

Algorithmic Contiguity from Low-Degree Conjecture and Applications in Correlated Random Graphs

Zhangsong Li   

School of Mathematical Sciences, Peking University, Beijing, China

Abstract

In this paper, assuming a natural strengthening of the low-degree conjecture, we provide evidence of computational hardness for two problems: (1) the (partial) matching recovery problem in the sparse correlated Erdős-Rényi graphs $\mathcal{G}(n, q; \rho)$ when the edge-density $q = n^{-1+o(1)}$ and the correlation $\rho < \sqrt{\alpha}$ lies below the Otter's threshold, this resolves a remaining problem in [15]; (2) the detection problem between a pair of correlated sparse stochastic block model $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$ and a pair of independent stochastic block models $\mathcal{S}(n, \frac{\lambda s}{n}; k, \epsilon)$ when $\epsilon^2 \lambda s < 1$ lies below the Kesten-Stigum (KS) threshold and $s < \sqrt{\alpha}$ lies below the Otter's threshold, this resolves a remaining problem in [9].

One of the main ingredient in our proof is to derive certain forms of *algorithmic contiguity* between two probability measures based on bounds on their low-degree advantage. To be more precise, consider the high-dimensional hypothesis testing problem between two probability measures \mathbb{P} and \mathbb{Q} based on the sample \mathbf{Y} . We show that if the low-degree advantage $\text{Adv}_{\leq D}(\frac{d\mathbb{P}}{d\mathbb{Q}}) = O(1)$, then (assuming the low-degree conjecture) there is no efficient algorithm \mathcal{A} such that $\mathbb{Q}(\mathcal{A}(\mathbf{Y}) = 0) = 1 - o(1)$ and $\mathbb{P}(\mathcal{A}(\mathbf{Y}) = 1) = \Omega(1)$. This framework provides a useful tool for performing reductions between different inference tasks.

2012 ACM Subject Classification Theory of computation \rightarrow Random network models; Mathematics of computing \rightarrow Random graphs

Keywords and phrases Algorithmic Contiguity, Low-degree Conjecture, Correlated Random Graphs

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2025.30

Category RANDOM

Related Version *Full Version*: <https://arxiv.org/abs/2502.09832> [41]

Funding *Zhangsong Li*: Partially supported by National Key R&D program of China (Project No. 2023YFA1010103) and NSFC Key Program (Project No. 12231002)

Acknowledgements Z.L. thanks Jian Ding and Jingqiu Ding for helpful discussions, Alexander S. Wein for helpful comments on the revised low-degree conjecture, and Hang Du for pointing him to the reference [56].

1 Introduction

Graph matching, also referred to as network alignment, is the problem of identifying a bijection between the vertex sets of two graphs that maximizes the number of common edges. When the two graphs are exactly isomorphic to each other, this problem reduces to the classical graph isomorphism problem, for which the best known algorithm runs in quasi-polynomial time [1]. In general, graph matching is an instance of the *quadratic assignment problem* [7], which is known to be NP-hard to solve or even approximate [42]. Motivated by real-world applications (such as social network de-anonymization [49] and computational biology [55]) as well as the need to understand the average-case computational complexity, recent research has focused on developing theoretical foundations and efficient algorithms for graph matching under statistical models. These models assume that the two graphs



© Zhangsong Li;
licensed under Creative Commons License CC-BY 4.0

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2025).

Editors: Alina Ene and Eshan Chattopadhyay; Article No. 30; pp. 30:1–30:18



Leibniz International Proceedings in Informatics

LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

are randomly generated with correlated edges under a hidden vertex correspondence, and a canonical model among them is the following *correlated random graph model*. For any integer n , denote by $U = U_n$ the set of unordered pairs (i, j) with $1 \leq i \neq j \leq n$.

► **Definition 1** (Correlated random graph model). *Given an integer $n \geq 1$, for $(i, j) \in U$ let $J_{i,j}$ and $K_{i,j}$ be independent Bernoulli variables with parameter s . In addition, let π_* be an independent uniform permutation on $[n] = \{1, \dots, n\}$. Then, we define a triple of correlated random graphs (G, A, B) as follows: first we generate G independently with $\{I_{i,j}, K_{i,j}\}$ and π_* from a specific probability distribution over all graphs on $[n]$, and then (conditioned on G) we define for each $(i, j) \in U$ that (note that we identify a graph with its adjacency matrix)*

$$A_{i,j} = G_{i,j}J_{i,j}, \quad B_{i,j} = G_{i,j}K_{\pi_*^{-1}(i), \pi_*^{-1}(j)}.$$

In short, we will subsample A, B from G with subsampling probability s and then permute the vertices of B by a uniform permutation.

Of particular interest in our paper are the following two special cases, namely the correlated Erdős-Rényi model and the correlated stochastic block model (SBM).

► **Definition 2** (Correlated Erdős-Rényi graph model). *Given an integer $n \geq 1$ and two parameters $p, s \in (0, 1)$, we generate a triple of correlated random graphs (G, A, B) such that we first generate G according to an Erdős-Rényi graph distribution $\mathcal{G}(n, p)$ (i.e., for each $(i, j) \in U$ we connect (i, j) in G independently with probability p), and then generate (A, B) from G according to Definition 1. For ease of presentation, we shall reparameterize such that $q = ps$ and $\rho = \frac{s(1-p)}{1-ps}$ respectively. We will denote the marginal law of (A, B) as $\mathcal{G}(n, q; \rho)$.*

► **Definition 3** (Stochastic block model). *Given an integer $n \geq 1$ and three parameters $k \in \mathbb{N}, \lambda > 0, \epsilon \in (0, 1)$, we define a random graph G as follows: (1) sample a labeling $\sigma_* \in [k]^n = \{1, \dots, k\}^n$ uniformly at random; (2) for every distinct pair $(i, j) \in U_n$, we let $G_{i,j}$ be an independent Bernoulli variable such that $G_{i,j} = 1$ (which represents that there is an undirected edge between i and j) with probability $\frac{(1+(k-1)\epsilon)\lambda}{n}$ if $\sigma_*(i) = \sigma_*(j)$ and with probability $\frac{(1-\epsilon)\lambda}{n}$ if $\sigma_*(i) \neq \sigma_*(j)$. In this case, we say that G is sampled from a stochastic block model $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon)$.*

► **Definition 4** (Correlated stochastic block model). *Given an integer $n \geq 1$ and four parameters $k \in \mathbb{N}, \lambda > 0, \epsilon, s \in (0, 1)$, we define a triple of correlated random graphs (G, A, B) as follows: we first sample G according to the law of a stochastic block model $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon)$ and then generate (A, B) from G according to Definition 1. We will denote the marginal law of (A, B) as $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$.*

Two fundamental problems in the study of correlated random graph model are as follows: (1) the *detection* problem, which involves determining whether a given pair of graphs (A, B) is sampled from a pair of correlated random graphs or from a pair of independent random graphs; (2) the *matching* problem, which focuses on recovering the latent matching π_* from a sample (A, B) from the distribution of correlated random graphs. In recent years, significant progress has been made in understanding these problems for both the correlated Erdős-Rényi model and correlated stochastic block models. Through the collective efforts of the community, the information-theoretic thresholds for detection and matching have been fully characterized for correlated Erdős-Rényi model [10, 31, 58, 59, 27, 13, 14, 21] and partially characterized for correlated SBMs [51, 30]. Additionally, various efficient detecting and matching algorithms have been developed with performance guarantees [11, 4, 18, 22, 23, 26, 28, 29, 43, 44, 46, 45, 16, 17, 8, 9, 40]. Notably, like many other inference tasks in

high-dimensional statistics [61, 52, 39, 24], these problems appear to exhibit *information-computation gaps*. Specifically, for certain ranges of the correlation strength, detection or matching is information theoretically possible but no efficient algorithm is known to achieve these tasks. We now focus on the algorithmic side of these problems as they are more relevant to our work. Indeed, it has been shown that many inference tasks in the correlated random graph models exhibits sharp algorithmic phase transitions, as we elaborate below:

- For the detection problem between a pair of correlated Erdős-Rényi models $\mathcal{G}(n, q; \rho)$ and two independent Erdős-Rényi models $\mathcal{G}(n, q)$, we focus on the sparse regime where $q = n^{-1+o(1)}$. In this regime, on the one hand, it was shown in [46] that when $\rho > \sqrt{\alpha}$ where $\alpha \approx 0.338$ is the Otter's constant [50], there is an efficient algorithm that strongly distinguish these two models; on the other hand, it was shown in [15] that when $s < \sqrt{\alpha}$ there are evidence suggesting that all algorithms based on *low-degree polynomials* fail to strongly distinguish these two models.
- For the detection problem between a pair of correlated SBMs $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$ and two independent Erdős-Rényi graphs $\mathcal{G}(n, \frac{\lambda s}{n})$, we focus on the constant degree regime where $\lambda = O(1)$. In this regime, on the one hand, it was shown in [9] that when $s > \min\{\sqrt{\alpha}, \frac{1}{\epsilon^2 \lambda}\}$ where $\alpha \approx 0.338$ is the Otter's constant and $\frac{1}{\epsilon^2 \lambda}$ is the Kesten-Stigum threshold [35], there is an efficient algorithm that strongly distinguish these two models; on the other hand, it was also shown in [9] that when $s < \min\{\sqrt{\alpha}, \frac{1}{\epsilon^2 \lambda}\}$ there are evidence suggesting that all algorithms based on *low-degree polynomials* fail to strongly distinguish these two models.

