.

Proof. The proof is a slight modification of the proof of Theorem 3.1. Essentially it reduces to prove that

$$\mathbb{P}\left(X_a \leq -\frac{2\log n}{\log\log n}\right) \geq n^{-I_2(\alpha,\beta)-o(1)}.$$

where

$$X_a = \sum_{e: e\cap U=\{a\}} (A_{\mathcal{H}})_e\left(\sum_{\{i,j\}\subseteq e} (\mathbf{x}_0)_i(\mathbf{x}_0)_j - (\mathbf{x}_0^{(a)})_i(\mathbf{x}_0^{(a)})_j\right)$$

and $U = U_A \cup U_B$ and $a \in U_A$ are defined as in the proof of Theorem 3.1, and this tail bound follows from the Theorem 3.17. Details are deferred to Section 3.6.3. □

3.5 Discussion

Let us first recapitulate the main results of this paper. In the stochastic block model for k-uniform hypergraphs where the (hyper)edge probabilities are given as

$$p = \frac{\alpha \log n}{\binom{n-1}{k-1}} \quad \text{and} \quad q = \frac{\beta \log n}{\binom{n-1}{k-1}}$$

for some constants α and β such that $\alpha > \beta > 0$, we observed the following phase transition behaviours on exact recovery problem:

(i) If $I(\alpha, \beta) < 1$, then exact recovery is not possible. Conversely, if $I(\alpha, \beta) < 1$ then the ML estimator recovers the correct partition (up to a global sign flip) with probability $1 - o(1)$.

(ii) If $I_{sdp}(\alpha, \beta) > 1$, then the truncate-and-relax algorithm recovers the partition (up to a global sign flip) with probability $1 - o(1)$.

(iii) If $I_2(\alpha, \beta) < 1$, then the truncate-and-relax algorithm fails with probability $1 - o(1)$.

Here I, I_2 and I_{sdp} are functions depending on α and β (and implicitly depending on k, which we assumed to be a constant) defined as

$$\begin{aligned}
I(\alpha, \beta) &= \frac{1}{2^{k-1}} \left(\sqrt{\alpha} - \sqrt{\beta} \right)^2 \\
I_2(\alpha, \beta) &= \max_{t \geq 0} \frac{\alpha - \beta}{2^{k-1}} (1 - e^{-(k-1)t}) + \beta \left(1 - \left(\frac{e^t + e^{-t}}{2} \right)^{k-1} \right) \\
I_{sdp}(\alpha, \beta) &= \frac{k-1}{2^{2k}} \frac{(\alpha - \beta)^2}{\left(\frac{k}{2^k} \alpha + \left(1 - \frac{k}{2^k} \right) \beta \right)}.
\end{aligned}$$

We first note that sharp phase transition occurs at $I(\alpha, \beta) = 1$ for exact recovery. Indeed, it can be *efficiently* achieved, by spectral algorithms with a local refinement step as suggested in [34]. Specifically authors of [34] prove that their algorithm achieves exact recovery whenever $I(\alpha, \beta) > 1$ and conjectured that $I(\alpha, \beta) = 1$ is the sharp threshold. We confirmed their conjecture in this work. On the other hand, there is a gap between the guaranteed performance of the truncate-and-relax algorithm and the impossibility region of the algorithm as shown in Figure 3-1 and Figure 3-2. We are yet to show how the algorithm works in between, which is when α and β satisfies $I_{sdp}(\alpha, \beta) < 1$ but $I_2(\alpha, \beta) > 1$. We propose that the line $I_2(\alpha, \beta) = 1$ is the correct threshold for the performance guarantee of the algorithm.

Conjecture 3.5.1. *If $I_2(\alpha, \beta) > 1$, then the truncate-and-relax algorithm successfully recovers $\mathbf{x}_0\mathbf{x}_0^T$ with probability $1 - o(1)$.*

There are a few reasons to believe this conjecture. First, if we look deeper into the proof of Theorem 3.2 then the main obstacle to prove the conjecture arises from when we use the matrix Bernstein inequality to bound $\|\Pi(L_\Gamma - \mathbb{E}\, L_\Gamma)\Pi\|$. The matrix Bernstein inequality gives us that

$$\mathbb{E}\, \|\Pi(L_\Gamma - \mathbb{E}\, L_\Gamma)\Pi\| \lesssim \sigma\sqrt{\log n}$$

where

$$\sigma^2 = \left\|\mathbb{E}(\Pi(L_\Gamma - \mathbb{E}\, L_\Gamma)\Pi)^2\right\|.$$

In the case of $k = 2$, the random matrix Γ has independent entries and one can obtain a tighter bound for $\|L_\Gamma - \mathbb{E}\, L_\Gamma\|$, via combinatorial method [44], stochastic comparison argument [56], or trace method [22]. Also, in [21] the following bound for Laplacian random matrices was proved.

Theorem 3.16. *Let L be a $n \times n$ symmetric random Laplacian matrix (i.e. satisfying $L\mathbf{1} = 0$) with centered independent off-diagonal entries such that $\sum_{j\in[n]\setminus\{i\}} \mathbb{E}\, L_{ij}^2$ is equal for all i, and*

$$\sum_{j\in[n]\setminus\{i\}} \mathbb{E}\, L_{ij}^2 \gtrsim \max_{i\neq j} \|L_{ij}\|_\infty^2 \log n.$$

Then, with high probability,

$$\|L\| \lesssim \left(1 + \frac{1}{\sqrt{\log n}}\right) \max_i L_{ii}.$$

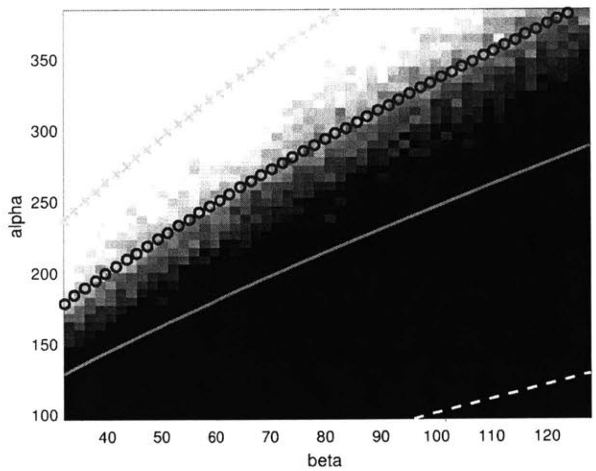

Figure 3-2: Result of simulation of the truncate-and-relax algorithm for $k = 6$ and $n = 500$. Each gray-scale block corresponds to a pair (α, β), and its color denotes the success rate over 30 trials (black corresponds to 0 success, and brighter color correspond to higher success rate). The solid line represents $I(\alpha, \beta) = 1$, the circled line represents $I_2(\alpha, \beta) = 1$, and the x-marked line represents $I_{sdp}(\alpha, \beta) = 1$.

This bound cannot be used for $k > 2$ as entries of Γ are not independent to each other. We ask whether the bound could be extended to our setting: Can we obtain a similar bound when L can be expressed as

$$L = \sum_{S \subseteq [n], |S| = k} \xi_S L^{(S)},$$

where $L^{(S)}$ is $n \times n$ symmetric Laplacian matrix such that $L^{(S)}_{ij}$ is non-zero only if $i, j \in S$?

We ran a simulation to support our conjecture. For each α and β, we generated 30 random hypergraphs according to the model, and constructed the dual certificate for each hypergraph as in the proof of Theorem 3.2. When the constructed dual solution is positive-semidefinite, it was counted as a success. Figure 3-2 shows the result of the simulation and it suggests that the true phase transition occurs at $I_2(\alpha, \beta) = 1$ as we proposed.

3.6 Tail probability of weighted sum of binomial variables

In this section, we investigate the precise asymptotics of the tail probability of weighted sum of independent binomial variables. Using Theorem 3.17, we derive the formulas which were used to prove information-theoretic limits in section 3.3 and 3.4.

Theorem 3.17. *Let r and s be positive integers. Let $c_1, c_2, \cdots, c_r$ be non-zero real numbers. Let $h(n)$ be a non-decreasing function which is $\Omega(1)$ and $o(n^{s/2}/\log n)$. For each $i \in [r]$, let Y_i be the random variable distributed as the binomial distribution* $\mathrm{Bin}(N_i, p_i)$ *where*

$$N_i = (1+o(1))\rho_i \cdot h(n)\binom{n}{s} \quad \text{and} \quad p_i = (1+o(1))\alpha_i \cdot \frac{\log n}{\binom{n}{s}},$$

for some positive constant (not depending on n) ρ_i and α_i. Let $X = \sum_{i=1}^r c_i Y_i$. Suppose that (i) not all c_i are positive, and (ii) $\sum_{i=1}^r c_i \alpha_i \rho_i > 0$. Then, for any $\delta \in (-\infty, \sum_{i=1}^r c_i \alpha_i \rho_i)$, we have

$$\mathbb{P}(X \leq (1+o(1))\delta \cdot h(n)\log n) = \exp\left(-(1+o(1))I^* \cdot h(n)\log n\right),$$

where

$$I^* = \max_{t \geq 0}\left(-t\delta + \sum_{i=1}^r \alpha_i \rho_i (1 - e^{-tc_i})\right).$$

Remark. The condition $h(n) = o(n^{s/2}/\log n)$ is not required for the upper bound.

Proof. Let us first prove the upper bound on the tail probability. Let $x = (1+o(1))\delta \cdot$

$h(n)\log n$. Then, by Chebyshev-type inequality, for any $t \geq 0$ we have

$$\begin{aligned}
\mathbb{P}(X \leq x) &\leq \frac{\mathbb{E}\, e^{-tX}}{e^{-tx}} \\
&= e^{tx} \prod_{i=1}^{r} \mathbb{E}\, e^{-tc_i Y_i} \\
&= e^{tx} \prod_{i=1}^{r} \left(1 - p_i(1 - e^{-c_i t})\right)^{N_i} \\
&\leq \exp\left(tx - \sum_{i=1}^{r} N_i p_i (1 - e^{-c_i t}) \right) \\
&= \exp\left(-(1 + o(1)) h(n) \log n \cdot \left(-t\delta + \sum_{i=1}^{r} \alpha_i \rho_i (1 - e^{-c_i t}) \right) \right).
\end{aligned}$$

Here the fourth inequality follows from $1 - x \leq e^{-x}$. By optimizing over $t \geq 0$, we get the desired bound.

To prove the lower bound, note that

$$\mathbb{P}(X \leq x) \geq \prod_{i=1}^{r} \mathbb{P}(Y_i = y_i) = \prod_{i=1}^{r} \binom{N_i}{y_i} p_i^{y_i} (1 - p_i)^{N_i - y_i}$$

for any positive integers $y_1, \cdots, y_r$ satisfying $\sum_{i=1}^{r} c_i y_i \leq x$.

Let $\phi(t) = -\delta t + \sum_{i=1}^{r} \alpha_i \rho_i (1 - e^{-c_i t})$ and let t^* be the maximizer of $\phi(t)$. Note that $\phi(t)$ is strictly convex, as

$$\phi''(t) = \sum_{i=1}^{r} c_i^2 \alpha_i \rho_i e^{-c_i t} > 0$$

for any $t \geq 0$. Moreover, $\phi'(0) = \sum_{i=1}^{r} c_i \alpha_i \rho_i - \delta > 0$ and $\lim_{t \to \infty} \phi'(t) = -\infty$. Hence, there exists unique t^* satisfying $\phi'(t^*) = 0$, which is the maximizer of $\phi(t)$.

Let $\tau_i = \alpha_i \rho_i e^{-c_i t^*}$ for $i \in [r]$ and let $y_1, \cdots, y_r$ be integers such that $\sum_{i=1}^{r} c_i y_i \leq x$ and $y_i = (1 - o(1)) \tau_i \cdot h(n) \log n$. Such y_i's exist because

$$\sum_{i=1}^{r} c_i \tau_i = \sum_{i=1}^{r} c_i \alpha_i \rho_i e^{-c_i t^*} = \delta.$$

We are going to use the following bound for the binomial coefficient for $\ell \le \sqrt{N}$

$$\binom{N}{\ell} \ge \frac{N^\ell}{4 \cdot \ell!}.$$

By Stirling's approximation, we have $\ell! \le e\sqrt{\ell} \cdot \left(\frac{\ell}{e}\right)^\ell$ so

$$\log \binom{N}{\ell} \ge \ell \log\left(\frac{eN}{\ell}\right) - \log(4e\sqrt{\ell}).$$

Note that $y_i \ll \sqrt{N_i}$ since $h(n) = o(n^{s/2}/\log n)$. Hence,

$$\log\left(\binom{N_i}{y_i} p_i^{y_i}(1-p_i)^{N_i - y_i}\right) \ge y_i \log\left(\frac{eN_i p_i}{(1-p_i)y_i}\right) + N_i \log(1-p_i) - \log(4e\sqrt{y_i}).$$

Moreover,

$$\begin{aligned} y_i \log\left(\frac{eN_i p_i}{(1-p_i)y_i}\right) &= (1-o(1))h(n)\log n \cdot \tau_i \log\left(\frac{e\alpha_i\rho_i}{\tau_i}\right) \\ N_i \log(1-p_i) &= -(1+o(1))\alpha_i\rho_i \cdot h(n)\log n \\ \log(4e\sqrt{y_i}) &= o(h(n)\log n). \end{aligned}$$

We get

$$\mathbb{P}(X \le x) \ge \exp\left(-(1+o(1))h(n)\log n \cdot \sum_{i=1}^{r}\left(\alpha_i\rho_i - \tau_i \log\left(\frac{e\alpha_i\rho_i}{\tau_i}\right)\right)\right).$$

Plugging in $\tau_i = \alpha_i\rho_i e^{-c_i t^*}$, we get

$$\begin{aligned} \sum_{i=1}^{r}\left(\alpha_i\rho_i - \tau_i \log\left(\frac{e\alpha_i\rho_i}{\tau_i}\right)\right) &= \sum_{i=1}^{r}\left(\alpha_i\rho_i - \alpha_i\rho_i(1+c_i t^*)e^{-c_i t^*}\right) \\ &= -t^* \sum_{i=1}^{r} c_i\alpha_i\rho_i e^{-c_i t^*} + \sum_{i=1}^{r}\alpha_i\rho_i(1-e^{-c_i t^*}) \\ &= -\delta t^* + \sum_{i=1}^{r}\alpha_i\rho_i(1-e^{-c_i t^*}) \\ &= I^*, \end{aligned}$$

where the third equality follows from that $\sum_{i=1}^{r} c_i\alpha_i\rho_i e^{-c_i t^*} = \delta$. Hence,

$$\mathbb{P}(X \le x) \ge \exp\left(-(1+o(1))I^* \cdot h(n)\log n\right)$$

as desired. $\square$

3.6.1 Proof of Lemma 3.6

Let us restate the lemma for readers.

Lemma 3.18 (Lemma 3.6, restated). *Let X be a sum of independent Bernoulli variables such that $\mathbb{E}\,X = \Theta(\gamma \log n)$ where $\gamma = o_n(\log^{-1} n)$. Let δ be a positive number which decays to 0 as n grows, with $\delta = \omega_n(\log^{-1} n)$. Then,*

$$\mathbb{P}\left(X > \delta \log n\right) \le n^{-\delta \log \frac{\delta}{\gamma} + o(1)}.$$

Proof. A standard Chernoff's bound implies that

$$\mathbb{P}(X \ge t\,\mathbb{E}\,X) \le \left(\frac{e^{t-1}}{t^t}\right)^{\mathbb{E}\,X}$$

for any $t \ge 0$. Let $t = \frac{\delta \log n}{\mathbb{E}\,X}$. Then,

$$\begin{aligned}
\mathbb{P}(X \ge \delta \log n) \le \quad & \exp\left((t-1-t\log t)\,\mathbb{E}\,X\right) \\
= \quad & \exp\left(\left(1 - \log\frac{\delta \log n}{\mathbb{E}\,X}\right)\delta \log n - \mathbb{E}\,X\right) \\
= \quad & \exp\left(\left(1 - \log\frac{\delta}{\gamma} + O(1)\right)\delta \log n - o(1)\right) \\
= \quad & n^{-\delta \log \frac{\delta}{\gamma} + o(1)}
\end{aligned}$$

as desired. $\square$

3.6.2 Proof of Lemma 3.7

Let $a \in U_A$. Recall that

$$X_a = \sum_{e : e \cap U = \{a\}} c_e (A_{\mathcal{H}})_e,$$

where $c_e = \Sigma_e - \Sigma_e^{(ab)}$ for any $b \in U_B$. Concretely, the value of c_e for e satisfying $e \cap U = \{a\}$ is determined by the size of intersection of $e \setminus \{a\}$ and $A \setminus U_A$ as follows:

$$c_e = \begin{cases} +1 & \text{if } e \setminus \{a\} \subseteq A \setminus U_A \\ -1 & \text{if } (e \setminus \{a\}) \cap (A \setminus U_A) = \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

Hence, $X_a = Y_1 - Y_2$ where Y_1 and Y_2 are independent random variables such that $Y_1 \sim \mathrm{Bin}(N_1, p)$ and $Y_2 \sim \mathrm{Bin}(N_2, q)$ with

$$N_1 = \binom{n/2 - |U_A|}{k-1} \quad \text{and} \quad N_2 = \binom{n/2 - |U_B|}{k-1}.$$

Using Theorem 3.17 with

$$c_1 = 1, \quad c_2 = -1, \quad \alpha_1 = \alpha, \quad \alpha_2 = \beta,$$
$$\rho_1 = \rho_2 = \frac{1}{2^{k-1}}, \quad h(n) = 1, \quad \text{and} \quad \delta = 0,$$

we get

$$\mathbb{P}\left(X_a \le \frac{\log n}{\log \log n} \right) = \exp\left(-(1 + o(1)) I^* \cdot \log n \right),$$

where

$$\begin{aligned} I^* &= \max_{t \ge 0} \sum_{i=1}^{2} \alpha_i \rho_i (1 - e^{-c_i t}) \\ &= \max_{t \ge 0} \frac{1}{2^{k-1}} \left(\alpha(1 - e^{-t}) + \beta(1 - e^{t}) \right). \end{aligned}$$

The maximum is attained at $t^* = \frac{1}{2} \log\left(\frac{\alpha}{\beta}\right) > 0$ and

$$I^* = \frac{1}{2^{k-1}} \left(\sqrt{\alpha} - \sqrt{\beta} \right)^2 = 1 - \epsilon.$$

Hence,

$$\mathbb{P}\left(X_a \leq \frac{\log n}{\log \log n}\right) \geq n^{-1+\epsilon-o(1)}$$

as desired.

3.6.3 Proof of the tail bound in Theorem 3.15

We recall that X_a is defined as

$$X_a = \sum_{e: e \cap U = \{a\}} (A_{\mathcal{H}})_a \left(\sum_{ij \subseteq e} (\mathbf{x}_0)_i (\mathbf{x}_0)_j - (\mathbf{x}_0^{(a)})_i (\mathbf{x}_0^{(a)})_j \right).$$

By definition of $(\mathbf{x}_0)^{(a)}$, we have

$$\begin{aligned} \left(\sum_{ij \subseteq e} (\mathbf{x}_0)_i (\mathbf{x}_0)_j - (\mathbf{x}_0^{(a)})_i (\mathbf{x}_0^{(a)})_j \right) &= 2(\mathbf{x}_0)_a \sum_{i \in e \setminus \{a\}} (\mathbf{x}_0)_i \\ &= 2\left(k-1-2|e \cap B|\right). \end{aligned}$$

Hence, $X_a = \sum_{r=0}^{k-1} c_r Y_r$ where $c_r = 2(k-1-2r)$ and $Y_r \sim \mathrm{Bin}(N_r, p_r)$ with

$$N_r = \binom{n/2 - |U_A|}{k-1-r}\binom{n/2-|U_B|}{r} = (1+o(1))\frac{1}{2^{k-1}}\binom{k-1}{r}\binom{n}{k-1}$$

and $p_r = p$ if $r = 0$ and $p_r = q$ otherwise.

Using Theorem 3.17 with

$$c_r = 2(k-1-2r), \quad \rho_r = \frac{1}{2^{k-1}}\binom{k-1}{r}, \quad \alpha_r = \begin{cases} \alpha & \text{if } r = 0 \\ \beta & \text{otherwise,} \end{cases}$$

and $h(n) = 1$ and $\delta = 0$, we have

$$\mathbb{P}\left(X_a \leq -\frac{2\log n}{\log\log n}\right) = \exp\left(-(1+o(1))I_2 \cdot \log n\right)$$

where

$$\begin{aligned} I_2 = & \max_{t\geq 0} \sum_{r=0}^{k-1} \alpha_i \rho_i (1 - e^{-c_i t}) \\ = & \max_{t\geq 0} \frac{1}{2^{k-1}} \left(\alpha(1 - e^{-(k-1)t}) + \beta \sum_{r=1}^{k-1} \binom{k-1}{r} (1 - e^{-(k-1-2r)t}) \right), \end{aligned}$$

as desired.

