

On Statistical-Computational Gaps in Graph Combinatorial Testing Problems

Ikjun Choi *

ikjunchoi@utexas.edu

Pratik Patil *

pratikpatil@utexas.edu

Alessandro Rinaldo *

alessandro.rinaldo@austin.utexas.edu

Abstract

We study high-dimensional graph combinatorial hypothesis testing problems in which a signal is supported on the edge set of an unknown subgraph drawn from a class invariant under vertex relabeling. For Gaussian observations, we characterize sharp statistical and computational thresholds (at polynomial order) for the signal magnitude required for strong detection across various combinatorial graph alternatives. Our analysis is based on bounding the second moments of the full and low-degree likelihood ratios in terms of the overlap of two independent latent structures and comparing the resulting lower bounds with scan, average, and coordinatewise maximum tests. Under an appropriate low-degree conjecture for computational hardness, we obtain planted-clique thresholds matching those known for Gaussian submatrix detection and identify new statistical-computational gaps for multiple planted cliques and hypercliques, thick cycles and thick paths. We also establish a novel transfer principle across natural exponential family observation models: matching the one-coordinate chi-square divergence preserves full likelihood-ratio second moments exactly and makes low-degree second moments equivalent up to constant factors under a degree condition. Under local regularity, we obtain corresponding testing upper bounds in terms of an effective signal strength. When these upper and lower bounds match, they yield universal threshold scales in the local alternative regime. As a consequence, statistical-computational gaps for planted-clique detection in Erdős–Rényi settings can be deduced from the Gaussian model.

1 Introduction

Signal detection problems are central to modern statistics and arise in applications ranging from network analysis and neuroscience to genomics and spatial statistics. A recurring feature of these problems is that a weak signal is embedded in a high-dimensional ambient space, but is present only on an unknown subset of coordinates with prescribed combinatorial structure. In graph-valued data, for example, the signal may be supported on the edge set of an unknown clique, collection of cliques, path, cycle, or other structured subgraph. A scan over all admissible structures, or an equivalent generalized likelihood-ratio test, often achieves rate-optimal detection but can be computationally infeasible when the class of possible supports is large. This raises a basic question: when can statistically optimal detection be achieved by computationally efficient tests?

A growing body of work shows that the answer can depend sharply on the geometry of the latent structure. For many high-dimensional testing problems, there is a range of signal strengths in which detection is information-theoretically possible but no known polynomial-time method succeeds. Such a *statistical-computational gap* is most familiar from planted clique and related subgraph detection problems, but analogous phenomena appear across sparse principal component analysis, submatrix detection, tensor problems, and structured normal means models. Despite this broad picture, the mechanisms that determine the size of the gap remain only partially understood for many graph combinatorial testing problems, especially under observation models beyond the classical Erdős–Rényi setting.

*Department of Statistics and Data Sciences, University of Texas at Austin, TX 78712, USA.

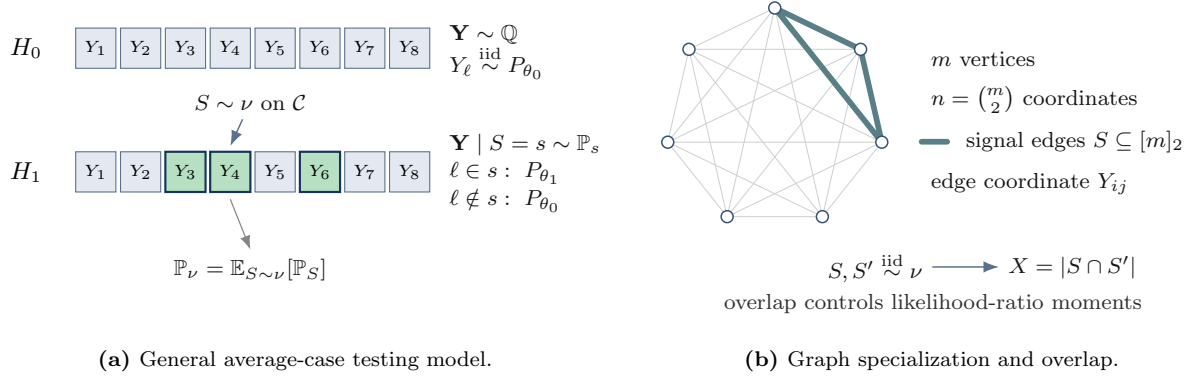


Figure 1: Testing framework. Panel (a) shows the general mixture model: under H_0 , the coordinates follow P_{θ_0} , whereas under H_1 , a latent structure $S \sim \nu$ determines which coordinates follow P_{θ_1} instead of P_{θ_0} , and mixing over S yields $\mathbb{P}_\nu$. Panel (b) specializes the framework to edge-indexed observations on K_m ; the overlap $X = |S \cap S'|$, for $S, S' \stackrel{\text{iid}}{\sim} \nu$, controls the likelihood-ratio moments used in our lower bounds.

This paper studies these questions through a common testing framework for signals supported on unknown graph-indexed combinatorial structures. Our starting point is the Gaussian observation model, in which the full and low-degree likelihood-ratio second moments can both be expressed in terms of the overlap distribution of two independent latent structures. The former yields information-theoretic lower bounds, while the latter provides evidence for computational lower bounds under the low-degree heuristic. We combine these lower bounds with standard scan, average, and coordinatewise maximum tests to identify statistical and conjectural computational thresholds for several graph testing problems. We also show that the same overlap calculations extend beyond Gaussian noise to natural exponential family models by matching the one-coordinate chi-square divergence.

1.1 Summary of results

We study sequences of combinatorial hypothesis-testing problems indexed by the number $n = \binom{m}{2}$ of edges in K_m , the complete graph on m vertices. The null hypothesis represents unstructured noise, whereas the alternative is a mixture of structured distributions indexed by a class of edge-induced subgraphs of K_m ; Figure 1 illustrates this distributional hierarchy and its graph specialization. Let $\{P_\theta : \theta \in \Theta\}$ be a parametric family of distributions on a sample space $\Omega \subseteq \mathbb{R}$, where $\Theta \subseteq \mathbb{R}$ is open. We observe a random vector $\mathbf{Y} = (Y_1, \dots, Y_n) \in \Omega^n$ whose coordinates are indexed by the edges of K_m . Under the null hypothesis, the coordinates of $\mathbf{Y}$ are independent draws from a reference distribution P_{θ_0} , for $\theta_0 \in \Theta$. Under the alternative, a latent edge set is sampled from a combinatorial class $\mathcal{C} \subseteq 2^{[n]}$, and the coordinates indexed by that set instead follow P_{θ_1} for some $\theta_1 \neq \theta_0$. More precisely, for each fixed structure $s \in \mathcal{C}$, let $\mathbb{P}_s$ denote the distribution under which the coordinates of $\mathbf{Y}$ are independent and

$$Y_\ell \sim \begin{cases} P_{\theta_1}, & \ell \in s, \\ P_{\theta_0}, & \ell \notin s, \end{cases} \quad \ell \in [n]. \quad (1)$$

Let ν be a mixing measure on $\mathcal{C}$, taken to be uniform in the basic fixed-size models. The alternative distribution is the mixture $\mathbb{P}_\nu$ of $\{\mathbb{P}_s : s \in \mathcal{C}\}$ with respect to ν , as illustrated in Figure 1a; that is,

$$\mathbb{P}_\nu(A) := \mathbb{E}_{S \sim \nu}[\mathbb{P}_S(A)], \quad A \subseteq \Omega^n \text{ measurable}. \quad (2)$$

The class $\mathcal{C}$ controls the combinatorial complexity of the alternative and can be viewed as a collection of subgraphs of K_m indexing the possible anomalous coordinates of $\mathbf{Y}$. The corresponding edge-indexed graph model and the overlap variable used in our lower-bound arguments appear in Figure 1b.

We study the resulting graph combinatorial testing problem

$$H_0 : \mathbf{Y} \sim \mathbb{Q} \quad \text{versus} \quad H_1 : \mathbf{Y} \sim \mathbb{P}_\nu, \quad (3)$$

where $\mathbb{Q} = P_{\theta_0}^{\otimes n}$ is the n -fold product of the reference distribution. We consider asymptotic regimes in which the dimension n diverges and the combinatorial class $\mathcal{C}$ and signal parameter θ_1 may vary with n . The reference parameter θ_0 is fixed in the Gaussian model and in the local alternative universality results of Section 4.3. Our general bounds for natural exponential family models also allow θ_0 to vary with n , as in the sparse Erdős–Rényi applications in Section 4.3.1. Our primary interest is in regimes where the separation between the null and alternative vanishes as n grows. We suppress dependence on n when it is clear from context.

A test is a measurable function $\Delta : \Omega^n \rightarrow \{0, 1\}$, with $\Delta(\mathbf{Y}) = 1$ indicating rejection of the null hypothesis. We measure its performance by the average-case risk

$$R_\nu(\Delta) := \mathbb{Q}\{\Delta(\mathbf{Y}) = 1\} + \mathbb{P}_\nu\{\Delta(\mathbf{Y}) = 0\}, \quad (4)$$

the sum of its type I error and its type II error averaged over the latent structure.

Definition 1 (Strong detection, statistical and computational distinguishability). *Consider a sequence of testing problems $H_0 : \mathbf{Y} \sim \mathbb{Q}_n$ versus $H_1 : \mathbf{Y} \sim \mathbb{P}_{\nu_n}$. A sequence of tests $\Delta_n : \Omega^n \rightarrow \{0, 1\}$ achieves strong detection if*

$$R_{\nu_n}(\Delta_n) \rightarrow 0.$$

The sequence of testing problems is statistically distinguishable if there exists a sequence of tests achieving strong detection, and computationally distinguishable if such a sequence can be evaluated in time polynomial in n . If no sequence of tests achieves strong detection, the testing problems are statistically indistinguishable; if no polynomial-time sequence does so, they are computationally indistinguishable.

Remark 1 (Strong vs weak detection). For comparison, a sequence of tests Δ_n achieves *weak detection* if $\limsup_{n \rightarrow \infty} R_{\nu_n}(\Delta_n) < 1$, meaning that it achieves a nonvanishing improvement over the trivial risk of one. Our distinguishability statements and threshold results concern strong detection.

We review the second-moment and low-degree likelihood-ratio methods used to establish statistical and computational lower bounds in Section 2.1.

Our goal is to characterize, as a function of n and the complexity of $\mathcal{C}$, the signal scale at which statistical and computational distinguishability emerge. Most of the paper studies the Gaussian sequence model of [Addario-Berry et al. \(2010\)](#), in which $P_{\theta_0} = N(0, 1)$ and $P_{\theta_1} = N(\mu, 1)$ for some nonzero μ . In this model, the separation between the hypotheses is measured by the signal strength $|\mu|$, and we denote the statistical and computational thresholds by μ_{stat}^* and μ_{comp}^* , respectively. These thresholds, defined formally in Definition 3, identify the signal scales for strong detection up to subpolynomial factors. Our low-degree lower bounds rule out strong separation by polynomial statistics of polylogarithmic degree (Definition 2); under Conjecture 1, we interpret low-degree indistinguishability as evidence that no polynomial-time test achieves strong detection.

Our main results for the Gaussian sequence model are collected in Table 1. The threshold scales summarize the polynomially separated parameter regimes (see Section 1.3 for the definition of polynomial separation) specified in the corresponding results. As in [Addario-Berry et al. \(2010\)](#), we focus on combinatorial classes $\mathcal{C}$ that are invariant under permutations of the vertices of K_m . For the nonclique classes studied there, the full and low-degree likelihood-ratio lower bounds match polynomial-time testing upper bounds at polynomial order, and hence no statistical-computational gap arises; the corresponding results appear in the appendix sections listed in the first block of Table 1. We obtain analogous no-gap results for k -stars and for paths and cycles of length k . By contrast, Section 3 identifies statistical-computational gaps for planted cliques, multiple planted cliques and hypercliques, and thick cycles and paths.

For k -vertex planted cliques, Theorem 1 identifies the statistical threshold $\mu_{\text{stat}}^* \asymp k^{-1/2} \wedge m/k^2$ and the conjectural computational threshold $\mu_{\text{comp}}^* \asymp 1 \wedge m/k^2$; see Figure 2. These thresholds agree with those

Table 1: Summary of problems, with thresholds and test optimality stated at polynomial order. The threshold scales summarize the polynomially separated parameter regimes (see Section 1.3 for the definition of polynomial separation) specified in the corresponding results. [†] denotes the problems whose statistical thresholds were studied by Addario-Berry et al. (2010). The statistical and computational thresholds for k -vertex cliques match those for Gaussian submatrix detection, where Ma and Wu (2015) obtained computational lower bounds through a planted-clique reduction. A checkmark indicates a statistical-computational gap (conditional on Conjecture 1) in some polynomial parameter regime.

Problem	Section	μ_{stat}^*	μ_{comp}^*	Rate-optimal test	Polytime test	Comp-stat gap
K -Disjoint sets [†]	Appendix C.1.1	$\sqrt{\frac{1}{K}}$	$\sqrt{\frac{1}{K}}$	scan	scan	×
K -sets [†]	Appendix C.1.2	$\sqrt{\frac{n}{K^2}} \wedge 1$	$\sqrt{\frac{n}{K^2}} \wedge 1$	scan, average	scan, average	×
Perfect matching [†]	Appendix C.1.3	1	1	average	average	×
Star [†]	Appendix C.1.4	$\sqrt{\frac{1}{m}}$	$\sqrt{\frac{1}{m}}$	scan	scan	×
Spanning tree [†]	Appendix C.1.5	1	1	average	average	×
k -vertex star	Appendix C.2.1	$\sqrt{\frac{m}{k^2}} \wedge 1$	$\sqrt{\frac{m}{k^2}} \wedge 1$	degree, outlier	degree, outlier	×
k -length path, cycle	Appendix C.2.2	1	1	outlier	outlier	×
k -vertex clique [†]	Section 3.1	$\frac{m}{k^2} \wedge \sqrt{\frac{1}{k}}$	$\frac{m}{k^2} \wedge 1$	scan, average	outlier, average	✓
γ multiple k -vertex cliques	Section 3.2	$\frac{m}{\gamma k^2} \wedge \sqrt{\frac{1}{k}}$	$\frac{m}{\gamma k^2} \wedge 1$	scan, average	outlier, average	✓
γ multiple k -vertex hypercliques	Section 3.2	$\frac{m^{t/2}}{k^t \gamma} \wedge \sqrt{\frac{1}{k^{t-1}}}$	$\frac{m^{t/2}}{k^t \gamma} \wedge 1$	scan, average	outlier, average	✓
length- ℓ , Δ -thick cycle, path	Section 3.3	$\frac{m}{\ell \Delta} \wedge \sqrt{\frac{1}{\Delta}}$	$\frac{m}{\ell \Delta} \wedge 1$	scan, average	outlier, average	✓

established for Gaussian submatrix detection by Butucea and Ingster (2013); Ma and Wu (2015), with computational lower bounds obtained through a polynomial-time reduction from planted clique under the planted-clique conjecture (Alon et al., 2007). In contrast, our derivation uses the low-degree likelihood-ratio method. Across the Gaussian families summarized in Table 1, the low-degree lower bounds match the best polynomial-time upper bounds at polynomial order, covering both the no-gap examples and the models that exhibit statistical-computational gaps. Corollary 1 shows that the same thresholds hold when the planted clique has a binomially distributed vertex set with mean size k . In the intermediate regime, the exhaustive scan test succeeds below the conjectural computational threshold, while the low-degree obstruction predicts, under Conjecture 1, that no polynomial-time test achieves strong detection.

Theorem 2 extends the analysis to a union of γ disjoint planted k -vertex cliques and gives $\mu_{\text{stat}}^* \asymp k^{-1/2} \wedge m/(\gamma k^2)$ and $\mu_{\text{comp}}^* \asymp 1 \wedge m/(\gamma k^2)$ for both fixed-size and random-size constructions. Corollary 2 treats the corresponding t -uniform hypergraph model, with statistical threshold of order $k^{-(t-1)/2} \wedge m^{t/2}/(\gamma k^t)$ and computational threshold of order $1 \wedge m^{t/2}/(\gamma k^t)$ in the regimes stated there. For the random-size multiple-clique and hyperclique models, we derive the computational lower bounds by comparing the overlap moments with those of auxiliary single-clique and single-hyperclique models, respectively.

For length- ℓ , Δ -thick cycles, Theorem 3 identifies $\mu_{\text{stat}}^* \asymp \Delta^{-1/2} \wedge m/(\ell \Delta)$ and $\mu_{\text{comp}}^* \asymp 1 \wedge m/(\ell \Delta)$. Corollary 3 shows that the same thresholds hold for the corresponding thick path model. The lower-bound argument reduces the relevant overlap calculation to the multiple-clique analysis by decomposing each thick cycle or path into local patches.

Finally, Section 4 extends the framework beyond Gaussian observations to natural exponential families. Theorem 4 shows that two natural exponential family models with matching one-coordinate chi-square divergence have identical full likelihood-ratio second moments and equivalent low-degree second moments under the stated degree condition. Proposition 3 shows that the average and scan upper bounds depend on the observation model through its effective standardized mean shift. Together, these results transfer Gaussian thresholds to models such as Erdős-Rényi observations after the appropriate chi-square calibration and under the stated regularity conditions.

1.2 Related work

Gaussian combinatorial testing and graph scans. The Gaussian framework for testing structured means in this paper is closest to [Addario-Berry et al. \(2010\)](#), who analyze combinatorial classes through scan and averaging tests and through lower bounds based on the intersections of latent supports. Related work on anomalous paths and connected clusters includes [Arias-Castro et al. \(2008, 2011\)](#), while [Sharpnack et al. \(2013, 2016\)](#) develop computationally tractable relaxations of graph scan statistics. Those graph-scan models generally place the signal on vertices or vertex-indexed observations, whereas our principal examples place the signal on the edge set of an unknown subgraph. Distribution-free permutation and rank-based scans were studied by [Arias-Castro et al. \(2018\)](#), including performance guarantees for one-parameter natural exponential families.