The lower bound in the aforementioned results explored inherent computational barriers from the perspective of the *low-degree polynomial framework*. Indeed, it has been proved that the class of low-degree polynomial algorithms is a useful proxy for computationally efficient algorithms, in the sense that the best-known polynomial-time algorithms for a wide variety of high-dimensional inference problems are captured by the low-degree class; see e.g. [33, 32, 54, 39]. However, these aforementioned results suffer from two significant limitations, which we now discuss. Firstly, the aforementioned result cannot provide the evidence that *partial matching recovery* (that is, recover a positive fraction of the coordinates of the latent matching π_*) is impossible by efficient algorithms although efficient detection has already been ruled out. Secondly, (in the case of correlated SBM) they are only able to establish the computation threshold on the detection problem between the correlated model and a pair of independent Erdős-Rényi graphs. These limitations motivate two natural questions:

- (1) Can we provide the evidence that partial matching recovery is impossible (by efficient algorithms) in the same parameter regime where there are evidence suggesting that detection is impossible (by efficient algorithms)?
- (2) What can we say about the (arguably more natural) detection problem between a pair of correlated SBMs $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$ and a pair of independent SBMs $\mathcal{S}(n, \frac{\lambda s}{n}; k, \epsilon)$?

The aim of this paper is to (partially) answer these two problems. Our main result can be informally summarized as follows:

► **Theorem 5 (Informal).** *We have the following results.*

- (1) *Assuming a natural strengthening of the low-degree conjecture (see Conjecture 12 for its precise meaning), then for the correlated Erdős-Rényi model $\mathcal{G}(n, q, \rho)$, when $q = n^{-1+o(1)}$ and $\rho < \sqrt{\alpha}$ it is impossible to recover a positive fraction of the coordinates of π_* with constant probability by efficient algorithms.*

- (2) Assuming a natural strengthening of the low-degree conjecture (see Conjecture 12 for its precise meaning), then for the correlated stochastic block models $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$, when $\lambda = O(1)$ and $s < \min\{\sqrt{\alpha}, \frac{1}{\epsilon^2 \lambda}\}$ it is impossible to recover a positive fraction of the coordinates of π_* with probability tending to 1 as $n \rightarrow \infty$ by efficient algorithms.
- (3) Assuming the low-degree conjecture (see Conjecture 9 for its precise meaning), then for the correlated stochastic block models $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$, when $s < \min\{\sqrt{\alpha}, \frac{1}{\epsilon^2 \lambda}\}$ it is impossible to strongly distinguish this model and a pair of independent SBMs $\mathcal{S}(n, \frac{\lambda s}{n}; k, \epsilon)$ by efficient algorithms, provided that the average degree λs is sufficiently large.

► **Remark 6.** Note that when $\rho > \sqrt{\alpha}$, the results in [45, 26, 28, 29] show that there exists an efficient algorithm that achieves partial recovery of π_* in a pair of correlated Erdős-Rényi models with probability tending to 1 as $n \rightarrow \infty$. Thus, Item (1) in Theorem 5 is tight in some sense and the algorithmic partial recovery threshold is indeed given by $\rho = \sqrt{\alpha}$.

Similarly, the results in [45, 26, 28, 29] naturally extends to show that there exists an efficient algorithm that achieves partial recovery of π_* in a pair of correlated SBMs when $s > \sqrt{\alpha}$ with probability tending to 1 as $n \rightarrow \infty$. Thus, Item (2) in Theorem 5 is also tight in the subcritical regime (i.e., when marginally two graphs are below the KS threshold).

► **Remark 7.** From Item (3) in Theorem 5 we see that for a pair of correlated SBMs (A, B) , when marginally both A and B are below the KS threshold (i.e., when $\epsilon^2 \lambda s < 1$), there is evidence suggesting that no efficient algorithm can strongly distinguish (A, B) from a pair of independent stochastic block models when the correlation $s < \sqrt{\alpha}$. Since the result in [46] extends naturally to the case of stochastic block models which provides an efficient algorithm that strongly distinguish these two models when $s > \sqrt{\alpha}$, we see that in the subcritical regime (i.e., when marginally both A and B are below the KS threshold) the algorithmic correlation detection threshold is given by $\sqrt{\alpha}$. On the contrary, in the supercritical regime where A and B are above the KS threshold (i.e., when $\epsilon^2 \lambda s > 1$), we believe that the algorithmic correlation detection threshold should be strict lower than $\sqrt{\alpha}$.

1.1 Key challenges and innovations

In this subsection, we briefly discuss our approach of proving Theorem 5 and some conceptual innovations behind it. Our idea can be informally summarized as follows:

- (1) As for Items (1) and (2), for simplicity we take correlated Erdős-Rényi model for example. Denote \mathbb{P} to be the law of $\mathcal{G}(n, q, \rho)$ and \mathbb{Q} to be the law of two independent $\mathcal{G}(n, q)$. We will argue by contradiction and assume that there is a partial recovery algorithm. Then we show that we can use this algorithm to efficiently construct a family of statistics $\{g_i : 1 \leq i \leq n\}$ such that g_i approximates $\mathbf{1}_{\{\pi_*(1)=i\}}$ under \mathbb{P} in a certain sense.
- (2) We will show that the low-degree advantage (see (1) for its precise meaning) between $\mathbb{P}(\cdot \mid \pi_*(1) = i)$ and \mathbb{Q} is bounded by an absolute constant. Then, from the standard low-degree conjecture (see Conjecture 9 for details) these two measures cannot be strongly distinguished by efficient algorithms. Thus, since g_i is “of positive constant order” under the measure $\mathbb{P}(\cdot \mid \pi_*(1) = 1)$ (as it should approximate $\mathbf{1}_{\{\pi_*(1)=1\}}$ in some sense) we expect that g_i should also be “of positive constant order” under \mathbb{Q} .
- (3) Define $g = g_1 + \dots + g_n$. We then have a statistics g that can be efficiently computed, and (I) under \mathbb{Q} we expect that g is “large” as it is the sum of n “of positive constant order” terms; (II) under \mathbb{P} we expect that g is “not large” as each g_i should approximate $\mathbf{1}_{\{\pi_*(1)=i\}}$ in some sense. Thus, the statistics g accumulates more signals than all low-degree polynomials, which (non-rigorously speaking) violates the low-degree conjecture.

- (4) As for Item (3), denote \mathbb{P} to be the law of correlated SBMs and \mathbb{Q} of independent SBMs. Also denote $\tilde{\mathbb{Q}}$ to be the law of independent Erdős-Rényi graphs. As it was already shown in [9] that \mathbb{P} and $\tilde{\mathbb{Q}}$ cannot be strongly distinguished, (non-rigorously speaking) we only need to show that \mathbb{Q} and $\tilde{\mathbb{Q}}$ are also “indistinguishable” in some sense.

However, there are certain obstacles when implementing the above ideas, as we shall discuss below. In Step (2) we need to “transfer” the behavior of a statistics g_i under $\mathbb{P}(\cdot \mid \pi_*(1) = i)$ to its behavior under \mathbb{Q} . If the low-degree advantage between $\mathbb{P}(\cdot \mid \pi_*(1) = i)$ and \mathbb{Q} is bounded by $1 + o(1)$, then the standard low-degree conjecture implies that we cannot distinguish between $\mathbb{P}(\cdot \mid \pi_*(1) = i)$ and \mathbb{Q} better than random by efficient algorithms. Consequently, we expect the behavior of g_i to be “almost identical” under both $\mathbb{P}(\cdot \mid \pi_*(1) = i)$ and \mathbb{Q} (as g_i can be efficiently computed). However, in our case the low-degree advantage is just bounded by a large (but fixed) constant. This weaker condition means that the standard low-degree conjecture only rules out the possibility to distinguish these two measures efficiently with vanishing errors, leaving room for non-negligible distinctions. Similar issues also arise in Step (4), in which simply showing the low-degree advantage between \mathbb{P} and $\tilde{\mathbb{Q}}$ and the low-degree advantage between \mathbb{Q} and $\tilde{\mathbb{Q}}$ is bounded is not strong enough for our goal.

To address these issues, one of the main conceptual contributions in our work is to give a more refined characterization of the limitations of efficient algorithms when the low-degree advantage is $O(1)$. Specifically, we show that (assuming low-degree conjecture) bounded low-degree advantage does not only excludes all algorithms that strongly distinguish \mathbb{P} and \mathbb{Q} , but also suggest a certain kind of *algorithmic contiguity*. To be more precise, if the low-degree advantage $\text{Adv}_{\leq D}(\frac{d\mathbb{P}}{d\mathbb{Q}})$ is bounded, then there is no efficient algorithm \mathcal{A} such that $\mathbb{Q}(\mathcal{A} = 0) = 1 - o(1)$ and $\mathbb{P}(\mathcal{A} = 1) = \Omega(1)$. This framework allows us to transfer the behavior of the behavior of a efficiently computable statistics g under different probability measures more easily. For example, if we know that $\mathbb{P}(g \geq c) \geq \Omega(1)$ it immediately holds that $\mathbb{Q}(g \geq c) \geq \Omega(1)$.

Another difficulty is that reducing between different inference tasks requires a framework that better constrains the behavior of efficiently computable estimators. To address this, we introduce a natural refinement of the low-degree conjecture: we posit that low-degree polynomials achieve signal-to-noise ratios that are at least as good as those of any efficient algorithm (see Conjecture 12 for a more precise description). While similar modifications to the low-degree conjecture have been explored in [47], our motivations differ significantly. Their work focused on characterizing the limits of precise error rate for all computationally feasible algorithms, whereas our goal is to enable reductions between distinct inference tasks.

1.2 Comparison to concurrent work

In a concurrent work [19], the authors investigate the computational hardness of the weak recovery problem in the stochastic block model $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon)$ when $k = n^{o(1)}$ and $\epsilon^2 \lambda < 1$ lies below the KS threshold. Notably, their proof is also based on a “recovery-to-detection reduction” approach, which argues that if there exists an efficient algorithm that achieves weak recovery, then we can use this algorithm to efficiently construct a statistics for detection (in a certain sense). Their results also relies on a strengthening of the low-degree conjecture, which is similar to the assumptions made in this paper.

However, the reduction technique employed in this paper differs significantly from that of [19]. The authors of [19] construct their statistic for detection using a *correlation preserving projection* technique established in [34]. This technique, which is designed specifically for weak recovery in block models, enables them to “regularize” the statistics and thus bound

its moments under the null hypothesis. In contrast, our construction is more straightforward as we leverage the universal framework described in Section 1.1 to “transfer” the behavior of statistics across different probability measures. It seems that in comparison to [19], our approach is easier to implement, less dependent on problem-specific details, and potentially applicable to a broader range of problems.