3.7 Miscellaneous proofs

3.7.1 Proof of Proposition 3.4

Recall that the maximum-likelihood estimator $\widehat{\mathbf{x}}_{MLE}(H)$ is defined as

$$\widehat{\mathbf{x}}_{MLE}(H) = \operatorname*{argmax}_{\mathbf{x}\in\{\pm 1\}^V : \mathbf{1}^T\mathbf{x}=0} f_H(\mathbf{x}),$$

where $f_H(\mathbf{x}) = \log \mathbb{P}_{(\mathbf{x}_0, \mathcal{H})}(\mathcal{H} = H | \mathbf{x}_0 = \mathbf{x})$. Note that

$$\begin{aligned} f_H(\mathbf{x}) &= \log \mathbb{P}(\mathcal{H} = H | \mathbf{x}_0 = x) \\ &= \log \left(\prod_{e\in\binom{V}{k}} \mathbb{P}(e \in E(\mathcal{H}) | \mathbf{x}_0 = \mathbf{x})^{(A_H)_e} \, \mathbb{P}(e \notin E(\mathcal{H}) | \mathbf{x}_0 = \mathbf{x})^{1-(A_H)_e} \right) \\ &= \sum_{\substack{e\in\binom{V}{k} \\ e\text{: in-cl. w.r.t. } x}} (A_H)_e \log p + (1 - (A_H)_e) \log(1-p) \\ &\quad + \sum_{\substack{e\in\binom{V}{k} \\ e\text{: cr.-cl. w.r.t. } x}} (A_H)_e \log q + (1 - (A_H)_e) \log(1-q) \\ &= C + \log\left(\frac{p}{1-p}\right) \langle A_H, \mathbf{x}^{\ominus k} \rangle + \log\left(\frac{q}{1-q}\right) \langle A_H, 1 - x^{\ominus k} \rangle, \end{aligned}$$

with

$$\begin{aligned} C &= \log(1-p)\cdot \#(e: \text{ in-cl. w.r.t. } x) + \log(1-q)\cdot \#(e: \text{ cr.-cl. w.r.t. } x) \\ &= 2\binom{n/2}{k}\log(1-p) + \left(\binom{n}{k} - 2\binom{n/2}{k}\right)\log(1-q). \end{aligned}$$

We note that C is a constant not depending on x. Also, $\langle A_H, 1\rangle$ is independent of x. We get

$$\widehat{\mathbf{x}}_{MLE}(H) = \underset{\mathbf{x}\in\{\pm 1\}^V : \mathbf{1}^T\mathbf{x}=0}{\operatorname{argmax}} \log\left(\frac{p(1-q)}{q(1-p)}\right)\left\langle A_H, \mathbf{x}^{\ominus k}\right\rangle.$$

It implies that

$$\widehat{\mathbf{x}}_{MLE}(H) = \begin{cases} \underset{\mathbf{x}\in\{\pm 1\}^V : \mathbf{1}^T\mathbf{x}=0}{\operatorname{argmax}} \left\langle A_H, \mathbf{x}^{\ominus k}\right\rangle & \text{if } p > q \\ \underset{\mathbf{x}\in\{\pm 1\}^V : \mathbf{1}^T\mathbf{x}=0}{\operatorname{argmin}} \left\langle A_H, \mathbf{x}^{\ominus k}\right\rangle & \text{if } p < q \end{cases}$$

since $\log\left(\frac{p(1-q)}{q(1-p)}\right)$ is positive if $p > q$ and it is negative if $p < q$.

3.7.2 Proof of Lemma 3.14

We recall that

$$Y_e = \left(\mathbf{1}_e^T(\mathbf{x}_0)_e \cdot \operatorname{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)\Pi\left(\mathbf{1}_e^T(\mathbf{x}_0)_e \cdot \operatorname{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right),$$

where $\Pi = I - \frac{1}{n}\mathbf{x}_0\mathbf{x}_0^T - \frac{1}{n}\mathbf{1}\mathbf{1}^T$. Also, recall that

$$\Sigma = \sum_e \mathbb{E}(A_{\mathcal{H}})_e \cdot Y_e = p\left(\sum_{e: e \text{ is in-cl.}} Y_e\right) + q\left(\sum_{e: e \text{ is cross-cl.}} Y_e\right).$$

Lemma 3.19 (Lemma 3.14, restated). $\Sigma = c_1 \cdot \frac{1}{n}\mathbf{x}_0\mathbf{x}_0^T + c_2\cdot\Pi$ *where*

$$c_1 = (k-1)\beta\log n + O\left(\frac{\log n}{n}\right)$$

and

$$c_2 = 2(k-1)\left(\frac{k}{2^k}\alpha + \left(1-\frac{k}{2^k}\right)\beta\right)\log n + O\left(\frac{\log n}{n}\right).$$

We note that

$$\left(\sum_{e:e \text{ is in-cl.}} Y_e\right) \quad \text{and} \quad \left(\sum_{e:e \text{ is cross-cl.}} Y_e\right)$$

are invariant under any permutation on V preserving $\mathbf{x}_0$. It implies that the both matrices and Σ are in $\text{span}\big(\{\Pi,\, \mathbf{x}_0\mathbf{x}_0^T,\, \mathbf{1}\mathbf{1}^T\}\big)$. Moreover,

$$\left(\mathbf{1}_e^T(\mathbf{x}_0)_e \cdot \text{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)\mathbf{1} = (\mathbf{1}_e^T(\mathbf{x}_0)_e)(\mathbf{x}_0)_e - (\mathbf{1}^T(\mathbf{x}_0)_e)(\mathbf{x}_0)_e = 0$$

so $\langle Y_e, \mathbf{1}\mathbf{1}^T\rangle = 0$. Hence,

$$\Sigma = \frac{\langle \Sigma, \Pi\rangle}{\langle \Pi, \Pi\rangle}\Pi + \left\langle \Sigma, \frac{1}{n}\mathbf{x}_0\mathbf{x}_0^T\right\rangle \cdot \frac{1}{n}\mathbf{x}_0\mathbf{x}_0^T.$$

It implies that

$$c_1 = \left\langle \Sigma, \frac{1}{n}\mathbf{x}_0\mathbf{x}_0^T\right\rangle = \frac{1}{n}\mathbf{x}_0^T\Sigma\mathbf{x}_0$$

and

$$c_2 = \frac{\langle \Sigma, \Pi\rangle}{\langle \Pi, \Pi\rangle} = \frac{1}{n-2}\left(tr(\Sigma) - \frac{1}{n}\mathbf{x}_0^T\Sigma\mathbf{x}_0\right).$$

Now, let us first compute $\mathbf{x}_0^T\Sigma\mathbf{x}_0$. For simplicity, let $r = \frac{1}{2}(k - \mathbf{1}_e^T(\mathbf{x}_0)_e)$. Then,

$$\left((k-2r)\text{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)\mathbf{x}_0 = (k-2r)\mathbf{1}_e - k(\mathbf{x}_0)_e$$

and

$$\begin{aligned}
\mathbf{x}_0^T Y_e \mathbf{x}_0 &= ((k-2r)\mathbf{1}_e - k(\mathbf{x}_0)_e)^T\,\Pi\,((k-2r)\mathbf{1}_e - k(\mathbf{x}_0)_e) \\
&= \|(k-2r)\mathbf{1}_e - k(\mathbf{x}_0)_e\|_2^2 - \frac{1}{n}\left(((k-2r)\mathbf{1}_e - k(\mathbf{x}_0)_e)^T\mathbf{x}_0\right)^2 \\
&= (k^3 - k(k-2r)^2) - \frac{1}{n}\left((k-2r)^2 - k^2\right)^2 \\
&= 4kr(k-r) - \frac{16}{n}r^2(k-r)^2.
\end{aligned}$$

In particular, $\mathbf{x}_0^T Y_e \mathbf{x}_0 = 0$ if e is in-cluster with respect to $\mathbf{x}_0$. Hence,

$$\begin{aligned}
\mathbf{x}_0^T \Sigma \mathbf{x}_0 &= q \sum_{e: e \text{ is cross-cl.}} \mathbf{x}_0^T Y_e \mathbf{x}_0 \\
&= q \sum_{r=1}^{k-1} \left(4kr(k-r) - \frac{16}{n} r^2 (k-r)^2 \right) \binom{n/2}{r} \binom{n/2}{k-r}.
\end{aligned}$$

We note that for any $s, t \in \{1, 2\}$, we have

$$\begin{aligned}
\sum_{r=1}^{k-1} \binom{r}{s} \binom{k-r}{t} \binom{n/2}{r} \binom{n/2}{k-r} &= \binom{n/2}{s} \binom{n/2}{t} \binom{n-s-t}{k-s-t} \\
&= \binom{n/2}{s} \binom{n/2}{t} \binom{n}{k} \binom{k}{s+t} \binom{n}{s+t}^{-1}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\sum_{r=1}^{k-1} r(k-r) \binom{n/2}{r} \binom{n/2}{k-r} &= \binom{n}{k} \frac{\binom{k}{2} \binom{n/2}{1}^2}{\binom{n}{2}} \\
&= \binom{n}{k} \frac{k(k-1)n^2}{4n(n-1)} = \frac{k(k-1)}{4} \binom{n}{k} + O(n^{k-1}).
\end{aligned}$$

and

$$\frac{1}{n} \sum_{r=1}^{k-1} r^2 (k-r)^2 \binom{n/2}{r} \binom{n/2}{k-r} = O(n^{k-1}).$$

Hence,

$$c_1 = \frac{1}{n} \mathbf{x}_0^T \Sigma \mathbf{x}_0 = (k-1)\beta \log n + O\left(\frac{\log n}{n} \right).$$

On the other hand,

$$
\begin{aligned}
tr(Y_e) = \;& \operatorname{tr}\left(\left((k-2r)\operatorname{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)^2 \Pi\right) \\
= \;& \operatorname{tr}\left(\left((k-2r)\operatorname{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)^2\right) \\
& - \frac{1}{n}\left\|\left((k-2r)\operatorname{diag}((\mathbf{x}_0)_e) - (\mathbf{x}_0)_e(\mathbf{x}_0)_e^T\right)\mathbf{x}_0\right\|_2^2 \\
= \;& \left((k-2)(k-2r)^2 + k^2\right) - \frac{1}{n}\|(k-2r)\mathbf{1}_e - k(\mathbf{x}_0)_e\|_2^2 \\
= \;& \left((k-2)(k-2r)^2 + k^2\right) - \frac{1}{n}(k^3 - k(k-2r)^2) \\
= \;& (k^3 - k^2) - 4(k - 2 + \frac{k}{n})r(k-r),
\end{aligned}
$$

so we have

$$
\begin{aligned}
tr(\Sigma) &= (k^3 - k^2)\left(q\binom{n}{k} + 2(p-q)\binom{n/2}{k}\right) - 4q\left(k - 2 + \frac{k}{n}\right)\cdot\frac{n^2}{4}\binom{n-2}{k-2} \\
&= \left[(k^2 - k)\left(\beta + \frac{\alpha - \beta}{2^{k-1}}\right) - (k-1)(k-2)\beta\right] n\log n + O(\log n).
\end{aligned}
$$

and hence

$$
\begin{aligned}
c_2 &= \frac{1}{n-2}\left(tr(\Sigma) - \frac{1}{n}\mathbf{x}_0^T\Sigma\mathbf{x}_0\right) \\
&= 2(k-1)\left(\frac{k}{2^k}\alpha + \left(1 - \frac{k}{2^k}\right)\beta\right)\log n + O\left(\frac{\log n}{n}\right).
\end{aligned}
$$

Chapter 4

Statistical Limits of Graphical Channel Models

We have considered the spiked tensor model generalizing the spiked Wigner model to k-tensors in Chapter 2, and the stochastic block model for k-uniform hypergraphs generalizing the stochastic block model for graphs in Chapter 3. In this chapter, we consider a class of statistical models which are called *graphical channel models* which was briefly introduced in Chapter 1. We remark that this class of models encompasses both the spiked tensor model and the stochastic block model for k-uniform hypergraphs.

4.1 Exact recovery in Graphical Channel Model

4.1.1 Description of the model

Let us first formally define the graphical channel model. Let n and k be integers satisfying $n \geq k \geq 2$.

- Let $\mathcal{H} = (V, E)$ be a k-uniform hypergraph on the vertex set V of size n, *i.e.*, the (hyper)edge set E is a collection of size k subsets of V.

- Let $\mathcal{X}$ be a finite set and $\mathcal{Y}$ be a measure space equipped with a reference measure λ. We call $\mathcal{X}$ and $\mathcal{Y}$ input and output alphabets, respectively.

- Let p be a probability distribution on $\mathcal{X}$, called prior distribution. We often regard p as a nonnegative vector whose entries sum to 1, since $\mathcal{X}$ is finite.

- Let Q be a probability kernel with source $\mathcal{X}^k$ and target $(\mathcal{Y}, \lambda)$, *i.e.*, for each $z \in \mathcal{X}^k$ the function $Q(\cdot|z) : y \mapsto Q(y|z)$ is a probability density function with respect to λ. In other words, Q is a stochastic function which gets $z \in \mathcal{X}^k$ as an input and outputs a random value y sampled from $Q(\cdot|z)$[1].

- Let $\mathbf{x} \in \mathcal{X}^V$ and $\mathbf{y} \in \mathcal{Y}^E$. Here $\mathbf{x}$ and $\mathbf{y}$ are called vectors of vertex-variables and edge-variables respectively. Let us define $\mathbb{P}_{\mathcal{M}(p,Q)}$ to be a probability measure on $\mathcal{X}^V \times \mathcal{Y}^E$ such that

$$\mathbb{P}_{\mathcal{M}(p,Q)}(\mathbf{x}) = \prod_{v \in V} p(\mathbf{x}_v) \quad \text{and} \quad \mathbb{P}_{\mathcal{M}(p,Q)}(\mathbf{y}|\mathbf{x}) = \prod_{e \in E} Q(\mathbf{y}_e|\mathbf{x}[e]),$$

  where $\mathbf{x}[e]$ is a shorthand notation for $(\mathbf{x}_{e_1}, \cdots, \mathbf{x}_{e_k})$ when $e = \{e_1 < e_2 < \cdots < e_k\}$.

We call the generative model obeying $\mathbb{P}_{\mathcal{M}(p,Q)}$ a *graphical channel model (of graph $\mathcal{H}$) with prior p and kernel Q*, denoted $\mathcal{M}(p, Q)$. We omitted $\mathcal{H}$ since it would be clear throughout the chapter. We will often write $(\mathbf{x}, \mathbf{y}) \sim \mathcal{M}(p, Q)$ to mean that $(\mathbf{x}, \mathbf{y})$ is a random vector having the model-defining distribution $\mathbb{P}_{\mathcal{M}(p,Q)}$.

Throughout the chapter, we will further restrict our focus on the models satisfying the following.

- We assume that $\mathcal{H}$ is the complete k-uniform hypergraph, *i.e.*, the edge set $E(\mathcal{H})$ is equal to $\binom{V}{k}$, the collection of all size k-subsets of V.

- We assume that $|\mathcal{X}| = 2$. In block models, it corresponds to the case that there are only two communities. Without loss of generality, we let $\mathcal{X} = \{0, 1\}$.

- We assume that the kernel Q is symmetric, *i.e.*, $Q(\cdot|z)$ is invariant under permutations of the indices of z.

[1]We often abuse the notation by denoting both the distribution and the density function by $Q(\cdot|z)$.

For brevity, let us denote the vertex set and the edge set of the complete k-uniform hypergraph on n vertices by V_n and E_n, respectively. We will often drop the subscripts when the context is clear.

We remark that indeed both the SBM with two symmetric communities and the spiked Wigner model are instances of graphical channel models. We discuss further relations in Section 4.5.

Let us introduce a few more notations for the sake of simplicity. Let $z \in \{0,1\}^m$. The length of z (denoted $\ell(z)$) is defined as $\ell(z) = m$. We denote the number of ones in z by $|z|$, *i.e.*, $|z| = \sum_{i=1}^{m} z_i$ and call $|z|$ the weight of z.

Note that $|z|$ is invariant under any permutation of the indices of z. Conversely, if a tuple z' has the same length and the same weight as z, then z' can be obtained by permuting the indices of z. Since Q is symmetric, we have $Q(\cdot|z) \equiv Q(\cdot|z')$ if $|z| = |z'|$. We denote $Q(y|z)$ by $q_{|z|}(y)$ for the sake of simplicity (note that this notation is well-defined).

4.1.2 Recovery requirements and exact recovery

Let $(\mathbf{x}^0, \mathbf{y}^0)$ be a sample from $\mathcal{M}(p, Q)$. We will often call $\mathbf{x}^0$ *prior* and $\mathbf{y}^0$ *posterior* to emphasize that $\mathbf{y}^0$ is a posterior value which we observe and $\mathbf{x}^0$ is a prior latent parameters which we would like to recover from $\mathbf{y}^0$. Since the model is probabilistic, we ask whether a recovery can be done with high probability with respect to the randomness of the sample.

There are three representative notions of recovery regarding the quality of the solution: exact recovery, almost exact recovery and partial recovery. Informally,

- *Exact recovery*: We aim to recover the ground truth $\mathbf{x}^0$.
- *Almost exact recovery*: We aim to find a vertex labeling $\mathbf{x}'$ which agrees with the ground truth $\mathbf{x}^0$ except in a vanishing fraction of the vertices.
- *Partial recovery*: We aim to find a labeling $\mathbf{x}'$ which is better than a random guess.

In some literatures, they are called strong consistency, weak consistency and detection, respectively. We focus on exact recovery, and recommend [1] for readers who are interested in other types of recovery.

Let $\widehat{\mathbf{x}}$ be an estimator of $\mathbf{x}^0$ given $\mathbf{y}^0$. We may not expect the estimator to be perfect because of the intrinsic randomness of probabilistic models. Instead, we measure the performance of $\widehat{\mathbf{x}}$ on exact recovery by the probability that $\widehat{\mathbf{x}}$ successfully recover the ground truth $\mathbf{x}^0$, *i.e.*,

$$\underset{(\mathbf{x}^0,\mathbf{y}^0)\sim\mathcal{M}(p,Q)}{\mathbb{P}}\left(\widehat{\mathbf{x}}(\mathbf{y}^0)=\mathbf{x}^0\right).$$

We are interested in a regime that k is fixed, but n grows to infinity. Specifically, we consider a *sequence* of models $\{\mathcal{M}(p,Q_n)\}_{n\geq k}$ where $\mathcal{M}(p,Q_n)$ is the graphical channel model with prior distribution p (which does not depend on n) and the symmetric kernel Q_n, such that whose graph is the complete k-uniform hypergraph on n vertices. For brevity, we will often write $\mathcal{M}(p,Q)$ to denote the sequence $\{\mathcal{M}(p,Q_n)\}$ of the models. Under this regime, we formally define the requirements for exact recovery as follows.

Definition 4.1. *For each $n\geq k$, let $D_n:\mathcal{Y}^{E_n}\to\mathcal{X}^{V_n}$ be an estimator of $\mathbf{x}^0$ given $\mathbf{y}^0$. Let D be the sequence of estimators $\{D_n\}_{n\geq k}$. We say that D* achieves exact recovery *for $\mathcal{M}(p,Q)$ if*

$$\underset{(\mathbf{x}^0,\mathbf{y}^0)\sim\mathcal{M}(p,Q_n)}{\mathbb{P}}\left(D_n(\mathbf{y}^0)=\mathbf{x}^0\right)=1-o(1).$$

We say that exact recovery for $\mathcal{M}(p,Q)$ is achievable *if there exists D achieving exact recovery, and otherwise exact recovery is* impossible*. Moreover, if some D achieves exact recovery and D_n is computable in polynomial time, then we say that exact recovery is* efficiently achievable.

For example, let us recall the SBM with two symmetric communities, *i.e.*, the model in which a prior $\mathbf{x}^0$ is sampled from $\{0,1\}^n$ uniformly at random, and a posterior $\mathbf{y}^0\in\{0,1\}^{\binom{n}{2}}$ is sampled in the way that for each $\{i,j\}\in\binom{[n]}{2}$, entry $\mathbf{y}^0_{\{i,j\}}$ is sampled independently and

$$\mathbf{y}^0_{\{i,j\}} \sim \begin{cases} \mathrm{Ber}(p) & \text{if } \mathbf{x}_i = \mathbf{x}_j \\ \mathrm{Ber}(q) & \text{otherwise.} \end{cases}$$

Now, let us consider a parameter regime that $p = \frac{a \log n}{n}$ and $q = \frac{b \log n}{n}$ for some $a, b > 0$. In the language of graphical channel models, this corresponds to $\mathcal{M}(U(\{0,1\}), Q)$ where $U(\{0,1\})$ is the uniform distribution on $\mathcal{X} = \{0,1\}$ and

$$Q_n(\cdot|x_1, x_2) = \begin{cases} \mathrm{Ber}\left(\frac{a \log n}{n}\right) & \text{if } x_1 = x_2 \\ \mathrm{Ber}\left(\frac{b \log n}{n}\right) & \text{otherwise.} \end{cases}$$

We remark that exact recovery shows a sharp phase transition behavior around the threshold $(\sqrt{a} - \sqrt{b})^2 = 2$, *i.e.*,

- if $(\sqrt{a} - \sqrt{b})^2 < 2$ then exact recovery is achievable (up to a global switch of two community labels), and
- if $(\sqrt{a} - \sqrt{b})^2 > 2$ then exact recovery is not achievable.

This is proved in [4]. Moreover, we can efficiently achieve exact recovery up to the information-theoretic threshold, via semidefinite relaxation technique [56] or via almost exact recovery algorithm with additional local refinement steps [4].

We would like to note that exact recovery for the SBM with two symmetric communities is achievable only "up to a global switch of two community labels" which means that we cannot distinguish $\mathbf{x}$ and $\mathbf{x}'$ obtained by exchanging labels of two communities, *i.e.*, $\mathbf{x}'_v = 1 - \mathbf{x}_v$ for all $v \in V$. It implies that exact recovery is intrinsically not achievable if we regard $\mathbf{x}'$ as an incorrect solution. However, it is natural to regard $\mathbf{x}'$ as a correct solution since two communities are not distinguishable. For this reason, we relax the requirement by allowing a global switch of the community labels in such cases.