Planted cliques and submatrices. The planted-clique problem is a canonical source of statistical-computational gaps, with early evidence for algorithmic limitations due to [Jerrum \(1992\)](#). Classical and subsequent algorithmic upper bounds include the spectral method of [Alon et al. \(1998\)](#) and the nearly linear-time method of [Deshpande and Montanari \(2015a\)](#). Lower bounds for restricted computational models include spectral, sum-of-squares, and statistical-query results ([Deshpande and Montanari, 2015b](#); [Montanari et al., 2017](#); [Barak et al., 2019](#); [Feldman et al., 2017](#)), and [Hirahara and Shimizu \(2024\)](#) establish equivalences among many common formulations of the planted-clique conjecture. For weighted observations, [Chatterjee et al. \(2025\)](#) study a general hidden-clique model and derive information-theoretic bounds and spectral guarantees, while our clique results focus on vanishing signals and on matching overlap calculations for both the full and low-degree likelihood ratios. Sparse submatrix detection has sharp statistical thresholds ([Butucea and Ingster, 2013](#); [Goswami and Mukherjee, 2026](#)), planted-clique-based computational lower bounds ([Ma and Wu, 2015](#)), and distributional universality results based on average-case reductions ([Brennan et al., 2019](#)). The closest submatrix comparison for our multiple-clique results is [Dadon et al. \(2024\)](#), who study detection and recovery for one or multiple hidden Gaussian submatrices using information-theoretic and low-degree methods. Related work addresses localization and heterogeneous multiple submatrices ([Chen and Xu, 2016](#); [Cai et al., 2017](#); [Oren-Loberman et al., 2026](#)). These models allow distinct row and column supports, whereas our clique models impose symmetric principal edge sets.

Multiple structures and hypergraphs. Multiple planted structures have also been studied through growing-cluster recovery guarantees ([Chen and Xu, 2016](#)), low-degree lower bounds for planted colorings or disjoint cliques covering most vertices ([Kothari et al., 2023](#)), and recent sum-of-squares and statistical-query lower bounds for multiple planted cliques and bicliques ([Mosievskiy and Reyzin, 2026](#)). Our multiple-clique problem instead concerns detecting a prescribed number of vertex-disjoint planted cliques and tracks how the number of planted components changes both the statistical and low-degree thresholds. For higher-order observations, related statistical-computational phenomena have been studied in planted tensor clustering ([Luo and Zhang, 2022](#)), planted dense subhypergraph recovery ([Corinzia et al., 2022](#)), and low-degree detection of dense subhypergraphs ([Dhawan et al., 2025](#)). The matrix-only observation of a planted hypergraph clique considered by [Alaluusua and Kumar \(2026\)](#) is complementary to our setting, where the full collection of independent hyperedge weights is observed.

General planted subgraphs, cycles, and paths. For dense and sparse planted communities in Erdős-Rényi graphs, [Arias-Castro and Verzelen \(2014\)](#); [Verzelen and Arias-Castro \(2015\)](#) establish information-theoretic thresholds, while [Hajek et al. \(2015\)](#) derive planted-clique-based computational lower bounds. More general planted-subgraph detection has been studied in [Huleihel \(2022\)](#); [Elimelech and Huleihel \(2025\)](#), with specialized results for bipartite graphs ([Rotenberg et al., 2024](#)), constant-degree polynomial tests ([Yu et al., 2024](#)), and planted matchings and spanning trees ([Addario-Berry et al., 2026](#)). The corresponding recovery literature includes sharp thresholds for planted subgraphs ([Mossel et al., 2025](#); [Huleihel, 2026](#)) and for hidden weighted sparse graphs ([Hou and Liu, 2026](#)). For path- and cycle-like structures, [Arias-Castro et al. \(2008\)](#) analyze anomalous trails, while [Mao et al. \(2023, 2025\)](#) establish computational and information-theoretic

thresholds for planted dense cycles. The random dense-cycle prior considered in these works coincides, after rescaling the bandwidth, with our random-size thick-cycle prior when $\ell = m$. They consider Bernoulli observations and an Erdős–Rényi null whose edge density matches the marginal edge density under the alternative. In our testing framework, the null coordinate distribution instead agrees with the distribution outside the planted structure. Mean-matched formulations also arise in the planted matching and spanning-tree problems studied by [Addario-Berry et al. \(2026\)](#); we discuss this distinction further in Section 5.

Frameworks for statistical and computational lower bounds. The second moment and conditional second moment methods provide information-theoretic lower bounds in many planted models ([Banks et al., 2017](#); [Perry et al., 2018, 2020](#)). Evidence for computational hardness has been developed through reductions from planted clique ([Brennan et al., 2018](#); [Brennan and Bresler, 2020](#)), sum-of-squares lower bounds ([Hopkins et al., 2017](#); [Barak et al., 2019](#)), statistical-query lower bounds ([Feldman et al., 2017](#)), and the low-degree polynomial framework ([Hopkins, 2018](#); [Kunisky et al., 2019](#); [Schramm and Wein, 2022](#); [Carpentier et al., 2025](#); [Wein, 2025](#)). The Franz–Parisi criterion gives a related overlap-based perspective ([Bandeira et al., 2022](#)), and [Chen et al. \(2025\)](#) connect an optimized version of this criterion to statistical-query lower bounds. Recent work also develops sharp low-degree thresholds beyond the planted-versus-null setting ([Skeja et al., 2026](#)). At the same time, counterexamples show that a general implication from low-degree indistinguishability to polynomial-time hardness requires additional assumptions ([Holmgren and Wein, 2021](#); [Buhai et al., 2025](#); [Mao, 2026](#)). Accordingly, our low-degree lower bounds rule out strong separation by low-degree polynomial statistics and are interpreted as conjectural evidence, rather than unconditional lower bounds, for arbitrary polynomial-time algorithms.

Distributional universality and natural exponential families. Distributional extensions of planted detection problems have been obtained through Gaussianization and related reductions ([Brennan et al., 2018, 2019](#); [Brennan and Bresler, 2020](#)), through distribution-free scan statistics ([Arias-Castro et al., 2018](#)), and through general weighted planted-clique models ([Chatterjee et al., 2025](#)). More generally, [Kunisky \(2025\)](#) develops low-coordinate-degree lower bounds for broad latent-variable channels and proves channel-universality results governed by the scalar Fisher information of the observation channel. Closely related to our natural exponential family results, [Kunisky \(2021\)](#) derives full and low-degree likelihood-ratio formulas and channel-monotonicity inequalities for the Morris class of natural exponential families with quadratic variance functions. For the models studied here, whose alternatives mix over latent supports, our comparison extends beyond the quadratic-variance setting. With a common latent prior, matching the one-coordinate chi-square divergence gives identical full likelihood-ratio second moments and low-degree second moments equivalent up to constant factors, subject to the degree condition in Theorem 4. We also analyze the average, scan, and degree tests under local regularity conditions in Proposition 3. When these upper bounds match the second-moment lower bounds, our local alternative result transfers the resulting statistical and conjectural computational thresholds across observation families under the conditions of Theorem 5.

1.3 Organization and notation

Organization. The rest of the paper is organized as follows. Section 2 develops the lower-bound framework based on the full and low-degree likelihood ratios, specializes it to Gaussian graph combinatorial testing, and records the overlap identities and the outlier, average, and scan upper bounds used throughout the paper. Section 3 applies these tools to planted cliques, multiple planted cliques and hypercliques, and thick cycles and paths, for both fixed-size and random-size constructions. Section 4 extends the framework to natural exponential family noise models through chi-square matching, develops corresponding upper bounds, and studies the universality of the resulting thresholds in the local alternative regime. Section 5 summarizes the main conclusions, limitations, and directions for future work. The appendices collect technical lemmas and complete proofs (Appendices A and B) and analyze additional latent structures that complement the main examples (Appendix C).

Notation. For a positive integer n , we write $[n] := \{1, \dots, n\}$. For $r \in [n]$, we write $[n]_r := \{(i_1, \dots, i_r) \in [n]^r : i_1 < \dots < i_r\}$, which we identify with the collection of r -element subsets of $[n]$. For a finite set A , we write $|A|$ for the cardinality of A . For real numbers a and b , we write $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. The indicator of an event $\mathcal{E}$ is denoted by $\mathbf{1}\{\mathcal{E}\}$. Expectations and variances are denoted by $\mathbb{E}$ and Var , respectively. Probabilities are denoted by Pr , with the source of randomness displayed in a subscript when useful; for example, $\text{Pr}_{S \sim \nu}\{\ell \in S\}$. For a named probability measure, we instead write its probability of an event directly, as in $\mathbb{Q}(A)$, $\mathbb{P}_s(A)$, and $\mathbb{P}_\nu(A)$. We use plain capital letters such as P_{θ_0} and P_{θ_1} for one-coordinate probability measures and blackboard-bold letters such as $\mathbb{Q}$ and $\mathbb{P}$ for joint probability measures on observation vectors. Calligraphic letters are used for classes, families, events, and function spaces, such as the latent-structure class $\mathcal{C}$. The notation $S, S' \sim \nu$ means that S and S' are independent draws from ν ; a fixed realization of S is denoted by $s \in \mathcal{C}$. We write $H \sim \text{Hypergeometric}(N, M, r)$ when H is the number of marked items in a sample of size r drawn uniformly without replacement from a population of N items, of which M are marked. We use boldface symbols for vectors and multi-indices and nonbold symbols for scalars and sets; in particular, $\mathbf{Y}$ denotes a random observation vector and $\mathbf{y}$ denotes one of its realizations.

Unless stated otherwise, all asymptotic statements are taken as the ambient dimension tends to infinity. In the general testing framework, the ambient dimension is n , whereas in the graph models, m denotes the number of vertices and $n = \binom{m}{2}$ denotes the number of observed edges. Model parameters, latent-structure classes, priors, and probability distributions may depend on n , or on m in the graph setting, even when this dependence is suppressed. Unless stated otherwise, constants such as $c, C > 0$ do not depend on the ambient dimension and may change from line to line.

For two positive sequences $\{a_n\}$ and $\{b_n\}$, we write $a_n \lesssim b_n$ if $a_n \leq Cb_n$ eventually for a constant $C > 0$ independent of n . We write $a_n \gtrsim b_n$ if $b_n \lesssim a_n$. We write $a_n \asymp b_n$ if both $a_n \lesssim b_n$ and $a_n \gtrsim b_n$ hold. We write $a_n \ll b_n$ if there exists a constant $c > 0$ such that $a_n \lesssim n^{-c}b_n$, and write $a_n \gg b_n$ if $b_n \ll a_n$. Thus, $\ll$ and $\gg$ denote separation by a polynomial factor, whereas $o(\cdot)$ and $\omega(\cdot)$ retain their usual asymptotic meanings.

2 Gaussian graph combinatorial testing

Throughout this section, we study the isotropic Gaussian specialization of the graph combinatorial testing problem in Equation (3), a model that has been studied extensively in the signal detection literature; see, e.g., [Arias-Castro et al. \(2008, 2011\)](#); [Addario-Berry et al. \(2010\)](#); [Ma and Wu \(2015\)](#); [Montanari et al. \(2017\)](#); [Dadon et al. \(2024\)](#). In our principal graph models, the data consist of the vector $\mathbf{Y} = (Y_{ij} : (i, j) \in [m]_2) \in \mathbb{R}^{\binom{m}{2}}$, whose coordinates are indexed by the edges of the complete graph K_m . For a latent structure $S \subseteq [m]_2$ randomly selected from a class $\mathcal{C}$ of edge sets of K_m , the null and alternative hypotheses are

$$H_0 : Y_{ij} \stackrel{iid}{\sim} N(0, 1) \quad \text{versus} \quad H_1 : Y_{ij} | S \stackrel{ind}{\sim} \begin{cases} N(\mu, 1), & \text{if } (i, j) \in S, \\ N(0, 1), & \text{if } (i, j) \notin S, \end{cases} \quad (5)$$

where $\mu > 0$ is the signal strength. Recall that $\mathbb{Q}$ and $\mathbb{P}_\nu$ denote the distributions of $\mathbf{Y}$ under the null and alternative hypotheses, respectively, where ν is an appropriate distribution on $\mathcal{C}$.

We consider two related types of latent structures: a structure drawn uniformly from a class $\mathcal{C}$ of fixed-size subgraphs of K_m , and a randomized counterpart whose size is not fixed. The former includes the class of all k -vertex cliques, while the latter includes cliques with random vertex sets. Inspired by the permutation and independent sampling schemes of [Carpentier et al. \(2025\)](#), we construct these classes and their associated priors using auxiliary random vectors $\hat{\mathbf{A}} = (\hat{A}_1, \dots, \hat{A}_m)$ and $\tilde{\mathbf{A}} = (\tilde{A}_1, \dots, \tilde{A}_m)$ whose coordinates determine vertex membership in the latent structure. Specifically, $\hat{\mathbf{A}}$ is a uniformly random permutation of $[m]$, while $\tilde{\mathbf{A}}$ has i.i.d. coordinates drawn uniformly from $[0, m]$. The latent structure is then defined as a problem-specific function of $\hat{\mathbf{A}}$ or $\tilde{\mathbf{A}}$. We denote the corresponding distributions of S by $\hat{\nu}$ and $\tilde{\nu}$ and their supports by $\hat{\mathcal{C}}$ and $\tilde{\mathcal{C}}$, respectively. Both constructions are symmetric: the distribution of S is invariant under permutations of the vertex labels.

2.1 Lower bounds for testing

Below, we briefly review the second-moment and low-degree likelihood-ratio methods used to establish lower bounds for the Gaussian testing problem above; see, e.g., [Banks et al. \(2017\)](#); [Perry et al. \(2018\)](#); [Hopkins \(2018\)](#); [Kunisky et al. \(2019\)](#); [Schramm and Wein \(2022\)](#); [Wein \(2025\)](#); [Carpentier et al. \(2025\)](#).

For each n , we assume throughout that the mixture alternative $\mathbb{P}_{\nu_n}$ is absolutely continuous with respect to the null distribution $\mathbb{Q}_n$. This condition holds when the one-coordinate signal distribution and the null distribution belong to the same regular exponential family with common support, as they do in every model considered in this paper. Define the likelihood ratio

$$\mathbf{y} \in \Omega^n \mapsto L_{\nu_n}(\mathbf{y}) := \frac{d\mathbb{P}_{\nu_n}}{d\mathbb{Q}_n}(\mathbf{y}).$$

For a fixed structure $s \in \mathcal{C}_n$, similarly write $L_{s,n} := d\mathbb{P}_{s,n}/d\mathbb{Q}_n$. When the dependence on n is clear, we suppress it and write $\mathbb{P}_\nu$, $\mathbb{Q}$, L_ν , and L_s .

2.1.1 Statistical lower bounds

For simple-versus-simple testing, it is well known that the minimum risk equals one minus the total variation distance between the two distributions. Directly analyzing this distance, or the optimal likelihood-ratio test, is often difficult in high-dimensional models. The second moment method instead controls the risk through the null second moment of L_{ν_n} ([Arias-Castro et al., 2008](#); [Addario-Berry et al., 2010](#); [Montanari et al., 2017](#); [Banks et al., 2017](#); [Perry et al., 2018, 2020](#)). The proof of the following lemma is given in [Appendix B.1](#).

Lemma 1 (Second moment method). *Suppose that, for some constant $C > 1$ and all sufficiently large n ,*

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}(\mathbf{Y})^2] \leq C.$$

Then every test Δ_n satisfies

$$R_{\nu_n}(\Delta_n) \geq 1 - \sqrt{1 - C^{-1}},$$

and hence the sequence of testing problems is statistically indistinguishable.

2.1.2 Low-degree lower bounds

We next describe the low-degree likelihood-ratio framework, which characterizes the ability of polynomial statistics to distinguish the null and alternative in a sense made precise below. This framework has been widely used to predict computational hardness in testing problems; see, e.g., ([Hopkins et al., 2017](#); [Hopkins, 2018](#); [Barak et al., 2019](#); [Kunisky et al., 2019](#); [Wein, 2025](#)). For this subsection, suppose that $\Omega \subseteq \mathbb{R}$ and that the polynomials under consideration are square-integrable under the null distribution. We recall the notion of strong separation from [Bandeira et al. \(2022, Definition 1.8\)](#).

Definition 2 (Strong separation). *A sequence of functions $f_n : \Omega^n \rightarrow \mathbb{R}$ strongly separates $\mathbb{P}_{\nu_n}$ and $\mathbb{Q}_n$ if*

$$\frac{|\mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_{\nu_n}} [f_n(\mathbf{Y})] - \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [f_n(\mathbf{Y})]|}{\sqrt{\max\{\text{Var}_{\mathbf{Y} \sim \mathbb{P}_{\nu_n}} [f_n(\mathbf{Y})], \text{Var}_{\mathbf{Y} \sim \mathbb{Q}_n} [f_n(\mathbf{Y})]\}}} \rightarrow \infty.$$

If f_n strongly separates the hypotheses, thresholding f_n at the midpoint of its two expectations achieves strong detection by Chebyshev's inequality.

Let $\mathcal{V}_n^{\leq D} \subseteq L^2(\mathbb{Q}_n)$ be the linear space of polynomials in the coordinates of $\mathbf{Y}$ that have total degree at most D and are square-integrable under the null distribution, and let $\text{Proj}_n^{\leq D}$ denote the orthogonal projection onto this space. Define the degree- D likelihood ratio by

$$L_{\nu_n}^{\leq D} := \text{Proj}_n^{\leq D} L_{\nu_n}.$$

The following lemma shows that boundedness of its second moment rules out strong separation by degree- D polynomial statistics; see [Bandeira et al. \(2022, Proposition 6.2\)](#) and its proof.