1.3 Notations

In this subsection, we record a list of notations that we shall use throughout the paper. Denote \mathfrak{S}_n the set of permutations over $[n]$ and denote μ the uniform distribution on \mathfrak{S}_n . In addition, denote ν to be the uniform distribution on $[k]^n$. We will use the following notation conventions on graphs.

- *Labeled graphs.* Denote by \mathcal{K}_n the complete graph with vertex set $[n]$ and edge set U_n . For any graph H , let $V(H)$ denote the vertex set of H and let $E(H)$ denote the edge set of H . We say H is a subgraph of G , denoted by $H \subset G$, if $V(H) \subset V(G)$ and $E(H) \subset E(G)$. Define the excess of the graph $\tau(H) = |E(H)| - |V(H)|$.
- *Isolated vertices.* For $u \in V(H)$, we say u is an isolated vertex of H , if there is no edge in $E(H)$ incident to u . Denote $\mathcal{I}(H)$ the set of isolated vertices of H . For two graphs H, S , we denote $H \times S$ if $H \subset S$ and $\mathcal{I}(S) \subset \mathcal{I}(H)$, and we denote $H \Subset S$ if $H \subset S$ and $\mathcal{I}(H) = \emptyset$. For any graph $H \subset \mathcal{K}_n$, let \tilde{H} be the subgraph of H induced by $V(H) \setminus \mathcal{I}(H)$.
- *Graph intersections and unions.* For $H, S \subset \mathcal{K}_n$, denote by $H \cap S$ the graph with vertex set given by $V(H) \cap V(S)$ and edge set given by $E(H) \cap E(S)$. Denote by $S \cup H$ the graph with vertex set given by $V(H) \cup V(S)$ and edge set $E(H) \cup E(S)$. In addition, denote by $S \pitchfork H$, $S \cupdot H$ and $S \triangle H$ the graph induced by the edge set $E(S) \cap E(H)$, $E(S) \cup E(H)$ and $E(S) \triangle E(H)$, respectively (in particular, these induced graphs have no isolated points).
- *Paths.* We say a subgraph $H \subset \mathcal{K}_n$ is a path with endpoints u, v (possibly with $u = v$), if there exist distinct $w_1, \dots, w_m \neq u, v$ such that $V(H) = \{u, v, w_1, \dots, w_m\}$ and $E(H) = \{(u, w_1), (w_1, w_2), \dots, (w_m, v)\}$. We say H is a simple path if its endpoints $u \neq v$. Denote $\text{EndP}(P)$ as the set of endpoints of a path P .
- *Cycles and independent cycles.* We say a subgraph H is an m -cycle if $V(H) = \{v_1, \dots, v_m\}$ and $E(H) = \{(v_1, v_2), \dots, (v_{m-1}, v_m), (v_m, v_1)\}$. For a subgraph $K \subset H$, we say K is an independent m -cycle of H , if K is an m -cycle and no edge in $E(H) \setminus E(K)$ is incident to $V(K)$. Denote by $\mathfrak{C}_m(H)$ the set of m -cycles of H and denote by $\mathfrak{C}_m(H)$ the set of independent m -cycles of H . For $H \subset S$, we define $\mathfrak{C}_m(S, H)$ to be the set of independent m -cycles in S whose vertex set is disjoint from $V(H)$. Define $\mathfrak{C}(S, H) = \cup_{m \geq 3} \mathfrak{C}_m(S, H)$.
- *Leaves.* A vertex $u \in V(H)$ is called a leaf of H , if the degree of u in H is 1; denote $\mathcal{L}(H)$ as the set of leaves of H .
- *Graph isomorphisms and unlabeled graphs.* Two graphs H and H' are isomorphic, denoted by $H \cong H'$, if there exists a bijection $\pi : V(H) \rightarrow V(H')$ such that $(\pi(u), \pi(v)) \in E(H')$ if and only if $(u, v) \in E(H)$. Denote by \mathcal{H} the isomorphism class of graphs; it is customary to refer to these isomorphic classes as unlabeled graphs. Let $\text{Aut}(H)$ be the number of automorphisms of H (graph isomorphisms to itself).

For two real numbers a and b , we let $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$. We use standard asymptotic notations: for two sequences a_n and b_n of positive numbers, we write $a_n = O(b_n)$, if $a_n < Cb_n$ for an absolute constant C and for all n (similarly we use the notation O_h is the constant C is not absolute but depends only on h); we write $a_n = \Omega(b_n)$, if $b_n = O(a_n)$; we write $a_n = \Theta(b_n)$, if $a_n = O(b_n)$ and $a_n = \Omega(b_n)$; we write $a_n = o(b_n)$ or

$b_n = \omega(a_n)$, if $a_n/b_n \rightarrow 0$ as $n \rightarrow \infty$. In addition, we write $a_n \stackrel{\circ}{=} b_n$ if $a_n = [1 + o(1)]b_n$. For a set A , we will use both $\#A$ and $|A|$ to denote its cardinality. For two probability measures \mathbb{P} and \mathbb{Q} , we denote the total variation distance between them by $\text{TV}(\mathbb{P}, \mathbb{Q})$.

1.4 Organization of this paper

The rest part of this paper is organized as follows: in Section 2 we state the precise framework of low-degree polynomials and how we relate it to the notion of algorithmic contiguity (see Theorem 11). In Section 3 we use this framework to deduce the hardness of partial matching in correlated Erdős-Rényi models and correlated SBMs, thus verifying Items (1) and (2) in Theorem 5 (see Theorem 14). In Section 4 we again use this framework to deduce the hardness of testing correlated SBMs against independent SBMs, thus verifying Item (3) in Theorem 5 (see Corollary 23).

2 Low-degree framework and algorithmic contiguity

The low-degree polynomial framework first emerged from the works of [5, 34, 33, 32] and has since been refined and extended in various directions. It has found applications across a broad spectrum of problems, including detection tasks such as planted clique, planted dense subgraph, community detection, sparse PCA, and tensor PCA (see [34, 33, 32, 39, 54, 12, 3, 20, 48, 15, 38]), optimization problems like finding maximal independent sets in sparse random graphs [25, 57], and constraint satisfaction problems such as random k -SAT [6]; see also the survey [39]. Additionally, it is conjectured in [32] that the failure of degree- D polynomials implies the failure of all “robust” algorithms with running time $n^{\tilde{O}(D)}$ (here \tilde{O} means having at most this order up to a poly log n factor). In the remaining of this paper, we will focus on applying this framework in the context of high-dimensional hypothesis testing problems.

To be more precise, consider the hypothesis testing problem between two probability measures \mathbb{P} and \mathbb{Q} based on the sample $\mathbf{Y} \in \mathbb{R}^N$. We will be especially interested in asymptotic settings where $N = N_n, \mathbb{Q} = \mathbb{Q}_n, \mathbb{P} = \mathbb{P}_n, \mathbf{Y} = \mathbf{Y}_n$ scale with n as $n \rightarrow \infty$ in some prescribed way. The standard low-degree polynomial framework primarily focus on the following notions on strong and weak detection.

► **Definition 8** (Strong/weak detection). *We say an algorithm \mathcal{A} that takes \mathbf{Y} as input and outputs either 0 or 1 achieves*

- **strong detection**, if the sum of type-I and type-II errors $\mathbb{Q}(\mathcal{A}(\mathbf{Y}) = 1) + \mathbb{P}(\mathcal{A}(\mathbf{Y}) = 0)$ tends to 0 as $n \rightarrow \infty$.
- **weak detection**, if the sum of type-I and type-II errors is uniformly bounded above by $1 - \epsilon$ for some fixed $\epsilon > 0$.

Roughly speaking, the low-degree polynomial approach focuses on understanding the capabilities and limitations of algorithms that can be expressed as low-degree polynomial functions of the input variables (in our case, the entries of \mathbf{Y}). To be more precise, let $\mathcal{P}_D = \mathcal{P}_{n,D}$ denote the set of polynomials from \mathbb{R}^N to \mathbb{R} with degree no more than D . With a slight abuse of notation, we will often refer to “a polynomial” to mean a sequence of polynomials $f = f_n \in \mathcal{P}_{n,D}$, where each f_n corresponds to problem size n ; the degree $D = D_n$ of such a polynomial may scale with n . As suggested by [32], the key quantity is the low-degree advantage

$$\text{Adv}_{\leq D} \left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right) := \sup_{f \in \mathcal{P}_D} \frac{\mathbb{E}_{\mathbb{P}}[f]}{\sqrt{\mathbb{E}_{\mathbb{Q}}[f^2]}}. \quad (1)$$

Note that if we denote the likelihood ratio $L(Y) = \frac{d\mathbb{P}}{d\mathbb{Q}}(Y)$, it is well-known (see, e.g., [32]) the right hand side equals the L_2 -norm of the projection of $L(Y)$ into the subspace spanned by all polynomials of degree bounded by D (the norm is induced by natural inner product under \mathbb{Q}) and thus we characterize it with $\frac{d\mathbb{P}}{d\mathbb{Q}}$. The low degree conjecture, proposed in [32], can be summarized as follows.

► **Conjecture 9** (Low-degree conjecture). *For “natural” high-dimensional hypothesis testing problems between \mathbb{P} and \mathbb{Q} , the following statements hold.*

- (1) *If $\text{Adv}_{\leq D}(\frac{d\mathbb{P}}{d\mathbb{Q}}) = O(1)$ as $n \rightarrow \infty$, then there exists a constant C such that there is no algorithm with running time $n^{D/(\log n)^C}$ that achieves strong detection between \mathbb{P} and \mathbb{Q} .*
- (2) *If $\text{Adv}_{\leq D}(\frac{d\mathbb{P}}{d\mathbb{Q}}) = 1 + o(1)$ as $n \rightarrow \infty$, then there exists a constant C such that there is no algorithm with running time $n^{D/(\log n)^C}$ that achieves weak detection between \mathbb{P} and \mathbb{Q} .*

Motivated by [32, Hypothesis 2.1.5 and Conjecture 2.2.4] as well as the fact that low-degree polynomials capture the best known algorithms for a wide variety of statistical inference tasks, Conjecture 9 appears to hold for distributions of a specific form that frequently arises in high-dimensional statistics. For further discussion on what types of distributions are suitable for this framework, we refer readers to [32, 39, 37, 60]. In addition, we point out that although in most applications (and in the statement of [32, Hypothesis 2.1.5]) it is typically to take \mathbb{Q} to be a “null” measure and \mathbb{P} to be a “planted” measure (which makes (1) more tractable), several recent works [53, 36] showed that this framework might also be applicable for many “planted-versus-planted” problems. Nevertheless, in this paper, we adopt a more conservative view and will explicitly indicate whenever \mathbb{Q} is treated as a planted measure.