Definition 4.2. *Let D be defined as in Definition 4.1. We say D* achieves exact recovery up to a global switch of the community labels *if*

$$\mathbb{P}_{(\mathbf{x}^0,\mathbf{y}^0)\sim\mathcal{M}(p,Q_n)}\left(D_n(\mathbf{y}^0)\in\{\mathbf{x}^0,\mathbf{1}-\mathbf{x}^0\}\right)=1-o(1).$$

Definitions for achievability and non-achievability are extended as well.

4.1.3 Maximum *a posteriori* estimator

The *maximum a posteriori (MAP) estimator* for $\mathbf{x}^0$ given $\mathbf{y}^0$ is defined as the maximizer of posterior probability over all possible priors, *i.e.*,

$$\widehat{\mathbf{x}}_{map}(\mathbf{y}^0):=\underset{\mathbf{x}\in\mathcal{X}^V}{\operatorname{argmax}}\ \mathbb{P}_{\mathcal{M}(p,Q)}(\mathbf{x}^0=\mathbf{x}|\mathbf{y}^0)$$

where ties are broken arbitrarily. Since $\mathbb{P}(\mathbf{x}|\mathbf{y})=\frac{\mathbb{P}(\mathbf{y}|\mathbf{x})\,\mathbb{P}(\mathbf{x})}{\mathbb{P}(\mathbf{y})}$, we have

$$\begin{aligned}\widehat{\mathbf{x}}_{map}(\mathbf{y}^0)&=\underset{\mathbf{x}\in\mathcal{X}^V}{\operatorname{argmax}}\ \log\left(\mathbb{P}_{\mathcal{M}(p,Q)}(\mathbf{y}^0|\mathbf{x}^0=\mathbf{x})\,\mathbb{P}_{\mathcal{M}(p,Q)}(\mathbf{x}^0=\mathbf{x})\right)\\&=\underset{\mathbf{x}\in\mathcal{X}^V}{\operatorname{argmax}}\ \sum_{v\in V}\log p(\mathbf{x}_v)+\sum_{e\in E}\log Q(\mathbf{y}^0_e|\mathbf{x}[e]).\end{aligned}$$

For the sake of simplicity, let us define

$$\ell(\mathbf{x})=\sum_{v\in V}\log p(\mathbf{x}_v)\quad\text{and}\quad\ell(\mathbf{y}|\mathbf{x})=\sum_{e\in E}\log Q(\mathbf{y}_e|\mathbf{x}[e]),$$

the log-likelihood of $\mathbf{x}$ and the log-likelihood of $\mathbf{y}$ conditioned on $\mathbf{x}$. With this notation,

$$\widehat{\mathbf{x}}_{map}(\mathbf{y}^0)=\underset{\mathbf{x}\in\mathcal{X}^V}{\operatorname{argmax}}\ \ell(\mathbf{x})+\ell(\mathbf{y}^0|\mathbf{x}).$$

Proposition 4.1 (folklore). *The MAP estimator outperforms all other estimators in the sense that the probability of returning an incorrect solution is minimized by the MAP estimator. In other words,*

$$\mathbb{P}\left(\widehat{\mathbf{x}}_{map}(\mathbf{y}^0)\neq\mathbf{x}^0\right)\leq\mathbb{P}\left(\widehat{\mathbf{x}}(\mathbf{y}^0)\neq\mathbf{x}^0\right)$$

for any estimator $\widehat{\mathbf{x}}$.

It implies that exact recovery is achievable if only if the MAP estimator achieves exact recovery. Let

$$p_{fail,global} = \mathbb{P}\left(\widehat{\mathbf{x}}_{map}(\mathbf{y}^0) \neq \mathbf{x}^0\right),$$

the probability that the MAP estimator fails to recover $\mathbf{x}^0$.

Proposition 4.2. *Exact recovery is achievable if and only if* $p_{fail,global} = o(1)$.

4.2 Local recovery: Binary hypothesis testing

4.2.1 Genie-aided local recovery

We note that our argument in this section closely follows and generalizes the genie-aided approach in [1] to the setting of graphical channel models.

Imagine that in addition to the observation of $\mathbf{y}^0$, a magical genie reveals us the true labels of all vertices except a vertex v. In other words, we are given an extra observation of

$$\mathbf{x}^0_{-v} := \{\mathbf{x}^0_u : u \in V \setminus \{v\}\}.$$

Let us call the problem of recovering $\mathbf{x}^0_v$ given $\mathbf{y}^0$ and $\mathbf{x}^0_{-v}$, a (genie-aided) local recovery at v.

Let $\widehat{\mathbf{x}}_{v,map}$ be the MAP estimator for $\mathbf{x}^0_v$ given $\mathbf{y}^0$ and $\mathbf{x}^0_{-v}$, *i.e.*,

$$\begin{aligned}
\widehat{\mathbf{x}}_{v,map}(\mathbf{y}^0, \mathbf{x}^0_{-v}) &= \operatorname*{argmax}_{x \in \{0,1\}} \mathbb{P}(\mathbf{x}^0_v = x | \mathbf{y}^0, \mathbf{x}^0_{-v}) \\
&= \operatorname*{argmax}_{x \in \{0,1\}} p(\mathbf{x})\, \mathbb{P}(\mathbf{y}^0 | \mathbf{x}^0_v = x, \mathbf{x}^0_{-v}).
\end{aligned}$$

We note that

$$p_{fail,global} \geq \mathbb{P}\left(\exists v \in V : \ \widehat{\mathbf{x}}_{v,map}(\mathbf{y}^0, \mathbf{x}^0_{-v}) \neq \mathbf{x}^0_v\right),$$

since failing at one of the local recoveries implies that we must fail to recover the whole $\mathbf{x}^0$.

Let E_v the event that $\widehat{\mathbf{x}}_{v,map}$ fails to recover $\mathbf{x}_v^0$. Let $\mathbf{x}^v$ be a vector in $\{0,1\}^V$ where $\mathbf{x}_u^v = \mathbf{x}_u^0$ for $u \in V \setminus \{v\}$ and $\mathbf{x}_v^v = 1 - \mathbf{x}_v^0$. Then,

$$E_v \text{ happens} \quad \Leftrightarrow \quad p(\mathbf{x}_v^0)\,\mathbb{P}(\mathbf{y}^0|\mathbf{x}^0) \leq p(\mathbf{x}_v^v)\,\mathbb{P}(\mathbf{y}^0|\mathbf{x}^v)$$

by definition of the MAP estimator. Taking log on both sides, we get

$$\log p(\mathbf{x}_v^0) + \ell(\mathbf{y}^0|\mathbf{x}^0) \leq \log p(\mathbf{x}_v^v) + \ell(\mathbf{y}^0|\mathbf{x}^v),$$

which is equivalent to

$$\ell(\mathbf{x}^0) + \ell(\mathbf{y}^0|\mathbf{x}^0) \leq \ell(\mathbf{x}^v) + \ell(\mathbf{y}^0|\mathbf{x}^v).$$

Let

$$p_{fail,local} = \mathbb{P}\left(\bigcup_{v\in V} E_v\right) \quad \text{and} \quad p_{fail,v} = \mathbb{P}(E_v).$$

We get the following relations between $p_{fail,global}$, $p_{fail,local}$ and $p_{fail,v}$.

Proposition 4.3. $p_{fail,global} \geq p_{fail,local}$ *and* $p_{fail,local} \leq \sum_{v\in V} p_{fail,v}$.

Remark. We can interpret $p_{fail,global}$, $p_{fail,local}$ and $p_{fail,v}$ in terms of the type of $\mathbf{x}^0$ of the stochastic function $\mathbf{x} \mapsto \ell(\mathbf{x}) + \ell(\mathbf{y}^0|\mathbf{x})$ as follows:

- $p_{fail,global}$ is the probability that $\mathbf{x}^0$ is not the unique global maximum.
- $p_{fail,local}$ is the probability that $\mathbf{x}^0$ is not strictly locally maximal.
- $p_{fail,v}$ is the probability that $\mathbf{x}^0$ is not the unique maximum in the direction of $\mathbf{x}_v^0$.

4.2.2 Local-to-global amplification

Roughly speaking, local-to-global amplification is a phenomenon that the success of global recovery can be approximated by independent successes of all local recoveries.

In other words, when local-to-global amplification holds we expect

$$p_{fail,global} \approx p_{fail,local} \approx 1 - \prod_{v \in V} (1 - p_{fail,v})$$

in terms of $p_{fail,global}$, $p_{fail,local}$ and $p_{fail,v}$. In such cases, achievability of exact recovery can be reduced to whether $p_{fail,v}$ is over or below some threshold.

Lemma 4.4 (Informal). *If local-to-global amplification holds, then*

$$p_{fail,global} = \begin{cases} 1 - o(1) & \text{if } \sum_{v \in V} p_{fail,v} = \omega_n(1) \\ o(1) & \text{if } \sum_{v \in V} p_{fail,v} = o(1). \end{cases}$$

In particular, exact recovery can be achieved when $p_{fail,v} = o_n(n^{-1})$ for all $v \in V$ and is not achievable when $p_{fail,v} = \omega_n(n^{-1})$ for all $v \in v$.

Informal proof. Since $1 - x \leq e^{-x}$ for all x, we have

$$p_{fail,global} \approx 1 - \prod_{v \in V} (1 - p_{fail,v}) \geq 1 - \exp\left(-\sum_{v \in V} p_{fail,v}\right).$$

On the other hand,

$$p_{fail,global} \approx p_{fail,local} \leq \sum_{v \in V} p_{fail,v}.$$

Hence, $p_{fail,global}$ can be approximated as

$$1 - \exp\left(-\sum_{v \in V} p_{fail,v}\right) \leq p_{fail,global} \leq \sum_{v \in V} p_{fail,v},$$

and we get the desired result. □

4.2.3 Binary hypothesis testing

Suppose that local-to-global amplification holds. To characterize the threshold for exact recovery, we now need to analyze $p_{fail,v}$ and understand when this probability happens to be large. Note that local recovery at v can be treated as a binary hypothesis testing between two hypotheses H_0 and H_1 given $\mathbf{y}^0$ and $\mathbf{x}^0_{-v}$, where H_0 and H_1 corresponds to $\mathbf{x}^0_v = 0$ and $\mathbf{x}^0_v = 1$, respectively. In the next two sections, we analyze the optimal error probability in generic binary hypothesis tests, and explain how it applies to local recovery problems.

In the simplest form, binary hypothesis testing can be thought as the problem of deciding between two hypotheses

$$H_0 : Y \sim \mu_0 \quad \text{and} \quad H_1 : Y \sim \mu_1,$$

given a sample Y which is sampled from either μ_0 or μ_1. Suppose that we already know $\mathbb{P}(H_0) = 1 - \rho$ and $\mathbb{P}(H_1) = \rho$ for some positive $\rho > 0$.

Let $\widehat{H}$ be a decision rule between H_0 and H_1 given Y, and let

$$A = \{y : \widehat{H}(y) = H_0\}.$$

Then, the probability for $\widehat{H}$ making a mistake is

$$\begin{aligned}
\mathbb{P}(error) = &\quad \mathbb{P}(H_0)\,\mathbb{P}(\widehat{H}(Y) = H_1|H_0) + \mathbb{P}(H_1)\,\mathbb{P}(\widehat{H}(Y) = H_0|H_1) \\
= &\quad (1-\rho)\mu_0(A^c) + \rho\mu_1(A) \\
= &\quad (1-\rho) + (\rho\mu_1(A) - (1-\rho)\mu_0(A)).
\end{aligned}$$

Hence,

$$\min_{\widehat{H}} \mathbb{P}(error) = (1-\rho) + \min_{A} \left[\rho\mu_1(A) - (1-\rho)\mu_0(A)\right].$$

The total variation (TV) distance between two (not necessarily probability) measures ν_1 and ν_2 is defined as

$$d_{TV}(\nu_1, \nu_2) := \frac{1}{2} \sup_A |\nu_1(A) - \nu_2(A)| + |\nu_1(A^c) - \nu_2(A^c)|.$$

If ν_1 and ν_2 have densities $\frac{d\nu_1}{d\lambda}$ and $\frac{d\mu_2}{d\lambda}$ with respect to a reference measure λ, then

d_{TV} is equal to half the L^1-norm of the difference $\frac{d\nu_1}{d\lambda} - \frac{d\nu_2}{d\lambda}$, *i.e.*,

$$d_{TV}(\nu_1, \nu_2) = \frac{1}{2} \int \left| \frac{d\nu_1}{d\lambda} - \frac{d\nu_2}{d\lambda} \right| d\lambda.$$

Proposition 4.5. *The minimum probability for a decision rule making a mistake is equal to*

$$\frac{1}{2} - d_{TV}((1-\rho)\mu_0, \rho\mu_1) = \frac{1}{2} - \frac{1}{2} \int \left| (1-\rho)\frac{d\mu_0}{d\lambda} - \rho\frac{d\mu_1}{d\lambda} \right| d\lambda$$

and the minimum is attained by $\widehat{H}_{map}$ *defined as*

$$\widehat{H}_{map}(Y) = \begin{cases} H_0 & \text{if } (1-\rho)\frac{d\mu_0}{d\lambda}(Y) \geq \rho\frac{d\mu_1}{d\lambda}(Y) \\ H_1 & \text{otherwise.} \end{cases}$$

We call $\widehat{H}_{map}$ *the MAP decision rule.*

Proof. Folklore. □

In the context of local recovery, $\mathbb{P}(E_v \,|\, \mathbf{x}^0_{-v})$ is the probability that $\widehat{\mathbf{x}}_{map,v}$ making a mistake (when $\mathbf{x}^0_{-v}$ is fixed), so we have

$$\mathbb{P}(E_v \,|\, \mathbf{x}^0_{-v}) = \frac{1}{2} - d_{TV}((1-\rho)\mu_0, \rho\mu_1)$$

where $p(0) = 1 - \rho$, $p(1) = \rho$ and

$$d\mu_0(\mathbf{y}) = \prod_{e \in E: e \ni v} Q\left(\mathbf{y}_e \middle| \mathbf{x}^0_{-v}\left[e \setminus \{v\}\right], 0\right) \cdot d\lambda(\mathbf{y}_e), \text{ and}$$

$$d\mu_1(\mathbf{y}) = \prod_{e \in E: e \ni v} Q\left(\mathbf{y}_e \middle| \mathbf{x}^0_{-v}\left[e \setminus \{v\}\right], 1\right) \cdot d\lambda(\mathbf{y}_e).$$

Moreover, $p_{fail,v}$ is the expected value of $\mathbb{P}(E_v \,|\, \mathbf{x}^0_{-v})$ over the randomness of $\mathbf{x}^0_{-v}$. It implies that estimating $d_{TV}((1-\rho)\mu_0, \rho\mu_1)$ for "typical" $\mathbf{x}^0_{-v}$ is a crucial step to get a tight estimate of $p_{fail,v}$.

4.2.4 Chernoff α-divergences

Recall that

$$\begin{aligned}\mathbb{P}(\text{MAP rule makes a mistake}) &= \frac{1}{2} - d_{TV}((1-\rho)\mu_0, \rho\mu_1) \\ &= \frac{1}{2} - \frac{1}{2}\int \left|(1-\rho)\frac{d\mu_0}{d\lambda} - \rho\frac{d\mu_1}{d\lambda}\right| d\lambda.\end{aligned}$$

Since μ_0 and μ_1 are probability distributions, it is equal to

$$\begin{aligned}&\frac{1}{2}\int \left((1-\rho)\frac{d\mu_0}{d\lambda} + \rho\frac{d\mu_1}{d\lambda} - \left|\rho\frac{d\mu_0}{d\lambda} - \rho\frac{d\mu_1}{d\lambda}\right|\right) d\lambda \\ &= \int \min\left((1-\rho)\frac{d\mu_0}{d\lambda}, \rho\frac{d\mu_1}{d\lambda}\right) d\lambda.\end{aligned}$$

We get

$$\int \min\left((1-\rho)\frac{d\mu_0}{d\lambda}, \rho\frac{d\mu_1}{d\lambda}\right) d\lambda \le \min_{\alpha\in[0,1]} (1-\rho)^\alpha \rho^{1-\alpha} \int \left(\frac{d\mu_0}{d\lambda}\right)^\alpha \left(\frac{d\mu_1}{d\lambda}\right)^{1-\alpha} d\lambda,$$

since $\min(a,b) \le a^\alpha b^{1-\alpha}$ for any $a, b > 0$ and $\alpha \in [0,1]$.

Definition 4.3. *Let μ_0 and μ_1 be probability distributions on the measure space $\mathcal{U}$ equipped with a reference measure λ. Suppose that μ_i has a density $\frac{d\mu_i}{d\lambda}$ with respect to λ for $i \in \{0,1\}$ and μ_0 and μ_1 are absolutely continuous to each other. For $\alpha \in \mathbb{R}$, let*

$$D_\alpha(\mu_0 : \mu_1) = -\log \int \left(\frac{d\mu_0}{d\lambda}\right)^\alpha \left(\frac{d\mu_1}{d\lambda}\right)^{1-\alpha} d\lambda.$$

$D_\alpha(\mu_0 : \mu_1)$ is called the Chernoff α-divergence.

In summary, $\mathbb{P}(\text{MAP rule makes a mistake})$ is at most

$$\exp\left(-\max_{\alpha\in[0,1]} \alpha\log(1-\rho)^{-1} + (1-\alpha)\log\rho^{-1} + D_\alpha(\mu_0 : \mu_1)\right).$$

We remark that D_α is defined for all $\alpha \in \mathbb{R}$ not restricted on $[0,1]$. Indeed, as long as μ_0 and μ_1 are absolutely continuous to each other D_α is well-defined for all α, possibly equal to $-\infty$. Chernoff α-divergence has many nice properties such as concavity, smoothness and additivity.

Proposition 4.6. *Let μ_0 and μ_1 be probability distributions which are absolutely continuous to each other. Then, $D_\alpha = D_\alpha(\mu_0 : \mu_1)$ has the following properties:*

(i) D_α is concave.

(ii) D_α is smooth on the interior of the set $\{\alpha : D_\alpha(\mu_0 : \mu_1) > -\infty\}$.

(iii) $D_0 = D_1 = 0$. Hence, $D_\alpha < \infty$ everywhere and the maximum of D_α is attained at some $\alpha^* \in [0, 1]$.

(iv) D_α is strictly concave, unless μ_0 and μ_1 are almost surely identical.

(v) Suppose that $\mathcal{U} = \mathcal{U}_1 \otimes \mathcal{U}_2$ and μ_0 and μ_1 factorizes as $\mu_0 = \nu_1 \otimes \nu_2$ and $\mu_1 = \nu_1' \otimes \nu_2'$ where ν_i, ν_i' are probability measures on $\mathcal{U}_i$ for $i \in \{1, 2\}$. Then,

$$D_\alpha(\nu_1 \otimes \nu_2 : \nu_1' \otimes \nu_2') = D_\alpha(\nu_1 : \nu_1') + D_\alpha(\nu_2 : \nu_2')$$

for any α.

We omit the proof of Proposition 4.6, as it can be obtained in a general context of Csiszár f-divergences. We recommend the book by Amari and Nagaoka [16] to readers who are interested in the notions of α-divergence and its variants.

In summary, we have $\mathbb{P}(\text{MAP rule makes a mistake}) \le e^{-I(\rho)}$ where

$$I(\rho) = \max_{\alpha \in [0,1]} \alpha \log(1-\rho)^{-1} + (1-\alpha)\log \rho^{-1} + D_\alpha(\mu_0 : \mu_1).$$

Since D_α is strictly concave (if $\mu_0 \not\equiv \mu_1$), the maximum is attained at the unique $\alpha^* \in (0, 1)$ satisfying

$$\frac{dD_\alpha}{d\alpha} = \log \frac{1-\rho}{\rho},$$

if such an α^* exists.

Binary hypothesis testing given many observations

One may wonder how tight the inequality

$$\mathbb{P}(\text{MAP rule makes a mistake}) \leq e^{-I(\rho)}$$

is. This seems somewhat crude as it only exploits the inequality $\min(a,b) \leq a^\alpha b^{1-\alpha}$. However, this estimate becomes tighter when μ_i is a product of N i.i.d. distributions, *i.e.*,

$$\mu_0 = \nu_0^{\otimes N} \quad \text{and} \quad \mu_1 = \nu_1^{\otimes N}$$

for some ν_0 and ν_1 not depending on N.

In other words, this corresponds to a binary hypothesis testing problem with hypotheses H_0 with prior probability $1-\rho$ and H_1 with prior probability ρ where

$$H_0 : Y_i \sim \nu_0 \ \forall i \in [N] \quad \text{and} \quad H_1 : Y_i \sim \nu_1 \ \forall i \in [N],$$

given i.i.d. samples $Y_1, \cdots, Y_N$.

Theorem 4.7 (Chernoff). *Let us consider the binary hypothesis testing problem defined as above. Let p_N be the probability for the MAP decision rule making a mistake. Then,*

$$\lim_{N\to\infty} -\frac{1}{N} \log p_N = I(0) = \max_{\alpha\in[0,1]} D_\alpha(\nu_0 : \nu_1).$$

In other words, $p_N = e^{-I(0)\cdot N + o_N(N)}$.

We refer readers to Chapter 12 in [37] for the proof and further explanations. Note the similarity between this binary hypothesis testing with N samples and local recovery problem. In the latter case, we are given a vector $\{\mathbf{y}_e^0\}_{e\in E}$ with independent entries, and the goal is to decide between two hypotheses H_0 and H_1 such that

$$\begin{aligned} H_0 :\ & \mathbf{y}_e^0 \sim q_{|\mathbf{x}_{-v}^0[e\setminus v]|} && \forall e \in E \text{ s.t. } e \ni v, \\ H_1 :\ & \mathbf{y}_e^0 \sim q_{|\mathbf{x}_{-v}^0[e\setminus v]|+1} && \forall e \in E \text{ s.t. } e \ni v \end{aligned}$$

and prior is chosen with probability $\mathbb{P}(H_0) = 1-\rho$ and $\mathbb{P}(H_1) = \rho$. Here we recall that q_t is the short-hand notation for $Q(\cdot \mid z_1, \cdots, z_k)$ where $\sum_{i=1}^k z_i = t$ (See Section 4.1).