Lemma 2 (Low-degree second moment method). *For every $D \geq 1$,*

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D}(\mathbf{Y})^2] - 1 = \sup_{\substack{f \in \mathcal{V}_n^{\leq D} \\ \text{Var}_{\mathbf{Y} \sim \mathbb{Q}_n}[f(\mathbf{Y})] > 0}} \frac{(\mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_{\nu_n}}[f(\mathbf{Y})] - \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n}[f(\mathbf{Y})])^2}{\text{Var}_{\mathbf{Y} \sim \mathbb{Q}_n}[f(\mathbf{Y})]}. \quad (6)$$

Consequently, if $D = D_n$ and

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D_n}(\mathbf{Y})^2] = O(1),$$

then no sequence of degree- D_n polynomials strongly separates $\mathbb{P}_{\nu_n}$ and $\mathbb{Q}_n$.

The lemma is a statement about degree- D polynomial statistics. Interpreting it as evidence against all polynomial-time tests requires an additional hypothesis. For the planted models studied here, we use the following working hypothesis from the low-degree literature (Hopkins, 2018; Kunisky et al., 2019; Wein, 2025).

Conjecture 1 (Working low-degree hypothesis). *Consider the sequences of planted and null distributions $\{\mathbb{P}_{\nu_n}\}$ and $\{\mathbb{Q}_n\}$ defined above. If there exist $\epsilon > 0$ and degrees $D_n \geq (\log n)^{1+\epsilon}$ such that*

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D_n}(\mathbf{Y})^2] = O(1),$$

then no polynomial-time test achieves strong detection.

This hypothesis should not be interpreted as a valid general principle for arbitrary testing problems. Earlier formulations admit counterexamples (Holmgren and Wein, 2021), the quasi-polynomial-time version is false in general (Buhai et al., 2025), and Mao (2026) gives a counterexample to a permutation-invariant polynomial-time formulation even after independent noise is added. Recent results obtain rigorous consequences from low-degree indistinguishability under additional structural or algorithmic assumptions (Hsieh et al., 2026; Chen and Li, 2026). Accordingly, our low-degree lower bounds apply to polynomial statistics of degree at most D , in the sense of strong separation, while their interpretation as a lower bound against all polynomial-time tests is conditional on the working hypothesis above. The code-based counterexamples of Buhai et al. (2025) and Mao (2026), which target a specific formulation in Hopkins (2018, Conjecture 2.2.4), do not directly fit the one-sided Gaussian model studied here, whose alternative mixes over latent supports. Indeed, consider the natural attempted embedding in which the 1-entries of a counterexample form the support S and the observations are $Y_\ell = \mu \mathbb{1}\{\ell \in S\} + Z_\ell$, where $Z_\ell \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$. Exact low-order uniformity of these constructions gives $\Pr\{\ell \in S\} = 1/2$ and pairwise independence, so that $\mathbb{E}[|S|] = n/2$ and $|S|/n \rightarrow 1/2$ in probability. The degree-one likelihood-ratio projection under this embedding satisfies $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq 1}(\mathbf{Y})^2] = 1 + n\mu^2/4$. If $n\mu^2 \rightarrow \infty$, the average statistic strongly detects, whereas if $n\mu^2 = O(1)$, convexity of relative entropy gives $D_{\text{KL}}(\mathbb{P}_{\nu_n} \parallel \mathbb{Q}_n) \leq \mathbb{E}_{S \sim \nu_n} [D_{\text{KL}}(\mathbb{P}_{S,n} \parallel \mathbb{Q}_n)] = \mu^2 \mathbb{E}[|S|]/2 = n\mu^2/4 = O(1)$, and hence strong detection is statistically impossible. Thus, this natural embedding is either detectable by a degree-one statistic or statistically indistinguishable, and cannot reproduce the counterexamples' combination of low-degree indistinguishability and efficient detection. The sign-based Gaussian extension of Buhai et al. (2025) uses codeword signs rather than one-sided support shifts, while their spectral construction uses a nonproduct null, so both fall outside the present model.

Remark 2 (Conditional second moment method). Boundedness of the full and low-degree likelihood-ratio second moments is sufficient, but not necessary, to rule out strong detection and low-degree strong separation, respectively (Lemmas 1 and 2). Rare events under the alternative can make the second moment diverge even when strong detection remains impossible. The conditional second moment method addresses this issue by restricting the alternative to a high-probability event that excludes such rare configurations before applying the second moment argument (Banks et al., 2017; Perry et al., 2018, 2020; Bandeira et al., 2022; Dhawan et al., 2025). In our analysis, we primarily use the conditional second moment method to transfer lower bounds from fixed-size models to their random-size counterparts through high-probability truncation; see, e.g., Corollary 1 and Theorems 2 and 3 and their proofs.

2.2 Statistical and computational thresholds

To characterize the statistical and conjectural computational thresholds for the Gaussian combinatorial testing problem, we combine upper bounds on the testing risk with lower bounds from the full and low-degree likelihood-ratio second moments. Both second moments admit exact representations in terms of the overlap $|S \cap S'|$ between two independent draws $S, S' \sim \nu_n$. The following proposition records these identities; see [Arias-Castro et al. \(2008\)](#); [Banks et al. \(2017\)](#) for the full second moment and [Hopkins \(2018\)](#); [Kunisky et al. \(2019\)](#) for its low-degree counterpart.

Proposition 1. *Consider the hypothesis testing problem defined in Equation (5) with a prior distribution ν_n for S . Then,*

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu_n} \left[\exp(\mu^2 |S \cap S'|) \right]$$

and

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^D \frac{\mu^{2d} |S \cap S'|^d}{d!} \right].$$

This proposition reduces the derivation of statistical and computational lower bounds to the analysis of the overlap distribution induced by the alternative. In Section 3, we carry out this analysis for several combinatorial classes $\mathcal{C}$.

For the upper-bound analysis, we consider three tests: the *outlier test*, the *average test*, and the *scan test*, each obtained by thresholding the corresponding test statistic

$$T_{\text{out}}(\mathbf{y}) := \max_{(i,j) \in [m]_2} y_{ij}, \quad T_{\text{avg}}(\mathbf{y}) := \sum_{(i,j) \in [m]_2} y_{ij}, \quad \text{and} \quad T_{\text{scan}}(\mathbf{y}) := \max_{s \in \mathcal{C}} \sum_{(i,j) \in s} y_{ij}, \quad (7)$$

respectively. These tests, or variants of them, are widely used to establish upper bounds in signal detection problems; see, e.g., ([Donoho and Jin, 2004](#); [Arias-Castro et al., 2008, 2011](#); [Addario-Berry et al., 2010](#); [Butucea and Ingster, 2013](#); [Arias-Castro and Verzelen, 2014](#); [Ma and Wu, 2015](#); [Sharpnack et al., 2016](#); [Dadon et al., 2024](#); [Elimelech and Huleihel, 2025](#)). The outlier test uses the fact that the maximum of n standard normal random variables is approximately $\sqrt{2 \log n}$ and can therefore detect anomalous coordinates regardless of the size or shape of the latent structure S . The average test uses the accumulated signal $\mu|S|$: the statistic $T_{\text{avg}}(\mathbf{Y})$ follows $N(0, n)$ under the null and $N(\mu|S|, n)$ conditional on S under the alternative. Finally, the scan test exploits the combinatorial structure of S by maximizing the total observed signal over all candidates in $\mathcal{C}$. The outlier and average tests have computational complexity of order m^2 regardless of the alternative. A direct implementation of the scan test instead requires enumerating $\mathcal{C}$ and is therefore superpolynomial for many examples considered in this paper.

The next result characterizes the statistical performance of these tests; see, e.g., [Addario-Berry et al. \(2010\)](#) for related bounds and Appendix B.2 for a proof. For simplicity, the proposition assumes that $|S| = K$ deterministically; extensions to the random-size models considered here are justified in their respective proofs.

Proposition 2 (Outlier, average, and scan tests). *Consider the test statistics defined in Equation (7) for the hypothesis testing problem defined in Equation (5), where every $s \in \mathcal{C}$ satisfies $|s| = K \geq 1$, and let ν be a prior distribution on $\mathcal{C}$. Let $\delta \in (0, 1)$ be a target risk level. Then the following holds:*

- Consider the threshold $t_{\text{out}} := \sqrt{2 \log(2n/\delta)}$. The average risk of the outlier test defined by $\Delta_{\text{out}}(\mathbf{y}) := \mathbb{1}\{T_{\text{out}}(\mathbf{y}) > t_{\text{out}}\}$ satisfies $R_\nu(\Delta_{\text{out}}) \leq \delta$ if

$$\mu \geq \sqrt{2 \log \frac{2n}{\delta}} + \sqrt{2 \log \frac{2}{\delta}}.$$

- Consider the threshold $t_{\text{avg}} := \mu K/2$. The average risk of the average test defined by $\Delta_{\text{avg}}(\mathbf{y}) := \mathbb{1}\{T_{\text{avg}}(\mathbf{y}) > t_{\text{avg}}\}$ satisfies $R_\nu(\Delta_{\text{avg}}) \leq \delta$ if

$$\mu \geq \sqrt{\frac{8n}{K^2} \log \frac{2}{\delta}}.$$

- Consider the threshold $t_{\text{scan}} := (\mu K + \sqrt{2K \log |\mathcal{C}|})/2$. The average risk of the scan test defined by $\Delta_{\text{scan}}(\mathbf{y}) := \mathbb{1}\{T_{\text{scan}}(\mathbf{y}) > t_{\text{scan}}\}$ satisfies $R_\nu(\Delta_{\text{scan}}) \leq \delta$ if

$$\mu \geq \sqrt{\frac{2}{K} \log |\mathcal{C}|} + 2\sqrt{\frac{2}{K} \log \frac{2}{\delta}}.$$

Remark 3 (Degree test). Another commonly used computationally efficient procedure, also with complexity of order m^2 , is the degree test (Arias-Castro and Verzelen, 2014; Hajek et al., 2018; Elimelech and Huleihel, 2025), based on the statistic

$$T_{\text{deg}}(\mathbf{y}) := \max_{i \in [m]} T_i^{\text{deg}}(\mathbf{y}),$$

where $T_i^{\text{deg}}(\mathbf{y}) := \sum_{\{e \in [m]_2 : i \in e\}} y_e$ is the weighted degree of vertex $i \in [m]$. For $s \subseteq [m]_2$, let $\mathfrak{d}_s(i) := |\{e \in s : i \in e\}|$ denote the degree of vertex i in the graph with vertex set $[m]$ and edge set s . Suppose that $S \sim \nu$ satisfies $\max_{i \in [m]} \mathfrak{d}_S(i) = \mathfrak{d}_{\max} \geq 1$ almost surely. The following bound summarizes the performance of the degree test; see Appendix B.2 for a proof.

- Let $\delta \in (0, 1)$ be a target risk level. Consider the threshold $t_{\text{deg}} := \mathfrak{d}_{\max} \mu / 2$. Then, the average risk of the degree test defined by $\Delta_{\text{deg}}(\mathbf{y}) := \mathbb{1}\{T_{\text{deg}}(\mathbf{y}) > t_{\text{deg}}\}$ satisfies $R_\nu(\Delta_{\text{deg}}) \leq \delta$ if

$$\mu \geq \sqrt{\frac{8(m-1)}{\mathfrak{d}_{\max}^2} \log \frac{2m}{\delta}}.$$

The degree test is not used for the problems considered in Section 3, as it yields weaker upper bounds than the other tests in those settings. It can nevertheless provide a sharper upper bound for other structures, including the planted star studied in Appendix C.2.1.

To derive upper bounds for the statistical and computational thresholds, we use an appropriate combination of the tests in Proposition 2, depending on the scaling of μ . Since the scan test can require superpolynomial time in m , the statistical and computational upper bounds may differ in regimes where scanning is more powerful than the available polynomial-time tests. We now formalize the two thresholds. Because $\ll$ and $\gg$ denote polynomial separation, these definitions identify thresholds at polynomial order; unless stated otherwise, the main results therefore suppress polylogarithmic and other subpolynomial factors.

Definition 3 (Statistical and computational thresholds). *Consider a sequence of Gaussian graph testing problems defined in Equation (5), with prior $\nu = \nu_n$. A positive sequence μ_{stat}^* is called a statistical threshold if the following holds for every signal strength sequence $\mu = \mu_n$:*

- if $\mu \gg \mu_{\text{stat}}^*$, then there exists a sequence of tests Δ_n that achieves strong detection;
- if $\mu \ll \mu_{\text{stat}}^*$, then no sequence of tests can achieve strong detection.

Similarly, a positive sequence μ_{comp}^* is called a computational threshold if

- if $\mu \gg \mu_{\text{comp}}^*$, then there exists a sequence of polynomial-time tests Δ_n that achieves strong detection;
- if $\mu \ll \mu_{\text{comp}}^*$, then, under Conjecture 1, no sequence of polynomial-time tests can achieve strong detection.

Establishing either threshold requires matching upper and lower bounds. An upper bound μ^{UB} follows by demonstrating a test that achieves strong detection whenever $\mu \gg \mu^{\text{UB}}$. For a statistical lower bound μ^{LB} , it is sufficient to show that the full likelihood-ratio second moment remains bounded whenever $\mu \ll \mu^{\text{LB}}$. For a computational lower bound, we instead bound the low-degree likelihood-ratio second moment and invoke Conjecture 1 to interpret the resulting obstruction as evidence against all polynomial-time tests. If $\mu^{\text{UB}}/\mu^{\text{LB}} = n^{o(1)}$, either bound provides a valid representative of the polynomial-order threshold. We say that a statistical-computational gap occurs when $\mu_{\text{stat}}^* \ll \mu_{\text{comp}}^*$, with the computational lower bound understood in the context of Conjecture 1.

3 Statistical-computational gaps in Gaussian graph testing

We establish statistical-computational gaps, under Conjecture 1, for three classes of Gaussian combinatorial testing problems: planted cliques (Section 3.1), multiple planted cliques and hypercliques (Section 3.2), and thick cycles and paths (Section 3.3). Throughout, we consider symmetric alternatives induced by combinatorial classes of edge or hyperedge sets, with latent structures of either fixed or random size. The appendix, beginning with Appendix C.1.1, analyzes additional classes for which no gap occurs: K -disjoint sets, K -sets, perfect matchings, stars, spanning trees, k -stars, and paths and cycles of length k ; Addario-Berry et al. (2010) derived statistical thresholds for the first five classes. All results are summarized in Table 1.

3.1 A canonical example: the planted clique

We begin with the case in which the latent structure S is a clique, a canonical benchmark for studying the limitations of polynomial-time algorithms in high-dimensional statistics and theoretical computer science. Whereas the Erdős–Rényi model has been studied extensively, its Gaussian counterpart is less well understood. We therefore investigate the statistical-computational gap for planted clique detection under Gaussian observations.

Formally, we call a subset s of the edges of K_m a k -vertex clique if s is the edge set of a complete subgraph on k vertices. We consider two scenarios: (i) the fixed-size case, in which S is uniformly distributed over all k -vertex cliques in an m -vertex graph, and (ii) the random-size case, in which each of the m vertices is included independently with probability k/m . As indicated in Section 2, the fixed-size model can be formulated using the auxiliary random vector $\dot{\mathbf{A}} = (\dot{A}_1, \dots, \dot{A}_m)$, a uniformly random permutation of $[m]$. The random structure S is then defined by

$$S := \{(i, j) \in [m]_2 : \mathbf{1}(\dot{A}_i \leq k) \mathbf{1}(\dot{A}_j \leq k) = 1\} \quad (8)$$

This construction induces the prior ν on the set of all k -vertex cliques. To model a random-size clique with expected vertex count k , we define S analogously using the random vector $\tilde{\mathbf{A}} = (\tilde{A}_1, \dots, \tilde{A}_m)$, whose coordinates are i.i.d. uniform random variables on $[0, m]$.

We start with the fixed-size case. For this canonical benchmark, we state the statistical and low-degree lower bounds separately in Lemmas 3 and 4 and then combine them with the tests in Proposition 2 to obtain Theorem 1. For the subsequent models, we streamline the main text by stating the resulting threshold theorems directly and deferring the analogous intermediate bounds and proofs to the appendix.

Our first result gives a sufficient condition for the likelihood-ratio second moment to remain bounded, yielding a lower bound $\mu_{\text{stat}}^{\text{LB}}$ for the statistical threshold μ_{stat}^* . The analysis follows the approach of Addario-Berry et al. (2010) but uses Finner’s inequality (Finner, 1992); see Lemma 6 in Appendix A. This inequality is well suited to our setting and extends naturally to the multiple-clique problem considered later.