The framework in Conjecture 9 provides a useful tool for probing the computational feasibility of strong or weak detection. However, as discussed in Section 1.1, it turns out that the failure of strong detection is not enough in our cases, especially when we hope to perform some reductions between statistical models in a regime where weak (but not strong) detection is possible. Thus, in this regime, we aim to characterize a stronger framework that rules out all *one-sided test*. This motivates thus to propose the following notion of *algorithmic contiguity*.

► **Definition 10.** *For “natural” high-dimensional hypothesis testing problems between \mathbb{P} and \mathbb{Q} , we say an algorithm \mathcal{A} that takes Y as input and outputs either 0 or 1 is a \mathbb{Q} -based one-sided test, if*

$$\mathbb{Q}(\mathcal{A}(Y) = 0) = 1 - o(1) \text{ and } \mathbb{P}(\mathcal{A}(Y) = 1) = \Omega(1). \quad (2)$$

We say that \mathbb{P} is time- n^D algorithmic contiguous with respect to \mathbb{Q} , denoted as $\mathbb{P} \triangleleft_{\leq D} \mathbb{Q}$, if no \mathbb{Q} -based one-sided testing algorithm runs in time n^D . We say that \mathbb{Q} and \mathbb{P} are degree- D algorithmic mutually contiguous, denoted as $\mathbb{Q} \triangleleft \triangleright_{\leq D} \mathbb{P}$, if both $\mathbb{Q} \triangleleft_{\leq D} \mathbb{P}$ and $\mathbb{P} \triangleleft_{\leq D} \mathbb{Q}$ hold.

Recall that in probability theory we say a sequence of probability measure $\mathbb{P} = \mathbb{P}_n$ is contiguous with respect to $\mathbb{Q} = \mathbb{Q}_n$, if for all sequence of events $\{A_n\}$ we have $\mathbb{Q}_n(A_n) \rightarrow 0$ implies that $\mathbb{P}_n(A_n) \rightarrow 0$. Thus, our definition can be regarded as the generalization of contiguity in algorithmic view. Our main result in this section can be stated as follows.

► **Theorem 11.** *Assuming the Conjecture 9, for the high-dimensional hypothesis testing problem between \mathbb{P} and \mathbb{Q} , if $\text{Adv}_{\leq D}(\frac{d\mathbb{P}'}{d\mathbb{Q}'}) = O(1)$ for some \mathbb{P}', \mathbb{Q}' such that $\text{TV}(\mathbb{P}, \mathbb{P}') = o(1)$ and $\text{TV}(\mathbb{Q}, \mathbb{Q}') = o(1)$, then we have $\mathbb{P} \triangleleft_{\leq D/(\log n)^C} \mathbb{Q}$ for some constant C .*

In the remaining part of this work we will also need a strengthening of Conjecture 9, as introduced in Section 1.1. As discussed in the beginning of this section, the low-degree conjecture asserts that (for certain testing problems) low-degree polynomials are at least

as powerful as all algorithms of the corresponding runtime (where the correspondence is described in Conjecture 9). For our purposes, we introduce a natural refinement of the low-degree conjecture, which posits that low-degree polynomials perform at least as well as all algorithms of the corresponding runtime in terms of the value of the ratio on the right-hand side of (1).

► **Conjecture 12** (Revised low-degree conjecture). *For “natural” high-dimensional hypothesis testing problems between \mathbb{P} and \mathbb{Q} , the following statements hold:*

- (1) *If $\text{Adv}_{\leq D}(\frac{d\mathbb{P}'}{d\mathbb{Q}'}) = O(1)$ as $n \rightarrow \infty$ for some \mathbb{P}', \mathbb{Q}' such that $\text{TV}(\mathbb{P}, \mathbb{P}'), \text{TV}(\mathbb{Q}, \mathbb{Q}') = o(1)$, then there is a constant C such that there is no algorithm with running time $n^{D/(\log n)^C}$ that achieves strong detection between \mathbb{P} and \mathbb{Q} .*
- (2) *If there is no algorithm with running time n^D that achieves strong detection between \mathbb{P} and \mathbb{Q} , then for all statistics $f = f(\mathbf{Y})$ that can be computed in running time n^D , there exists some \mathbb{P}', \mathbb{Q}' (we allow \mathbb{P}', \mathbb{Q}' depend on f) such that $\text{TV}(\mathbb{P}, \mathbb{P}'), \text{TV}(\mathbb{Q}, \mathbb{Q}') = o(1)$ such that $\text{Adv}(f) = O(1)$, where*

$$\text{Adv}(f) = \frac{\mathbb{E}_{\mathbb{P}'}[f]}{\sqrt{\mathbb{E}_{\mathbb{Q}'}[f^2]}}.$$

- (3) *If $\text{Adv}_{\leq D}(\frac{d\mathbb{P}'}{d\mathbb{Q}'}) = 1 + o(1)$ as $n \rightarrow \infty$ for some \mathbb{P}', \mathbb{Q}' such that $\text{TV}(\mathbb{P}, \mathbb{P}'), \text{TV}(\mathbb{Q}, \mathbb{Q}') = o(1)$, then there is a constant C such that there is no algorithm with running time $n^{D/(\log n)^C}$ that achieves weak detection between \mathbb{P} and \mathbb{Q} .*
- (4) *If there is no algorithm with running time n^D that achieves weak detection between \mathbb{P} and \mathbb{Q} , then for all statistics $f = f(\mathbf{Y})$ that can be computed in running time n^D , there exists some \mathbb{P}', \mathbb{Q}' (we allow \mathbb{P}', \mathbb{Q}' depend on f) such that $\text{TV}(\mathbb{P}, \mathbb{P}'), \text{TV}(\mathbb{Q}, \mathbb{Q}') = o(1)$ such that $\text{Adv}(f) = 1 + o(1)$.*

Note that in Conjecture 12 we are allowed to replace \mathbb{P}, \mathbb{Q} with some \mathbb{P}', \mathbb{Q}' that are *statistically indistinguishable* with \mathbb{P}, \mathbb{Q} . This modification enables us to avoid some situations where the low-degree advantage explodes due to some “rare events” (see [2, 15, 12] for example).¹ This revised low-degree conjecture was first proposed in [47] (in a slightly different manner) for a different purpose where they aimed to study the limits of precise error of all computationally feasible algorithms. However, we will show in the next two sections that this conjecture is also useful in performing reductions between different inference tasks.

3 Partial recovery in correlated random graphs

In this section, we will use the framework we established in Section 2 to show the hardness of partial matching in correlated random graphs, thus justifying Items (1) and (2) in Theorem 5. To this end, we first state the precise meaning that an algorithm achieves partial matching.

¹ In fact, as pointed out by Alexander S. Wein (who learned the following example from Ansh Nagda) in personal communication with the author, such truncation appears necessary for the conjecture to hold. Consider the sparse-PCA problem where the observation $Y = \frac{\lambda v v^T}{\|v\|^2} + W$ where W is an $n \times n$ GOE matrix and $v \in \mathbb{R}^n$ having coordinates i.i.d. drawn from the sparse Rademacher prior with parameter $\rho = o(1)$. In this case, when $0 < \lambda < 1$ we have $\mathbb{E}_{\mathbb{P}}[f]/\sqrt{\mathbb{E}_{\mathbb{Q}}[f^2]} = O(1)$ for all low-degree polynomials f but $\mathbb{E}_{\mathbb{P}}[f]/\sqrt{\mathbb{E}_{\mathbb{Q}}[f^2]} = \omega(1)$ if f is a thresholding function on $v_0^T Y v_0$ where v_0 is a fixed sparse $\{0, 1\}$ -vector. However, our truncated conjecture still holds as we can take \mathbb{P}' to be the conditional measure of \mathbb{P} given $v_0^T Y v_0$ is small.