As in Theorem 4.7, we might expect that

$$\mathbb{P}(\widehat{\mathbf{x}}_{map,v} \neq \mathbf{x}_v^0 | \mathbf{x}_{-v}^0) \approx \exp\left(-\max_{\alpha\in[0,1]} D_\alpha(\mu_0 : \mu_1)\right)$$

for sufficiently large n. However, there are two main obstacles to generalize Theorem 4.7 to our case:

- $\mathbf{y}_e^0$ are independent but not identically distributed.

- More importantly, Q depends on n as opposed to that ν_0 and ν_1 are independent of N in the binary hypothesis testing problem with N samples.

For this reason, we need a large-deviation type estimate on the triangular array of random variables to prove an analogous statement. We impose this condition as an assumption in Section 4.3.

As a final note, we unconditionally have

$$\begin{aligned}
&\mathbb{P}(\widehat{\mathbf{x}}_{map,v} \neq \mathbf{x}_v^0 | \mathbf{x}_{-v}^0) \leq \\
&\quad \exp\left(-\max_{\alpha\in[0,1]} \alpha \log \rho^{-1} + (1-\alpha)\log(1-\rho)^{-1} + D_\alpha(\mu_0 : \mu_1)\right).
\end{aligned}$$

By additivity of D_α, we get

$$D_\alpha(\mu_0 : \mu_1) = \sum_{e\in E: e\ni v} D_\alpha\left(q_{|\mathbf{x}_{-v}^0[e\setminus v]|} : q_{|\mathbf{x}_{-v}^0[e\setminus v]|+1}\right).$$

Hence,

$$\begin{aligned}
D_\alpha(\mu_0 : \mu_1) = & \sum_{t=0}^{k-1} D_\alpha(q_t : q_{t+1}) \cdot \#(e \ni v : |\mathbf{x}_{-v}^0[e \setminus \{v\}]| = t) \\
= & \sum_{t=0}^{k-1} \binom{|\mathbf{x}_{-v}^0|}{t} \binom{n-1-|\mathbf{x}_{-v}^0|}{k-1-t} D_\alpha(q_t : q_{t+1}).
\end{aligned}$$

Observe that $\mathbb{E}\,|\mathbf{x}^0_{-v}| = \rho(n-1)$ for typical $\mathbf{x}^0_{-v}$. So,

$$\begin{aligned}
D_\alpha(\mu_0 : \mu_1) &\approx \sum_{t=0}^{k-1} \binom{\rho n}{t}\binom{(1-\rho)n}{k-1-t} D_\alpha(q_t : q_{t+1}) \\
&\approx \binom{n}{k-1} \sum_{t=0}^{k-1} \binom{k-1}{t} \rho^t (1-\rho)^{k-1-t} \cdot D_\alpha(q_t : q_{t+1}).
\end{aligned}$$

Putting all together, if there exists a constant I such that

$$I \approx \max_{\alpha\in[0,1]} \left[\frac{\binom{n}{k-1}}{\log n} \sum_{t=0}^{k-1} \binom{k-1}{t} \rho^t (1-\rho)^{k-1-t} \cdot D_\alpha(q_t : q_{t+1}) \right]$$

for sufficiently large n, then $p_{fail,v} \le n^{-I+o(1)}$.

4.3 Local-to-global amplification

In this section, we present our main result in terms of local-to-global amplification.

Definition 4.4. Local-to-global amplification holds for $\mathcal{M}(p, Q)$ *if*

$$p_{fail,global} = (1 \pm o(1)) p_{fail,local}$$

and

$$1 - o(1) - \left(\sum_{v\in V} p_{fail,v} \right)^{-1} \le p_{fail,local} \le \sum_{v\in V} p_{fail,v}.$$

Recall that in Section 4.2.4 we (informally) proved that $p_{fail,v} \le n^{-I+o(1)}$ where $I > 0$ is a constant not depending on n such that

$$I \approx \max_{\alpha\in[0,1]} \left[\frac{\binom{n}{k-1}}{\log n} \sum_{t=0}^{k-1} \binom{k-1}{t} \rho^t (1-\rho)^{k-1-t} \cdot D_\alpha(q_t : q_{t+1}) \right]$$

when n is sufficiently large. This motivates us to consider a parameter regime where

$$D_\alpha(q_s : q_t) = O\left(\frac{\log n}{\binom{n-1}{k-1}} \right)$$

for any fixed α and $s, t \in \{0, 1, \cdots, k\}$.

4.3.1 Weak amplification

Recall that q_t is defined as $q_t(y) = Q(y|z)$ for any $y \in \mathcal{Y}$ and $z \in \{0,1\}^k$ with $|z| = t$. If necessary, we will write $q_t^{(n)}$ instead of q_t to emphasize the dependence of q_t on n.

Formally, we make the following assumption on graphical channel models under consideration.

Assumption (A_1). *For any fixed $\alpha \in \mathbb{R}$ and $s, t \in \{0, \cdots, k\}$, there exists the limit*

$$d_{s:t}(\alpha) := \lim_{n\to\infty} \frac{\binom{n-1}{k-1}}{\log n} D_\alpha(q_s^{(n)} : q_t^{(n)}),$$

and $d_{s:t}(\alpha)$ is finite. We further assume that for any $A > 0$ and n,

$$\sup_{\alpha\in[-A,A]} \frac{\binom{n-1}{k-1}}{\log n} \left(|D'_\alpha(q_s^{(n)} : q_t^{(n)})| + |D''_\alpha(q_s^{(n)} : q_t^{(n)})|\right) < \infty,$$

where D'_α and D''_α are the first and second derivative of D_α with respect to α. Finally, we assume that $d_{s:t}$ is not identically zero for some $s, t \in \{0, \cdots, k\}$ to avoid the trivial case.

Since $d_{s:t}(\alpha)$ is a scaled limit of $D_\alpha(q_s : q_t)$, we expect that $d_{s:t}(\alpha)$ inherits many properties of D_α such as concavity.

Proposition 4.8. *Let $r, s, t \in \{0, 1, \cdots, k\}$. Then,*

(i) $d_{s:t}(\alpha)$ is concave, $d_{s:t}(0) = d_{s:t}(1) = 0$, and $d_{s:t}(\alpha) = d_{t:s}(1-\alpha)$.

(ii) If $d_{r:s}$ and $d_{r:t}$ are identically zero, then $d_{s:t}$ is identically zero as well. Hence, $d_{s:t} \equiv 0$ defines a equivalence relation.

(iii) If $d_{s:t}$ is identically zero, then $d_{r:s} \equiv d_{r:t}$.

Proof. $d_{s:t}$ is concave since it is a pointwise limit of concave functions. We have $d_{s:t}(0) = d_{s:t}(1) = 0$ because $D_0 = D_1 = 0$, and we have $d_{s:t}(\alpha) = d_{t:s}(1-\alpha)$ since $D_\alpha(\mu_1 : \mu_2) = D_{1-\alpha}(\mu_2 : \mu_1)$.

Before we prove (ii) and (iii), we claim that

$$D_\alpha(\mu_1 : \mu_2) \geq \frac{1}{p} D_{p\alpha}(\mu_1 : \mu_3) + \frac{1}{q} D_{q(1-\alpha)}(\mu_2 : \mu_3)$$

for any $p, q \in (1, \infty)$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$. This follows from

$$\begin{aligned} e^{-D_\alpha(\mu_1:\mu_2)} &= \int \left(\frac{d\mu_1}{d\mu_3}\right)^\alpha \left(\frac{d\mu_2}{d\mu_3}\right)^{1-\alpha} d\mu_3 \\ &\leq \left(\int \left(\frac{d\mu_1}{d\mu_3}\right)^{p\alpha} d\mu_3\right)^{1/p} \left(\int \left(\frac{d\mu_2}{d\mu_3}\right)^{q(1-\alpha)} d\mu_3\right)^{1/q} \\ &= \exp\left(-\frac{1}{p} D_{p\alpha}(\mu_1 : \mu_3) - \frac{1}{q} D_{q(1-\alpha)}(\mu_2 : \mu_3)\right). \end{aligned}$$

Here the second inequality follows from Hölder's inequality.

Suppose that $d_{r:s}$ and $d_{r:t}$ are identically zero. For any α, we have

$$D_\alpha(q_s^{(n)} : q_t^{(n)}) \geq \frac{1}{2} D_{2\alpha}(q_s^{(n)} : q_r^{(n)}) + \frac{1}{2} D_{2(1-\alpha)}(q_t^{(n)} : q_r^{(n)})$$

and we get

$$\begin{aligned} d_{s:t}(\alpha) \geq & \frac{1}{2} d_{s:r}(2\alpha) + \frac{1}{2} d_{t:r}(2(1-\alpha)) \\ = & \frac{1}{2} d_{r:s}(1-2\alpha) + \frac{1}{2} d_{r:t}(2\alpha - 1) = 0. \end{aligned}$$

It implies that $d_{s:t}$ is identically zero, since $d_{s:t}$ is concave and $d_{s:t}(0) = d_{s:t}(1) = 0$.

Finally, let us prove (iii). Suppose $d_{s:t}$ is identically zero. Note that

$$d_{r:s}(\alpha) = d_{s:r}(1-\alpha) \geq \frac{1}{p} d_{s:t}(p(1-\alpha)) + \frac{1}{q} d_{r:t}(q\alpha) = \frac{1}{q} d_{r:t}(q\alpha)$$

and

$$d_{r:t}(\alpha) = d_{t:r}(1-\alpha) \geq \frac{1}{p} d_{t:s}(p(1-\alpha)) + \frac{1}{q} d_{r:s}(q\alpha) = \frac{1}{q} d_{r:s}(q\alpha)$$

for any $\alpha \in \mathbb{R}$ and $p, q \in (1, \infty)$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$. We get

$$d_{r:s}(\alpha) \geq \frac{1}{q} d_{r:t}(q\alpha) \geq \frac{1}{q^2} d_{r:s}(q^2\alpha)$$

for any $q \in (1, \infty)$. Taking $q \to 1$, we have $d_{r:s} \equiv d_{r:t}$ since $d_{r:s}$ and $d_{r:t}$ are continuous.

□

Assuming (A_1), we define $I(\alpha)$ as

$$I(\alpha) = \sum_{t=0}^{k-1} \binom{k-1}{t} \rho^{k-1-t}(1-\rho)^t d_{t:t+1}(\alpha) \tag{4.1}$$

and $I = \max_\alpha I(\alpha)$. Note that the maximum is attained at some $\alpha^* \in (0,1)$, since I is concave and $I(0) = I(1) = 0$.

Let us justify that (A_1) is a natural assumption to make, from two examples we discussed in the introduction: the SBM with two symmetric communities and the spiked Wigner model with Rademacher prior. We remark that (A_1) holds in both examples.

Example 4.1. ***The SBM with two symmetric communities.*** *Let a and b be positive constants. The SBM with two symmetric communities (in a parameter regime that degrees grow logarithmically) is a graphical channel model $\mathcal{M}(p, Q)$ with uniform prior $p = U(\{0,1\})$, and kernel Q_n where*

$$q_t^{(n)} = Q(\cdot|0^{k-t}1^t) = \begin{cases} \mathrm{Ber}\left(\frac{a \log n}{n}\right) & \text{if } t = 0 \text{ or } t = 2 \\ \mathrm{Ber}\left(\frac{b \log n}{n}\right) & \text{if } t = 1. \end{cases}$$

We have

$$d_{s:t}(\alpha) = \begin{cases} \alpha a + (1-\alpha)b - a^\alpha b^{1-\alpha} & \text{if } s \neq 1 \text{ and } t = 1 \\ (1-\alpha)a + \alpha b - a^{1-\alpha} b^\alpha & \text{if } s = 1 \text{ and } t \neq 1 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\begin{aligned} I(\alpha) &= \frac{1}{2}(d_{0:1}(\alpha) + d_{1:2}(\alpha)) \\ &= \frac{1}{2}\left(a + b - a^\alpha b^{1-\alpha} - a^{1-\alpha} b^\alpha\right). \end{aligned}$$

Hence, we get $p_{fail,v} \le n^{-I+o(1)}$ *where*

$$I = \max_{\alpha \in [0,1]} I(\alpha) = \frac{1}{2}(\sqrt{a} - \sqrt{b})^2. \qquad \square$$

Example 4.2. ***Spiked Wigner model with Rademacher prior.*** *Let* β *be a positive constant, and let* $\sigma_n = \sqrt{\frac{n}{\beta \log n}}$. *The spiked Wigner model with Rademacher prior is a graphical channel model* $\mathcal{M}(p, Q)$ *with uniform prior* $p = U(\{0,1\})$, *and the kernel* Q_n *where*

$$q_t^{(n)} = Q(\cdot|0^{k-t}1^t) = \begin{cases} N(1, \sigma_n^2) & \text{if } t = 0 \text{ or } t = 2 \\ N(-1, \sigma_n^2) & \text{if } t = 1. \end{cases}$$

We have

$$d_{s:t}(\alpha) = \begin{cases} 2\alpha(1-\alpha)\beta & \text{if } |s-t| = 1 \\ 0 & \text{otherwise,} \end{cases}$$

and $I(\alpha) = \frac{1}{2}(d_{0:1}(\alpha) + d_{1:2}(\alpha)) = 2\alpha(1-\alpha)\beta$. *Hence, we get* $p_{fail,v} \le n^{-I+o(1)}$ *where*

$$I = \max_{\alpha \in [0,1]} I(\alpha) = \frac{\beta}{2}. \qquad \square$$

We prove that the upper bound $p_{fail,v} \le n^{-I+o(1)}$ can be amplified to an upper bound $p_{fail,global} \lesssim n \cdot n^{-I}$, which we call *weak amplification.*

Theorem 4.9 (Weak amplification). *Suppose* (A_1) *holds. Then,* $p_{fail,v} \le n^{-I+o(1)}$ *and* $p_{fail,global} \le n^{-(I-1)+o(1)}$. *In particular, exact recovery is achievable whenever* $I > 1$.

4.3.2 Strong amplification

We note that the inequality $p_{fail,v} \le n^{-I+o(1)}$ is tight in both examples. See [4] for the SBM with two symmetric communities, and see [21] and [61] for the spiked Wigner model. We remark that we implicitly used a Chernoff-type bound to get $p_{fail,v} \le n^{-I+o(1)}$, and the papers we mentioned above argue that this Chernoff-type bound is essentially tight. However, their approaches rely heavily on the explicit description of the channel.

Instead, we would like to impose a similar condition on generic graphical channel models, namely the "tightness of Chernoff-type bounds".

Chernoff-type bounds

Let us derive again $p_{fail,v} \leq n^{-I+o(1)}$ using a Chernoff-type method (which is essentially equivalent to our original derivation, just in different language).

Given a sequence of independent random variables $X_1, \cdots, X_N$, the tail probability of the sum $\sum_{i=1}^N X_i$ can be estimated as

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^N X_i \geq t\right) &\leq \min_{\lambda \geq 0} e^{-\lambda t} \prod_{i=1}^N \mathbb{E}\, e^{\lambda X_i} \\ &= \exp\left(-\max_{\lambda \geq 0}\left\{\lambda t - \sum_{i=1}^N \log \mathbb{E}\, e^{\lambda X_i}\right\}\right), \end{aligned}$$

which is often called a Chernoff-type bound. When $X_1, \cdots, X_N$ are identically distributed and $t = Nx$, we get

$$-\log \mathbb{P}\left(\sum_{i=1}^N X_i \geq Nx\right) \geq N \cdot \max_{\lambda \geq 0}\left(\lambda x - \log \mathbb{E}\, e^{\lambda X_1}\right).$$

Here $\kappa_{X_1}(\lambda) := \log \mathbb{E}\, e^{\lambda X_1}$ is called *cumulant-generating function* of X_1, and

$$\kappa^*_{X_1}(\mathbf{x}) := \max_{\lambda} \lambda x - \kappa_{X_1}(\lambda)$$

is called the *convex conjugate* or *Fenchel-Legendre transform* of κ_{X_1}. When $x > \mathbb{E}\, X_1$, we have

$$\mathbb{P}\left(\sum_{i=1}^N X_i \geq Nx\right) \leq e^{-N\kappa^*_{X_1}(\mathbf{x})}.$$

Cramér's Theorem tells us that this inequality is essentially tight.

Theorem 4.10 (Cramér's Theorem). *Suppose that $\kappa_{X_1}(\lambda)$ is finite everywhere. Let* *us fix* $x > \mathbb{E} X_1$. *Then,*

$$\lim_{N\to\infty} -\frac{1}{N} \log \mathbb{P}\left(\sum_{i=1}^{N} X_i \geq Nx\right) = \kappa^*_{X_1}(\mathbf{x}).$$

We recommend a book by Dembo and Zeitouni [39] to readers for further information on Cramér's theorem and its variants.

Now, let us return to our case of $p_{fail,v}$. Let $\mathbf{x}^0 \in \{0,1\}^V$ with $|\mathbf{x}^0| = \rho n$ and $\mathbf{x}^0_v = 0$. Then, we have

$$\begin{aligned} p_{fail,v} \approx\ & \mathbb{P}\left(\ell(\mathbf{x}^0) + \ell(\mathbf{y}^0|\mathbf{x}^0) \leq \ell(\mathbf{x}^v) + \ell(\mathbf{y}^0|\mathbf{x}^v)\right) \\ =\ & \mathbb{P}\left(\sum_{e\in E} \log \frac{Q(\mathbf{y}^0_e|\mathbf{x}^v[e])}{Q(\mathbf{y}^0_e|\mathbf{x}^0[e])} \geq \log \frac{p(\mathbf{x}^0_v)}{1-p(\mathbf{x}^0_v)}\right) \end{aligned}$$

For $e \in E$, let

$$L_e := \log \frac{Q(\mathbf{y}^0_e|\mathbf{x}^v[e])}{Q(\mathbf{y}^0_e|\mathbf{x}^0[e])} \quad \text{where } \mathbf{y}^0_e \sim Q(\cdot|\mathbf{x}^0[e]).$$

Let $T = \{v \in V : \mathbf{x}^0_v = 1\}$. Then, $|\mathbf{x}^0[e]| = |e \cap T|$ and

$$|\mathbf{x}^v[e]| = \begin{cases} |e\cap T| & \text{if } e \not\ni v \\ |e \cap T| + 1 & \text{if } e \ni v. \end{cases}.$$

We have $L_e \equiv 0$ if $e \not\ni v$, and

$$L_e \equiv \log \frac{q_{t+1}(y)}{q_t(y)} \quad \text{where } y \sim q_t$$

if $v \in e$ and $|e \cap T| = t$.

Let κ_t be the cumulant-generating function of L_e where $|e \cap T| = t$. By Chernoff-type bound, we get

$$-\log \mathbb{P}\left(\sum_{e\in E} L_e \geq \log \frac{1-\rho}{\rho}\right) \geq \max_{\lambda \geq 0} \lambda \log \frac{1-\rho}{\rho} - \sum_{e \ni v} \kappa_{|e\cap T|}(\lambda).$$

Moreover,

$$
\begin{aligned}
\sum_{e \ni v} \kappa_{|e \cap T|}(\lambda) = & \sum_{t=0}^{k-1} \kappa_t(\lambda) \cdot \binom{|T|}{t} \binom{|V \setminus (T \cup \{v\})|}{k-1-t} \\
\approx & \binom{n-1}{k-1} \sum_{t=0}^{k-1} \binom{k-1}{t} \rho^t (1-\rho)^{k-1-t} \kappa_t(\lambda).
\end{aligned}
$$

We note that

$$
\begin{aligned}
\kappa_t(\lambda) = & \log \mathop{\mathbb{E}}_{y \sim q_t} \exp\left(\lambda \log \frac{q_{t+1}(y)}{q_t(y)}\right) \\
= & \log \mathop{\mathbb{E}}_{y \sim q_t} \left(\frac{q_{t+1}(y)}{q_t(y)}\right)^{\lambda} = -D_\lambda(q_{t+1} : q_t).
\end{aligned}
$$

Thus we have

$$
-\log p_{fail,v} \geq (1 - o(1)) \log n \cdot \max_{\lambda \geq 0} I(\lambda).
$$

Since I is concave and $I(0) = I(1) = 0$, we recover $p_{fail,v} \leq n^{-I+o(1)}$ where $I = \max_{\alpha \in [0,1]} I(\alpha)$.

Tightness of Chernoff-type bounds and strong amplification

In the previous section, we obtained

$$
I \leq \liminf_{n \to \infty} -\frac{\log p_{fail,v}}{\log n}
$$

via a Chernoff-type bound. As in Cramer's Theorem, it would be nice to have

$$
I \geq \limsup_{n \to \infty} -\frac{\log p_{fail,v}}{\log n},
$$

which means that the Chernoff-type bound we exploited in the previous section is essentially tight. Formally, we make the following assumption.