Lemma 3 (Statistical lower bound). *Fix any constant $\delta \in (0, 0.1)$. The second moment of the likelihood ratio is bounded as long as μ satisfies: (i) if $k^3 m^{-2} \leq m^{-\delta}$, then $\mu \leq \sqrt{\delta k^{-1} \log m}$; (ii) if $k^3 m^{-2} = \Omega(1)$, then $\mu = O(mk^{-2})$.*

Proof sketch. The overlap $|S \cap S'|$ has the same distribution as $\binom{H}{2}$, where H is a hypergeometric random variable with parameters (m, k, k) . By the convex ordering of the hypergeometric distribution by the binomial distribution (Hoeffding, 1963), it suffices to work with $B \sim \text{Bin}(k, k/m)$ instead of H . Thus, the lower-bound analysis reduces to showing that $\mathbb{E}[\exp(\mu^2 \binom{B}{2})]$ is bounded. Writing $B \stackrel{d}{=} \sum_{i=1}^k B_i$, where the B_i are i.i.d. Bernoulli random variables with success probability k/m , we have $\binom{B}{2} \stackrel{d}{=} \sum_{i < j} B_i B_j$. Applying Finner’s inequality (Lemma 6) gives

$$\begin{aligned} \mathbb{E}_B \left[\exp \left(\mu^2 \sum_{i < j} B_i B_j \right) \right] &\leq \prod_{i < j} \mathbb{E} \left[\exp(\mu^2 B_i B_j)^{k-1} \right]^{\frac{1}{k-1}} \\ &= \mathbb{E} \left[\exp\{\mu^2 (k-1) B_1 B_2\} \right]^{k/2} \end{aligned}$$

$$\leq \exp\left(\frac{k^3}{2m^2}(\exp(\mu^2 k) - 1)\right),$$

where the last inequality follows from the moment generating function of a Bernoulli random variable $B_1 B_2$ with success probability $(k/m)^2$ and the fact that $1 + x \leq e^x$ for any $x \in \mathbb{R}$. Analyzing the asymptotic behavior of the final expression gives the conditions in Lemma 3; the details appear in Appendix B.3. $\square$

The preceding lower bound suggests two regimes for the statistical threshold, separated at $k = m^{2/3}$. We next give matching upper bounds. Recall that $n = \binom{m}{2}$, $K = \binom{k}{2}$, and $|\dot{\mathcal{C}}| = \binom{m}{k}$ for the fixed-size planted clique problem. By Proposition 2, the scan and average tests achieve strong detection when $\mu = \omega(\sqrt{k^{-1} \log m})$ and $\mu = \omega(mk^{-2})$, respectively. The scan test attains the matching upper bound in the first regime, but its direct implementation searches all $\binom{m}{k}$ candidates in $\dot{\mathcal{C}}$ and therefore requires superpolynomial time. The average test runs in polynomial time and attains the matching upper bound in the second regime.

The following lemma provides a lower bound $\mu_{\text{comp}}^{\text{LB}}$ for the computational threshold μ_{comp}^* by analyzing the second moment of the low-degree likelihood ratio. The argument uses moment bounds for sub-Poissonian random variables due to Ahle (2022).

Lemma 4 (Computational lower bound). *Fix any constant $\delta \in (0, 0.1)$. For the sequence $D_m = \log^{1+\epsilon} m$ with $\delta > \epsilon > 0$, the second moment of the low-degree likelihood ratio is bounded as long as μ satisfies: (i) if $k^2 m^{-1} \leq m^{-\delta}$, then $\mu \leq C(\log m)^{(1-\epsilon)/2}$ for some sufficiently small constant $C > 0$; (ii) if $k^2 m^{-1} \geq \log^{1+\delta} m$, then $\mu = O(mk^{-2})$.*

Proof sketch. The convex ordering of the hypergeometric distribution by the binomial distribution gives

$$\mathbb{E}_{S, S' \sim \nu}[|S \cap S'|^d] = \mathbb{E}\left[\left(\frac{H}{2}\right)^d\right] \leq \mathbb{E}\left[\left(\frac{H^2}{2}\right)^d\right] \leq \mathbb{E}\left[\left(\frac{B^2}{2}\right)^d\right],$$

where $B \sim \text{Bin}(k, k/m)$. Thus, the moments of the binomial random variable B up to order $2D$ must be controlled carefully to bound the second moment of the low-degree likelihood ratio. Applying the moment bound of Ahle (2022), we obtain

$$\mathbb{E}_{S, S' \sim \nu}\left[\sum_{d=0}^D \frac{\mu^{2d} |S \cap S'|^d}{d!}\right] \leq \mathbb{E}\left[\sum_{d=0}^D \frac{\mu^{2d} B^{2d}}{2^d d!}\right] \leq 1 + \sum_{d=1}^D \frac{2^d \mu^{2d}}{d!} \left(\frac{d}{\log(1 + 2dm/k^2)}\right)^{2d}.$$

It remains to determine when this expression is bounded as a function of μ , k , m , and D . The details are given in Appendix B.4. $\square$

Together with matching polynomial-time upper bounds, this result identifies the polynomial order of μ_{comp}^* in both regimes of k . Specifically, the outlier test achieves strong detection if $\mu = \omega(\sqrt{\log m})$, while the average test achieves strong detection if $\mu = \omega(mk^{-2})$.

With Lemmas 3 and 4 in place, we now state the statistical and computational thresholds for the fixed-size planted clique problem. Although the preceding lower and upper bounds retain subpolynomial factors, we focus here on the setting in which k has positive polynomial order in m and report the resulting threshold scales up to subpolynomial factors. This convention keeps the statement transparent while preserving the polynomial-order thresholds. The proof is given in Appendix B.5, and Figure 2 visualizes the result.

Theorem 1. *Consider the Gaussian graph testing problem with the latent structure S defined in Equation (8) and prior distribution $\dot{\nu}$. Suppose that k is of positive polynomial order in m . Then, the statistical and computational thresholds are*

$$\mu_{\text{stat}}^* = \begin{cases} \sqrt{\frac{1}{k}}, & k \ll m^{2/3}, \\ \frac{m}{k^2}, & k \gg m^{2/3} \end{cases} \quad \text{and} \quad \mu_{\text{comp}}^* = \begin{cases} 1, & k \ll m^{1/2}, \\ \frac{m}{k^2}, & k \gg m^{1/2}. \end{cases}$$

The statistical-computational phase diagram in Figure 2 is not new, but is consistent with existing results in the literature on Gaussian submatrix detection. In particular, Ma and Wu (2015) established the corresponding computational lower bounds through reductions from planted clique under the planted-clique

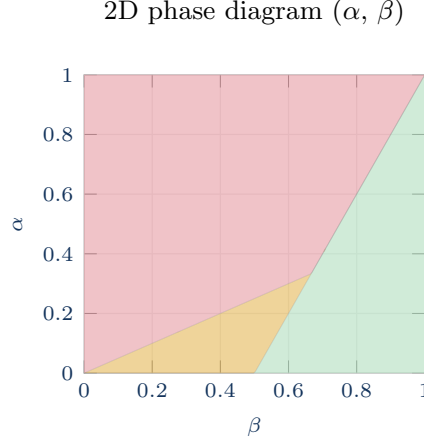


Figure 2: Phase diagram for the planted clique problem under the parameterization $\mu = m^{-\alpha}$ and $k = m^{\beta}$. Green, yellow, and pink denote the easy, hard, and impossible regions, respectively.

conjecture (see, e.g., [Alon et al. 2007](#), Conjecture 4.13 and [Berthet and Rigollet 2013](#), Hypothesis B_{PC}), and the subsequent work of [Brennan et al. \(2019\)](#) extended these reduction-based bounds to submatrix detection with matching row and column supports. The latter results imply the same computational lower bounds for the Gaussian planted clique model by restricting the observations to the strictly upper-triangular entries of the matrix. In contrast to these reduction-based approaches, our analysis recovers equivalent computational lower bounds more directly within the low-degree polynomial framework by studying the second moment of the low-degree likelihood ratio. This provides a complementary perspective on computational hardness and a unified approach that extends naturally to the more general combinatorial structures considered below.

We next consider the random-size case, which has the same statistical and computational thresholds as the fixed-size model. The key fact is that the number of planted vertices concentrates around k with high probability. The detailed proof of the following corollary is given in [Appendix B.6](#).

Corollary 1. *Consider the Gaussian graph testing problem with the latent structure S defined in Equation (8) and prior distribution $\tilde{\nu}$. Suppose that k is of positive polynomial order in m . Then, the statistical and computational thresholds μ_{stat}^* and μ_{comp}^* are identical to those in Theorem 1.*

The fixed-size guarantees for the average and scan tests do not apply directly when the clique size is random, but the tests can be modified to detect a clique on ck vertices for any fixed constant $c \in (0, 1)$. The planted clique has at least ck vertices with high probability, and every larger clique contains a clique on ck vertices. The complementary event, on which the clique has fewer than ck vertices, contributes only a negligible amount to the risk.

For the statistical lower bound, the ordinary second-moment analysis used in the fixed-size case yields a suboptimal result: on a rare event, the overlap is large enough to make $\mathbb{E}[\exp(\mu^2 |S \cap S'|)]$ diverge even when strong detection is impossible. See [Appendix B.6](#) for details. To bypass this limitation, we use the conditional second moment method ([Remark 2](#)) and condition on a high-probability event that excludes realizations with a disproportionate contribution to the second moment. We then show through stochastic dominance that the random-size clique detection problem is at least as hard as the fixed-size problem. In particular, we truncate and renormalize the prior $\tilde{\nu}$ to the event that the clique size is at most Ck for a fixed constant $C > 1$. The resulting loss in risk is negligible because the induced alternative distribution changes by a vanishing amount in total variation distance. Under the modified prior, we show that $|S \cap S'|$ is stochastically dominated by $\binom{H}{2}$, where H is a hypergeometric random variable with parameters (m, Ck, Ck) . This is the overlap distribution in the fixed-size model with clique size Ck . Intuitively, two random subsets of size at most Ck tend to share fewer vertices than two subsets of fixed size Ck .

The same truncation argument yields the matching computational lower bound for the random-size model.

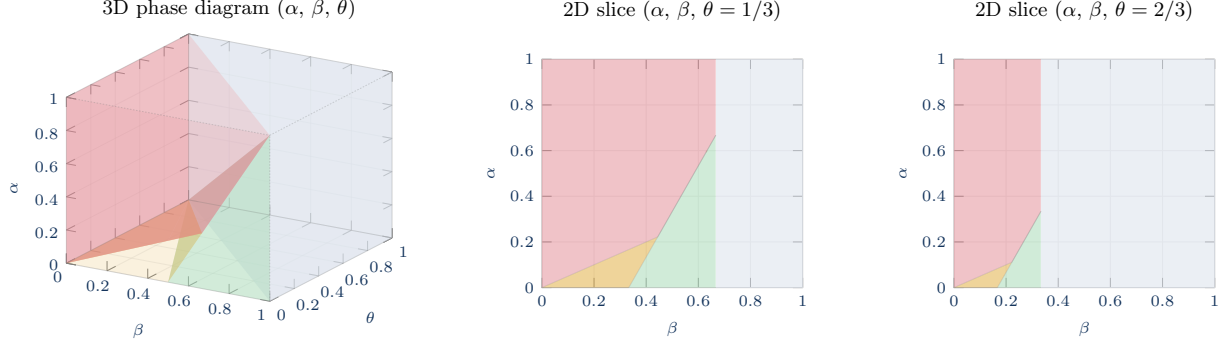


Figure 3: Phase diagram for the multiple planted cliques problem under the parameterization $\mu = m^{-\alpha}$, $k = m^\beta$, and $\gamma = m^\theta$. The left panel shows the three-dimensional phase diagram, and the two right panels show two-dimensional slices at $\theta = 1/3$ and $\theta = 2/3$. Green, yellow, and pink denote the easy, hard, and impossible regions, respectively; blue-gray denotes parameter values outside the admissible range.

For this low-degree bound, however, truncation is not necessary: direct analysis of the ordinary low-degree likelihood-ratio second moment gives the same result.

3.2 Multiple planted cliques and hypercliques

We treat the single-clique problem separately as a canonical benchmark that introduces the main proof ideas before extending them to several interacting planted components. We now generalize the planted clique setting to alternatives consisting of multiple disjoint planted cliques. Let m, k, γ be positive integers satisfying $\gamma k \leq m$, and S be the union of γ vertex-disjoint cliques, each with either a fixed vertex count k or an expected vertex count k . For the fixed-size model, let $\dot{\mathbf{A}} = (\dot{A}_1, \dots, \dot{A}_m)$ be a uniformly random permutation of $[m]$ and define

$$S := \bigcup_{r=1}^{\gamma} \left\{ (i, j) \in [m]_2 : (r-1)k < \dot{A}_i, \dot{A}_j \leq rk \right\}. \quad (9)$$

A closely related model is the stochastic block model, whose statistical-computational gap has been studied extensively under the Erdős–Rényi observation model; see, e.g., [Hopkins and Steurer \(2017\)](#); [Hopkins \(2018\)](#); [Bandeira et al. \(2021\)](#); [Carpentier et al. \(2025\)](#); [Ding et al. \(2025\)](#); [Sohn and Wein \(2025\)](#). When $\gamma k = m$, the latent structure partitions all vertices into γ communities, as in the stochastic block model. The following theorem analyzes the statistical and computational thresholds in both the fixed-size and random-size cases. In the random-size model, S is defined as in Equation (9), with $\dot{\mathbf{A}}$ replaced by $\tilde{\mathbf{A}} = (\tilde{A}_1, \dots, \tilde{A}_m)$, whose entries are i.i.d. uniform random variables on $[0, m]$. The thresholds match those in Theorem 1 after replacing the ambient vertex count m by m/γ , and the resulting phase diagram appears in Figure 3.

Theorem 2. *Consider the Gaussian graph testing problem with the latent structure S defined in Equation (9) and prior distribution $\dot{\nu}$ or $\tilde{\nu}$. Suppose that k is of positive polynomial order in m . Then, in both cases, the statistical and computational thresholds are*

$$\mu_{\text{stat}}^* = \begin{cases} \sqrt{\frac{1}{k}}, & \gamma^2 k^3 \ll m^2, \\ \frac{m}{\gamma k^2}, & \gamma^2 k^3 \gg m^2 \end{cases} \quad \text{and} \quad \mu_{\text{comp}}^* = \begin{cases} 1, & \gamma k^2 \ll m, \\ \frac{m}{\gamma k^2}, & \gamma k^2 \gg m. \end{cases}$$

We next sketch the argument; the full proof is given in Appendix B.7. For the statistical lower bound, a key step is to prove that the likelihood-ratio second moment in the multiple-clique setting is bounded by the γ^2 power of an appropriate binomial moment, using the conditional negative association of the clique overlaps.

In the fixed-size case, let S_r and S'_r denote the r th planted cliques in S and S' , respectively. We show that

$$\begin{aligned}\mathbb{E}_{S, S' \sim \tilde{\nu}} \left[\exp \left(\mu^2 |S \cap S'| \right) \right] &= \mathbb{E}_{S, S' \sim \tilde{\nu}} \left[\exp \left(\mu^2 \sum_{a, b=1}^{\gamma} |S_a \cap S'_b| \right) \right] \\ &\leq \mathbb{E}_B \left[\exp \left\{ \mu^2 \binom{B}{2} \right\} \right]^{\gamma^2}\end{aligned}$$

where $B \sim \text{Binomial}(k, q_{\gamma, k})$ and $q_{\gamma, k} := k/(m - (\gamma - 1)k)$. We bound the last expectation using Finner's inequality, following the same argument as in the proof for the single-clique case (Lemma 3). For the random-size model, as in the proof of Corollary 1, we apply the conditional second moment method after truncating the prior to the event that all γ cliques have size at most Ck for some constant $C > 1$.

The computational lower bound requires a more delicate comparison. One possible route is to represent the overlap as a U-statistic and control its moments using decoupling inequalities (de la Peña and Giné, 1999; Giné et al., 2000), as in the analysis of a different testing problem by Mao et al. (2025). We provide a proof based on this approach in Appendix B.15. In our main analysis, however, we use a direct graph-theoretic argument. We first analyze the random-size model, where the overlap is the number of edges in a random γ^2 -multiple-clique model, and bound its moments by those of an auxiliary single-clique model. Specifically, we show that for every integer $d \geq 1$,

$$\mathbb{E}_{S, S' \sim \tilde{\nu}} [|S \cap S'|^d] \leq \mathbb{E} \left[\binom{B}{2}^d \right],$$

where $B \sim \text{Binomial}(m, \gamma k^2/m^2)$. This comparison lets us apply the tools developed for the single-clique problem, and a truncation argument then transfers the resulting bound to the fixed-size model.

Multiple planted hypercliques. We further generalize the multiple planted clique model to t -uniform planted hypercliques. Formally, a t -uniform *hyperclique* is the edge set of a complete t -uniform hypergraph. Each hyperedge is represented by an increasing t -tuple rather than a pair of vertices. For $\mathbf{I} = (i_1, \dots, i_t) \in [m]_t$, define the latent structure S by

$$S := \bigcup_{r=1}^{\gamma} \left\{ \mathbf{I} \in [m]_t : (r-1)k < \dot{A}_{i_1}, \dots, \dot{A}_{i_t} \leq rk \right\},$$

where, as usual, $(\dot{A}_1, \dots, \dot{A}_m)$ is a random permutation of $[m]$. The random-size case is defined analogously by replacing $\dot{\mathbf{A}}$ with $\bar{\mathbf{A}} = (\bar{A}_1, \dots, \bar{A}_m)$.

The proof strategy for multiple planted cliques extends to this hypergraph setting, yielding the following corollary. The detailed proof is given in Appendix B.8. For $t = 3$, the resulting phase diagram is shown in Figure 4.