30:10 Algorithmic Contiguity from Low-Degree Conjecture

► **Definition 13** (Partial recovery algorithm in correlated Erdős-Rényi model). *Given a sample (A, B) from the law a pair of correlated random graphs in Definition 1 (we denote this law of \mathbb{P}_*). We say an algorithm \mathcal{A} achieves strong partial matching, if it takes (A, B) as input and outputs a family of estimators $\{h_{i,j} : 1 \leq i, j \leq n\}$ such that*

- (1) $h_{i,j} \in \{0, 1\}$ for all $1 \leq i, j \leq n$ a.s. under \mathbb{P}_* ;
- (2) $h_{i,1} + \dots + h_{i,n} = 1$ for all $1 \leq i \leq n$ a.s. under \mathbb{P}_* ;
- (3) There exists a fixed constant $\iota > 0$ such that

$$\mathbb{P}_*(h_{1,\pi_*(1)} + \dots + h_{n,\pi_*(n)} \geq \iota n) = 1 - o(1).$$

We say an algorithm \mathcal{A} achieves weak partial matching, if it takes (A, B) as input and outputs a family of estimators $\{h_{i,j} : 1 \leq i, j \leq n\}$ satisfying Items (1), (2) above and

$$\mathbb{P}_*(h_{1,\pi_*(1)} + \dots + h_{n,\pi_*(n)} \geq \iota n) = \Omega(1).$$

We point out that in the above definition $h_{i,j}$ can be thought as the estimator of $\mathbf{1}_{\{\pi_*(i)=j\}}$, where Item (3) implies this algorithm correctly matches a positive fraction of $\pi_*(i)$ with probability $1 - o(1)$ (or with positive probability) and Items (1) and (2) are some regularity requirements. Our result in this section can be stated as follows:

► **Theorem 14.** *Assuming Conjecture 12, we have the following:*

- (1) *For the correlated Erdős-Rényi graphs $\mathcal{G}(n, q, \rho)$ where $q = n^{-1+o(1)}$ and $\rho < \sqrt{\alpha} - \delta$ for a fixed constant $\delta > 0$. There exists constant C such that no algorithm with running time $n^{D/(\log n)^C}$ that achieves weak partial matching, provided that*

$$D = \exp\left(o\left(\frac{\log n}{\log(nq)} \wedge \sqrt{\log n}\right)\right). \quad (3)$$

- (2) *For the correlated SBMs $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon, s)$ where $\lambda = O(1)$ and $\epsilon^2 \lambda < 1 - \delta, s < \sqrt{\alpha} - \delta$ for a fixed constant $\delta > 0$. There exists constant C such that no algorithm with running time $n^{D/(\log n)^C}$ that achieves strong partial matching, provided that*

$$D = n^{o(1)}. \quad (4)$$

The main step of proving Theorem 14 is to show the following proposition:

► **Proposition 15.** *Assuming Conjecture 12, we have the following:*

- (1) *If $(G, A, B) \sim \mathcal{G}(n, q, \rho)$ and let \mathbb{P}_* to be the joint law of (π_*, G, A, B) where π_* is the latent matching. Suppose that q, ρ, D satisfy the assumptions in Item (1) of Theorem 14. Then there exists constant C such that for all $\{f_{i,j} : 1 \leq i, j \leq n\}$ that can be computed in time $n^{D/(\log n)^C}$, we have $\mathbb{E}_{\mathbb{P}_*}[f_{i,\pi_*(i)}] = o(1)$ for all $1 \leq i \leq n$.*
- (2) *If $(G, A, B) \sim \mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon, s)$ and let \mathbb{P}_* to be the joint law of (π_*, G, A, B) where π_* is the latent matching. Suppose that λ, k, ϵ, s satisfy the assumptions in Item (2) of Theorem 14. Then there exist a constant C and an event \mathcal{E} such that $\mathbb{P}_*(\mathcal{E}) = \Omega(1)$, and for all $\{f_{i,j} : 1 \leq i, j \leq n\}$ that can be computed in time $n^{D/(\log n)^C}$ we have $\mathbb{E}_{\mathbb{P}_*}[f_{i,\pi_*(i)} | \mathcal{E}] = o(1)$ for all $1 \leq i \leq n$.*

Clearly, based on Proposition 15, we can deduce Theorem 14 via a simple Markov inequality. The rest of this section is devoted to the proof of Proposition 15. In the following subsections, our main focus is on proving Item (1) of Proposition 15. Given the similarity between the proofs of Item (1) and Item (2), for Item (2) we will provide an outline with the main differences while adapting arguments from proving Item (1) without presenting full details.

3.1 Proof of Item (1) in Proposition 15

This subsection is devoted to the proof of Item (1) in Proposition 15. Throughout this subsection, we will denote \mathbb{P}_* to be the law of (π_*, G, A, B) where $(G, A, B) \sim \mathcal{G}(n, q; \rho)$. We will also denote \mathbb{P} to be the marginal law of (A, B) . In addition, we assume throughout this subsection that there exists a small constant $0 < \delta < 0.01$ such that

$$\rho^2 < \alpha - \delta, \quad q = n^{-1+o(1)}, \quad \log D = o\left(\frac{\log n}{\log(nq)} \wedge \sqrt{\log n}\right). \quad (5)$$

We first introduce some notations used in [15].

► **Definition 16.** *Given a graph $H = H(V, E)$, define*

$$\Phi(H) = (n^{1+4/D} D^{20})^{|V(H)|} (qD^6)^{|E(H)|}, \quad (6)$$

and the graph H is said to be bad if $\Phi(H) < (\log n)^{-1}$. Furthermore, we say a graph is admissible if it contains no bad subgraph, and we say it is inadmissible otherwise.

Denote \mathcal{E} for the event that G does not contain any bad subgraph with no more than d^2 vertices. In addition, let $\bar{\mathbb{P}}_*$ be the conditional version of \mathbb{P}_* given \mathcal{E} , and let $\bar{\mathbb{P}}$ be the corresponding marginal distribution of $\bar{\mathbb{P}}_*$ on (A, B) .

We remark here that our definition of “bad” amounts to an atypically large edge density, with a carefully chosen quantitative threshold on “large”. Roughly speaking, we expect that any subgraph with size no more than $D^2 = n^{o(1)}$ of a sparse Erdős-Rényi graph has edge-to-vertex ratio $1 + o(1)$. In the definition of Φ , the term $n^{1+4/D} D^{20}$ should be thought as $n^{1+o(1)}$, and qD^6 as $n^{-1+o(1)}$. The $o(1)$ terms are tuned carefully so that for a typical subgraph H of a sparse Erdős-Rényi graph, $\Phi(H)$ is much bigger than 1. The choice of $(\log n)^{-1}$ as the Φ -threshold for bad graph is somewhat arbitrary, which we will only need to be vanishing as $n \rightarrow \infty$. In [15], the authors showed that one the one hand, we have $\mathbb{P}_*(\mathcal{E}) = 1 - o(1)$ and thus $\text{TV}(\mathbb{P}_*, \bar{\mathbb{P}}_*)$, $\text{TV}(\mathbb{P}, \bar{\mathbb{P}}) = o(1)$; on the other hand, we have $\text{Adv}_{\leq D}\left(\frac{d\bar{\mathbb{P}}}{d\mathbb{Q}}\right) = O_\delta(1)$, thus verifying the low-degree hardness for the detection problem. The first step of our proof is to generalize the result in [15], as incorporated in the following lemma.

► **Lemma 17.** *For all $1 \leq i, j \leq n$, we have $\text{Adv}_{\leq D}\left(\frac{d\bar{\mathbb{P}}(\cdot|\pi_*(i)=j)}{d\mathbb{Q}}\right) = O_\delta(1)$.*

Now, based on Lemma 17, we establish the following result, which basically suggests that it is impossible to obtain a “good approximation” of $\mathbf{1}_{\{\pi_*(i)=j\}}$ in some sense.

► **Lemma 18.** *Assuming Conjecture 12, there exists a constant C such that for all $1 \leq i \leq n$ and all statistics $\{g_{i,j} = g_{i,j}(A, B) : 1 \leq j \leq n\}$ that can be computed in time $n^{D/(\log n)^C}$, it holds that*

$$\sum_{j=1}^n \mathbb{E}_{\mathbb{P}_*} \left[\left(\mathbf{1}_{\{\pi_*(i)=j\}} - g_{i,j} \right)^2 \right] \geq 1 - o(1). \quad (7)$$

Now we can finish the proof of Item (1) in Proposition 15.

Proof of Item (1) in Proposition 15. We choose C as in the statement of Lemma 18. Suppose on the contrary there are statistics $\{f_j : 1 \leq j \leq n\}$ that can be computed in time $n^{D/(\log n)^C}$ with $\mathbb{E}_{\mathbb{P}_*}[f_{\pi_*(i)}] \geq 1 - c$ for some fixed constant $0 < c < 0.01$. Using Items (1) and (2) in Definition 13, we see that $f_j f_k = 0$ for all $k \neq j$, and thus

$$1 = \mathbb{E}_{\mathbb{P}} \left[(f_1 + \dots + f_n)^2 \right] = \sum_{j=1}^n \mathbb{E}_{\mathbb{P}} \left[f_j^2 \right]. \quad (8)$$

30:12 Algorithmic Contiguity from Low-Degree Conjecture

In addition, we have

$$\mathbb{E}_{\mathbb{P}_*} [f_{\pi_*(i)}] = \frac{1}{n} \sum_{j=1}^n \mathbb{E}_{\mathbb{P}_*} [f_j \mid \pi_*(i) = j] \geq c. \quad (9)$$

Thus, for all $\lambda \in [0, 1]$ we have

$$\sum_{j=1}^n \mathbb{E}_{\mathbb{P}_*} \left[\left(\mathbf{1}_{\{\pi_*(i)=j\}} - \frac{1-\lambda}{n} - \lambda f_j \right)^2 \right] \stackrel{(8),(9)}{\leq} 1 + \lambda^2 - 2c\lambda + O\left(\frac{1}{n}\right).$$

Thus, by choosing $\lambda = \lambda(c)$ to be a sufficiently small positive constant we get that

$$\sum_{j=1}^n \mathbb{E}_{\mathbb{P}_*} \left[\left(\mathbf{1}_{\{\pi_*(i)=j\}} - g_j \right)^2 \right] = 1 - \Omega(1) \text{ where } g_j = \frac{1-\lambda}{n} + \lambda f_j,$$

contradicting to Lemma 18. This leads to the desired result. \blacktriangleleft

3.2 Proof of Item (2) in Proposition 15

This subsection is devoted to the proof of Item (2) in Proposition 15. Recall Definitions 3 and 4. Throughout this subsection, we will denote \mathbb{P}_* to be the joint law of $(\pi_*, \sigma_*, G, A, B)$ where $(G, A, B) \sim \mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$ and \mathbb{P} the marginal law of (A, B) . In addition, we denote \mathbb{Q} to be the law of a pair of independent Erdős-Rényi models $\mathcal{G}(n, \frac{\lambda \epsilon}{n})$. Also, we assume throughout this subsection that there exists a small constant $0 < \delta < 0.01$ such that

$$s < \sqrt{\alpha} - \delta, \quad \epsilon^2 \lambda s < 1 - \delta. \quad (10)$$

We start again by recalling some notations introduced in [9]. We choose a sufficiently large constant $N = N(k, \lambda, \delta, \epsilon, s) \geq 2/\delta$ such that

$$\begin{aligned} (\sqrt{\alpha} - \delta)(1 + \epsilon^N k) &\leq \sqrt{\alpha} - \delta/2; & 10k(1 - \delta)^N &\leq (1 - \delta/2)^N; \\ (\sqrt{\alpha} - \delta/2)(1 + (1 - \delta/2)^N)^2 &\leq \sqrt{\alpha} - \delta/4; & (1 - \delta/2)^N(N + 1) &\leq 1. \end{aligned} \quad (11)$$

We first show how to construct the event \mathcal{E} in Item (2) in Proposition 15.