Assumption (A_2). *For $t \in \{0, \cdots, k-1\}$, let $\eta^{(n)}_{t+1:t}$ be the distribution of $\log \frac{q^{(n)}_{t+1}(Y)}{q^{(n)}_t(Y)}$ where $Y \sim q^{(n)}_t$. Let $N_0, \cdots, N_{k-1}$ be nonnegative integers such that $N_t = c_t \binom{n-1}{k-1} + o(n^{k-1})$ for some constants $c_0, \cdots, c_{k-1}$, and let $X^{(t)}_i$ for $t = 0, \cdots, k-1$ and $i = 1, \cdots, N_t$ be independent random variables where $X^{(t)}_i \sim \eta^{(n)}_{t+1:t}$. Let $L = \sum_{t=0}^{k-1} \sum_{i=1}^{N_t} X^{(t)}_i$*

and let

$$\widetilde{I}(\alpha) = \sum_{t=0}^{k-1} c_t d_{t+1:t}(\alpha) \quad \textit{and} \quad \widetilde{I}^*(\delta) = \max_{\alpha \geq 0} \alpha\delta + I(\alpha).$$

We assume that for any constant $\delta > -\frac{d\widetilde{I}}{d\lambda}(0)$ *the following large deviation estimate holds:*

$$\mathbb{P}\left(L \geq \delta \log n\right) \geq \exp\left(-\widetilde{I}^*(\delta) \cdot \log n - o(\log n)\right).$$

We further require

$$\lim_{n\to\infty} \frac{\mathbb{P}\left(L \geq (\delta + \epsilon_n)\log n\right)}{\mathbb{P}\left(L \geq \delta \log n\right)} = 1$$

for any sequence ϵ_n *satisfying* $\lim_{n\to\infty} \epsilon_n \log n = 0$.

Now we are ready to present the main result of this chapter.

Theorem 4.11 (Strong amplification). *Local-to-global amplification holds for* $\mathcal{M}(p, Q)$ *assuming* (A_1) *and* (A_2).

Hence, exact recovery in $\mathcal{M}(p, Q)$ exhibits a sharp phase transition around $I = 1$.

Corollary 4.12. *Assume* (A_1) *and* (A_2). *If* $I > 1$, *then exact recovery is achievable. On the other hand, exact recovery is impossible if* $I < 1$.

As a final remark, we ask whether (A_2) is a necessary assumption to have a strong amplification. In particular, if (A_2) can be deduced unconditionally (or with milder set of conditions) from (A_1), then it would imply that local-to-global amplification holds for any model in the parameter regime under consideration. For instance, Gärtner-Ellis theorem and its strengthening [31] implies that tight large-deviation type estimate is possible as long as the scaled cumulant generating function converges nicely (See [39] for more information on large-deviation theory). This may be applicable to the context of graphical channel model since α-divergence is essentially the cumulant generating function, but we leave this part for future work.

Nevertheless, we note that the assumption (A_1) and (A_2) are both a condition on the *local recovery*, hence the exact recovery in graphical channel model can be characterized solely by the "capacity" of the channel which corresponds to I in the context of this chapter.

We note that for detection problem a similar result was obtained under the name of *channel universality* [66].

4.3.3 Exact recovery up to a global switch of labels

Recall that exact recovery up to a global switch of labels is an alternative recovery requirement in which we want to recover either $\mathbf{x}^0$ or $\mathbf{1} - \mathbf{x}^0$ from given $\mathbf{y}^0$. In particular, exact recovery is only possible up to a global switch of labels when the model satisfy $q_t \equiv q_{k-t}$ for any t.

Definition 4.5. *$\mathcal{M}(p, Q)$ is said to be* asymptotically symmetric up to a global switch of labels, *or simply* symmetric, *if p is uniform and $d_{t:k-t} \equiv 0$ for all t. We say $\mathcal{M}(p, Q)$* asymmetric *if it is not symmetric. Moreover, we say $\mathcal{M}(p, Q)$ is* strongly asymmetric *if $d_{t:k-t} \not\equiv 0$ for some t.*

Note that this definition of symmetry allows that $\mathcal{M}(p, Q_n)$ being almost symmetric but not symmetric. For example, let us consider a SBM with uniform prior $\mathbf{x}^0 \in \{0, 1\}^V$ and posterior

$$\mathbf{y}_{ij}^0 = \begin{cases} \frac{a \log n}{n} & \text{if } \mathbf{x}_i^0 = \mathbf{x}_j^0 = 0 \\ \frac{b \log n}{n} & \text{if } \mathbf{x}_i^0 \neq \mathbf{x}_j^0 \\ \frac{(a+\epsilon_n) \log n}{n} & \text{if } \mathbf{x}_i^0 = \mathbf{x}_j^0 = 1. \end{cases}$$

When $\epsilon_n = o(1)$, this model is not exactly symmetric since $\mathbb{P}(\mathbf{y}^0|\mathbf{x}^0) \not\equiv \mathbb{P}(\mathbf{y}^0|\mathbf{1} - \mathbf{x}^0)$, but we cannot distinguish $\mathbf{x}^0$ and $\mathbf{1} - \mathbf{x}^0$ for any instances of a and b. Hence, it makes sense to call this model symmetric as our definition suggests.

When the model is symmetric, we may want to modify our definition of $p_{fail,global}$ accordingly:

$$\begin{aligned} p_{fail,global} &= \mathbb{P}\left(\widehat{\mathbf{x}}_{map}(\mathbf{y}^0) \notin \{\mathbf{x}^0, \mathbf{1} - \mathbf{x}^0\}\right) \\ &= \mathbb{P}\left(\bigcup_{S \notin \{\emptyset, V\}} E_S\right). \end{aligned}$$

We have the corresponding amplification results for symmetric models.

Theorem 4.13. *Suppose that* (A_1) *holds and the model is symmetric. Then, (i) weak amplification holds and (ii) strong amplification holds if* (A_2) *holds. We remark that amplifications are based on the modified definition of* $p_{fail,global}$ *for symmetric models.*

4.3.4 Proof overview

We close this section by outlining the proof of Theorem 4.11 and 4.9.

Weak amplification: Achievability

To prove the weak amplification, we need to analyze $p_{fail,global}$ which is the probability for the (global) MAP estimator failing to return the correct prior. For $\mathbf{x}^0 \in \{0,1\}^V$ and $S \subseteq V$, let us define $\mathbf{x}^S \in \{0,1\}^V$ with entries

$$\mathbf{x}_v^S = \begin{cases} \mathbf{x}_v & \text{if } v \notin S \\ 1 - \mathbf{x}_v & \text{if } v \in S. \end{cases}$$

Let E_S be the event defined as

$$E_S := \{\ell(\mathbf{x}^0) + \ell(\mathbf{y}^0|\mathbf{x}^0) \le \ell(\mathbf{x}^S) + \ell(\mathbf{y}^0|\mathbf{x}^S)\}.$$

Then, we can write $p_{fail,global}$ as the probability that E_S happens for some $S \neq \emptyset$ (or $S \notin \{\emptyset, V\}$ in symmetric case). Hence,

$$\begin{aligned} p_{fail,global} = \mathbb{P}\left(\bigcup_{S\neq\emptyset} E_S\right) \le & \sum_{S\neq\emptyset} \mathbb{P}\left(E_S \setminus \bigcup_{v\in S} E_{S\setminus v}\right) \\ = & \sum_{v\in V} \mathbb{P}(E_v) + \sum_{S:|S|\ge 2} \mathbb{P}\left(E_S \setminus \bigcup_{v\in S} E_{S\setminus v}\right). \end{aligned}$$

Note that the first sum in the right-hand side is equal to $\sum_{v\in V} p_{fail,v}$. If we could prove that

$$\mathbb{P}\left(E_S \setminus \bigcup_{v\in S} E_{S\setminus v}\right) \lesssim \prod_{v\in S} p_{fail,v},$$

then it would imply one side of the strong amplification: $p_{fail,global} \lesssim \sum_{v \in V} p_{fail,v}$. Instead, we prove a weaker bound

$$\mathbb{P}\left(E_S \setminus \bigcup_{v \in S} E_{S \setminus v}\right) \lesssim n^{-(1+o(1))I|S|}$$

which implies that

$$\begin{aligned}\sum_{S:|S| \geq 2} \mathbb{P}\left(E_S \setminus \bigcup_{v \in S} E_{S \setminus v}\right) &\lesssim \sum_{d \geq 2} \binom{n}{d} n^{-dI} \\ &\approx \sum_{d \geq 2} n^{-d(I-1)} = o(n^{-(I-1)}).\end{aligned}$$

Hence,

$$p_{fail,global} \lesssim \sum_{v} p_{fail,v} + o(n^{-(I-1)}) \lesssim n^{-(I-1)}.$$

See Section 4.4.1 for the details.

Strong amplification: Impossibility

We note that our argument is a close generalization of a second-moment method used in [1]. Recall that $p_{fail,global} \geq p_{fail,local}$ where

$$p_{fail,local} = \mathbb{P}\left(\bigcup_{v \in V} E_v\right).$$

We would like to argue that E_v's have low pairwise correlations, and want that

$$1 - p_{fail,local} = \mathbb{P}\left(\bigcap_{v \in V} (E_v)^c\right) \approx \prod_{v \in V} \mathbb{P}((E_v)^c).$$

Let $Z = \sum_{v \in V} \mathbf{1}_{E_v}$ where $\mathbf{1}_{E_v}$ is the indicator random variable for the event E_v. By Chebyshev's inequality, we have

$$1 - p_{fail,local} = \mathbb{P}(Z \leq 0) \leq \frac{\mathbb{E}(Z - \mathbb{E}\, Z)^2}{(\mathbb{E}\, Z)^2} = \frac{\mathbb{E}\, Z^2}{(\mathbb{E}\, Z)^2} - 1.$$

Hence,

$$p_{fail,local} \geq 2 - \frac{\mathbb{E}\, Z^2}{(\mathbb{E}\, Z)^2}.$$

Note that $\mathbb{E}\, Z = \sum_{v \in V} p_{fail,v}$ and

$$\mathbb{E}\, Z^2 = \sum_{u,v \in V} \mathbb{E}\, \mathbf{1}_{E_v} \mathbf{1}_{E_v} = \sum_{v \in V} p_{fail,v} + \sum_{u \neq v} \mathbb{P}(E_u \cap E_v).$$

Suppose that for any $u \neq v$, we have

$$\frac{\mathbb{P}(E_u \cap E_v)}{\mathbb{P}(E_u)\, \mathbb{P}(E_v)} \leq 1 + o(1).$$

In other words, events E_u and E_v are almost independent. Then, we get

$$p_{fail,global} \geq p_{fail,local} \geq 1 - o(1) - \left(\sum_{v \in V} p_{fail,v} \right)^{-1}.$$

Moreover, (A_2) implies that $p_{fail,v} \gtrsim n^{-I}$. Together with the weak amplification, we get

$$\begin{aligned} 1 - o(1) - \left(n^{1-I}\right)^{-1} \lesssim\ & 1 - o(1) - \left(\sum_{v \in V} p_{fail,v} \right)^{-1} \\ \leq\ & p_{fail,local} \leq p_{fail,global} \lesssim n^{1-I} \lesssim \sum_{v \in V} p_{fail,v} \end{aligned}$$

as desired. See Section 4.4.2 for the details.

4.4 Proofs

In this section, we prove Theorem 4.9 and 4.11. We only focus on strongly asymmetric models, *i.e.*, $d_{t:k-t} \not\equiv 0$ for some $t \in \{0, \cdots, k\}$. We remark that the proofs for non-strongly asymmetric cases or symmetric cases is a simple adaptation of the proof for the strongly asymmetric cases.

Let us introduce and recall a few notations which are heavily used throughout the section. First of all, we will omit n in subscripts or superscripts if the context is clear. In particular, we let $V = [n]$ and $E = \binom{V}{k}$, and all asymptotics are based on $n \to \infty$.

We consider $\mathcal{M}(p, Q)$ where $p = \text{Ber}(\rho)$ for some $\rho \in (0, 1/2]$ and $Q(\cdot|z) \equiv q_{|z|}(\cdot)$ for $z \in \{0,1\}^k$. Here Q and q_t depend on n, implicitly.

For $S \subseteq V$ and $\mathbf{x}^0 \in \{0,1\}^V$, a vector $\mathbf{x}^S \in \{0,1\}^V$ is defined as

$$\mathbf{x}_u^S = \begin{cases} \mathbf{x}_u^0 & \text{if } u \notin S \\ 1 - \mathbf{x}_u^0 & \text{if } u \in S, \end{cases}$$

and E_S is defined as the event that

$$\ell(\mathbf{x}^0) + \ell(\mathbf{y}^0|\mathbf{x}^0) \leq \ell(\mathbf{x}^S) + \ell(\mathbf{y}^0|\mathbf{x}^S)$$

happens. We denote $E_{\{v\}}$ by E_v for brevity.

The global and local failure probability $p_{fail,global}$ and $p_{fail,local}$ can be expressed as

$$p_{fail,global} = \mathbb{P}\left(\bigcup_{S \neq \emptyset} E_S\right) \quad \text{and} \quad p_{fail,local} = \mathbb{P}\left(\bigcup_{v \in V} E_v\right)$$

where probabilities are taken over $(\mathbf{x}^0, \mathbf{y}^0) \sim \mathcal{M}(p, Q)$. The probability for failing local recovery at v is equal to $p_{fail,v} = \mathbb{P}(E_v)$.

We assume (A_1) throughout the section: For any $\alpha \in \mathbb{R}$ and $s, t \in \{0, \cdots, k\}$, the limit

$$d_{s:t}(\alpha) = \lim_{n \to \infty} \frac{\binom{n-1}{k-1}}{\log n} D_\alpha(q_s : q_t)$$

exists and finite. The goal of this section is to prove the following amplification results:

- **Weak amplification.** $p_{fail,v} \leq n^{-I+o(1)}$ and $p_{fail,global} \leq n^{-(I-1)+o(1)}$.

- **Strong amplifcation.** If (A_2) holds, then $p_{fail,global} \geq (1 + o(1)) p_{fail,local}$ and

$$p_{fail,local} \geq 1 - \left(\sum_{v \in V} p_{fail,v}\right)^{-1} - o(1).$$

4.4.1 Weak amplification: Proof of Theorem 4.9

We would like to focus our attention to the event that $|\mathbf{x}^0|$ is very close to $\mathbb{E}\,|\mathbf{x}^0|$. We call such $\mathbf{x}^0$ *typical.* A standard Chernoff's bound gives us the following concentration result.

Theorem 4.14. *Let $X_1, \cdots, X_N$ be i.i.d. distribution with $X_1 \sim \mathrm{Ber}(\rho)$. Let δ be a real number such that $0 \le \delta \le 1$. Then,*

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{i=1}^{N} X_i - \rho\right| \ge \delta\rho\right) \le 2e^{-\frac{\delta^2 \rho N}{3}}.$$

Let $\mathsf{Typ} = \{\rho n - \sqrt{n}\log n \le |\mathbf{x}^0| \le \rho n + \sqrt{n}\log n\}$. Letting $\delta = \frac{\log n}{\rho\sqrt{n}}$, we get $\mathbb{P}(\mathsf{Typ}) \ge 1 - e^{-\Omega(\log^2 n)}$ by Theorem 4.14. Thus,

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{S\neq\emptyset} E_S\right) &\le \mathbb{P}(\mathsf{Typ}) \max_{\mathbf{x}^0\in\mathsf{Typ}} \mathbb{P}\left(\bigcup_{S\neq\emptyset} E_S \middle| \mathbf{x}^0\right) + \mathbb{P}(\neg\mathsf{Typ}) \\
&\le \max_{\mathbf{x}^0\in\mathsf{Typ}} \mathbb{P}\left(\bigcup_{S\neq\emptyset} E_S \middle| \mathbf{x}^0\right) + n^{-\Omega(\log n)}.
\end{aligned}$$

This implies that it is sufficient to prove that

$$\mathbb{P}\left(\bigcup_{S\neq\emptyset} E_S \middle| \mathbf{x}^0\right) \le n^{-(I-1)+o(1)}$$

for any typical $\mathbf{x}^0$, to prove $p_{fail,global} \le n^{-(I-1)+o(1)}$,

Let

$$E'_S = E_S \setminus \left(\bigcup_{v\in S} E_{S\setminus\{v\}}\right).$$

By a union bound, we have

$$\mathbb{P}\left(\bigcup_{S\neq\emptyset} E_S \middle| \mathbf{x}^0\right) \le \sum_{S\neq\emptyset} \mathbb{P}(E'_S|\mathbf{x}^0).$$

We are going to argue that when $|S| \le \epsilon n$ (where $\epsilon = o(1)$ will be chosen later), we have

$$\mathbb{P}(E'_S|\mathbf{x}^0) \leq n^{-(I-o(1))|S|}$$

and so

$$\begin{aligned}\sum_{S:1\leq|S|\leq\epsilon n} \mathbb{P}(E'_S|\mathbf{x}^0) &\leq \sum_{m=1}^{\epsilon n} \binom{n}{m} (n^{-I+o(1)})^m \\ &\leq \sum_{m\geq 1} (n^{-(I-1)+o(1)})^m \leq O(n^{-(I-1)+o(1)}).\end{aligned}$$

Lemma 4.15. *Let* $\epsilon = (\log\log n)^{-1}$. *If* $|S| \leq \epsilon n$, *then*

$$\mathbb{P}(E'_S|\mathbf{x}^0) \leq \exp\left(-|S| \cdot I \log n + o(|S|\log n)\right).$$

In particular,

$$\mathbb{P}(E_v|\mathbf{x}^0) = \mathbb{P}(E'_v|\mathbf{x}^0) \leq n^{-I+o(1)},$$

hence $p_{fail,v} \leq n^{-I+o(1)}$.

Moreover, we show that if $|S| > \epsilon n$ then

$$\mathbb{P}(E'_S|\mathbf{x}^0) \leq \mathbb{P}(E_S|\mathbf{x}^0) \leq e^{-\Omega(\epsilon^k n \log n)},$$

hence

$$\sum_{S:|S|>\epsilon n} \mathbb{P}(E'_S|\mathbf{x}^0) \leq 2^n e^{-\Omega(\epsilon^k n \log n)}.$$

The right-hand side is $e^{-\omega(n)}$ as long as $\epsilon \gg (\log n)^{-1/k}$, thus we get

$$p_{fail,global} \leq n^{-(I-1)+o(1)},$$

as $e^{-\omega(n)}$ and $e^{-\Omega(\log^2 n)}$ decays much faster than $n^{-(I-1)}$.

Lemma 4.16. *If* $\epsilon n < |S| \leq \frac{n}{2}$, *then* $\mathbb{P}(E_S|\mathbf{x}^0) \leq e^{-\Omega(\epsilon^k n \log n)}$. *If* $|S| > \frac{n}{2}$ *and the model is strongly asymmetric, then* $\mathbb{P}(E_S|\mathbf{x}^0) \leq e^{-\Omega(\epsilon^k n \log n)}$.

As we argued above, Lemma 4.15 and Lemma 4.16 together implies the weak amplification for strongly asymmetric models. We note that similar argument can be applied for models which are not strongly asymmetric. For example, in symmetric models we can estimate $\mathbb{P}(E'_S|\mathbf{x}^0)$ as

$$\mathbb{P}(E'_S|\mathbf{x}^0) \approx \mathbb{P}(E'_{V\setminus S}|\mathbf{x}^0)$$

when $|S| > \frac{n}{2}$, hence we get

$$p_{fail,global} \lesssim 2 \sum_{S:1\leq|S|\leq\frac{n}{2}} \mathbb{P}(E'_S|\mathbf{x}^0)$$

and weak amplification as well.