Corollary 2. *Consider the Gaussian hypergraph testing problem with the latent structure S defined above and prior distribution $\tilde{\nu}$ or $\tilde{\nu}$. Suppose that $t \geq 2$ is a fixed integer and k is of positive polynomial order in m . Then, in both cases, the statistical and computational thresholds are*

$$\mu_{\text{stat}}^* = \begin{cases} \sqrt{\frac{1}{k^{t-1}}}, & \gamma^2 k^{t+1} \ll m^t, \\ \frac{m^{t/2}}{\gamma k^t}, & \gamma^2 k^{t+1} \gg m^t \end{cases} \quad \text{and} \quad \mu_{\text{comp}}^* = \begin{cases} 1, & \gamma^2 k^{2t} \ll m^t, \\ \frac{m^{t/2}}{\gamma k^t}, & \gamma^2 k^{2t} \gg m^t. \end{cases}$$

Remark 4 (Multiple p -partite hypercliques). Another possible generalization is the detection of multiple p -partite hypercliques in p -partite hypergraphs, for any fixed integer $p \geq 2$. The case $p = 2$ is equivalent to detecting multiple equal-sized submatrices with disjoint row supports and disjoint column supports, while $p \geq 3$ corresponds to multiple tensor detection. The submatrix detection problem and its statistical-computational gap have been studied extensively; see, e.g., Butucea and Ingster (2013); Ma and Wu (2015);

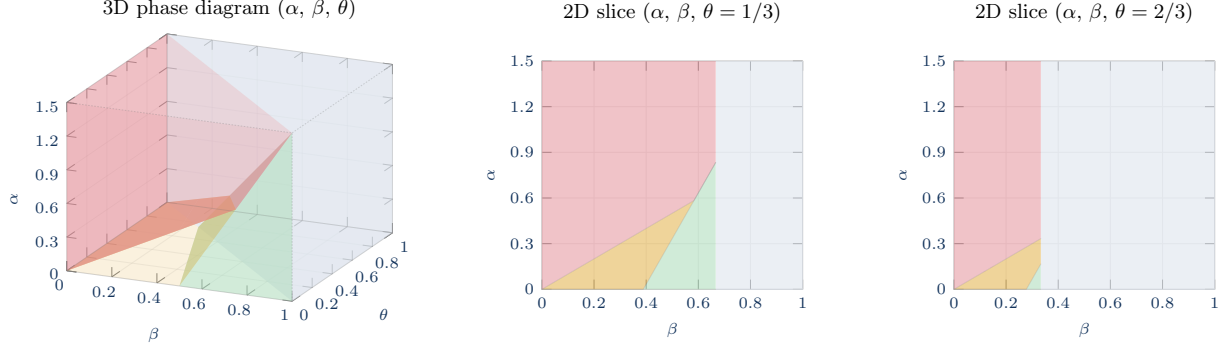


Figure 4: Phase diagram for the multiple planted hypercliques problem with $t = 3$ under the parameterization $\mu = m^{-\alpha}$, $k = m^\beta$, and $\gamma = m^\theta$. The left panel shows the three-dimensional phase diagram, and the two right panels show two-dimensional slices at $\theta = 1/3$ and $\theta = 2/3$. Green, yellow, and pink denote the easy, hard, and impossible regions, respectively; blue-gray denotes parameter values outside the admissible range.

Brennan et al. (2019); Dadon et al. (2024). It turns out that the statistical and computational thresholds for this problem coincide with those in Corollary 2 with t replaced by p , since the key arguments used to establish that result also apply to the p -partite setting; see Appendix B.8.3 for details. The closest work is by Dadon et al. (2024), who study multiple fixed-size Gaussian submatrix detection and derive statistical and low-degree computational strong detection thresholds. Their statistical proof invokes negative association of the full block-overlap arrays, a property that need not hold; see Equation (56). Their low-degree proof uses a bound on overlap moments whose justification in the cited preprint makes the same assumption. Our proofs avoid this global assumption: the statistical lower bound uses conditional negative association followed by iterated conditioning, while the computational lower bound uses a separate moment comparison.

3.3 Thick cycles and paths

We next consider *thick cycles* and *thick paths*, which connect nearby vertices along an underlying cycle or path. Specifically, two vertices are adjacent when their distance in the underlying ordering is at most Δ . For positive integers ℓ and Δ satisfying $2\Delta < \ell \leq m$, define a length- ℓ , Δ -thick cycle by the edge set

$$S := \left\{ (i, j) \in [m]_2 : \dot{A}_i, \dot{A}_j \in (0, \ell], d_{\text{cyc}}^{(\ell)}(\dot{A}_i, \dot{A}_j) \leq \Delta \right\}, \quad (10)$$

where $d_{\text{cyc}}^{(\ell)}(a, b) := \min\{|a - b|, \ell - |a - b|\}$ is the cyclic distance on a circle of circumference ℓ , and $(\dot{A}_1, \dots, \dot{A}_m)$ is a random permutation of $[m]$. We denote the distribution of S by $\dot{\nu}$. A length- ℓ , Δ -thick path is defined analogously by replacing the cyclic distance with the linear distance $d_{\text{lin}}^{(\ell)}(a, b) := |a - b|$. For random-size thick cycles and paths, we replace the coordinates of the random permutation in Equation (10) by i.i.d. uniform random variables on $[0, m]$. We write $\tilde{\nu}$ for the resulting distribution of S .

Related testing problems have been studied in Carpentier et al. (2025); Mao et al. (2025), primarily under Erdős–Rényi observations or with recovery rather than detection as the objective. We begin by analyzing the statistical and computational thresholds for thick cycles.

Theorem 3. *Consider the Gaussian graph testing problem with the latent structure S defined above and prior distribution $\dot{\nu}$ or $\tilde{\nu}$. Suppose that Δ is of positive polynomial order in m . Then, the statistical and computational thresholds are*

$$\mu_{\text{stat}}^* = \begin{cases} \sqrt{\frac{1}{\Delta}}, & \ell\sqrt{\Delta} \ll m, \\ \frac{m}{\ell\Delta}, & \ell\sqrt{\Delta} \gg m \end{cases} \quad \text{and} \quad \mu_{\text{comp}}^* = \begin{cases} 1, & \ell\Delta \ll m, \\ \frac{m}{\ell\Delta}, & \ell\Delta \gg m. \end{cases}$$

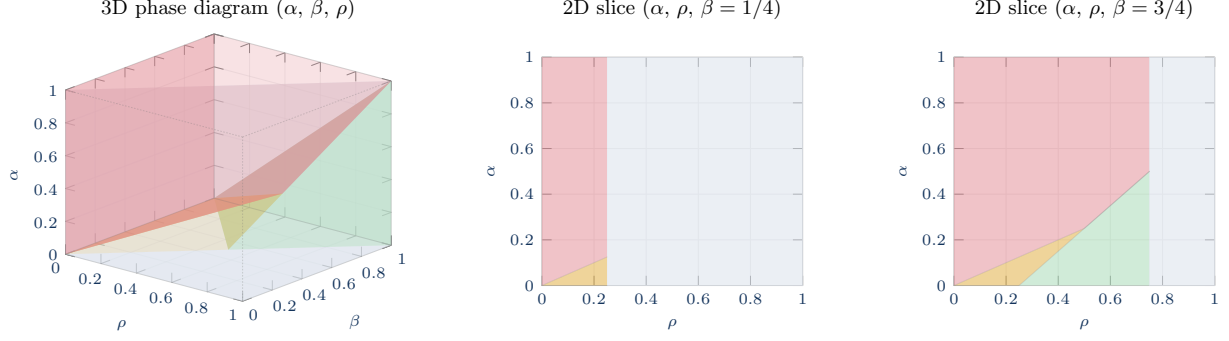


Figure 5: Phase diagram for the thick cycle problem under the parameterization $\mu = m^{-\alpha}$, $\ell = m^{\beta}$, and $\Delta = m^{\rho}$. The left panel shows the three-dimensional phase diagram, and the two right panels show two-dimensional slices at $\beta = 1/4$ and $\beta = 3/4$. Green, yellow, and pink denote the easy, hard, and impossible regions, respectively; blue-gray denotes parameter values outside the admissible range.

The phase diagram associated to the thresholds described in Theorem 3 is shown in Figure 5.

We briefly describe the proof strategy for the fixed-size model; the full proof is given in Appendix B.9. The random-size analysis then follows because the numbers of neighbors on either side of each vertex concentrate around Δ with high probability. The upper bounds follow from the fact that a fixed-size thick cycle has $\ell\Delta$ edges and contains a clique on $\Delta + 1$ consecutive vertices. The average and outlier tests, together with a scan over all $(\Delta + 1)$ -vertex cliques, therefore yield the claimed upper bounds.

To establish the lower bounds, we control $|S \cap S'|$ using the overlap between patchwork covers of S and S' . For illustration, assume that ℓ is divisible by 2Δ . In the fixed-size model, set $q := \ell/(2\Delta)$ and let $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ be a patchwork with two components

$$\mathcal{S}_1 = \cup_{i=1}^q \mathcal{S}_{1,i} \quad \text{and} \quad \mathcal{S}_2 = \cup_{i=1}^q \mathcal{S}_{2,i},$$

each of which is a union of $\gamma = \ell/(2\Delta)$ disjoint cliques of size $k = 2\Delta$; see Figure 6 for an illustration. Since S is a subset of $\mathcal{S}$, the overlap can be upper bounded as

$$\begin{aligned} |S \cap S'| &\leq |(\mathcal{S}_1 \cup \mathcal{S}_2) \cap (\mathcal{S}'_1 \cup \mathcal{S}'_2)| \\ &\leq |\mathcal{S}_1 \cap \mathcal{S}'_1| + |\mathcal{S}_1 \cap \mathcal{S}'_2| + |\mathcal{S}_2 \cap \mathcal{S}'_1| + |\mathcal{S}_2 \cap \mathcal{S}'_2|, \end{aligned} \tag{11}$$

where $\mathcal{S}' = \mathcal{S}'_1 \cup \mathcal{S}'_2$ is the patchwork for S' defined in the same way as $\mathcal{S}$. Each term on the right-hand side of Equation (11) is the overlap between two independent copies of the multiple-clique model analyzed in Section 3.2. It remains to combine the four contributions. For the statistical lower bound, Hölder's inequality yields

$$\mathbb{E} \left[\prod_{i,j=1}^2 \exp(\mu^2 |\mathcal{S}_i \cap \mathcal{S}'_j|) \right] \leq \prod_{i,j=1}^2 \left(\mathbb{E} [\exp(4\mu^2 |\mathcal{S}_i \cap \mathcal{S}'_j|)] \right)^{1/4} = \mathbb{E} [\exp(4\mu^2 |\mathcal{S}_1 \cap \mathcal{S}'_1|)],$$

where the last equality holds because $\mathcal{S}_i \cap \mathcal{S}'_j$, for $i, j \in \{1, 2\}$, are identically distributed. For the computational lower bound, the inequality $(\sum_{i=1}^4 x_i)^d \leq 4^{d-1} \sum_{i=1}^4 x_i^d$ for $x_i \geq 0$ gives

$$\mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} \left\{ \sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j| \right\}^d \right] \leq \mathbb{E} \left[\sum_{d=1}^D \frac{4^{d-1} \mu^{2d}}{d!} \sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j|^d \right] = \mathbb{E} \left[\sum_{d=1}^D \frac{4^d \mu^{2d}}{d!} |\mathcal{S}_1 \cap \mathcal{S}'_1|^d \right].$$

Applying the bounds on overlap moments proved for the multiple-clique model (Theorem 2), with $\gamma = \ell/(2\Delta)$ and $k = 2\Delta$, yields the desired lower bounds.

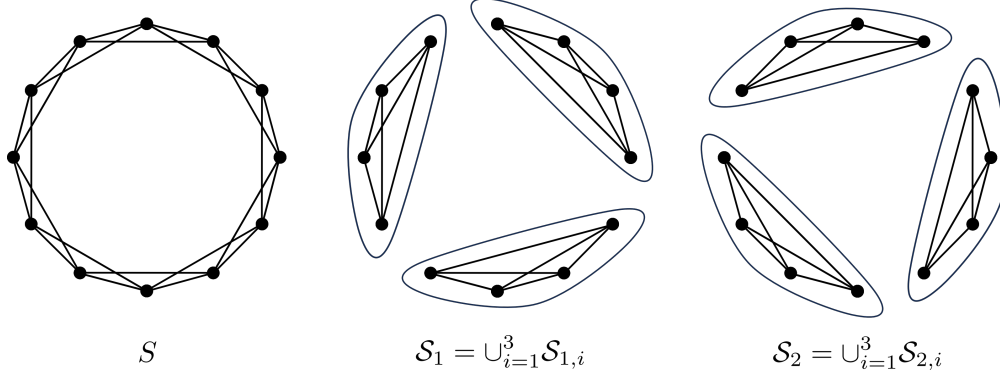


Figure 6: Illustration of the patchwork cover of a thick cycle. The left panel shows the thick cycle S for $\ell = 12$ and $\Delta = 2$. The middle and right panels show the two patchwork components $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, whose union $S = \mathcal{S}_1 \cup \mathcal{S}_2$ covers S . Each patch $\mathcal{S}_{1,i}$ or $\mathcal{S}_{2,i}$ is a clique on 2Δ vertices.

We conclude by analyzing the statistical and computational thresholds for thick paths. Write $\mathbf{A} = \dot{\mathbf{A}}$ in the fixed-size model and $\mathbf{A} = \tilde{\mathbf{A}}$ in the random-size model. A length- ℓ , Δ -thick path is defined by the edge set

$$S := \left\{ (i, j) \in [m]_2 : A_i, A_j \in (0, \ell], d_{in}^{(\ell)}(A_i, A_j) \leq \Delta \right\},$$

where $d_{in}^{(\ell)}(a, b) := |a - b|$ is the linear distance on a line of length ℓ . The upper- and lower-bound arguments for thick cycles extend to thick paths, with minor endpoint adjustments for the upper bounds, yielding the same statistical and computational thresholds. The following corollary records the result; its proof is given in Appendix B.10.

Corollary 3. *Consider the Gaussian graph testing problem with the latent structure S defined above and prior distribution $\dot{\nu}$ or $\tilde{\nu}$. Suppose that Δ is of positive polynomial order in m . Then, in both cases, the statistical and computational thresholds μ_{stat}^* and μ_{comp}^* are identical to those in Theorem 3.*

4 Equivalence across natural exponential family models

In this section, we extend the Gaussian framework for graph testing to natural exponential families (NEFs). Under appropriate conditions, a canonical rescaling of the signal strength transfers statistical and computational thresholds from one NEF model to another.

We consider a NEF model with a nondegenerate reference measure ρ and exponentially tilted distribution P_θ with natural parameter θ . The density of P_θ with respect to ρ is given by

$$dP_\theta(y) = \exp(\theta y - \psi(\theta))\rho(dy), \quad \theta \in \Theta$$

Here, $\psi(\theta) = \log \int \exp(\theta y)\rho(dy)$ is the log-partition function, and $\Theta = \text{int}\{\theta \in \mathbb{R} : \psi(\theta) < \infty\}$ is the natural parameter space. Representative NEF models and their natural parameterizations are listed in Table 2. We consider an n -dimensional observation vector $\mathbf{Y} = (Y_1, \dots, Y_n)$ and a latent structure $S \sim \nu_n$, where ν_n is supported on a class $\mathcal{C} \subseteq 2^{[n]}$; the elements of $\mathcal{C}$ need not correspond to edge sets of subgraphs of K_m . Given a triplet $(\theta_0, \theta_1, \rho)$ with $\theta_0, \theta_1 \in \Theta$, we consider the testing problem:

$$H_0 : Y_l \stackrel{iid}{\sim} P_{\theta_0} \quad \text{versus} \quad H_1 : Y_l | S \stackrel{ind}{\sim} \begin{cases} P_{\theta_1}, & \text{if } l \in S, \\ P_{\theta_0}, & \text{if } l \notin S. \end{cases} \quad (12)$$

The latent structure may have fixed or random size; fixed-size assumptions are stated explicitly where needed for the upper bounds. The exponential tilts θ_0 and θ_1 , corresponding to the null and alternative hypotheses,

may depend on n . In the Gaussian graph testing problem of Equation (5), the signal strength μ is the mean difference between $P_{\theta_0} = N(0, 1)$ and $P_{\theta_1} = N(\mu, 1)$. For the NEF model in Equation (12), the analogous quantity is the *effective signal strength*

$$\tau_n := \frac{\psi'(\theta_1) - \psi'(\theta_0)}{\sqrt{\psi''(\theta_0)}}, \quad (13)$$

which is the standardized mean difference between P_{θ_0} and P_{θ_1} because $\psi'(\theta)$ and $\psi''(\theta)$ are the mean and variance of P_θ , respectively. The following subsections develop lower bounds through the one-coordinate chi-square divergence and upper bounds in terms of the effective signal strength τ_n .

4.1 Lower bounds

Let $\chi^2(P_{\theta_1} \| P_{\theta_0}) = \int (\frac{dP_{\theta_1}}{dP_{\theta_0}} - 1)^2 dP_{\theta_0}$ denote the chi-square divergence between P_{θ_0} and P_{θ_1} . In the following theorem, we show that the asymptotic behavior of the full and low-degree likelihood-ratio second moments for NEF models can be characterized by this one-coordinate divergence, without further dependence on the specific form of the distributions under the stated degree condition. This result allows us to transfer statistical and computational lower bounds for the Gaussian graph testing problem to other NEF models by matching the chi-square divergence between the two models. The proof is given in Appendix B.12.

Theorem 4 (χ^2 -characterization of lower bounds for NEF models). *Consider the testing problem in Equation (12) with $\theta_0, \theta_1 \in \Theta$ satisfying $\theta_0 \neq \theta_1$ and $\psi''(\theta_0) > 0$. Assume that $\psi(2\theta_1 - \theta_0) < \infty$, so that the chi-square divergence $\chi_n^2 := \chi^2(P_{\theta_1} \| P_{\theta_0})$ is finite. Then, the second moment of the likelihood ratio $L_{\nu_n}(\mathbf{Y})$ is given by*

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu_n} \left[(1 + \chi_n^2)^{|S \cap S'|} \right].$$

Moreover, for any sequence of degrees $D = D_n$ satisfying $D_n \log(\chi_n^2/\tau_n^2) = O(1)$, the second moment of the low-degree likelihood ratio $L_{\nu_n}^{\leq D}(\mathbf{Y})$ satisfies the asymptotic equivalence

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D}(\mathbf{Y})^2] \asymp \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D \wedge |S \cap S'|} \binom{|S \cap S'|}{d} (\chi_n^2)^d \right].$$

The implicit constants depend only on the upper bound for $D_n \log(\chi_n^2/\tau_n^2)$.