► **Definition 19.** Denote $\tilde{\lambda} = \lambda \vee 1$. Given a graph $H = H(V, E)$, define

$$\Upsilon(H) = \left(\frac{2\tilde{\lambda}^2 k^2 n}{D^{50}} \right)^{|V(H)|} \left(\frac{1000\tilde{\lambda}^{20} k^{20} D^{50}}{n} \right)^{|E(H)|}. \quad (12)$$

Then we say the graph H is bad if $\Upsilon(H) < (\log n)^{-1}$, and we say a graph H is self-bad if H is bad and $\Upsilon(H) < \Upsilon(K)$ for all $K \subset H$. Furthermore, we say that a graph H is admissible if it contains no bad subgraph and $\mathcal{C}_j(H) = \emptyset$ for $j \leq N$; we say H is inadmissible otherwise. Denote $\mathcal{E} = \mathcal{E}^{(1)} \cap \mathcal{E}^{(2)}$, where $\mathcal{E}^{(1)}$ is the event that G does not contain any bad subgraph with no more than D^3 vertices, and $\mathcal{E}^{(2)}$ is the event that G does not contain any cycles with length at most N .

► **Definition 20.** List all self-bad subgraphs of \mathcal{K}_n with at most D^3 vertices and all cycles of \mathcal{K}_n with lengths at most N in an arbitrary but prefixed order (B_1, \dots, B_M) . Define a stochastic block model with “bad graphs” removed as follows: (1) sample $G \sim \mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon)$; (2) for each $1 \leq i \leq M$ such that $B_i \subset G$, we independently uniformly remove one edge in B_i . The unremoved edges in G constitute a graph G' , which is the output of our modified

stochastic block model. Clearly, from this definition G' does not contain any cycle of length at most N nor any bad subgraph with at most D^3 vertices. Conditioned on G' and π_* , we define

$$A'_{i,j} = G'_{i,j} J'_{i,j}, B'_{i,j} = G'_{\pi_*^{-1}(i), \pi_*^{-1}(j)} K'_{i,j},$$

where J' and K' are independent Bernoulli variables with parameter s . Let $\tilde{\mathbb{P}}_* = \tilde{\mathbb{P}}_{*,n}$ be the law of $(\sigma_*, \pi_*, G, G', A', B')$ and denote $\tilde{\mathbb{P}} = \tilde{\mathbb{P}}_n$ the marginal law of (A', B') .

It was shown in [9, Lemmas 4.2 and 4.4] that

$$\mathbb{P}_*(\mathcal{E}) = \Omega(1) \text{ and } \text{TV}(\tilde{\mathbb{P}}, \mathbb{P}(\cdot | \mathcal{E})) = o(1).$$

Similarly as in Section 3.1, our first step is to show the following lemma.

► **Lemma 21.** *We have $\text{Adv}_{\leq D}(\frac{\tilde{\mathbb{P}}(\cdot | \pi_*(i)=j)}{\text{d}\mathbb{Q}}) = O_{\delta,k}(1)$.*

Based on Lemma 21, we can deduce our main result just as how we deduce Theorem 14 from Lemma 17. The only difference is that we will replace all $\bar{\mathbb{P}}$ with $\tilde{\mathbb{P}}$ and replace all \mathbb{P} with $\mathbb{P}(\cdot | \mathcal{E})$ so we omit further details here.

4 Detection in correlated SBMs

In this section, we will use the framework we established in Section 2 in stochastic block models below KS-threshold. The main results of this section is incorporated as follows.

► **Theorem 22.** *For any constant $K \in \mathbb{N}$, denote \mathbb{P} to be the law of K independent stochastic block models $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon)$ and denote \mathbb{Q} to be the law of K independent Erdős-Rényi graphs $\mathcal{G}(n, \frac{\lambda}{n})$. Then, assuming Conjecture 9, for any $\delta > 0$ there exists $\lambda_0 = \lambda_0(\delta, k)$ to be a sufficiently large constant such that when $\epsilon^2 \lambda < 1 - \delta$ and $\lambda > \lambda_0$, we have $\mathbb{P} \triangleleft \triangleright_{\leq D} \mathbb{Q}$ for any $D = n^{o(1)}$.*

Our result has an immediate corollary in the detection problem between a pair of correlated SBMs and a pair of independent SBMs, as incorporated in the following corollary.

► **Corollary 23.** *Assuming Conjecture 9, when $\epsilon^2 \lambda s < 1 - \delta$, $s < \sqrt{\alpha} - \delta$ and $\lambda > \lambda_0(\delta, k)$ there is no algorithms with polynomial running time that can strongly distinguish $\mathcal{S}(n, \frac{\lambda}{n}; k, \epsilon; s)$ and two independent $\mathcal{S}(n, \frac{\lambda s}{n}; k, \epsilon)$.*

The rest part of this section is devoted to the proof of Theorem 22. For notational simplicity, in the following we will only prove the case where $K = 1$ and the proof for general K is similar. Note that $\epsilon^2 \lambda < 1 - \delta$ and $\lambda > \lambda_0$ implies that $\epsilon < \epsilon_0 = \lambda_0^{-1/2}$. We choose $\lambda_0 = \lambda_0(\delta, k)$ to be a sufficient large constant such that

$$(1 - \delta)^{-1} \cdot \frac{(k-1)\sqrt{1-\epsilon} + \sqrt{1+\epsilon(k-1)}}{k} \geq (1 - \delta/2)^{-1} \text{ for all } \epsilon < \epsilon_0 = \lambda_0^{-1/2}. \quad (13)$$

In the rest part of this section we will always assume that

$$D = n^{o(1)}, \lambda > \lambda_0 \text{ and } \epsilon^2 \lambda < 1 - \delta \text{ for some constant } 0 < \delta < 0.01. \quad (14)$$

Clearly, using Theorem 11, it suffices to show that under (14) and λ_0 we have

$$\text{Adv}_{\leq D}(\frac{\text{d}\mathbb{P}}{\text{d}\mathbb{Q}}) = O_{\delta,k}(1) \text{ and } \text{Adv}_{\leq D}(\frac{\text{d}\mathbb{Q}}{\text{d}\mathbb{P}}) = O_{\delta,k}(1).$$

30:14 Algorithmic Contiguity from Low-Degree Conjecture

Indeed, it has been shown in [34] that $\text{Adv}_{\leq D}(\frac{d\mathbb{P}}{d\mathbb{Q}}) = O_{\delta,k}(1)$ provided with (14). It remains to show that under (14) we have (note that now \mathbb{P} is the planted measure)

$$\text{Adv}_{\leq D}(\frac{d\mathbb{Q}}{d\mathbb{P}}) = \sup_{f \in \mathcal{P}_D} \frac{\mathbb{E}_{\mathbb{Q}}[f]}{\sqrt{\mathbb{E}_{\mathbb{P}}[f^2]}} = O_{\delta,k}(1). \quad (15)$$

We point out that our approach to proving (15) is based on the the work [56]. To this end, define $\omega(\sigma_i, \sigma_j) = k-1$ for $\sigma_i = \sigma_j$ and $\omega(\sigma_i, \sigma_j) = -1$ otherwise. In addition, for all $S \in \mathcal{K}_n$ define

$$\phi_S(\{G_{i,j}\}) = \prod_{(i,j) \in E(S)} \frac{G_{i,j} - \frac{\lambda}{n}}{\sqrt{\frac{\lambda}{n}(1 - \frac{\lambda}{n})}}. \quad (16)$$

It is well known in [34] that $\{\phi_S : S \in \mathcal{K}_n, |E(S)| \leq D\}$ constitutes a standard orthogonal basis of \mathcal{P}_D under \mathbb{Q} . Thus, each $f \in \mathcal{P}_D$ can be written as

$$f(G) = \sum_{S \in \mathcal{K}_n, |E(S)| \leq D} \hat{f}_S \cdot \phi_S(G), \quad (17)$$

which means that f is uniquely characterized by a vector \hat{f} indexed by $\{S \in \mathcal{K}_n : |E(S)| \leq D\}$. In addition, direct calculation yields that $\mathbb{E}_{\mathbb{Q}}[\phi_S(G)] = \mathbf{1}_{\{S=\emptyset\}}$. Thus, we have $\mathbb{E}_{\mathbb{Q}}[f] = \hat{f}_{\emptyset} = \langle \hat{f}, c \rangle$, where c is a vector indexed by $\{S \in \mathcal{K}_n : |E(S)| \leq D\}$ with $c_S = \mathbf{1}_{\{S=\emptyset\}}$. We now turn to $\mathbb{E}_{\mathbb{P}}[f^2]$. For any $\sigma \in [k]^n$ and $S \in \mathcal{K}_n$, define

$$\psi_{\sigma,S}(\{G_{i,j}\}) = k^{\frac{n}{2}} \mathbf{1}_{\sigma_* = \sigma} \cdot \prod_{(i,j) \in E(S)} \frac{G_{i,j} - \frac{(1+\epsilon\omega(\sigma_i, \sigma_j))\lambda}{n}}{\sqrt{\frac{(1+\epsilon\omega(\sigma_i, \sigma_j))\lambda}{n}(1 - \frac{(1+\epsilon\omega(\sigma_i, \sigma_j))\lambda}{n})}} \quad (18)$$

We can check that $\{\psi_{\sigma,S} : \sigma \in [k]^n, S \in \mathcal{K}_n\}$ is standard orthogonal under \mathbb{P}_* , i.e., we have

$$\mathbb{E}_{\mathbb{P}_*}[\psi_{\sigma,S} \psi_{\sigma',S'}] = \mathbf{1}_{\{\sigma=\sigma', S=S'\}}. \quad (19)$$