Proof of Lemma 4.15

Note that E'_S is the event that

$$\max_{v\in S} \ell(\mathbf{x}^{S\setminus\{v\}}) + \ell(\mathbf{y}^0|\mathbf{x}^{S\setminus\{v\}}) < \ell(\mathbf{x}^0) + \ell(\mathbf{y}^0|\mathbf{x}^0) \leq \ell(\mathbf{x}^S) + \ell(\mathbf{y}^0|\mathbf{x}^S)$$

holds. Let $F_{S,v}$ be the event that

$$\ell(\mathbf{x}^{S\setminus\{v\}}) + \ell(\mathbf{y}^0|\mathbf{x}^{S\setminus\{v\}}) \leq \ell(\mathbf{x}^S) + \ell(\mathbf{y}^0|\mathbf{x}^S)$$

happens. Since E'_S implies $\bigcap_{v\in S} F_{S,v}$, we have

$$\mathbb{P}(E'_S|\mathbf{x}^0) \leq \mathbb{P}\left(\bigcap_{v\in S} F_{S,v}\middle|\, \mathbf{x}^0\right).$$

Let us take a look at $F_{S,v}$ by expanding ℓ. We have $F_{S,v}$ if and only if

$$\sum_{e\in E:e\ni v} \log \frac{Q(\mathbf{y}^0_e|\mathbf{x}^S[e])}{Q(\mathbf{y}^0_e|\mathbf{x}^{S\setminus\{v\}}[e])} \geq \log \frac{p(\mathbf{x}^0_v)}{p(1-\mathbf{x}^0_v)}$$

since $|\mathbf{x}^S[e]| \neq |\mathbf{x}^{S\setminus\{v\}}[e]|$ only if e contains v. Let $\mathcal{F}_{S,v}$ and $\mathcal{F}'_S$ be subsets of E where

$$\mathcal{F}_{S,v} = \{e \in E : e \cap S = \{v\}\} \quad \text{and} \quad \mathcal{F}'_S = \{e \in E : |e \cap S| \geq 2\}.$$

Note that if $e \in \mathcal{F}_{S,v}$, then we have $\mathbf{x}^S[e] = \mathbf{x}^v[e]$ and $\mathbf{x}^{S\setminus\{v\}}[e] = \mathbf{x}^0[e]$. Hence, if $F_{S,v}$ happens then $L_{S,v} + L'_{S,v} \geq c$ where

$$c := \log \frac{\rho}{1-\rho} \leq \min_{v \in S} \log \frac{p(\mathbf{x}_v^0)}{1 - p(\mathbf{x}_v^0)},$$

and

$$L_{S,v} = \sum_{e \in \mathcal{F}_{S,v}} \log \frac{q_{|\mathbf{x}^v[e]|}(\mathbf{y}_e^0)}{q_{|\mathbf{x}^0[e]|}(\mathbf{y}_e^0)} \quad \text{and} \quad L'_{S,v} = \sum_{e \in \mathcal{F}'_S : e \ni v} \log \frac{Q(\mathbf{y}_e^0 | \mathbf{x}^S[e])}{Q(\mathbf{y}_e^0 | \mathbf{x}^{S\setminus\{v\}}[e])}.$$

Claim. *$\{L_{S,v} : v \in S\}$ is a family of jointly independent random variables, when conditioned on $\mathbf{x}^0$. Two families $\{L_{S,v} : v \in S\}$ and $\{L'_{S,v} : v \in S\}$ are independent of each other when conditioned on $\mathbf{x}^0$.*

Proof of Claim. It follows from that $\mathcal{F}_{S,v}$ are mutually disjoint and that $\mathcal{F}_{S,v}$ does not intersect $\mathcal{F}'_S$ for any $v \in S$. □

This implies that

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{v \in S} F_{S,v} \,\middle|\, \mathbf{x}^0 \right) &\leq \mathbb{P}\left(\bigcap_{v \in S} \{L_{S,v} + L'_{S,v} \geq c\} \,\middle|\, \mathbf{x}^0 \right) \\
&= \mathop{\mathbb{E}}_{\{\mathbf{y}_e^0\}_{e \in \mathcal{F}'_S}} \left[\mathbb{P}\left(\bigcap_{v \in S} \{L_{S,v} \geq -L'_{S,v} + c\} \,\middle|\, \mathbf{x}^0, \{\mathbf{y}_e^0\}_{e \in \mathcal{F}'_S} \right) \right] \\
&= \mathop{\mathbb{E}}_{\{\mathbf{y}_e^0\}_{e \in \mathcal{F}'_S}} \left[\prod_{v \in S} \mathbb{P}\left(L_{S,v} \geq -L'_{S,v} + c \mid \mathbf{x}^0, \{\mathbf{y}_e^0\}_{e \in \mathcal{F}_{S,v}} \right) \right] \\
&\leq \mathop{\mathbb{E}}_{\{\mathbf{y}_e^0\}_{e \in \mathcal{F}'_S}} \left[\prod_{v \in S} \exp\left(- \max_{\alpha \in [0,1]} \alpha(c - L'_{S,v}) - \log \mathbb{E}\, e^{\alpha L_{S,v}} \right) \right].
\end{aligned}$$

Now, note that

$$-\log \mathbb{E}\, e^{\alpha L_{S,v}} = \sum_{e \in \mathcal{F}_{S,v}} D_\alpha \left(q_{|\mathbf{x}^v[e]|} : q_{|\mathbf{x}^0[e]|} \right).$$

When $\mathbf{x}^0_v = 0$, we have

$$
\begin{aligned}
-\log \mathbb{E}\, e^{\alpha L_{S,v}} &= \sum_{t=0}^{k-1} D_\alpha(q_{t+1} : q_t) \cdot \#(e : (e \setminus \{v\}) \cap S = \emptyset,\ |\mathbf{x}^0[e \setminus \{v\}]| = t) \\
&\leq \sum_{t=0}^{k-1} D_\alpha(q_{t+1} : q_t) \binom{|\mathbf{x}^0| - |S|}{t} \binom{n - |\mathbf{x}^0| - |S|}{k-1-t} \\
&= (1 + o(1)) \log n \sum_{t=0}^{k-1} \binom{k-1}{t} \rho^t (1-\rho)^{k-1-t} d_{t+1:t}(\alpha) \\
&= \log n \cdot (I(\alpha) + o(1))
\end{aligned}
$$

since $|S| \leq \epsilon n = o(n)$ and $|\mathbf{x}^0| = (\rho + o(1))n$. Similarly, if $\mathbf{x}^0_v = 1$, then

$$
-\log \mathbb{E}\, e^{\alpha L_{S,v}} = \log n \cdot (I(1-\alpha) + o(1)).
$$

Let $\alpha^* \in (0,1)$ be where the maximum of $I(\alpha)$ is attained. Then, we have

$$
\max_{\alpha \in [0,1]} \alpha(c - L'_{S,v}) - \log \mathbb{E}\, e^{\alpha L_{S,v}} \geq (I + o(1)) \log n + \begin{cases} \alpha^* & \text{if } \mathbf{x}^0_v = 0 \\ 1 - \alpha^* & \text{if } \mathbf{x}^0_v = 1. \end{cases}
$$

Hence, $\mathbb{P}\left(\bigcap_{v \in S} F_{S,v} \;\middle|\; \mathbf{x}^0\right)$ is at most

$$
n^{-(I+o(1))|S|} \cdot \mathop{\mathbb{E}}_{\{\mathbf{y}^0_e\}_{e \in \mathcal{F}'_S}} \left[\exp\left(\alpha^* \sum_{\substack{v \in S: \\ \mathbf{x}^0_v = 0}} (c - L'_{S,v}) + (1 - \alpha^*) \sum_{\substack{v \in S: \\ \mathbf{x}^0_v = 1}} (c - L'_{S,v}) \right) \right].
$$

It remains to show that the expected value on the right-hand side is small. We will use the independence of $\mathbf{y}^0_e$'s to break it into the form of

$$
C \cdot \prod_{e \in \mathcal{F}'_S} \mathbb{E}_{\mathbf{y}^0_e} f_e(\mathbf{y}^0_e),
$$

and prove that

$$
\log C = o(|S| \log n) \quad \text{and} \quad \log \mathbb{E}\, f_e(\mathbf{y}^0_e) = O\left(\frac{\log n}{n^{k-1}}\right).
$$

To see this, first note that C is equal to $e^{c(m_0\alpha^*+m_1(1-\alpha^*))}$ where $m_1 = |\mathbf{x}^0[S]|$ and $m_0 = |S| - m_1$. Hence, $\log C \le c|S| = o(|S|\log n)$ as desired. On the other hand, we claim the following:

Claim. *Fix $e \in \mathcal{F}'_S$. Let $r_0 = |e \cap S| - |\mathbf{x}^0[e \cap S]|$ and $r_1 = |\mathbf{x}^0[e \cap S]|$, and let $s = |\mathbf{x}^0[e]|$ and $t = |\mathbf{x}^S[e]|$. Then,*

$$f_e(y) = \left(\frac{q_{t+1}(y)}{q_t(y)}\right)^{r_0\alpha^*}\left(\frac{q_{t-1}(y)}{q_t(y)}\right)^{r_1(1-\alpha^*)}.$$

Moreover, there is a constant $M > 0$ not depending on n such that

$$-\log \mathop{\mathbb{E}}_{\mathbf{y}_e^0\sim q_s} f_e(\mathbf{y}_e^0) \ge -M \cdot \frac{\log n}{\binom{n-1}{k-1}}$$

for sufficiently large n.

As a result, we get

$$C \cdot \prod_{e\in\mathcal{F}'_S} \mathbb{E}_{\mathbf{y}_e^0} f_e(\mathbf{y}_e^0) \le \exp\left(|S|\log n \cdot \left(o(1) + O\left(\frac{|S|}{n}\right)\right)\right) = n^{o(|S|)},$$

since $|\mathcal{F}'_S| = O(|S|^2 n^{k-2})$ and $\frac{|S|}{n} \le \epsilon = o(1)$. It concludes the proof of Lemma 4.15, as we get

$$\begin{aligned}\mathbb{P}(E'_S \mid \mathbf{x}^0) \le \mathbb{P}\left(\bigcap_{v\in S} F_{S,v} \,\middle|\, \mathbf{x}^0\right) &\le n^{-(I+o(1))|S|} \cdot n^{o(|S|)} \\ &= n^{-(I+o(1))|S|}\end{aligned}$$

for any typical $\mathbf{x}^0$ and S with $|S| \le \epsilon n$.

Proof of Claim. By collecting the terms depending on $\mathbf{y}_e^0$ in

$$\exp\left(-\sum_{e\in\mathcal{F}'_S}\left(\alpha^*\sum_{\substack{v\in S:\\ \mathbf{x}_v^0=0}} L'_{S,v} + (1-\alpha^*)\sum_{\substack{v\in S:\\ \mathbf{x}_v^0=1}} L'_{S,v}\right)\right),$$

we get

$$f_e(\mathbf{y}_e^0) = \exp\left(- \sum_{v \in e \cap S} (\alpha^*(1-\mathbf{x}_v^0) + (1-\alpha^*)\mathbf{x}_v^0) \log \frac{Q(\mathbf{y}_e^0 | \mathbf{x}^S[e])}{Q(\mathbf{y}_e^0 | \mathbf{x}^{S\setminus\{v\}}[e])} \right)$$
$$= \prod_{v \in e \cap S} \left(\frac{q_{t+1}(\mathbf{y}_e^0)}{q_t(\mathbf{y}_e^0)} \right)^{\alpha^*(1-\mathbf{x}_v^0)} \left(\frac{q_{t-1}(\mathbf{y}_e^0)}{q_t(\mathbf{y}_e^0)} \right)^{(1-\alpha^*)\mathbf{x}_v^0}$$
$$= \left(\frac{q_{t+1}(\mathbf{y}_e^0)}{q_t(\mathbf{y}_e^0)} \right)^{r_0\alpha^*} \left(\frac{q_{t-1}(\mathbf{y}_e^0)}{q_t(\mathbf{y}_e^0)} \right)^{r_1(1-\alpha^*)}.$$

Let $\beta_1 = r_0\alpha^*$ and $\beta_2 = r_1(1-\alpha^*)$. Then,

$$\begin{aligned}
\mathbb{E}\, f_e(\mathbf{y}_e^0) &= \mathop{\mathbb{E}}_{Y\sim q_s} q_{t+1}(Y)^{\beta_1} q_{t-1}(Y)^{\beta_2} q_t(Y)^{-\beta_1-\beta_2} \\
&= \mathop{\mathbb{E}}_{Y\sim q_s} \left(\frac{q_{t+1}}{q_s}(Y) \right)^{\beta_1} \left(\frac{q_{t-1}}{q_s}(Y) \right)^{\beta_2} \left(\frac{q_{t+1}}{q_s}(Y) \right)^{-\beta_1-\beta_2} \\
&\le \left(\mathop{\mathbb{E}}_{Y\sim q_s} \left(\frac{q_{t+1}}{q_s}(Y) \right)^{3\beta_1} \mathop{\mathbb{E}}_{Y\sim q_s} \left(\frac{q_{t-1}}{q_s}(Y) \right)^{3\beta_2} \mathop{\mathbb{E}}_{Y\sim q_s} \left(\frac{q_{t+1}}{q_s}(Y) \right)^{-3\beta_1-3\beta_2} \right)^{1/3},
\end{aligned}$$

by Hölder's inequality. Hence,

$$-\log \mathbb{E}\, f_e(\mathbf{y}_e^0) \ge \frac{1}{3} \left(D_{3\beta_1}(q_{t+1} : q_s) + D_{3\beta_2}(q_{t-1} : q_s) + D_{-3(\beta_1+\beta_2)}(q_t : q_s) \right).$$

Since β_1 and β_2 lies in $[0, k]$, we have

$$-\log \mathbb{E}\, f_e(\mathbf{y}_e^0) \ge \inf_{s,t \in \{0,\cdots,k\}} \inf_{\alpha \in [-3k, 3k]} D_\alpha(q_s : q_t).$$

Hence, by choosing M satisfying

$$-M < \min_{s,t \in \{0,\cdots,k\}} \min_{\alpha \in [-3k,3k]} d_{s:t}(\alpha),$$

we get $-\log \mathbb{E}\, f_e(\mathbf{y}_e^0) > -M \frac{\log n}{\binom{n-1}{k-1}}$ for sufficiently large n. □

Proof of Lemma 4.16

By Chernoff-type bound, we have

$$\begin{aligned}
\mathbb{P}(E_S|\mathbf{x}^0) = \;& \mathbb{P}\left(\sum_{e:e\cap S\neq\emptyset}\log\frac{q_{|\mathbf{x}^S[e]|}(\mathbf{y}^0_e)}{q_{|\mathbf{x}^0[e]|}(\mathbf{y}^0_e)} \geq \sum_{v\in S}\log\frac{p(\mathbf{x}^0_v)}{1-p(\mathbf{x}^0_v)}\right)\\
\leq \;& \exp\left(-\max_{\alpha\in[0,1]}\left[\alpha\cdot|S|\log\frac{\rho}{1-\rho}+\sum_{e:e\cap S\neq\emptyset}D_\alpha(q_{|\mathbf{x}^S[e]|}:q_{|\mathbf{x}^0[e]|})\right]\right)\\
\leq \;& \exp\left(O(|S|)-\max_{\alpha\in[0,1]}\sum_{e:e\cap S\neq\emptyset}D_\alpha(q_{|\mathbf{x}^S[e]|}:q_{|\mathbf{x}^0[e]|})\right).
\end{aligned}$$

Let $T=\{v\in V:\mathbf{x}^0_v=1\}$. Since $\mathbf{x}^0$ is typical, we must have $|T|=(\rho\pm o(1))n$. Observe that $|\mathbf{x}^0[e]|=|e\cap T|$ and $|\mathbf{x}^S[e]|=|e\cap(S\oplus T)|$. Hence,

$$\sum_{e:e\cap S\neq\emptyset}D_\alpha(q_{|\mathbf{x}^S[e]|}:q_{|\mathbf{x}^0[e]|}) = \sum_{s,t\in\{0,\cdots,k\}}N_{s,t}D_\alpha(q_s:q_t),$$

where

$$\begin{aligned}
N_{s,t} = \;& \#(e:|e\cap T|=t,|e\cap(S\oplus T)|=s)\\
= \;& \sum_{r=0}^{k}\binom{|S\cap T|}{t-r}\binom{|S\setminus T|}{s-r}\binom{|T\setminus S|}{r}\binom{|V\setminus(S\cup T)|}{k-s-t+r}.
\end{aligned}$$

We are going to show that there are s and t such that $d_{s:t}$ is not identically zero and $N_{s,t}\gtrsim\epsilon^k n^k$ as long as $|S|>\epsilon n$. This implies Lemma 4.16 as we get

$$\begin{aligned}
\mathbb{P}(E_S|\mathbf{x}^0)\leq \;& \exp\left(O(|S|)-\max_{\alpha\in[0,1]}\sum_{s,t}N_{s,t}D_\alpha(q_s:q_t)\right)\\
\leq \;& \exp\left(O(|S|)-\Omega\left(\epsilon^k n^k\cdot\frac{\log n}{\binom{n-1}{k-1}}\right)\right)=e^{-\Omega(\epsilon^k n\log n)}.
\end{aligned}$$

Case 1: When $\epsilon n<|S|\leq\frac{n}{2}$.

Let

$$N^{(r)}_{s,t}=\binom{|S\cap T|}{t-r}\binom{|S\setminus T|}{s-r}\binom{|T\setminus S|}{r}\binom{|V\setminus(S\cup T)|}{k-s-t+r}.$$

By definition, $N_{s,t} \geq N_{s,t}^{(r)}$ for any r. In particular we get

$$\begin{aligned}
N_{0,t} \geq \ & N_{0,t}^{(0)} = \binom{|S \cap T|}{t}\binom{|V \setminus (S \cup T)|}{k-t}, \\
N_{t,0} \geq \ & N_{t,0}^{(0)} = \binom{|S \setminus T|}{t}\binom{|V \setminus (S \cup T)|}{k-t}, \\
N_{k,k-t} \geq \ & N_{k,k-t}^{(k-t)} = \binom{|S \setminus T|}{t}\binom{|T \setminus S|}{k-t}, \\
N_{k-t,k} \geq \ & N_{k-t,k}^{(k-t)} = \binom{|S \cap T|}{t}\binom{|T \setminus S|}{k-t}.
\end{aligned}$$

Note that we must have

$$\max(|S \cap T|, |S \setminus T|) \geq \frac{\epsilon n}{2} \quad \text{and} \max(|T \setminus S|, |V \setminus (S \cup T)|) \geq \frac{n}{4}$$

when $\epsilon n < |S| \leq \frac{n}{2}$. Without loss of generality, let us assume that $|S \cap T| \geq \frac{\epsilon n}{2}$ and $|V \setminus (S \cup T)| \geq \frac{n}{4}$. For any $t \in \{0, \cdots, k\}$,

$$N_{0,t} \geq \binom{\epsilon n/2}{t}\binom{n/4}{k-t} \gtrsim \epsilon^t n^k \geq \epsilon^k n^k.$$

We claim that $d_{0:t}$ is not identically zero for some t. Suppose this is not true, *i.e.*, $d_{0:t} \equiv 0$ for all t. Then, by Proposition 4.8, we must have $d_{s:t} \equiv 0$ for all s and t which contradicts the assumption that the model is not oblivious. This closes the case that $\epsilon n < |S| \leq \frac{n}{2}$.

Case 2: When $|S| \geq \frac{n}{2}$ and the model is strongly asymmetric.

We first note that if $|S| \leq (1-\epsilon)n$, then similar argument as in Case 1 shows that there exists t where $d_{0:t}$ is not identically zero and $N_{0,t} \gtrsim \epsilon^k n^k$. Suppose that $|S| \geq (1-\epsilon)n$. Then, since $|T| = (\rho \pm o(1))n$ and $\epsilon = o(1)$, we must have

$$|S \setminus T| = (1 - \rho \pm o(1))n \quad \text{and} \quad |S \cap T| = (\rho \pm o(1))n.$$

For any $t \in \{0, \cdots, k\}$, we have

$$N_{t,k-t} \geq N^{(0)}_{t,k-t} = \binom{|S \cap T|}{t}\binom{|S \setminus T|}{k-t} \gtrsim n^k = \Omega(\epsilon^k n^k),$$

since $\rho \in (0, \frac{1}{2})$. Furthermore, $d_{t:k-t}$ is not identically zero for some t as desired, since the model is strongly asymmetric.

4.4.2 Strong amplification: Proof of Theorem 4.11

As we discussed in Section 4.3.4, we are going to exploit a second moment method to prove strong amplification. Recall that $p_{fail,global} \geq p_{fail,local}$ and note that

$$\begin{aligned} p_{fail,local} = \mathbb{P}\left(\bigcup_{v\in V} E_v\right) &\geq \mathbb{P}(\mathsf{Typ})\,\mathbb{P}\left(\bigcup_{v\in V} E_v \,\middle|\, \mathsf{Typ}\right) \\ &\geq \max_{\mathbf{x}^0 \in \mathsf{Typ}} \mathbb{P}\left(\bigcup_{v\in V} E_v \,\middle|\, \mathbf{x}^0\right) - e^{-\Omega(\log^2 n)}. \end{aligned}$$

Our goal is to prove that

$$\mathbb{P}\left(\bigcup_{v\in V} E_v \,\middle|\, \mathbf{x}^0\right) \geq 1 - o(1) - \left(\sum_{v\in V} \mathbb{P}(E_v \mid \mathbf{x}^0)\right)^{-1} \tag{4.2}$$

for any typical $\mathbf{x}^0$. If $I > 1$, then the inequality trivially holds as $\mathbb{P}(E_v \mid \mathbf{x}^0) = o(n^{-1})$. For this reason, we are going to assume that $I < 1$. Here we exclude the borderline case $I = 1$ which is out of scope of this thesis.

Together with the bound $\mathbb{P}(E_v \mid \mathbf{x}^0) \geq n^{-I-o(1)}$ for typical $\mathbf{x}^0$, this implies that

$$p_{fail,local} \geq 1 - o(1) - \left(\sum_{v\in V} p_{fail,v}\right)^{-1}$$

since $p_{fail,v} = n^{-I-o(1)} - e^{-\Omega(\log^2 n)}$. On the other hand, by weak amplification we get

$$p_{fail,global} \leq n^{1-I+o(1)}.$$

Hence, $p_{fail,global} \lesssim \sum_{v \in V} p_{fail,v}$ so local-to-global amplification holds.

Lemma 4.17. *For any typical* $\mathbf{x}^0$ *and* $v \in V$*, the inequality* $\mathbb{P}(E_v \mid \mathbf{x}^0) \geq n^{-I-o(1)}$ *holds if we assume* (A_2).

Proof. Direct from the assumption (A_2). □

It remains to show that the inequality (4.2) holds. Note that we have

$$\mathbb{P}\left(\bigcup_{v\in V} E_v \,\middle|\, \mathbf{x}^0\right) \geq 1 - \left(\sum_{v\in V} \mathbb{P}(E_v \mid \mathbf{x}^0)\right)^{-1} + 1 - \frac{2\sum_{\{u,v\}\binom{V}{2}} \mathbb{P}(E_u \cap E_v \mid \mathbf{x}^0)}{\left(\sum_{v\in V} \mathbb{P}(E_v \mid \mathbf{x}^0)\right)^2},$$

by the same argument as in Section 4.3.4. Let

$$M = \max_{\{u,v\}\in\binom{V}{2}} \frac{\mathbb{P}(E_u \cap E_v \mid \mathbf{x}^0)}{\mathbb{P}(E_u \mid \mathbf{x}^0)\,\mathbb{P}(E_v \mid \mathbf{x}^0)}.$$

Then,

$$\begin{aligned} 2\sum_{\{u,v\}\in\binom{V}{2}} \mathbb{P}(E_u \cap E_v \mid \mathbf{x}^0) &\leq 2M \sum_{\{u,v\}\in\binom{V}{2}} \mathbb{P}(E_u \mid \mathbf{x}^0)\,\mathbb{P}(E_v \mid \mathbf{x}^0) \\ &\leq M\left(\sum_{v\in V} \mathbb{P}(E_v \mid \mathbf{x}^0)\right)^2, \end{aligned}$$

hence we get

$$\mathbb{P}\left(\bigcup_{v\in V} E_v \,\middle|\, \mathbf{x}^0\right) \geq 1 - \left(\sum_{v\in V} \mathbb{P}(E_v \mid \mathbf{x}^0)\right)^{-1} - (M-1).$$

If $M \leq 1 + o(1)$, then we have

$$\mathbb{P}\left(\bigcup_{v\in V} E_v \,\middle|\, \mathbf{x}^0\right) \geq 1 - o(1) - \left(\sum_{v\in V} \mathbb{P}(E_v \mid \mathbf{x}^0)\right)^{-1}$$

as desired.