The theorem implies that, for two different NEF models defined by $(\theta_{0,n}^A, \theta_{1,n}^A, \rho^A)$ and $(\theta_{0,n}^B, \theta_{1,n}^B, \rho^B)$ with the same latent prior ν_n and matching chi-square divergences χ_n^2 , the second moments of their likelihood ratios coincide exactly. Furthermore, if both models satisfy $D_n \log(\chi_n^2/\tau_n^2) = O(1)$, their low-degree likelihood-ratio second moments are equivalent up to constant factors. Consequently, the resulting statistical and computational lower bounds obtained from second moment methods (Lemmas 1 and 2) for one model can be transferred to the other model by matching the chi-square divergences between the two models. The degree condition $D_n \log(\chi_n^2/\tau_n^2) = O(1)$ can be parsed explicitly for specific NEF models. For example, in a Gaussian location model with mean difference μ and known variance σ^2 , we have $\chi_n^2 = \exp(\mu^2/\sigma^2) - 1$ and $\tau_n = \mu/\sigma$. Since $\log((e^x - 1)/x) \sim x/2$ for $x \rightarrow 0$, the degree condition is equivalent to $D_n \mu^2 = O(1)$ for fixed $\sigma^2 > 0$. In the Bernoulli model with success probabilities p_0 and p_1 , we have $\chi_n^2 = (p_1 - p_0)^2/p_0(1 - p_0) = \tau_n^2$, so the degree condition is always satisfied. More generally, for any NEF model in the local alternative regime where θ_0 is fixed and $\theta_1 \rightarrow \theta_0$, a simple sufficient condition is $D_n(\theta_1 - \theta_0)^2 = O(1)$. We formalize this observation in Proposition 4.

Remark 5 (Compatibility with conditional second moment methods). An astute reader may wonder whether the lower bounds for the random-size problems in Section 3 can be transferred through Theorem 4 directly, since their proofs rely not merely on the standard second moment method, but on the conditional second moment method (Remark 2) involving a truncation argument. Indeed, such conditional arguments, particularly

those based on high-probability events involving the latent variables that define S , are compatible with Theorem 4, since the χ^2 -characterization continues to hold for the second moments under the truncated prior. More specifically, given a prior ν , let $\nu_{\mathcal{E}}(\cdot) := \nu(\cdot | \mathcal{E})$ be a truncated prior defined by a high-probability event $\mathcal{E}$ for S under ν . Then, Theorem 4 ensures that the second moment of the likelihood ratio between the null distribution $\mathbb{Q}$ and the alternative distribution $\mathbb{P}_{\nu_{\mathcal{E}}} := \mathbb{E}_{S \sim \nu_{\mathcal{E}}}[\mathbb{P}_S]$ is given by $\mathbb{E}_{S, S' \sim \nu_{\mathcal{E}}}[(1 + \chi_n^2)^{|S \cap S'|}]$, simply by replacing ν with $\nu_{\mathcal{E}}$. Since this expectation involves only the overlap structure under $\nu_{\mathcal{E}}$ and the chi-square divergence χ_n^2 , the full likelihood-ratio second moments coincide across NEF models with the same conditional prior $\nu_{\mathcal{E}}$ and matching chi-square divergence χ_n^2 . The same argument yields equivalence up to constant factors for the low-degree second moments, and the degree condition $D_n \log(\chi_n^2/\tau_n^2) = O(1)$ is independent of the truncation event $\mathcal{E}$. Hence, under the degree condition, the lower bounds for the random-size problems in Section 3 can also be transferred across NEF models by matching χ^2 .

On the other hand, when the conditional second moment method involves conditioning on a high-probability event defined in terms of the observations $\mathbf{Y}$ (see e.g., [Arias-Castro and Verzelen \(2014\)](#); [Dhawan et al. \(2025\)](#)) rather than the latent structure S , the conditioning event may depend on model-specific features of the observation distribution. Consequently, transferring the resulting bounds through Theorem 4 may not be applicable.

One caveat is that the lower bounds obtained from Theorem 4 are not necessarily sharp across different NEF models unless matching upper bounds are available for each model; the full and low-degree (conditional) second moment methods provide only sufficient conditions for ruling out strong detection and low-degree strong separation, respectively. Sharper lower bounds may therefore be obtainable through more refined arguments, such as the observation-dependent conditional second moment methods mentioned in Remark 5. A concrete example of this is the planted clique problem under the Erdős–Rényi model, where the parameters p_0 and p_1 both go to zero as $n \rightarrow \infty$ at different rates. In this problem, transferring the corresponding result for the Gaussian model via chi-square matching only yields loose lower bounds. We will revisit this example in Section 4.3.1.

To establish equivalence of the statistical and computational thresholds, we must also obtain matching upper bounds for the NEF models. In the next subsection, we will show that the performance of the tests in Proposition 2 for general NEF models can be characterized in terms of the effective signal strength τ_n under appropriate regularity conditions. Importantly, when $D_n \rightarrow \infty$, the degree condition $D_n \log(\chi_n^2/\tau_n^2) = O(1)$ implies $\chi_n^2/\tau_n^2 \rightarrow 1$. Therefore, for two NEF models with matching chi-square divergences that both satisfy the degree condition, the magnitudes of their effective signal strengths are asymptotically equivalent, and the resulting upper bounds for the two models are of the same order. Consequently, given $S \sim \nu$, if the lower bounds from second moment methods and upper bounds achieved by such tests for one model are sharp in the sense that they match up to a constant factor, then the transferred lower bounds and the corresponding upper bounds also match up to a constant factor under the degree condition and regularity assumptions on the upper bound. We discuss when these assumptions hold in Section 4.3.

4.2 Upper bounds

We now revisit the average and scan tests from Proposition 2 and analyze their performance under NEF models in terms of the effective signal strength. The corresponding test statistics are

$$T_{\text{avg}}(\mathbf{y}) := \sum_{i=1}^n y_i, \quad \text{and} \quad T_{\text{scan}}(\mathbf{y}) := \max_{s \in \mathcal{C}} \sum_{l \in s} y_l, \quad (14)$$

respectively. The risks of these tests under the Gaussian model are controlled using standard Gaussian tail bounds. For a regular NEF, the centered moment generating function is finite in a neighborhood of the origin, so the observations are sub-exponential; more specifically, the following identity yields the local quadratic bound needed for our tests:

$$\log(\mathbb{E}_{Y \sim P_{\theta}}[\exp\{t[Y - \psi'(\theta)]\}]) = \psi(\theta + t) - \psi(\theta) - t\psi'(\theta)$$

Table 2: Representative regular natural exponential families under convenient reference measures ρ . The reference measure specifies the expression of θ in terms of the parameters of the family and is not unique. For the Gaussian, Binomial, negative binomial, Gamma, and inverse Gaussian families, respectively, σ^2 , N , r , α , and λ are held fixed. Here, β denotes the scale parameter for the Gamma and exponential families. For finite chi-square divergences, $p_1^2 < p_0$ is required for the negative binomial family, $2\beta_0 > \beta_1$ for the Gamma and exponential families, and $2\mu_0^2 \geq \mu_1^2$ for the inverse Gaussian family.

Name	Reference measure ρ	Support	Exponential tilt	Θ	$\chi^2(P_{\theta_1} \ P_{\theta_0})$
Gaussian	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx$	$x \in \mathbb{R}$	$\theta = \frac{\mu}{\sigma^2}$	$\mathbb{R}$	$\exp\left(\frac{(\mu_1 - \mu_0)^2}{\sigma^2}\right) - 1$
Bernoulli	$\delta_0 + \delta_1$	$x \in \{0, 1\}$	$\theta = \log \frac{p}{1-p}$	$\mathbb{R}$	$\frac{(p_1 - p_0)^2}{p_0(1-p_0)}$
Binomial	$\sum_{x=0}^N \binom{N}{x} \delta_x$	$x \in \{0, \dots, N\}$	$\theta = \log \frac{p}{1-p}$	$\mathbb{R}$	$\left(1 + \frac{(p_1 - p_0)^2}{p_0(1-p_0)}\right)^N - 1$
Poisson	$\sum_{x=0}^{\infty} \frac{1}{x!} \delta_x$	$x \in \mathbb{N}_0$	$\theta = \log \lambda$	$\mathbb{R}$	$\exp\left(\frac{(\lambda_1 - \lambda_0)^2}{\lambda_0}\right) - 1$
Negative binomial	$\sum_{x=0}^{\infty} \binom{x+r-1}{x} \delta_x$	$x \in \mathbb{N}_0$	$\theta = \log p$	$(-\infty, 0)$	$\left[\frac{p_0(1-p_1)^2}{(1-p_0)(p_0-p_1^2)}\right]^r - 1$
Gamma	$\frac{x^{\alpha-1}}{\Gamma(\alpha)} dx$	$x > 0$	$\theta = -\frac{1}{\beta}$	$(-\infty, 0)$	$\left[\frac{\beta_0^2}{\beta_1(2\beta_0 - \beta_1)}\right]^\alpha - 1$
Exponential	dx	$x > 0$	$\theta = -\frac{1}{\beta}$	$(-\infty, 0)$	$\frac{\beta_0^2}{\beta_1(2\beta_0 - \beta_1)} - 1$
Inverse Gaussian	$\sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda}{2x}\right) dx$	$x > 0$	$\theta = -\frac{\lambda}{2\mu^2}$	$(-\infty, 0)$	$\exp\left\{\lambda \left(\frac{2}{\mu_1} - \frac{1}{\mu_0} - \sqrt{\frac{2}{\mu_1^2} - \frac{1}{\mu_0^2}}\right)\right\} - 1$

$$= t^2 \int_0^1 (1-s) \psi''(\theta + st) ds,$$

where the last equality follows from Taylor's expansion with integral remainder. Therefore, if ψ'' is uniformly bounded near θ , the centered log moment generating function has a quadratic upper bound on a local interval; equivalently, the random variable satisfies a local sub-Gaussian concentration bound while remaining sub-exponential in general. This motivates a local regularity condition requiring the curvature of ψ near θ_0 and θ_1 to be bounded by a constant multiple of $v_0 := \psi''(\theta_0)$.

Definition 4 (Local regularity of the NEF models). *Let $r_n > 0$. We say that the NEF models P_{θ_0} and P_{θ_1} are locally regular at scale r_n if $[\theta_j - r_n, \theta_j + r_n] \subseteq \Theta$ for each $j \in \{0, 1\}$ and there exists a constant $C \geq 1$ such that*

$$\sup_{\substack{j \in \{0, 1\} \\ |u| \leq r_n}} \frac{\psi''(\theta_j + u)}{\psi''(\theta_0)} \leq C,$$

for all sufficiently large n .

Under local regularity, each coordinate Y_l , conditional on the latent structure, is locally sub-Gaussian with the common variance proxy Cv_0 , whether its distribution is P_{θ_0} or P_{θ_1} . Consequently, the average and scan tests attain risk guarantees analogous to those obtained under the Gaussian model. When the latent structure S is a subgraph of the complete graph K_m and each observation is indexed by an edge, we may also use the degree test introduced in Remark 3. The weighted degree of vertex i and the maximum-degree statistic are, respectively,

$$T_i^{\deg}(\mathbf{y}) := \sum_{\{e \in [m]_2 : i \in e\}} y_e \quad \text{and} \quad T_{\deg}(\mathbf{y}) := \max_{i \in [m]} T_i^{\deg}(\mathbf{y})$$

Conditional on the latent structure, each weighted degree is a sum of independent observations, so the same concentration argument gives an analogous guarantee under NEF models. The next proposition collects these bounds; its proof is given in Appendix B.11.

Proposition 3 (Upper bound for NEF models). *Consider the test statistics defined in Equation (14). Let $\delta = \delta_n \in (0, 1)$ be a target risk, and let $d_\theta := \psi'(\theta_1) - \psi'(\theta_0)$, $v_0 := \psi''(\theta_0)$, so that the effective signal strength is given by $\tau_n := d_\theta / \sqrt{v_0}$. Suppose that every $s \in \mathcal{C}$ satisfies $|s| = K \geq 1$, and let ν be a prior distribution on $\mathcal{C}$. Then,*

- Assume that the NEF models P_{θ_0} and P_{θ_1} are locally regular at scale $r_{\text{avg}} := \sqrt{2(nv_0)^{-1} \log(2/\delta)}$ with a constant $C_{\text{avg}} \geq 1$. For the threshold $t_{\text{avg}} := n\psi'(\theta_0) + Kd_\theta/2$, the average test $\Delta_{\text{avg}}(\mathbf{y}) := \mathbb{1}\{T_{\text{avg}}(\mathbf{y}) > t_{\text{avg}}\}$ satisfies $R_\nu(\Delta_{\text{avg}}) \leq \delta$ if

$$\tau_n \geq \sqrt{\frac{8C_{\text{avg}}n}{K^2} \log \frac{2}{\delta}}.$$

- Assume that the NEF models P_{θ_0} and P_{θ_1} are locally regular at scale $r_{\text{scan}} := \sqrt{2(Kv_0)^{-1} \log(2|\mathcal{C}|/\delta)}$ with a constant $C_{\text{scan}} \geq 1$. For the threshold $t_{\text{scan}} := K\psi'(\theta_0) + (Kd_\theta + \sqrt{2KC_{\text{scan}}v_0 \log |\mathcal{C}|})/2$, the scan test $\Delta_{\text{scan}}(\mathbf{y}) := \mathbb{1}\{T_{\text{scan}}(\mathbf{y}) > t_{\text{scan}}\}$ satisfies $R_\nu(\Delta_{\text{scan}}) \leq \delta$ if

$$\tau_n \geq \sqrt{\frac{2C_{\text{scan}}}{K}} \left(2\sqrt{\log \frac{2}{\delta}} + \sqrt{\log |\mathcal{C}|} \right).$$

Furthermore, suppose that the latent structure S is a subgraph of the complete graph K_m , so that $n = \binom{m}{2}$ and each observation is indexed by an edge $e \in [m]_2$ connecting two vertices. Let $T_{\text{deg}}(\mathbf{y}) := \max_{i \in [m]} T_i^{\text{deg}}(\mathbf{y})$, and suppose that the maximum number of signal edges incident to any vertex satisfies $\max_{i \in [m]} |\{e \in S : i \in e\}| = \mathfrak{d}_{\text{max}} \geq 1$ almost surely under $S \sim \nu$. Then,

- Assume that the NEF models P_{θ_0} and P_{θ_1} are locally regular with constant $C_{\text{deg}} \geq 1$ at scale

$$r_{\text{deg}} := \sqrt{2\{(m-1)v_0\}^{-1} \log(2m/\delta)}.$$

For the threshold $t_{\text{deg}} := (m-1)\psi'(\theta_0) + \mathfrak{d}_{\text{max}}d_\theta/2$, the degree test $\Delta_{\text{deg}}(\mathbf{y}) := \mathbb{1}\{T_{\text{deg}}(\mathbf{y}) > t_{\text{deg}}\}$ satisfies $R_\nu(\Delta_{\text{deg}}) \leq \delta$ if

$$\tau_n \geq \sqrt{\frac{8C_{\text{deg}}(m-1)}{\mathfrak{d}_{\text{max}}^2} \log \frac{2m}{\delta}}.$$

Applying the proposition generally requires verifying local regularity for the particular NEF model. For the Gaussian model with fixed variance, the condition holds globally with $C = 1$ because ψ'' is constant. For the Bernoulli model, if $p_0 < p_1$ and $p_1/p_0 = O(1)$, local regularity holds with some $C > 1$ for any sequence $r_n = o(1)$. If instead $p_1/p_0 \rightarrow \infty$, local regularity may fail because $\psi''(\theta_1)/\psi''(\theta_0) = p_1(1-p_1)/\{p_0(1-p_0)\}$ may diverge to infinity. Without local regularity, sums of observations are not guaranteed to satisfy the Gaussian-type tail bounds used above, so a different analysis may be needed.

Remark 6 (Generalized scan test). The scan test in Proposition 3 can be extended to random-size alternatives and scans targeting substructures of the planted support. Specifically, suppose that there exists a class $\mathcal{C}_\star$ of subsets of $[n]$, each of cardinality $K_\star$, such that

$$\Pr_{S \sim \nu_n} (\exists s_\star \in \mathcal{C}_\star : s_\star \subseteq S) \rightarrow 1.$$

Then, under the corresponding local regularity condition, the scan-test bound continues to hold with $(\mathcal{C}, K)$ replaced by $(\mathcal{C}_\star, K_\star)$, with an additional $o(1)$ term in the risk bound. Several scan tests considered in this paper are instances of this generalization. For example, in the random-size k -vertex clique problem, the planted support contains a clique on ck vertices with high probability for any fixed $c \in (0, 1)$, whereas in the fixed-size Δ -thick cycle problem, the planted support always contains a clique on $\Delta + 1$ vertices. In both cases, the scan test targets these cliques rather than the entire planted support; see Appendices B.6 and B.9. Similarly, the average and degree tests can be extended to random-size alternatives using high-probability lower bounds on the support size and maximum degree, respectively.

We next identify a regime in which local regularity holds automatically and need not be checked separately for each model.

4.3 Universality in the local alternative regime

We now identify a natural regime in which statistical and computational thresholds established by matching the preceding lower and upper bounds for one NEF model remain valid across other NEF models. Let

$$\theta_0 = c \quad \text{and} \quad \theta_1 = c + h_n,$$

where c is a fixed point in the interior of the natural parameter space Θ and $\{h_n\}$ is a vanishing positive sequence. We use $h_n > 0$ to keep the one-sided tests and the effective signal strength τ_n oriented consistently; the case $h_n < 0$ is analogous after reversing the tests and replacing τ_n by $|\tau_n|$. We refer to this setting as the *local alternative regime*. Here, “local” means that the one-coordinate alternative P_{θ_1} approaches the null P_{θ_0} ; it does not prescribe a particular contiguity scale for the full structured testing problem. For an n -fold product experiment in which every coordinate is shifted, differentiability in quadratic mean yields mutual contiguity of the sequences of product laws $P_{\theta_0}^{\otimes n}$ and $P_{\theta_1}^{\otimes n}$ when $h_n = O(n^{-1/2})$ (van der Vaart, 1998, Chapters 6–7), whereas strong detection requires $h_n = \omega(n^{-1/2})$. In our setting, however, only the coordinates in the random structure S are shifted, so the relevant detection scale depends on the size and geometry of S and on its prior. Nevertheless, convexity of the relative entropy, conditional independence given S , and $|S| \leq n$ give

$$D_{\text{KL}}(\mathbb{P}_{\nu_n} \parallel \mathbb{Q}_n) \leq nD_{\text{KL}}(P_{c+h_n} \parallel P_c) = O(nh_n^2).$$

The last equality follows from a second-order expansion at the fixed interior parameter c , and Equation (16) shows that any bounded subsequence of nh_n^2 prevents the risk from tending to zero. Although $h_n = \omega(n^{-1/2})$ is still necessary for strong detection, this condition alone does not guarantee distinguishability in our structured testing problems. We therefore define the local alternative regime simply through the vanishing condition $h_n \rightarrow 0$ and characterize distinguishability scales through the statistical and computational thresholds derived below.