► **Lemma 24.** *We have*

$$\mathbb{E}_{\mathbb{P}}[\phi_S(G) \psi_{\sigma,H}(G)] = \frac{\mathbf{1}_{H \subset S}}{k^{\frac{n}{2}}} \prod_{(i,j) \in E(H)} \mathfrak{h}(\sigma_i, \sigma_j) \prod_{(i,j) \in E(S) \setminus E(H)} \omega(\sigma_i, \sigma_j) \sqrt{\frac{\epsilon^2 \lambda}{n}}, \quad (20)$$

where

$$\mathfrak{h}(\sigma_i, \sigma_j) = \sqrt{\frac{(1 - \frac{(1+\epsilon\omega(\sigma_i, \sigma_j))\lambda}{n})(1 + \epsilon\omega(\sigma_i, \sigma_j))}{1 - \frac{\lambda}{n}}}. \quad (21)$$

Based on Lemma 24, we define a matrix M with rows indexed by $\{S : S \in \mathcal{K}_n, |E(S)| \leq D\}$ and columns indexed by $\{(\sigma, S) : \sigma \in [k]^n, S \in \mathcal{K}_n, |E(S)| \leq D\}$, and entries given by

$$M_{S;(\sigma,H)} = \frac{\mathbf{1}_{H \subset S}}{k^{\frac{n}{2}}} \prod_{(i,j) \in E(H)} \mathfrak{h}(\sigma_i, \sigma_j) \prod_{(i,j) \in E(S) \setminus E(H)} \omega(\sigma_i, \sigma_j) \sqrt{\frac{\epsilon^2 \lambda}{n}}. \quad (22)$$

From Parserval's inequality, we see that

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[f^2] &\geq \sum_{\substack{\sigma \in [k]^n, H \in \mathcal{K}_n \\ |E(H)| \leq D}} \mathbb{E}_{\mathbb{P}}[f \cdot \psi_{\sigma,H}]^2 \stackrel{(17)}{=} \sum_{\substack{\sigma \in [k]^n, H \in \mathcal{K}_n \\ |E(H)| \leq D}} \left(\sum_{\substack{S \in \mathcal{K}_n \\ |E(S)| \leq D}} \hat{f}_S \mathbb{E}_{\mathbb{P}}[\phi_S \cdot \psi_{\sigma,H}] \right)^2 \\ &= \sum_{\substack{\sigma \in [k]^n, H \in \mathcal{K}_n \\ |E(H)| \leq D}} \left(\sum_{\substack{S \in \mathcal{K}_n \\ |E(S)| \leq D}} \hat{f}_S M_{S;(\sigma,H)} \right)^2 = \|\hat{f}M\|^2. \end{aligned} \quad (23)$$

Thus, we have

$$\text{Adv}_{\leq D}(\frac{d\mathbb{Q}}{d\mathbb{P}}) = \sup_{f \in \mathcal{P}_D} \left\{ \frac{\mathbb{E}_{\mathbb{Q}}[f]}{\sqrt{\mathbb{E}_{\mathbb{P}}[f^2]}} \right\} \leq \sup_{\hat{f}} \left\{ \frac{\langle \hat{f}, c \rangle}{\|\hat{f}M\|} \right\} \leq \inf_{Mu^\top = c} \{\|u\|\}, \quad (24)$$

where the last inequality follows from the fact that for $Mu_0^\top = c$ we have

$$\langle \hat{f}, c \rangle = \langle \hat{f}, Mu_0^\top \rangle = \langle \hat{f}M, u_0 \rangle \leq \|u_0\| \cdot \|\hat{f}M\|.$$

Regarding (24), it suffices to show that there exists $Mu^\top = c$ and $\|u\| = O_{\delta,k}(1)$. We now construct the solution $\{u_{\sigma,H} : \sigma \in [k]^n, H \in \mathcal{K}_n, |E(H)| \leq D\}$ as follows: let $u_{\sigma,H} = \frac{1}{k^{\frac{\alpha}{2}}} \cdot \Xi(H)$, where

$$\Xi(\emptyset) = 1 \text{ and } \Xi(S) = 0 \text{ for } \mathcal{L}(S) \neq \emptyset \quad (25)$$

and then iteratively define for all $\mathcal{L}(S) = \emptyset$ by

$$\begin{aligned} \Xi(S) = & - \left(\mathbb{E}_{\sigma \sim \nu} \left[\prod_{(i,j) \in E(S)} \mathfrak{h}(\sigma_i, \sigma_j) \right] \right)^{-1} \sum_{\substack{H \subset S \\ \mathcal{L}(H) = \emptyset}} \left(\frac{\epsilon^2 \lambda}{n} \right)^{\frac{|E(S)| - |E(H)|}{2}} \Xi(H) \\ & \times \mathbb{E}_{\sigma \sim \nu} \left[\prod_{(i,j) \in E(H)} \mathfrak{h}(\sigma_i, \sigma_j) \prod_{(i,j) \in E(S) \setminus E(H)} \omega(\sigma_i, \sigma_j) \right], \end{aligned} \quad (26)$$

To prove Theorem 22, it suffices to show the following lemma.

► **Lemma 25.** *The vector $\{u_{\sigma,H} : \sigma \in [k]^n, H \in \mathcal{K}_n, |E(H)| \leq D\}$ satisfies $Mu^\top = c$ and $\|u\| = O_{\delta,k}(1)$.*

References

- 1 László Babai. Graph isomorphism in quasipolynomial time. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, pages 684–697. ACM, 2016.
- 2 Afonso S. Bandeira, Ahmed El Alaoui, Samuel B. Hopkins, Tselil Schramm, Alexander S. Wein, and Ilias Zadik. The Franz-Parisi criterion and computational trade-offs in high dimensional statistics. In *Advances in Neural Information Processing Systems (NIPS)*, volume 35, pages 33831–33844. Curran Associates, Inc., 2022.
- 3 Afonso S. Bandeira, Dmitriy Kunisky, and Alexander S. Wein. Computational hardness of certifying bounds on constrained PCA problems. In *Proceedings of the 11th Innovations in Theoretical Computer Science Conference (ITCS)*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.ITCS.2020.78.
- 4 Boaz Barak, Chi-Ning Chou, Zhixian Lei, Tselil Schramm, and Yueqi Sheng. (Nearly) efficient algorithms for the graph matching problem on correlated random graphs. In *Advances in Neural Information Processing Systems (NIPS)*, volume 32, pages 9190–9198. Curran Associates, Inc., 2019.
- 5 Boaz Barak, Samuel B. Hopkins, Jonathan Kelner, Pravesh K. Kothari, Ankur Moitra, and Aaron Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM Journal on Computing*, 48(2):687–735, 2019. doi:10.1137/17M1138236.
- 6 Guy Bresler and Brice Huang. The algorithmic phase transition of random k -SAT for low degree polynomials. In *Proceedings of the IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 298–309. IEEE, 2022. doi:10.1109/FOCS52979.2021.00038.
- 7 Rainer E. Burkard, Eranda Cela an Panos M. Pardalos, and Leonidas S. Pitsoulis. The quadratic assignment problem. *Handbook of combinatorial optimization*, pages 1713–1809, 1998.

- 8 Shuwen Chai and Miklós Z. Rácz. Efficient graph matching for correlated stochastic block models. In *Advances in Neural Information Processing Systems (NIPS)*, volume 38, pages 116388–116461. Curran Associates, Inc., 2024.
- 9 Guanyi Chen, Jian Ding, Shuyang Gong, and Zhangsong Li. A computational transition for detecting correlated stochastic block models by low-degree polynomials. *arXiv preprint*, 2024. [arXiv:2409.00966](https://arxiv.org/abs/2409.00966).
- 10 Daniel Cullina and Negar Kiyavash. Exact alignment recovery for correlated Erdős-Rényi graphs. *arXiv preprint*, 2024. [arXiv:1711.06783](https://arxiv.org/abs/1711.06783).
- 11 Osman E. Dai, Daniel Cullina, Negar Kiyavash, and Matthias Grossglauser. Analysis of a canonical labeling algorithm for the alignment of correlated Erdős-Rényi graphs. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 3(2):1–25, 2019. doi:10.1145/3341617.3326151.
- 12 Abhishek Dhawan, Cheng Mao, and Alexander S. Wein. Detection of dense subhypergraphs by low-degree polynomials. *arXiv preprint*, 2024. to appear in *Random Structures and Algorithms*. [arXiv:2304.08135](https://arxiv.org/abs/2304.08135).
- 13 Jian Ding and Hang Du. Detection threshold for correlated Erdős-Rényi graphs via densest subgraph. *IEEE Transactions on Information Theory*, 69(8):5289–5298, 2023. doi:10.1109/TIT.2023.3265009.
- 14 Jian Ding and Hang Du. Matching recovery threshold for correlated random graphs. *Annals of Statistics*, 51(4):1718–1743, 2023.
- 15 Jian Ding, Hang Du, and Zhangsong Li. Low-degree hardness of detection for correlated Erdős-Rényi graphs. *arXiv preprint*, 2023. to appear in *Annals of Statistics*. [arXiv:2311.15931](https://arxiv.org/abs/2311.15931).
- 16 Jian Ding and Zhangsong Li. A polynomial time iterative algorithm for matching Gaussian matrices with non-vanishing correlation. *arXiv preprint*, 2023. to appear in *Foundations of Computational Mathematics*. [arXiv:2212.13677](https://arxiv.org/abs/2212.13677).
- 17 Jian Ding and Zhangsong Li. A polynomial-time iterative algorithm for random graph matching with non-vanishing correlation. *arXiv preprint*, 2023. [arXiv:2306.00266](https://arxiv.org/abs/2306.00266).
- 18 Jian Ding, Zongming Ma, Yihong Wu, and Jiaming Xu. Efficient random graph matching via degree profiles. *Probability Theory and Related Fields*, 179(1-2):29–115, 2021.
- 19 Jingqiu Ding, Yiding Hua, Lucas Slot, and David Steurer. Low degree conjecture implies sharp computational thresholds in stochastic block model. *arXiv preprint*, 2025. [arXiv:2502.15024](https://arxiv.org/abs/2502.15024).
- 20 Yunzi Ding, Dmitriy Kunisky, Alexander S. Wein, and Afonso S. Bandeira. Subexponential-time algorithms for sparse PCA. *Foundations of Computational Mathematics*, 22(1):1–50, 2022.
- 21 Hang Du. Optimal recovery of correlated Erdős-Rényi graphs. *arXiv preprint*, 2025. [arXiv:2502.12077](https://arxiv.org/abs/2502.12077).
- 22 Zhou Fan, Cheng Mao, Yihong Wu, and Jiaming Xu. Spectral graph matching and regularized quadratic relaxations I: Algorithm and Gaussian analysis. *Foundations of Computational Mathematics*, 23(5):1511–1565, 2023. doi:10.1007/S10208-022-09570-Y.
- 23 Zhou Fan, Cheng Mao, Yihong Wu, and Jiaming Xu. Spectral graph matching and regularized quadratic relaxations II: Erdős-Rényi graphs and universality. *Foundations of Computational Mathematics*, 23(5):1567–1617, 2023.
- 24 David Gamarnik. The overlap gap property: A topological barrier to optimizing over random structures. *Proceedings of the National Academy of Sciences of the United States of America*, 118(41):e2108492118, 2021.
- 25 David Gamarnik, Aukosh Jagannath, and Alexander S. Wein. Low-degree hardness of random optimization problems. In *Proceedings of the IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 131–140. IEEE, 2020. doi:10.1109/FOCS46700.2020.00021.
- 26 Luca Ganassali and Laurent Massoulié. From tree matching to sparse graph alignment. In *Proceedings of the 33rd Conference on Learning Theory (COLT)*, pages 1633–1665. PMLR, 2020. URL: <http://proceedings.mlr.press/v125/ganassali20a.html>.