Lemma 4.18. *Let* u *and* v *be distinct vertices in* V*, and let* $\mathbf{x}^0$ *be typical. Then,*

$$\frac{\mathbb{P}(E_u \cap E_v \mid \mathbf{x}^0)}{\mathbb{P}(E_u \mid \mathbf{x}^0)\,\mathbb{P}(E_v \mid \mathbf{x}^0)} \leq 1 + o(1).$$

Hence, $M \leq 1 + o(1)$.

Informally speaking, it means that event E_u and E_v are very close to being independent. We conclude the proof of Theorem 4.11 by showing Lemma 4.18 in the following section.

Proof of Lemma 4.18: Almost independence of E_u and E_v

Recall that E_u is the event that

$$\sum_{e \ni u} \log \frac{Q(\mathbf{y}_e^0|\mathbf{x}^u[e])}{Q(\mathbf{x}_e^0|\mathbf{x}^0[e])} \geq \log\left(\frac{p(\mathbf{x}_u^0)}{1-p(\mathbf{x}_u^0)}\right)$$

holds, and similar for E_v. Let us define X_u, X_v, Y_u and Y_v be random variables depending on $\mathbf{y}^0$ where

$$\begin{aligned} X_u &= \sum_{e \cap \{u,v\} = \{u\}} \log \frac{Q(\mathbf{y}_e^0|\mathbf{x}^u[e])}{Q(\mathbf{y}_e^0|\mathbf{x}^0[e])}, & Y_u &= \sum_{e \supseteq \{u,v\}} \log \frac{Q(\mathbf{y}_e^0|\mathbf{x}^u[e])}{Q(\mathbf{y}_e^0|\mathbf{x}^0[e])}, \\ X_v &= \sum_{e \cap \{u,v\} = \{v\}} \log \frac{Q(\mathbf{y}_e^0|\mathbf{x}^v[e])}{Q(\mathbf{y}_e^0|\mathbf{x}^0[e])}, & Y_v &= \sum_{e \supseteq \{u,v\}} \log \frac{Q(\mathbf{y}_e^0|\mathbf{x}^v[e])}{Q(\mathbf{y}_e^0|\mathbf{x}^0[e])}. \end{aligned}$$

Let $c_u = \log \frac{p(\mathbf{x}_u^0)}{1-p(\mathbf{x}_u^0)}$ and $c_v = \log \frac{p(\mathbf{x}_v^0)}{1-p(\mathbf{x}_v^0)}$. Then, by definition

$$E_u \Leftrightarrow X_u + Y_u \geq c_u \quad \text{and} \quad E_v \Leftrightarrow X_v + Y_v \geq c_v.$$

Let $\mathcal{F}_u = \{e : e \cap \{u,v\} = \{u\}\}$, $\mathcal{F}_v = \{e : e \cap \{u,v\} = \{v\}\}$ and $\mathcal{F}_{uv} = \{e : e \supseteq \{u\}\}$. Clearly $\mathcal{F}_u$, $\mathcal{F}_v$ and $\mathcal{F}_{uv}$ are mutually disjoint and so X_u, X_v and $\{Y_u, Y_v\}$ are mutually independent (but Y_u and Y_v are not independent).

Since E_u and E_v are independent when conditioned on the value of $\mathbf{y}^0[\mathcal{F}_{uv}]$, we have

$$\begin{aligned} \mathbb{P}(E_u \cap E_v \mid \mathbf{x}^0) = \underset{\mathbf{y}^0[\mathcal{F}_{uv}]}{\mathbb{E}} &\left[\mathbb{P}\left(X_u + Y_u \geq c_u \mid \mathbf{y}^0[\mathcal{F}_{uv}]\right)\right. \\ &\left.\cdot \mathbb{P}\left(X_v + Y_v \geq c_v \mid \mathbf{y}^0[\mathcal{F}_{uv}]\right)\right]. \end{aligned}$$

Let $\eta = \frac{1}{(\log n)^2}$ and let Good be the event that $|Y_u| \leq \eta$ and $|Y_v| \leq \eta$. Note that if

$|Y_u| \leq \eta$, then

$$\mathbb{P}(X_u \geq c_u - \eta) \;\leq\; \mathbb{P}(X_u + Y_u \geq c_u \mid \mathbf{y}^0[\mathcal{F}_{u,v}]) \;\leq\; \mathbb{P}(X_u \geq c_u + \eta),$$

and vice versa for v. Hence,

$$\begin{aligned}
\mathbb{P}(E_u \cap E_v \mid \mathsf{Good}, \mathbf{x}^0) \leq &\quad \mathbb{P}(X_u \geq c_u - \eta)\,\mathbb{P}(X_v \geq c_v - \eta),\\
\mathbb{P}(E_u \mid \mathbf{x}^0) \geq &\quad \mathbb{P}(\mathsf{Good})\,\mathbb{P}(X_u \geq c_u + \eta),\\
\mathbb{P}(E_v \mid \mathbf{x}^0) \geq &\quad \mathbb{P}(\mathsf{Good})\,\mathbb{P}(X_v \geq c_v + \eta).
\end{aligned}$$

By the assumption (A_2), we have

$$\mathbb{P}(X_u \geq c_u + \eta) \geq n^{-I-o(1)}$$

and

$$\frac{\mathbb{P}(X_u \geq c_u - \eta)}{\mathbb{P}(X_u \geq c_u + \eta)} = 1 + o(1)$$

since $\eta = o((\log n)^{-1})$. Hence,

$$\frac{\mathbb{P}(E_u \cap E_v \mid \mathbf{x}^0)}{\mathbb{P}(E_u \mid \mathbf{x}^0)\,\mathbb{P}(E_v \mid \mathbf{x}^0)} \leq \quad \frac{1 + o(1)}{\mathbb{P}(\mathsf{Good})} \cdot \frac{\mathbb{P}(E_u \cap E_v \cap \neg\mathsf{Good} \mid \mathbf{x}^0)}{\mathbb{P}(\mathsf{Good})^2 n^{-2I-o(1)}}.$$

It remains to show that $\mathbb{P}(\mathsf{Good} \mid \mathbf{x}^0) = 1 - o(1)$ and

$$\frac{\mathbb{P}(E_u \cap E_v \cap \neg\mathsf{Good} \mid \mathbf{x}^0)}{n^{-2I+o(1)}} = o(1).$$

Let $\alpha \in (0,1)$ be a maximizer of $I(\alpha)$. By definition, $I(\alpha)$ is equal to I. We have

$$\begin{aligned}
\mathbb{P}(X_u + Y_u \geq c_u \mid (\mathbf{y}_e)_{e \in \mathcal{F}_{uv}}) \leq &\quad \exp\left(-\alpha(c_u - Y_u) + \log \mathbb{E}\, e^{\alpha X_u}\right)\\
= &\quad e^{\alpha Y_u} \exp\left(-\alpha c_u - I(\alpha)\log n + o(\log n)\right)\\
= &\quad e^{\alpha Y_u} \cdot n^{-I+o(1)}.
\end{aligned}$$

Thus we get

$$\begin{aligned}\mathbb{P}(E_u \cap E_v \cap \neg\mathsf{Good} \mid \mathbf{x}^0) \le\; & n^{-2I+o(1)} \mathop{\mathbb{E}}_{(\mathbf{y}_e^0)_{e\in\mathcal{F}_{uv}}} \left[\mathbf{1}\{\neg\mathsf{Good}\} e^{\alpha(Y_u+Y_v)}\right] \\ \le\; & n^{-2I+o(1)}\, \mathbb{P}(\neg\mathsf{Good})^{1/2} \left(\mathop{\mathbb{E}}_{(\mathbf{y}_e^0)_{e\in\mathcal{F}_{uv}}} e^{2\alpha(Y_u+Y_v)}\right)^{1/2},\end{aligned}$$

where the second inequality follows from Cauchy-Schwarz inequality.

Claim. $\mathbb{E}_{(\mathbf{y}_e^0)_{e\in\mathcal{F}_{uv}}} e^{2\alpha(Y_u+Y_v)} = 1 + o(1)$.

Proof. Expanding $e^{2\alpha(Y_u+Y_v)}$, we get

$$\mathbb{E}\, e^{2\alpha(Y_u+Y_v)} = \prod_{e\in\mathcal{F}_{uv}} \mathbb{E} \left(\frac{q_{|\mathbf{x}^u[e]|}}{q_{|\mathbf{x}^0[e]|}}\right)^{2\alpha} \left(\frac{q_{|\mathbf{x}^v[e]|}}{q_{|\mathbf{x}^0[e]|}}\right)^{2\alpha}.$$

Note that Cauchy-Schwarz inequality implies that for any r, s and t

$$\begin{aligned}\mathop{\mathbb{E}}_{Y\sim q_t} \left(\frac{q_r}{q_t}(Y)\right)^{2\alpha} \left(\frac{q_s}{q_t}(Y)\right)^{2\alpha} \le\; & \left(\mathop{\mathbb{E}}_{Y\sim q_t} \left(\frac{q_r}{q_t}(Y)\right)^{4\alpha}\right)^{1/2} \left(\mathop{\mathbb{E}}_{Y\sim q_t} \left(\frac{q_s}{q_t}(Y)\right)^{4\alpha}\right)^{1/2} \\ =\; & e^{-\frac{1}{2}D_{4\alpha}(q_r:q_t) - \frac{1}{2}D_{4\alpha}(q_s:q_t)}\end{aligned}$$

holds. Moreover, the right-hand side is bounded by $e^{c\frac{\log n}{\binom{n-1}{k-1}}}$ for some constant $c > 0$. Thus we get

$$\mathbb{E}\, e^{2\alpha(Y_u+Y_v)} \le \exp\left(c\frac{\log n}{\binom{n-1}{k-1}} \cdot |\mathcal{F}_{uv}|\right) = \exp\left(O\left(\frac{\log n}{n}\right)\right)$$

which is $1 + o(1)$. ☐

Claim. $\mathbb{P}(\neg\mathsf{Good}) = n^{-1+o(1)}$.

Proof. Note that

$$\begin{aligned}\mathbb{P}(\neg\mathsf{Good}) \le\; & \mathbb{P}(|Y_u| > \eta) + \mathbb{P}(|Y_v| > \eta) \\ \le\; & \eta^{-2}\, \mathbb{E}\, Y_u^2 + \eta^{-2}\, \mathbb{E}\, Y_v^2,\end{aligned}$$

by Markov's inequality. The second moment of Y_u can be expressed as

$$\begin{aligned}\mathbb{E}\, Y_u^2 = &\ \frac{d^2}{d\alpha^2} \log \mathbb{E}\, e^{\alpha Y_u}\bigg|_{\alpha=0} \\ = &\ \frac{d^2}{d\alpha^2} \sum_{e \in \mathcal{F}_{uv}} D_\alpha(q_{|\mathbf{x}^u[e]|} : q_{|\mathbf{x}^0[e]|})\bigg|_{\alpha=0},\end{aligned}$$

which is bounded by

$$|\mathcal{F}_{uv}| \cdot \max_{s,t} \left| \left(\frac{d^2}{d\alpha^2} D_\alpha(q_s : q_t)\bigg|_{\alpha=0} \right) \right| \lesssim n^{k-2} \cdot \frac{\log n}{n^{k-1}} = \frac{\log n}{n}.$$

Hence,

$$\mathbb{P}(\neg\mathsf{Good}) \lesssim \frac{\log n}{\eta^2 n} = \frac{\log^5 n}{n} = n^{-1+o(1)}$$

as desired. $\square$

In summary, we have

$$\frac{\mathbb{P}(E_u \cap E_v \cap \neg\mathsf{Good} \mid \mathbf{x}^0)}{n^{-2I-o(1)}} \lesssim n^{-1+o(1)},$$

so together with $\mathbb{P}(\mathsf{Good}) = 1 - o(1)$ we get

$$\frac{\mathbb{P}(E_u \cap E_v \mid \mathbf{x}^0)}{\mathbb{P}(E_u \mid \mathbf{x}^0)\, \mathbb{P}(E_v \mid \mathbf{x}^0)} \leq 1 + o(1).$$

4.5 Applications

In this section, we apply the strong amplification result (Theorem 4.11) to concrete examples. As a result, we obtain exact statistical thresholds of exact recovery in spiked tensor models and k-HSBMs, which includes the results we have presented in Chapter 2 and Chapter 3. We also apply Theorem 4.11 to the hypergraph version of binary censored block model and reproduce the result of [14].

4.5.1 Spiked tensor models

For illustration, we will focus on the single-spiked k-tensor model and the bisection-spiked k-tensor model in this section (See Section 2.2 for the definition of two models). In the language of graphical channel, we consider the model with parameters

$$\rho = \frac{1}{2} \quad \text{and} \quad Q(\cdot|z_1, \cdots, z_k) \sim N\left(\mu_{|z|}, \sigma^2/k!\right),$$

where $\sigma = \sigma_n$ is the noise scaling factor and $(\mu_t : t = 0, \cdots, k)$ is defined as follows in each model:

- (Single-spiked model) $\mu_t = (-1)^t$
- (Bisection-spiked model) $\mu_t = 1$ if $t \in \{0, k\}$, and $\mu_t = 0$ otherwise.

We claim that we must have $\sigma = \Theta(\frac{n^{\frac{k-1}{2}}}{\sqrt{\log n}})$ for the assumption $(\mathsf{A}_1)$to hold. To see this, we need the following proposition.

Proposition 4.19. *Let $\nu_0 = N(\mu_0, \sigma^2)$ and $\nu_1 = N(\mu_1, \sigma^2)$. Then, the Chernoff α-divergence D_α from ν_0 to ν_1 is equal to*

$$D_\alpha(\nu_1 : \nu_0) = \alpha(1-\alpha) \cdot \frac{(\mu_0 - \mu_1)^2}{2\sigma^2}.$$

Proof. Direct computation. □

Thus,

$$D_\alpha(q_s : q_t) = \alpha(1-\alpha)\frac{k!}{2\sigma^2}(\mu_s - \mu_t)^2,$$

and when $\sigma = \tau\sqrt{\frac{n^{k-1}}{2\log n}}$, we get

$$\begin{aligned} d_{s:t}(\alpha) &= \lim_{n\to\infty} \frac{\binom{n-1}{k-1}}{\log n} D_\alpha(q_s : q_t) \\ &= \alpha(1-\alpha)(\mu_s - \mu_t)^2 \lim_{n\to\infty} \frac{\binom{n-1}{k-1}}{\log n} \cdot \frac{k! \log n}{\tau^2 n^{k-1}} \\ &= k\alpha(1-\alpha)\frac{(\mu_s - \mu_t)^2}{\tau^2}. \end{aligned}$$

Hence, (A_1)holds assuming that $\sigma = \tau\sqrt{\frac{n^{k-1}}{2\log n}}$.

Recall that

$$I(\alpha) = \sum_{t=0}^{k-1} \binom{k-1}{t} \rho^{k-1-t}(1-\rho)^t d_{t:t+1}(\alpha).$$

Hence, we get

$$I_{single}(\alpha) = k\alpha(1-\alpha)\frac{4}{\tau^2} \quad \text{and} \quad I_{single} = \frac{k}{\tau^2}$$

for the single-spiked model, and

$$I_{bisec}(\alpha) = \frac{k}{2^{k-2}}\alpha(1-\alpha)\frac{1}{\tau^2} \quad \text{and} \quad I_{bisec} = \frac{k}{2^k\tau^2}$$

for the bisection-spiked model. Assuming that (A_2)holds as well, we get the following corollary, which is a restatement of Corollary 2.4 and Corollary 2.6.

Corollary 4.20. *The threshold for exact recovery in the single-spiked model is at*

$$\sigma = \sqrt{k \cdot \frac{n^{k-1}}{2\log n}},$$

and the threshold for exact recovery in the bisection-spiked model is at

$$\sigma = \sqrt{\frac{k}{2^k} \cdot \frac{n^{k-1}}{2\log n}}.$$

For completeness, let us argue that the assumption (A_2)holds in those two models (and in general, graphical channel models with Gaussian kernel).

Informal argument. We note that

$$\eta_{s:t} \sim \log\left(\frac{q_s(Y)}{q_t(Y)}\right) \quad \text{where} \quad Y \sim q_t$$

is also a Gaussian with the same variance $\sigma^2/k!$, as the collection $\{N(\mu, \sigma^2/k!) : \mu \in$

$\mathbb{R}\}$ forms an exponential family[2]. It implies that

$$L = \sum_{t=0}^{k-1} \sum_{i=1}^{c_t\binom{n-1}{k-1}} X_i^{(t)}$$

where $X_i^{(t)} \sim \eta_{t+1:t}$ will be again a Gaussian distribution with mean $\Theta(\log n)$ and variance $\Theta(\log n)$. We can directly estimate the tail probability $\mathbb{P}(L \geq t)$ by

$$\mathbb{P}(L \geq t) = \mathbb{P}\left(\frac{L - \mathbb{E}\,L}{\sqrt{\mathbb{E}(L - \mathbb{E}\,L)^2}} \geq \frac{t - \mathbb{E}\,L}{\sqrt{\mathbb{E}(L - \mathbb{E}\,L)^2}}\right) = \Phi\left(\frac{t - \mathbb{E}\,L}{\sqrt{\mathbb{E}(L - \mathbb{E}\,L)^2}}\right),$$

which can be tightly estimated as $n^{-\epsilon}$ with the certain value $\epsilon > 0$ determined by the parameters. The assumption (A_2) follows from the direct calculation, which we omit the details here. $\square$

4.5.2 k-HSBMs with two communities

Note that the stochastic block model for k-uniform hypergraph with two communities can be described as a graphical channel model with the kernel Q where

$$Q(\cdot \mid z_1, \cdots, z_k) \sim \mathrm{Ber}(p_{|z|})$$

with parameters $p_0, \cdots, p_k \in [0, 1]$. For illustration, let us restrict our focus to the model $\mathsf{HSBM}(n, p, q; k)$ which appears in Chapter 3. In this case, we have

$$\rho = \frac{1}{2} \quad \text{and} \quad p_t = \begin{cases} p & \text{if } t \in \{0, k\} \\ q & \text{otherwise.} \end{cases}$$

Let us compute the α-divergence between $\mathrm{Ber}(p)$ and $\mathrm{Ber}(q)$. When p and q are

[2]For a nice overview on exponential family distributions, see [81].

$o(1)$, we have

$$\begin{aligned} D_\alpha(\mathrm{Ber}(p) : \mathrm{Ber}(q)) = & \ -\log\left(p^\alpha q^{1-\alpha} + (1-p)^\alpha(1-q)^{1-\alpha}\right) \\ = & \ -\log\left(1 - \alpha p - (1-\alpha)q + p^\alpha q^{1-\alpha} + o(p+q)\right) \\ \approx & \ \alpha p + (1-\alpha)q - p^\alpha q^{1-\alpha}. \end{aligned}$$

In particular, when $p = \frac{a \log n}{\binom{n-1}{k-1}}$ and $q = \frac{b \log n}{\binom{n-1}{k-1}}$ then we get

$$D_\alpha(\mathrm{Ber}(p) : \mathrm{Ber}(q)) \approx \frac{\log n}{\binom{n-1}{k-1}} \left(\alpha a + (1-\alpha)b - a^\alpha b^{1-\alpha}\right).$$

Hence,

$$d_{s:t}(\alpha) = \begin{cases} \alpha a + (1-\alpha)b - a^\alpha b^{1-\alpha} & (s,t) = (0,1) \text{ or } (s,t) = (k, k-1), \\ (1-\alpha)a + \alpha b - a^{1-\alpha} b^\alpha & (s,t) = (1,0) \text{ or } (s,t) = (k-1, k), \\ 0 & \text{otherwise.} \end{cases}$$

and the assumption (A_1) holds.

We get

$$\begin{aligned} I(\alpha) = & \ \frac{1}{2^{k-1}} \left(d_{1:0}(\alpha) + d_{k:k-1}(\alpha)\right) \\ = & \ \frac{1}{2^{k-1}} \left(a + b - a^\alpha b^{1-\alpha} - a^{1-\alpha} b^\alpha\right), \end{aligned}$$

and

$$\begin{aligned} I = & \ \frac{1}{2^{k-1}} \max_\alpha \left(a + b - a^\alpha b^{1-\alpha} - a^{1-\alpha} b^\alpha\right) \\ = & \ \frac{1}{2^{k-1}} (a + b - 2\sqrt{ab}) \\ = & \ \frac{1}{2^{k-1}} \left(\sqrt{a} - \sqrt{b}\right)^2. \end{aligned}$$

Assuming that (A_2) holds, we recover the theorem 3.1.