We define τ_{stat}^* and τ_{comp}^* , analogously to μ_{stat}^* and μ_{comp}^* in Definition 3, as the statistical and computational thresholds in the local alternative regime. Throughout this subsection, we restrict attention to polynomially vanishing effective signal strength sequences τ_n , by which we mean sequences satisfying $\limsup_{n \rightarrow \infty} \log_n \tau_n < 0$.

Definition 5 (Statistical and computational thresholds in the local alternative regime). *Consider a sequence of testing problems defined in Equation (12) in the local alternative regime, and let τ_n denote the corresponding effective signal strength. A positive sequence τ_{stat}^* is called a statistical threshold in the local alternative regime if the following holds for every effective signal strength sequence τ_n :*

- if $\tau_n \gg \tau_{\text{stat}}^*$, then there exists a sequence of tests Δ_n that achieves strong detection;
- if $\tau_n \ll \tau_{\text{stat}}^*$, then no sequence of tests can achieve strong detection.

Similarly, a positive sequence τ_{comp}^* is called a computational threshold in the local alternative regime if

- if $\tau_n \gg \tau_{\text{comp}}^*$, then there exists a sequence of polynomial-time tests Δ_n that achieves strong detection;
- if $\tau_n \ll \tau_{\text{comp}}^*$, then, under Conjecture 1, no sequence of polynomial-time tests can achieve strong detection.

If strong detection is impossible for every polynomially vanishing effective signal sequence in the local alternative regime using arbitrary tests (respectively, polynomial-time tests), we adopt the convention $\tau_{\text{stat}}^* \asymp 1$ (respectively, $\tau_{\text{comp}}^* \asymp 1$), for which the achievability direction is vacuously satisfied.

We now state the main result of this subsection.

Theorem 5 (Universality of tightness in the local alternative regime via chi-square matching). *For $M \in \{A, B\}$, fix a nondegenerate NEF $\{P_{\theta}^M : \theta \in \Theta_M\}$ with reference measure ρ_M and a null parameter $c_M \in \text{int}(\Theta_M)$. Consider the corresponding families of testing problems in Equation (12), with*

$$\theta_0^M = c_M, \quad \theta_1^M = c_M + h_n^M, \quad h_n^M > 0, \quad h_n^M \rightarrow 0,$$

where h_n^M ranges over all positive sequences tending to zero for which the effective signal strength τ_n^M vanishes polynomially. Also suppose that every $s \in \mathcal{C}$ satisfies $|s| = K$. Both families use the same sequence of latent priors ν_n , fixed independently of the signal strength. Suppose that, for model A , the lower bounds obtained from the likelihood-ratio and degree- D_n likelihood-ratio second moments match the upper bounds in Proposition 3, achieved by the average and scan tests, together with the degree test when applicable, to establish statistical and computational thresholds $\tau_{stat}^{A,*}$ and $\tau_{comp}^{A,*}$ in the sense of Definition 5. If the scan test provides the upper bound, also assume that $\log |\mathcal{C}| = o(K)$. Then $\tau_{stat}^{A,*}$ and $\tau_{comp}^{A,*}$ are also valid statistical and computational thresholds, respectively, for model B , in the sense of Definition 5.

Although Theorem 5 is stated for fixed-size alternatives and relies on the upper bounds in Proposition 3, the same transfer argument extends to random-size alternatives through the generalized tests discussed in Remark 6. It also accommodates upper bounds obtained from substructure scans, provided that the corresponding upper and lower bounds match.

We prove Theorem 5 by simplifying the low-degree condition and verifying that local regularity holds automatically in the local alternative regime.

Lower bounds in the local alternative regime. The key to simplifying the lower-bound condition $D_n \log(\chi_n^2/\tau_n^2) = O(1)$ is the relation $\chi_n^2/\tau_n^2 = 1 + O(h_n^2)$. Indeed, the cumulant generating function $\log \mathbb{E}_{Y \sim P_{\theta_0}}[\exp(tY)] = \psi(\theta_0 + t) - \psi(\theta_0)$ gives

$$\begin{aligned} 1 + \chi^2(\theta_0, \theta_1) &= \mathbb{E}_{Y \sim P_{\theta_0}}[\exp\{2(\theta_1 - \theta_0)Y - 2(\psi(\theta_1) - \psi(\theta_0))\}] \\ &= \exp\{\psi(\theta_0 + 2h_n) - 2\psi(\theta_0 + h_n) + \psi(\theta_0)\}, \end{aligned}$$

while the effective signal strength is $\tau_n := (\psi'(\theta_1) - \psi'(\theta_0))/\psi''(\theta_0)^{1/2}$. Taylor expansions of ψ and its derivatives therefore yield

$$\frac{\chi_n^2}{\tau_n^2} = \frac{v_0 h_n^2 + c_3 h_n^3 + O(h_n^4)}{v_0 h_n^2 + c_3 h_n^3 + O(h_n^4)} = 1 + O(h_n^2),$$

where $v_0 := \psi''(\theta_0)$ and $c_3 := \psi^{(3)}(\theta_0)$. Combining this relation with $\log(1+x) = O(x)$ as $x \rightarrow 0$ proves the following proposition; further details appear in Appendix B.13.

Proposition 4 (Condition for the local alternative regime). *Let $\theta_0 = c$ for some constant $c \in \Theta$, and let $\theta_1 - \theta_0 =: h_n \rightarrow 0$. Then, the condition $D_n \log(\chi_n^2/\tau_n^2) = O(1)$ is satisfied as long as $D_n h_n^2 = O(1)$, or equivalently, $D_n \chi_n^2 = O(1)$ or $D_n \tau_n^2 = O(1)$.*

For two NEF models A and B with matching chi-square divergences, it is therefore enough to verify $D_n \chi_n^2 = O(1)$. By Proposition 4, this is equivalent to verifying any one of $D_n \tau_{A,n}^2 = O(1)$, $D_n \tau_{B,n}^2 = O(1)$, $D_n h_{A,n}^2 = O(1)$, or $D_n h_{B,n}^2 = O(1)$. Since we consider polylogarithmic degrees D_n , the condition $D_n \tau_n^2 = O(1)$ is automatically satisfied under the assumption that the effective signal strength τ_n vanishes polynomially with n .

Upper bounds in the local alternative regime. A second useful feature of the local alternative regime is that every NEF model automatically satisfies the local regularity condition in Definition 4. The following proposition formalizes this observation and allows us to apply the upper bounds in Proposition 3.

Proposition 5 (Local regularity in the local alternative regime). *Suppose that $\theta_0 \in \Theta$ is fixed and $\theta_1 \rightarrow \theta_0$. Then, for any sequence $r_n > 0$ satisfying $r_n \rightarrow 0$,*

$$\limsup_{n \rightarrow \infty} \sup_{\substack{j \in \{0,1\} \\ |u| \leq r_n}} \frac{\psi''(\theta_j + u)}{\psi''(\theta_0)} = 1.$$

In particular, for any fixed $C > 1$, the NEF models P_{θ_0} and P_{θ_1} are locally regular at scale r_n with constant C for sufficiently large n .

For fixed $\theta_0 \in \Theta$, continuity of ψ'' and the limits $\theta_1 \rightarrow \theta_0$ and $r_n \rightarrow 0$ prove the proposition. For the tests in Proposition 3, local regularity holds whenever $\log(2/\delta) = o(n)$ for the average test, $\log(2|\mathcal{C}|/\delta) = o(K)$ for the scan test, and $\log(2m/\delta) = o(m)$ for the degree test. Because the target risk δ may vanish arbitrarily slowly in the strong-detection problem, the average- and degree-test conditions hold automatically, while the scan-test condition reduces to $\log|\mathcal{C}| = o(K)$. The fixed-size scanning classes used for the models in Section 3, including the substructure scans described in Remark 6, satisfy this condition when k or Δ is polynomial in n .

Together, Propositions 4 and 5 show that, after matching chi-square divergences, both lower and upper bounds transfer across NEF models in the local alternative regime without model-specific analysis. Thus, thresholds that are tight for one NEF model remain tight for any other such model. The remaining technical details for Theorem 5 are deferred to Appendix B.14.

We conclude by relating τ_n to the usual parameters of two representative NEF models. For the Bernoulli model with success probabilities $p_0 < p_1$, we have $\tau_n = (p_1 - p_0)/\sqrt{p_0(1 - p_0)}$. For the Gamma model with fixed shape parameter α and scale parameters $\beta_0 < \beta_1$, we have $\tau_n = \sqrt{\alpha}(\beta_1 - \beta_0)/\beta_0$. In the local alternative regime, p_0 and β_0 are fixed, so $\tau_n \asymp p_1 - p_0$ for the Bernoulli model and $\tau_n \asymp \beta_1 - \beta_0$ for the Gamma model.

4.3.1 Applications to Erdős–Rényi models

In this subsection, we apply the preceding equivalence results to Erdős–Rényi observations, first for planted clique detection and then for arbitrary planted subgraphs.

Planted clique detection We revisit the planted clique problem in Section 3.1 to provide an explicit example of how our universality result operates in the local alternative regime, as well as of the possibilities and limitations of transferring lower and upper bounds beyond the local alternative regime. Specifically, we consider the Gaussian model in Section 2 and Erdős–Rényi models and examine how the applicability of our equivalence results depends on the asymptotic scaling of the success probability parameters in the Erdős–Rényi model. Consider the latent structure S defined in Equation (8) and the prior distribution ν for fixed-size case. Let model A and B be the standard Gaussian model and be the Erdős–Rényi, respectively. More specifically, we consider the model A

$$H_0^A : Y_l \stackrel{iid}{\sim} N(0, 1) \quad \text{v.s.} \quad H_1^A : Y_l | S \stackrel{ind}{\sim} \begin{cases} N(\mu, 1), & \text{if } l \in S, \\ N(0, 1), & \text{if } l \notin S, \end{cases}$$

with $\mu > 0$, and consider the model B

$$H_0^B : Y_l \stackrel{iid}{\sim} \text{Ber}(p_0) \quad \text{v.s.} \quad H_1^B : Y_l | S \stackrel{ind}{\sim} \begin{cases} \text{Ber}(p_1), & \text{if } l \in S, \\ \text{Ber}(p_0), & \text{if } l \notin S, \end{cases}$$

where $1 > p_1 > p_0 > 0$. Note that the two models can be expressed with exponential tilts with $(\theta_0^A, \theta_1^A) = (0, \mu)$ and $(\theta_0^B, \theta_1^B) = (\log(p_0/(1 - p_0)), \log(p_1/(1 - p_1)))$. The chi-square divergences and effective signal strengths of the two models are given by

$$\chi^2(P_{\theta_1^A}^A \| P_{\theta_0^A}^A) = \exp(\mu^2) - 1, \quad \chi^2(P_{\theta_1^B}^B \| P_{\theta_0^B}^B) = \frac{(p_1 - p_0)^2}{p_0(1 - p_0)}, \quad \tau_n^A = \mu, \quad \tau_n^B = \frac{p_1 - p_0}{\sqrt{p_0(1 - p_0)}}.$$

This motivates the scaling of the signal strength between the two models,

$$\mu = \sqrt{\log(1 + (p_1 - p_0)^2 / (p_0(1 - p_0)))},$$

which guarantees identical chi-square divergences, and thus the same $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}^M} [L_\nu^M(\mathbf{Y})^2]$ and equivalent $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}^M} [L_\nu^{M, \leq D}(\mathbf{Y})^2]$ for $M \in \{A, B\}$ under the assumption $D_n \chi_n^2 = O(1)$. We study how the equivalence result can be applied to the planted clique problem in the following three scenarios, depending on the polynomial orders of p_0 and p_1 in the Erdős–Rényi model. In particular, we analyze statistical and computational distinguishability in terms of the difference between the probability parameters $p_1 - p_0$ and define the

corresponding thresholds $(p_1 - p_0)_{\text{stat}}^*$ and $(p_1 - p_0)_{\text{comp}}^*$ analogously to Definition 3. Depending on the scenario, the model may fall within the local alternative regime or exhibit a polynomially vanishing effective signal strength. In such cases, a vacuous threshold may arise under the convention in Definition 5 or its extension. We explicitly indicate such cases when presenting the results.

- **Scenario 1:** $p_0 = c$ and $p_1 = c + m^{-\beta}$ for some fixed constants $c \in (0, 1)$ and $\beta > 0$. Note that this case corresponds to the local alternative regime. Although the assumptions for the lower and upper bounds can be verified directly for the Gaussian and Bernoulli models, we examine them through the local alternative framework to illustrate the universality of the tightness of the lower and upper bounds. For the lower bound, the condition $D_n \log(\chi_n^2/\tau_n^2) = O(1)$ boils down to $D_n \chi_n^2 = O(1)$ in Proposition 4, which is satisfied for sequence $D_n = (\log m)^{1+\epsilon}$ for any $\epsilon > 0$, since $\chi_n^2 \asymp m^{-2\beta}$. For the upper bound, the local regularity condition in Proposition 5 is automatically satisfied, and the condition $\log(|\mathcal{C}|) = o(K)$ is satisfied for the planted clique problem, since $|\mathcal{C}| = \binom{m}{k} \leq m^k$ and $\log m = o(k)$. Provided that $\chi^2/\tau_n^2 \rightarrow 1$ from $D_n \log(\chi_n^2/\tau_n^2) = O(1)$, $\tau_n^A/\tau_n^B \rightarrow 1$ is guaranteed, and we replace $\tau_n^A = \mu$ with $\tau_n^B = (p_1 - p_0)/\sqrt{p_0(1 - p_0)}$ in the statistical and computational thresholds for the Gaussian model displayed in Theorem 1 to obtain the corresponding thresholds for the Erdős–Rényi model:

$$(p_1 - p_0)_{\text{stat}}^* \asymp \begin{cases} \sqrt{\frac{1}{k}}, & k \ll m^{2/3}, \\ \frac{m}{k^2}, & k \gg m^{2/3}, \end{cases} \quad (p_1 - p_0)_{\text{comp}}^* \asymp \begin{cases} 1, & k \ll m^{1/2} \quad (\text{vacuous}), \\ \frac{m}{k^2}, & k \gg m^{1/2}, \end{cases}$$

provided that $\tau_n^B \asymp p_1 - p_0$. We note that the constant branch of the computational threshold for $k \ll m^{1/2}$ follows from our threshold convention in Definition 5.

- **Scenario 2:** $p_0 = m^{-\alpha}$ and $p_1 = c \cdot m^{-\alpha}$ for some fixed constants $c > 1$ and $\alpha \in (0, 2)$. First, we note that the case $\alpha > 2$ is not considered, because the expected number of edges in the planted clique under the alternative is $(n - K)p_0 + Kp_1 \leq cnp_0 = o(1)$. Consequently, for every fixed structure s , Markov's inequality gives $\mathbb{P}_s^B\{\mathbf{Y} = \mathbf{0}\} = 1 - o(1)$. Similarly, $\mathbb{Q}^B\{\mathbf{Y} = \mathbf{0}\} = 1 - o(1)$, and averaging the fixed-structure bound over $S \sim \nu$ gives $\mathbb{P}_\nu^B\{\mathbf{Y} = \mathbf{0}\} = 1 - o(1)$. It follows that the total variation distance between the null and alternative distributions converges to zero, so strong detection is impossible in this regime; hence, we restrict our attention to $\alpha \in (0, 2)$.

While this case does not fall into the local alternative regime, the model-specific analysis can be carried out to verify the assumptions for the lower and upper bounds. As pointed out earlier in Section 4.1, the condition $D_n \log(\chi_n^2/\tau_n^2) = O(1)$ in Theorem 4 is automatically satisfied for the Bernoulli model, and also is satisfied for the Gaussian model with $\mu^2 = O(D^{-1})$. Note that the chi-square matching between the two models yields $\mu^2 \asymp m^{-\alpha}$, and thus the condition $D_n \log(\chi_n^2/\tau_n^2) = O(1)$ is satisfied for sequence $D_n = (\log m)^{1+\epsilon}$ for any $\epsilon > 0$. For the upper bound, the local regularity condition in Proposition 3 is required. Note that the curvature $\psi''(\theta)$ is the variance of the model, and thus the Gaussian model always satisfies the local regularity condition with constant $C = 1$. For the Bernoulli model, we claim that for arbitrary sequence $r_n = O(1)$, the local regularity condition is satisfied with some constant $C > 1$. To see this, from $p(\theta) = e^\theta/(1 + e^\theta)$ and $\psi''(\theta) = p(\theta)(1 - p(\theta))$, observe that

$$\frac{\psi''(\theta_j + u)}{\psi''(\theta_j)} = \frac{e^u}{(1 - p(\theta_j) + p(\theta_j)e^u)^2} \leq e^{|u|},$$

for $j \in \{0, 1\}$ and $|u| \leq r_n$. Consequently, for any $R > 0$ such that $r_n \leq R$ for sufficiently large n , we have

$$\sup_{\substack{j \in \{0, 1\} \\ |u| \leq r_n}} \frac{\psi''(\theta_j + u)}{\psi''(\theta_0)} = \sup_{\substack{j \in \{0, 1\} \\ |u| \leq r_n}} \left(\frac{\psi''(\theta_j + u)}{\psi''(\theta_j)} \frac{\psi''(\theta_j)}{\psi''(\theta_0)} \right) \leq \max \left\{ 1, \frac{p(\theta_1)(1 - p(\theta_1))}{p(\theta_0)(1 - p(\theta_0))} \right\} \sup_{\substack{j \in \{0, 1\} \\ |u| \leq r_n}} \frac{\psi''(\theta_j + u)}{\psi''(\theta_j)}.$$

Therefore, if $p(\theta_1)/p(\theta_0) = c > 1$ for some constant c , the local regularity condition holds with $C = ce^R$ for sufficiently large n .