- 27 Luca Ganassali, Laurent Massoulié, and Marc Lelarge. Impossibility of partial recovery in the graph alignment problem. In *Proceedings of the 34th Conference on Learning Theory (COLT)*, pages 2080–2102. PMLR, 2021. URL: <http://proceedings.mlr.press/v134/ganassali21a.html>.
- 28 Luca Ganassali, Laurent Massoulié, and Marc Lelarge. Correlation detection in trees for planted graph alignment. *Annals of Applied Probability*, 34(3):2799–2843, 2024.
- 29 Luca Ganassali, Laurent Massoulié, and Guilhem Semerjian. Statistical limits of correlation detection in trees. *Annals of Applied Probability*, 34(4):3701–3734, 2024.
- 30 Julia Gaudio, Miklós Z. Rácz, and Anirudh Sridhar. Exact community recovery in correlated stochastic block models. In *Proceedings of the 35th Conference on Learning Theory (COLT)*, pages 2183–2241. PMLR, 2022. URL: <https://proceedings.mlr.press/v178/gaudio22a.html>.
- 31 Georgina Hall and Laurent Massoulié. Partial recovery in the graph alignment problem. *Operations Research*, 71(1):259–272, 2023. doi:10.1287/OPRE.2022.2355.
- 32 Samuel B. Hopkins. *Statistical inference and the sum of squares method*. Cornell University, 2018. PHD thesis.
- 33 Samuel B. Hopkins, Pravesh K. Kothari, Aaron Potetchin, Prasad Raghavendra, Tselil Schramm, and David Steurer. The power of sum-of-squares for detecting hidden structures. In *Proceedings of the IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 720–731. IEEE, 2017. doi:10.1109/FOCS.2017.72.
- 34 Samuel B. Hopkins and David Steurer. Efficient Bayesian estimation from few samples: community detection and related problems. In *Proceedings of the IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 379–390. IEEE, 2017. doi:10.1109/FOCS.2017.42.
- 35 Harry Kesten and Bernt P. Stigum. Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *Annals of Mathematical Statistics*, 37:1463–1481, 1966.
- 36 Pravesh K. Kothari, Santosh S. Vempala, Alexander S. Wein, and Jeff Xu. Is planted coloring easier than planted clique? In *Proceedings of the 36th Conference on Learning Theory (COLT)*, pages 5343–5372. PMLR, 2023. URL: <https://proceedings.mlr.press/v195/kothari23a.html>.
- 37 Dmitriy Kunisky. Hypothesis testing with low-degree polynomials in the Morris class of exponential families. In *Proceedings of the 34th Conference on Learning Theory (COLT)*, pages 2822–2848. PMLR, 2021. URL: <http://proceedings.mlr.press/v134/kunisky21a.html>.
- 38 Dmitriy Kunisky, Cristopher Moore, and Alexander S. Wein. Tensor cumulants for statistical inference on invariant distributions. In *Proceedings of the IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1007–1026. IEEE, 2024. doi:10.1109/FOCS61266.2024.00067.
- 39 Dmitriy Kunisky, Alexander S. Wein, and Afonso S. Bandeira. Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. In *Mathematical Analysis, its Applications and Computation: ISAAC*, pages 1–50. Springer, 2022.
- 40 Zhangsong Li. Robust random graph matching in dense graphs via vector approximate message passing. *arXiv preprint*, 2024. arXiv:2412.16457.
- 41 Zhangsong Li. Algorithmic contiguity from low-degree conjecture and applications in correlated random graphs. *arXiv preprint*, 2025. arXiv:arXiv:2502.09832.
- 42 Konstantin Makarychev, Rajsekar Manokaran, and Maxim Sviridenko. Maximum quadratic assignment problem: Reduction from maximum label cover and lp-based approximation algorithm. In *Proceedings of the International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 594–604. Springer, 2010. doi:10.1007/978-3-642-14165-2_50.
- 43 Cheng Mao, Mark Rudelson, and Konstantin Tikhomirov. Random graph matching with improved noise robustness. In *Proceedings of the 34th Conference on Learning Theory (COLT)*, pages 3296–3329. PMLR, 2021. URL: <http://proceedings.mlr.press/v134/mao21a.html>.

- 44 Cheng Mao, Mark Rudelson, and Konstantin Tikhomirov. Exact matching of random graphs with constant correlation. *Probability Theory and Related Fields*, 186(2):327–389, 2023.
- 45 Cheng Mao, Yihong Wu, Jiaming Xu, and Sophie H. Yu. Random graph matching at Otter’s threshold via counting chandeliers. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1345–1356. ACM, 2023. doi:10.1145/3564246.3585156.
- 46 Cheng Mao, Yihong Wu, Jiaming Xu, and Sophie H. Yu. Testing network correlation efficiently via counting trees. *Annals of Statistics*, 52(6):2483–2505, 2024.
- 47 Ankur Moitra and Alexander S. Wein. Precise error rates for computationally efficient testing. *Annals of Statistics*, 53(2):854–878, 2023.
- 48 Andrea Montanari and Alexander S. Wein. Equivalence of approximate message passing and low-degree polynomials in rank-one matrix estimation. *Probability Theory and Related Fields*, 191(1-2):181–233, 2025.
- 49 Arvind Narayanan and Vitaly Shmatikov. Robust de-anonymization of large sparse datasets. In *IEEE Symposium on Security and Privacy*, pages 111–125. IEEE, 2008. doi:10.1109/SP.2008.33.
- 50 Richard Otter. The number of trees. *Annals of Mathematics*, 49(3):583–599, 1948.
- 51 Miklós Z. Rácz and Anirudh Sridhar. Correlated stochastic block models: exact graph matching with applications to recovering communities. In *Advances in Neural Information Processing Systems (NIPS)*, volume 34, pages 22259–22273. Curran Associates, Inc., 2021. URL: <https://proceedings.neurips.cc/paper/2021/hash/baf4f1a5938b8d520b328c13b51ccf11-Abstract.html>.
- 52 Prasad Raghavendra, Tselil Schramm, and David Steurer. High-dimensional estimation via sum-of-squares proofs. In *Proceedings of the International Congress of Mathematicians (ICM 2018)*, pages 3389–3423, 2019.
- 53 Cynthia Rush, Fiona Skerman, Alexander S. Wein, and Dana Yang. Is it easier to count communities than find them? In *Proceedings of the 14th Innovations in Theoretical Computer Science Conference (ITCS)*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.ITCS.2023.94.
- 54 Tselil Schramm and Alexander S. Wein. Computational barriers to estimation from low-degree polynomials. *Annals of Statistics*, 50(3):1833–1858, 2022.
- 55 Rohit Singh, Jinbo Xu, and Bonnie Berger. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proceedings of the National Academy of Sciences of the United State of America*, 105(35):12763–12768, 2008. doi:10.1073/PNAS.0806627105.
- 56 Youngtak Sohn and Alexander S. Wein. Sharp phase transitions in estimation with low-degree polynomials. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC)*, pages 891–902. ACM, 2025. doi:10.1145/3717823.3718294.
- 57 Alexander S. Wein. Optimal low-degree hardness of maximum independent set. *Mathematical Statistics and Learning*, 4(3/4):221–251, 2022.
- 58 Yihong Wu, Jiaming Xu, and Sophie H. Yu. Settling the sharp reconstruction thresholds of random graph matching. *IEEE Transactions on Information Theory*, 68(8):5391–5417, 2022. doi:10.1109/TIT.2022.3169005.
- 59 Yihong Wu, Jiaming Xu, and Sophie H. Yu. Testing correlation of unlabeled random graphs. *Annals of Applied Probability*, 33(4):2519–2558, 2023.
- 60 Ilias Zadik, Min Jae Song, Alexander S. Wein, and Joan Bruna. Lattice-based methods surpass sum-of-squares in clustering. In *Proceedings of the 35th Conference on Learning Theory (COLT)*, pages 1247–1248. PMLR, 2022. URL: <https://proceedings.mlr.press/v178/zadik22a.html>.
- 61 Lenka Zdeborová and Florent Krzakala. Statistical physics of inference: thresholds and algorithms. *Advances in Physics*, 65(5):453–552, 2016.