Corollary 4.21. *The threshold for exact recovery in the* $\mathsf{HSBM}(n, p, q; k)$ *where* $p = \frac{a \log n}{\binom{n-1}{k=1}}$ *and* $q = \frac{b \log n}{\binom{n-1}{k-1}}$ *is at*

$$\frac{1}{2^{k-1}} \left(\sqrt{a} - \sqrt{b}\right)^2 = 1.$$

Let us justify why the assumption (A_2) holds in graphical channel models with

Bernoulli kernels. We first remark that when $q_s \sim \mathrm{Ber}(p)$ and $q_t \sim \mathrm{Ber}(q)$, the distribution

$$\eta_{s:t} \sim \log\left(\frac{q_s(Y)}{q_t(Y)}\right) \quad \text{where} \quad Y \sim q_t$$

can be written as

$$\eta_{s:t} = \log\left(\frac{1-p}{1-q}\right) + \log\left(\frac{p(1-q)}{q(1-p)}\right) \cdot \mathrm{Ber}(q).$$

Hence, in the sum

$$L = \sum_{t=0}^{k-1} \sum_{i=1}^{c_t\binom{n-1}{k-1}} X_i^{(t)},$$

each inner sum $\sum_{i=1}^{c_t\binom{n-1}{k-1}} X_i^{(t)}$ is an affine transformation of the sum of i.i.d. Bernoulli variables.

Theorem 4.22 (Le Cam's Theorem). *Let $B_1, B_2, \cdots, B_N$ be independent Bernoulli variables. Let $S_N = \sum_{i=1}^N B_i$, and let X be a Poisson random variable with mean $\mathbb{E}\, S_N$. Then,*

$$\sum_{k=0}^{\infty} |\mathbb{P}(S_N = k) - \mathbb{P}(X = k)| < \sum_{i=1}^{N} (\mathbb{E}\, B_i)^2.$$

In particular, when $X_i^{(t)} \sim \mathrm{Ber}(p)$ with $p = \frac{a \log n}{\binom{n-1}{k-1}}$, total variation distance between the sum

$$\sum_{i=1}^{c_t\binom{n-1}{k-1}} X_i^{(t)} \sim \mathrm{Bin}(c_t\tbinom{n-1}{k-1}, \frac{a\log n}{\binom{n-1}{k-1}})$$

and the Poisson distribution $\mathrm{Poisson}(c_t \cdot a \log n)$ is at most $n^{-(k-1)}$. Then, assumption (A_2) follows from direct computation on the tail probability of Poisson random variables.

4.5.3 Censored block model for k-uniform hypergraphs

Let us consider the binary censored block model for k-uniform hypergraphs (HCBM) which we briefly discussed in Section 1.2.3. We specifically consider the model suggested in [14].

An HCBM can be described as a graphical channel model with parameters

$$\rho = \frac{1}{2} \quad \text{and} \quad Q(\cdot \mid z_1, \cdots, z_k) = \begin{cases} 0 & \text{with probability } 1-p \\ \mu(|z|) & \text{with probability } p(1-\theta) \\ -\mu(|z|) & \text{with probability } p\theta, \end{cases}$$

where $p \in [0,1]$ stands for the edge probability, $\theta \in (0, 1/2]$ stands for the probability that a single measurement is corrupted, and $\mu : \{0, \cdots, k\} \to \{\pm 1\}$ stands for the binary measurement for the hyperedges. In particular, we are interested in the following two types of the measurements:

- **Homogeneity measurement.** $\mu(t) = 1$ if $t \in \{0, k\}$ and $\mu(t) = -1$ otherwise.

- **Parity measurement.** $\mu(t) = (-1)^t$.

The terms *homogeneity* and *parity* are borrowed from [14].

Suppose that $p = \frac{c \log n}{\binom{n-1}{k-1}}$ for some $c > 0$. Then, we get the following.

Proposition 4.23. *If $\mu(s) = \mu(t)$, then $D_\alpha(q_s : q_t) = 0$. Otherwise,*

$$D_\alpha(q_s : q_t) = \frac{\log n}{\binom{n-1}{k-1}} \cdot c\left(1 - \theta^{1-\alpha}(1-\theta)^\alpha - (1-\theta)^{1-\alpha}\theta^\alpha\right).$$

Proof. Direct calculation. □

For brevity, let $f(\alpha)$ be

$$f(\alpha) := c\left(1 - \theta^{1-\alpha}(1-\theta)^\alpha - (1-\theta)^{1-\alpha}\theta^\alpha\right).$$

Then, we get

$$d_{s:t}(\alpha) = \begin{cases} f(\alpha) & \text{if } \mu(s) \neq \mu(t) \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, (A_1)holds in this parameter regime.

Now, note that

$$I_{hom}(\alpha) = \frac{1}{2^{k-1}} \cdot 2f(\alpha) \quad \text{and} \quad I_{hom} = \frac{1}{2^{k-2}} \max_{\alpha} f(\alpha)$$

in the case of homogeneity measurement, and

$$I_{par}(\alpha) = f(\alpha) \quad \text{and} \quad I_{par} = \max_{\alpha} f(\alpha)$$

in the case of parity measurement. Since $f(\alpha)$ is maximized at $\alpha = 1/2$, we get

$$\max_{\alpha} f(\alpha) = c(1 - 2\sqrt{\theta(1-\theta)}) = c\left(\sqrt{\theta} - \sqrt{1-\theta}\right)^2.$$

By assuming that (A_2) holds as well, we get the following corollary.

Corollary 4.24. *The threshold for exact recovery in the homogeniety measurement HCBM model is at*

$$c\left(\sqrt{\theta} - \sqrt{1-\theta}\right)^2 = 2^{k-2}.$$

The threshold for exact recovery in the parity measurement HCBM model is at

$$c\left(\sqrt{\theta} - \sqrt{1-\theta}\right)^2 = 1.$$

Bibliography

[1] Emmanuel Abbe. Community detection and stochastic block models: Recent developments. *Journal of Machine Learning Research*, 18(177):1–86, 2018.

[2] Emmanuel Abbe, Afonso S Bandeira, Annina Bracher, and Amit Singer. Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery. *IEEE Transactions on Network Science and Engineering*, 1(1):10–22, 2014.

[3] Emmanuel Abbe, Afonso S Bandeira, Annina Bracher, and Amit Singer. Linear inverse problems on erdős-rényi graphs: Information-theoretic limits and efficient recovery. In *Information Theory (ISIT), 2014 IEEE International Symposium on*, pages 1251–1255. IEEE, 2014.

[4] Emmanuel Abbe, Afonso S Bandeira, and Georgina Hall. Exact recovery in the stochastic block model. *IEEE Transactions on Information Theory*, 62(1):471–487, 2016.

[5] Emmanuel Abbe and Andrea Montanari. Conditional random fields, planted constraint satisfaction and entropy concentration. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 332–346. Springer, 2013.

[6] Emmanuel Abbe and Colin Sandon. Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 670–688. IEEE, 2015.

[7] Emmanuel Abbe and Colin Sandon. Detection in the stochastic block model with multiple clusters: proof of the achievability conjectures, acyclic bp, and the information-computation gap. *arXiv preprint arXiv:1512.09080*, 2015.

[8] Emmanuel Abbe and Colin Sandon. Recovering communities in the general stochastic block model without knowing the parameters. In *Advances in neural information processing systems*, pages 676–684, 2015.

[9] Emmanuel Abbe and Colin Sandon. Achieving the ks threshold in the general stochastic block model with linearized acyclic belief propagation. In *Advances in Neural Information Processing Systems*, pages 1334–1342, 2016.

[10] Dimitris Achlioptas, Assaf Naor, and Yuval Peres. Rigorous location of phase transitions in hard optimization problems. *Nature*, 435(7043):759, 2005.

[11] Naman Agarwal, Afonso S Bandeira, Konstantinos Koiliaris, and Alexandra Kolla. Multisection in the stochastic block model using semidefinite programming. In *Compressed Sensing and its Applications*, pages 125–162. Springer, 2017.

[12] Sameer Agarwal, Kristin Branson, and Serge Belongie. Higher order learning with graphs. In *Proceedings of the 23rd international conference on Machine learning*, pages 17–24. ACM, 2006.

[13] Sameer Agarwal, Jongwoo Lim, Lihi Zelnik-Manor, Pietro Perona, David Kriegman, and Serge Belongie. Beyond pairwise clustering. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 838–845. IEEE, 2005.

[14] Kwangjun Ahn, Kangwook Lee, and Changho Suh. Community recovery in hypergraphs. In *Communication, Control, and Computing (Allerton), 2016 54th Annual Allerton Conference on*, pages 657–663. IEEE, 2016.

[15] Kwangjun Ahn, Kangwook Lee, and Changho Suh. Hypergraph spectral clustering in the weighted stochastic block model. *arXiv preprint arXiv:1805.08956*, 2018.

[16] Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191. American Mathematical Soc., 2007.

[17] Animashree Anandkumar, Rong Ge, Daniel Hsu, and Sham M Kakade. A tensor approach to learning mixed membership community models. *The Journal of Machine Learning Research*, 15(1):2239–2312, 2014.

[18] Animashree Anandkumar, Rong Ge, Daniel Hsu, Sham M Kakade, and Matus Telgarsky. Tensor decompositions for learning latent variable models. *The Journal of Machine Learning Research*, 15(1):2773–2832, 2014.

[19] Montanari Andrea. Estimating random variables from random sparse observations. *European Transactions on Telecommunications*, 19(4):385–403, 2008.

[20] Maria Chiara Angelini, Francesco Caltagirone, Florent Krzakala, and Lenka Zdeborová. Spectral detection on sparse hypergraphs. In *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on*, pages 66–73. IEEE, 2015.

[21] Afonso S Bandeira. Random laplacian matrices and convex relaxations. *Foundations of Computational Mathematics*, 18(2):345–379, 2018.

[22] Afonso S Bandeira and Ramon van Handel. Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Annals of Probability*, 44(4):2479–2506, 2016.

[23] Boaz Barak and David Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. In *Proceedings of the International Congress of Mathematicians*, 2014.

[24] Wolfgang Barthel, Alexander K Hartmann, Michele Leone, Federico Ricci-Tersenghi, Martin Weigt, and Riccardo Zecchina. Hiding solutions in random satisfiability problems: A statistical mechanics approach. *Physical review letters*, 88(18):188701, 2002.

[25] Michael Behrisch, Amin Coja-Oghlan, and Mihyun Kang. The order of the giant component of random hypergraphs. *Random Structures & Algorithms*, 36(2):149–184, 2010.

[26] Vijay Bhattiprolu, Venkatesan Guruswami, and Euiwoong Lee. Certifying Random Polynomials over the Unit Sphere via Sum of Squares Hierarchy. 2016.

[27] Béla Bollobás. Random graphs. In *Modern graph theory*, pages 215–252. Springer, 1998.

[28] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[29] Irineo Cabreros, Emmanuel Abbe, and Aristotelis Tsirigos. Detecting community structures in hi-c genomic data. In *Information Science and Systems (CISS), 2016 Annual Conference on*, pages 584–589. IEEE, 2016.

[30] T Tony Cai, Tengyuan Liang, and Alexander Rakhlin. Inference via message passing on partially labeled stochastic block models. *arXiv preprint arXiv:1603.06923*, 2016.

[31] Narasinga Rao Chaganty and Jayaram Sethuraman. Strong large deviation and local limit theorems. *The Annals of Probability*, pages 1671–1690, 1993.

[32] Jingchun Chen and Bo Yuan. Detecting functional modules in the yeast protein–protein interaction network. *Bioinformatics*, 22(18):2283–2290, 2006.

[33] Yudong Chen and Jiaming Xu. Statistical-computational tradeoffs in planted problems and submatrix localization with a growing number of clusters and submatrices. *Journal of Machine Learning Research*, 17(27):1–57, 2016.

[34] I Chien, Chung-Yi Lin, I Wang, et al. On the minimax misclassification ratio of hypergraph community detection. *arXiv preprint arXiv:1802.00926*, 2018.

[35] Amin Coja-Oghlan, Cristopher Moore, and Vishal Sanwalani. Counting connected graphs and hypergraphs via the probabilistic method. *Random Structures & Algorithms*, 31(3):288–329, 2007.

[36] Oliver Cooley, Mihyun Kang, and Christoph Koch. Evolution of high-order connected components in random hypergraphs. *Electronic Notes in Discrete Mathematics*, 49:569–575, 2015.

[37] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[38] Aurelien Decelle, Florent Krzakala, Cristopher Moore, and Lenka Zdeborová. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Physical Review E*, 84(6):066106, 2011.

[39] A Dembo and O Zeitouni. Large deviations techniques and applications. 1998.

[40] Yash Deshpande, Emmanuel Abbe, and Andrea Montanari. Asymptotic mutual information for the binary stochastic block model. In *Information Theory (ISIT), 2016 IEEE International Symposium on*, pages 185–189. IEEE, 2016.

[41] Olivier Duchenne, Francis Bach, In-So Kweon, and Jean Ponce. A tensor-based algorithm for high-order graph matching. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2383–2395, 2011.

[42] Peter Elias. Coding for noisy channels. *IRE Conv. Rec.*, 3:37–46, 1955.

[43] Uriel Feige, Elchanan Mossel, and Dan Vilenchik. Complete convergence of message passing algorithms for some satisfiability problems. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 339–350. Springer, 2006.

[44] Uriel Feige and Eran Ofek. Spectral techniques applied to sparse random graphs. *Random Structures & Algorithms*, 27(2):251–275, 2005.

[45] Delphine Féral and Sandrine Péché. The largest eigenvalue of rank one deformation of large wigner matrices. *Communications in mathematical physics*, 272(1):185–228, 2007.

[46] Laura Florescu and Will Perkins. Spectral thresholds in the bipartite stochastic block model. *arXiv preprint arXiv:1506.06737*, 2015.

[47] Ehud Friedgut, Jean Bourgain, et al. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American mathematical Society*, 12(4):1017–1054, 1999.

[48] Gourab Ghoshal, Vinko Zlatić, Guido Caldarelli, and MEJ Newman. Random hypergraphs and their applications. *Physical Review E*, 79(6):066118, 2009.

[49] Debarghya Ghoshdastidar and Ambedkar Dukkipati. Consistency of spectral partitioning of uniform hypergraphs under planted partition model. In *Advances in Neural Information Processing Systems*, pages 397–405, 2014.

[50] Debarghya Ghoshdastidar and Ambedkar Dukkipati. A provable generalized tensor spectral method for uniform hypergraph partitioning. In *International Conference on Machine Learning*, pages 400–409, 2015.

[51] Debarghya Ghoshdastidar and Ambedkar Dukkipati. Spectral clustering using multilinear svd: Analysis, approximations and applications. In *AAAI*, pages 2610–2616, 2015.

[52] Debarghya Ghoshdastidar, Ambedkar Dukkipati, et al. Consistency of spectral hypergraph partitioning under planted partition model. *The Annals of Statistics*, 45(1):289–315, 2017.

[53] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[54] Venu Madhav Govindu. A tensor decomposition for geometric grouping and segmentation. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 1150–1157. IEEE, 2005.

[55] Dongning Guo and Sergio Verdú. Randomly spread cdma: Asymptotics via statistical physics. *IEEE Transactions on Information Theory*, 51(6):1983–2010, 2005.

[56] Bruce Hajek, Yihong Wu, and Jiaming Xu. Achieving exact cluster recovery threshold via semidefinite programming. *IEEE Transactions on Information Theory*, 62(5):2788–2797, 2016.

[57] Bruce Hajek, Yihong Wu, and Jiaming Xu. Achieving exact cluster recovery threshold via semidefinite programming: Extensions. *IEEE Transactions on Information Theory*, 62(10):5918–5937, 2016.

[58] Sam Hopkins. Tensor principal component analysis via sum-of-squares proofs. *arXiv:1507.03269*, 40:1–51, 2015.

[59] Adel Javanmard, Andrea Montanari, and Federico Ricci-Tersenghi. Phase transitions in semidefinite relaxations. *Proceedings of the National Academy of Sciences*, 113(16):E2218–E2223, 2016.

[60] David Karger, Rajeev Motwani, and Madhu Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM (JACM)*, 45(2):246–265, 1998.

[61] Chiheon Kim, Afonso S Bandeira, and Michel X Goemans. Community detection in hypergraphs, spiked tensor models, and sum-of-squares. In *Sampling Theory and Applications (SampTA), 2017 International Conference on*, pages 124–128. IEEE, 2017.

[62] Chiheon Kim, Afonso S Bandeira, and Michel X Goemans. Stochastic block model for hypergraphs: Statistical limits and a semidefinite programming approach. *arXiv preprint arXiv:1807.02884*, 2018.

[63] Shrinivas Kudekar and Nicolas Macris. Sharp bounds for map decoding of general irregular ldpc codes. In *Information Theory, 2006 IEEE International Symposium on*, pages 2259–2263. IEEE, 2006.

[64] K Raj Kumar, Payam Pakzad, Amir Hesam Salavati, and Amin Shokrollahi. Phase transitions for mutual information. In *Turbo Codes and Iterative Information Processing (ISTC), 2010 6th International Symposium on*, pages 137–141. IEEE, 2010.

[65] Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[66] Thibault Lesieur, Florent Krzakala, and Lenka Zdeborová. Mmse of probabilistic low-rank matrix estimation: Universality with respect to the output channel. In *Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on*, pages 680–687. IEEE, 2015.

[67] Thibault Lesieur, Florent Krzakala, and Lenka Zdeborová. Constrained low-rank matrix estimation: Phase transitions, approximate message passing and applications. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(7):073403, 2017.

[68] Thibault Lesieur, Léo Miolane, Marc Lelarge, Florent Krzakala, and Lenka Zdeborová. Statistical and computational phase transitions in spiked tensor estimation. *arXiv preprint arXiv:1701.08010*, 2017.

[69] Chung-Yi Lin, I Eli Chien, and I-Hsiang Wang. On the fundamental statistical limit of community detection in random hypergraphs. In *Information Theory (ISIT), 2017 IEEE International Symposium on*, pages 2178–2182. IEEE, 2017.

[70] Greg Linden, Brent Smith, and Jeremy York. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*, 7(1):76–80, 2003.

[71] Edward M Marcotte, Matteo Pellegrini, Ho-Leung Ng, Danny W Rice, Todd O Yeates, and David Eisenberg. Detecting protein function and protein-protein interactions from genome sequences. *Science*, 285(5428):751–753, 1999.

[72] Laurent Massoulié. Community detection thresholds and the weak ramanujan property. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 694–703. ACM, 2014.

[73] Tom Michoel and Bruno Nachtergaele. Alignment and integration of complex networks by hypergraph-based spectral clustering. *Physical Review E*, 86(5):056111, 2012.

[74] Andrea Montanari. Tight bounds for ldpc and ldgm codes under map decoding. *IEEE Transactions on Information Theory*, 51(9):3221–3246, 2005.

[75] Andrea Montanari and Emile Richard. A statistical model for tensor PCA. page 30, nov 2014.

[76] Elchanan Mossel, Joe Neeman, and Allan Sly. Stochastic block models and reconstruction. *arXiv preprint arXiv:1202.1499*, 2012.

[77] Elchanan Mossel, Joe Neeman, and Allan Sly. A proof of the block model threshold conjecture. *arXiv preprint arXiv:1311.4115*, 2013.

[78] Elchanan Mossel, Joe Neeman, and Allan Sly. Consistency thresholds for the planted bisection model. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 69–75. ACM, 2015.

[79] Yurii Nesterov. Squared functional systems and optimization problems. In *High performance optimization*, pages 405–440. Springer, 2000.

[80] Mark EJ Newman, Duncan J Watts, and Steven H Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl 1):2566–2572, 2002.

[81] Frank Nielsen and Vincent Garcia. Statistical exponential families: A digest with flash cards. *arXiv preprint arXiv:0911.4863*, 2009.

[82] Pablo A Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, 2000.

[83] Amelia Perry and Alexander S Wein. A semidefinite program for unbalanced multisection in the stochastic block model. In *Sampling Theory and Applications (SampTA), 2017 International Conference on*, pages 64–67. IEEE, 2017.

[84] Amelia Perry, Alexander S. Wein, and Afonso S. Bandeira. Statistical limits of spiked tensor models. page 39, dec 2016.

[85] Amelia Perry, Alexander S Wein, Afonso S Bandeira, and Ankur Moitra. Optimality and sub-optimality of pca for spiked random matrices and synchronization. *arXiv preprint arXiv:1609.05573*, 2016.

[86] Shaghayegh Sahebi and William W Cohen. Community-based recommendations: a solution to the cold start problem. In *Workshop on recommender systems and the social web, RSWEB*, 2011.

[87] Alexander Schrijver. New code upper bounds from the terwilliger algebra and semidefinite programming. *IEEE Transactions on Information Theory*, 51(8):2859–2866, 2005.

[88] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8):888–905, 2000.

[89] Naum Zuselevich Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics*, 23(5):695–700, 1987.

[90] Paul Terwilliger. The subconstituent algebra of an association scheme,(part i). *Journal of Algebraic Combinatorics*, 1(4):363–388, 1992.

[91] Joel A Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of computational mathematics*, 12(4):389–434, 2012.

[92] Ramon van Handel. Probability in high dimension. Technical report, Lecture Notes (Princeton University), 2014.

[93] Greg Ver Steeg, Cristopher Moore, Aram Galstyan, and Armen Allahverdyan. Phase transitions in community detection: A solvable toy model. *EPL (Europhysics Letters)*, 106(4):48004, 2014.

[94] Van Vu. A simple svd algorithm for finding hidden partitions. *arXiv preprint arXiv:1404.3918*, 2014.

[95] Dengyong Zhou, Jiayuan Huang, and Bernhard Schölkopf. Learning with hypergraphs: Clustering, classification, and embedding. In *NIPS*, volume 19, pages 1633–1640, 2006.

[96] Vinko Zlatić, Gourab Ghoshal, and Guido Caldarelli. Hypergraph topological quantities for tagged social networks. *Physical Review E*, 80(3):036118, 2009.