It remains to verify $r_{\text{avg}} = O(1)$ and $r_{\text{scan}} = O(1)$. Note that $n \asymp m^2$ and $v_0 \asymp m^{-\alpha}$, and thus $r_{\text{avg}} = \sqrt{2(nv_0)^{-1} \log(2/\delta)} \asymp m^{-1+\alpha/2} \sqrt{\log(2/\delta)}$, and we can find $\delta \rightarrow 0$ which satisfies $r_{\text{avg}} = O(1)$, provided that $\alpha < 2$. On the other hand, we have

$$r_{\text{scan}}^2 = 2(Kv_0)^{-1} \log(2|\mathcal{C}|/\delta) \asymp \frac{\log \binom{m}{k} + \log(1/\delta)}{k^2 m^{-\alpha}}.$$

Therefore, $r_{\text{scan}} = O(1)$ is satisfied if $\log \binom{m}{k} = O(k^2 m^{-\alpha})$ and $\delta \rightarrow 0$ sufficiently slowly. Substituting μ with $\tau_n^B \asymp \sqrt{p_1 - p_0}$, the Gaussian statistical and computational detection bounds translate into the following thresholds for the Erdős–Rényi model

$$(p_1 - p_0)_{\text{stat}}^* \asymp \begin{cases} \frac{1}{k}, & k \ll m^{2/3}, \\ \frac{m^2}{k^4}, & k \gg m^{2/3}, \end{cases} \quad (p_1 - p_0)_{\text{comp}}^* \asymp \begin{cases} 1, & k \ll m^{1/2} \\ \frac{m^2}{k^4}, & k \gg m^{1/2}. \end{cases} \quad (\text{vacuous}),$$

We note that the additional condition $\log \binom{m}{k} = O(k^2 m^{-\alpha})$ required for local regularity is automatically satisfied whenever the scan test achieves strong detection at polynomial order. Thus, it imposes no further restriction on the scan-test upper bound. We also remark that, although this scenario does not fall within the local alternative regime, the effective signal strength τ_n^B always vanishes polynomially. When $k \ll m^{1/2}$, the transferred low-degree lower bound rules out computational strong detection for every such signal in this scenario. The constant branch of the computational threshold represents this impossibility under our convention.

- **Scenario 3:** $p_0 = m^{-\alpha}$ and $p_1 = m^{-\beta}$ for some fixed $\alpha > \beta > 0$. To illustrate how our equivalence result transfers and to point out its limitations, we focus on the regime $\alpha < 1$, where both lower and upper bounds have been carefully studied in the existing literature. For instance, the tight statistical lower and upper bounds are established in [Arias-Castro and Verzelen \(2014\)](#); [Verzelen and Arias-Castro \(2015\)](#), and the tight computational lower and upper bounds are studied in [Dhawan et al. \(2025\)](#). We rewrite their results in terms of $p_1 - p_0$, with our notation:

$$(p_1 - p_0)_{\text{stat}}^* \asymp \begin{cases} \frac{1}{k}, & k \ll m^{1-\alpha/2}, \\ \frac{m^{1-\alpha/2}}{k^2}, & k \gg m^{1-\alpha/2}, \end{cases} \quad (p_1 - p_0)_{\text{comp}}^* \asymp \begin{cases} \frac{1}{k^\alpha}, & k \ll m^{1/2}, \\ \frac{m^{1-\alpha/2}}{k^2}, & k \gg m^{1/2}. \end{cases}$$

With these existing thresholds, we observe below that the lower- and upper-bound transfers provided by Theorem 4 and Proposition 3 are either loose or inapplicable in this scenario.

We first restrict our attention to the case $\alpha < 2\beta$. In this case, the chi-square matching between the two models requires that $\mu^2 \asymp \log(1 + m^{\alpha-2\beta})$, which is of the polynomially vanishing order $m^{\alpha-2\beta}$, and thus the degree condition $\mu^2 = O(D^{-1})$ for the Gaussian model is satisfied for $D_n = (\log m)^{1+\epsilon}$. Consequently, both the lower bounds from the Gaussian model are transferrable to the Erdős–Rényi model through Theorem 4; the second moments of the likelihood ratio (LR) and the low-degree likelihood ratio (LDLR) are bounded if

$$\text{LR} : (p_1 - p_0) \ll \begin{cases} \frac{m^{-\alpha/2}}{\sqrt{k}}, & k \ll m^{2/3}, \\ \frac{m^{1-\alpha/2}}{k^2}, & k \gg m^{2/3}, \end{cases} \quad \text{LDLR} : (p_1 - p_0) \ll \begin{cases} \frac{1}{m^{\alpha/2}}, & k \ll m^{1/2}, \\ \frac{m^{1-\alpha/2}}{k^2}, & k \gg m^{1/2}, \end{cases}$$

respectively. Observe that the first branches of the transferred lower bounds are loose compared to the tight statistical and computational thresholds above. On the other hand, when $\alpha > 2\beta$, the chi-square matching yields $\mu^2 \asymp \log m$, causing the full likelihood ratio second moment of the matched Gaussian model to diverge. Consequently, although the second-moment equivalence remains valid, it provides no nontrivial statistical lower bound in this regime. In addition, the degree condition $\mu^2 = O(D^{-1})$ is now violated for any $D_n = (\log m)^{1+\epsilon}$, and thus the low-degree equivalence cannot be applied, and transferring the computational lower bound from the Gaussian model through Theorem 4 is not applicable.

This observation helps clarify the scope of applicability of Theorem 4: its lower-bound transfer applies to bounds obtained through the second moment method or conditional second moment arguments based on the latent structure $S \sim \nu$, rather than to arbitrary lower bounds. Indeed, the tight statistical and computational lower bounds derived in the literature for the Bernoulli model rely on model-specific truncation arguments that exploit properties of the observed graph, which are not covered by Theorem 4. For the statistical lower bound, Arias-Castro and Verzelen (2014); Verzelen and Arias-Castro (2015) use a truncated likelihood ratio based on a high-probability event that excludes atypical configurations in which some subset of the planted community contains unusually many edges. Similarly, for the computational lower bound, Dhawan et al. (2025) obtains a sharp bound by excluding rare realizations of $\mathbf{Y}$ in which the planted subgraph contains an unusually dense small subgraph.

For the upper bounds, the local regularity condition fails for the Bernoulli model, regardless of the choice of $r_n = O(1)$, since $\psi''(\theta_1)/\psi''(\theta_0) = p_1(1-p_1)/p_0(1-p_0) \asymp m^{\alpha-\beta} \rightarrow \infty$. Consequently, the Gaussian upper bounds cannot be directly transferred to the Bernoulli model. Nevertheless, statistical upper bounds can be established directly by analyzing the average and scan tests using Bernoulli tail bounds rather than sub-Gaussian approximations. Such bounds have been established in Arias-Castro and Verzelen (2014); Verzelen and Arias-Castro (2015); Dhawan et al. (2025). In particular, the sufficient conditions for these tests to achieve strong detection are

$$(p_1 - p_0) \gg \begin{cases} \frac{1}{k}, & k \ll m^{1-\alpha/2}, \\ \frac{m^{1-\alpha/2}}{k^2}, & k \gg m^{1-\alpha/2}. \end{cases}$$

where the first branch is achieved by the scan test and the second by the average test. These upper bounds match the statistical thresholds stated above. However, the average test alone does not attain the sharp computational threshold in the regime $k \ll m^{1/2}$. To close this gap, Dhawan et al. (2025) develop a polynomial-time algorithm based on counting copies of a suitably chosen constant-size balanced subgraph. In this regime, their test achieves vanishing risk whenever $(p_1 - p_0) \gg k^{-\alpha}$, thereby attaining the computational threshold.

Detecting arbitrary planted subgraphs. For Erdős–Rényi models, the statistical and computational thresholds in the fixed-size case are studied in Elimelech and Huleihel (2025) in terms of some graph-theoretic properties such as maximum degree, maximum subgraph density and the number of edges or vertices. Notably, in the local alternative regime under the assumptions of their Theorem 16, the computational threshold at polynomial order is determined by the detection boundaries of the average and degree tests. We remark that the computational lower bounds in their work are established using the low-degree second moment method¹. Moreover, their upper-bound analysis yields the same asymptotic polynomial-order signal requirements as our analysis. By the transfer argument underlying Theorem 5, with computational thresholds interpreted under a logarithmic-degree version of the low-degree conjecture, the corresponding lower and upper bounds can be transferred to arbitrary NEF models through chi-square matching. Consequently, the corresponding computational threshold in the local alternative regime is characterized as

$$\tau_{\text{comp}}^* \asymp 1 \wedge \frac{m}{K} \wedge \frac{\sqrt{m}}{\mathfrak{d}_{\max}(S)} \quad (15)$$

up to subpolynomial factors, under a version of the low-degree conjecture that allows $D = O(\log m)$.

Now we shift our focus to the statistical threshold. For the upper bound, their scan statistic differs from our naive scan statistic in that it scans copies of a maximum-density subgraph $H^* \subseteq S$, rather than copies of the entire planted graph S . Nevertheless, the performance of this modified scan statistic can be analyzed by essentially the same argument as in Appendix B.11.2. Let $\mathcal{C}_{H^*}$ denote the collection of edge sets of all copies of H^* in the ambient complete graph, and let $K_* = |H^*|$. The proof of our NEF scan bound applies to this modified statistic with $(\mathcal{C}, K)$ replaced by $(\mathcal{C}_{H^*}, K_*)$. The resulting sufficient condition for the modified scan test is then $\tau_n \geq \sqrt{2C_{\text{scan}}K_*^{-1}}(2\sqrt{\log(2/\delta)} + \sqrt{\log|\mathcal{C}_{H^*}|})$.

¹However, we note that their lower bounds (Theorem 16) assume $D = O(\log m)$ instead of $D = \log^{1+\epsilon} m$.

This condition is satisfied if $\liminf_{m \rightarrow \infty} |V(H^*)|^{-1} K_\star (\log m)^{-1} \tau_n^2 > 2$, provided that $|\mathcal{C}_{H^*}| \leq m^{|V(H^*)|}$ and δ vanishes sufficiently slowly. This is of the same asymptotic order as the scan detection boundary $\liminf_{m \rightarrow \infty} |V(H^*)|^{-1} K_\star (\log m)^{-1} D_{\text{KL}}(\text{Ber}(p_1) \parallel \text{Ber}(p_0)) > 1$ in [Elimelech and Huleihel, 2025](#), as can be verified from the fact that the Taylor expansion of the KL divergence in the local alternative regime yields $D_{\text{KL}}(\text{Ber}(p_1) \parallel \text{Ber}(p_0)) = \tau_n^2(1 + o(1))/2$. Moreover, since $\tau_n \rightarrow 0$ in the local alternative regime, the above sufficient condition necessarily implies $|V(H^*)| K_\star^{-1} \log m \rightarrow 0$. Together with $|\mathcal{C}_{H^*}| \leq m^{|V(H^*)|}$, the local regularity condition $\log |\mathcal{C}_{H^*}| = o(K_\star)$ required for our NEF scan bound is automatically satisfied.

Therefore, the upper bound of [Elimelech and Huleihel, 2025](#) which consists of the average, maximum-degree, and modified scan tests can be transferred to any other NEF model through chi-square matching, and this implies that strong detection is possible if

$$\tau_n \gg \frac{m}{K} \wedge \frac{\sqrt{m}}{\mathfrak{d}_{\max}(S)} \wedge \sqrt{\frac{|V(H^*)|}{K_\star}}$$

in the local alternative regime.

On the other hand, the lower bound in their work is obtained by controlling the second moment of the likelihood ratio, and hence can again be transferred to another NEF model through chi-square matching. In our notation, their statistical lower bound implies that statistical detection is impossible when

$$\tau_n \ll \begin{cases} \frac{m}{K} \wedge \frac{\sqrt{m}}{\mathfrak{d}_{\max}(S)} \sqrt{\frac{|V(H^*)|}{K_\star}} \wedge \sqrt{\frac{|V(H^*)|}{K_\star}}, & \lim_{m \rightarrow \infty} \frac{\log(K_\star/|V(H^*)|)}{\log |V(S)|} < 1, \\ \frac{m}{K} \wedge \frac{\sqrt{m}}{\mathfrak{d}_{\max}(S)} \sqrt{\frac{|V(H^*)|}{K_\star}}, & \lim_{m \rightarrow \infty} \frac{\log(K_\star/|V(H^*)|)}{\log |V(S)|} = 1. \end{cases}$$

Since these lower bounds are obtained through the likelihood-ratio second moment, our equivalence result extends the same statistical impossibility regions to arbitrary NEF models in the local alternative regime through chi-square matching.

While these results provide a useful structure-agnostic characterization of the statistical threshold, the corresponding lower and upper bounds do not match in all regimes, as noted by [Elimelech and Huleihel \(2025\)](#). Our multiple-clique problem provides a concrete example in which these general statistical lower and upper bounds do not match. Consider γ disjoint k -cliques with $k = m^{3/5}$ and $\gamma = m^{1/20}$. Then

$$K \asymp \gamma k^2 = m^{5/4}, \quad \mathfrak{d}_{\max}(S) \asymp k = m^{3/5},$$

and we may take H^* to be a single k -clique, for which $|V(H^*)| = k$ and $K_\star \asymp k^2$. Moreover, since

$$\frac{\log(K_\star/|V(H^*)|)}{\log |V(S)|} \rightarrow \frac{12}{13} < 1,$$

the first case of the statistical lower bound applies. The general upper bound is then of order

$$m^{-1/4} \wedge m^{-1/10} \wedge m^{-3/10} = m^{-3/10},$$

whereas the general lower bound is only of order

$$m^{-1/4} \wedge m^{-2/5} \wedge m^{-3/10} = m^{-2/5}.$$

Thus, the general lower and upper bounds leave a polynomial gap in this example. In contrast, our analysis of the lower bound captures the correct polynomial order and thereby closes this gap. We also point out that, while the statistical and computational thresholds in our main text are stated only up to polynomial order for clarity of presentation, the corresponding appendices retain sharper signal scales including logarithmic factors. These more refined rates do not follow directly from the general bounds of [Elimelech and Huleihel, 2025](#), particularly in regimes where those bounds do not match. Finally, we emphasize that the framework of [Elimelech and Huleihel, 2025](#) is formulated for ordinary graphs, where each edge is a pair of vertices, and does not directly cover hypergraph settings such as [Corollary 2](#). Accordingly, the statistical and computational thresholds for NEF models obtained above by transferring their results should likewise be understood as applying to the ordinary graph setting.

5 Discussion

Our results identify the overlap distribution of two independent latent structures as a common basis for analyzing statistical and computational thresholds in graph combinatorial testing. In the Gaussian model, both the full and low-degree likelihood-ratio second moments depend only on this overlap, yielding a unified analysis of planted cliques, multiple cliques and hypercliques, and thick cycles and paths. Comparing the resulting lower bounds with the scan, average, and coordinatewise maximum tests identifies regimes with statistical-computational gaps. Spectral methods provide another important approach to planted-structure detection, and, under suitable conditions on the spectral gap and moments, matrix powers or trace moments connect such methods to low-degree polynomial statistics (Kunisky et al., 2019; Wein, 2025). We do not analyze spectral tests separately because our benchmark polynomial-time tests already match the low-degree lower bounds at polynomial order for the Gaussian models considered. The low-degree lower bounds rule out strong separation by polynomial statistics of polylogarithmic degree; their interpretation as lower bounds for all polynomial-time tests relies on the working low-degree hypothesis in Conjecture 1.

The natural exponential family results separate the role of the latent structure from that of the observation distribution. Under the conditions of Theorem 4, matching the one-coordinate chi-square divergence preserves the full likelihood-ratio second moments and makes the low-degree second moments equivalent up to constant factors. The upper bounds depend on the observation model through an effective signal strength and a local regularity condition. In the local alternative regime, chi-square matching also matches the effective signal strengths asymptotically, and local regularity holds automatically; under its stated assumptions, Theorem 5 therefore transfers tightness of these bounds across natural exponential families. Outside this regime, chi-square matching alone need not yield sharp thresholds: the relevant signal scales may differ, local regularity may fail, and sharp lower bounds may require observation-dependent truncations. The sparse Erdős–Rényi planted-clique example in Section 4.3.1 illustrates these limitations.

Several questions remain beyond the polynomial-order thresholds established here, including sharp logarithmic factors, optimal constants, and behavior on critical phase boundaries. Our sharp results focus on strong detection with observations that are independent conditional on the latent structure, known noise parameters, and symmetric priors; weak detection, recovery, adaptation, and models with conditionally dependent or heterogeneous observations may exhibit different behavior. Specialized polynomial-time statistics may also improve on the benchmark tests outside the local alternative regime.

One direction for further work is to characterize broader classes of graphs and hypergraphs through structural properties of their overlap distributions. Another direction is to study mean-matched nulls, whose coordinate means agree with those of the mixture alternative. This removes the mean shift exploited by our average test and calls for a separate analysis of detection thresholds and distributional transfer. A transfer principle for observation-dependent conditional second moments could help extend the universality theory beyond the local alternative regime. Rigorous average-case reductions or other complexity-theoretic lower bounds would help determine when our low-degree predictions describe the limits of all polynomial-time algorithms.

Acknowledgments

We thank Wasim Huleihel and Matey Neykov for helpful discussions. Ikjun Choi and Alessandro Rinaldo acknowledge partial support from the NIH Award RF1NS121913.

References

- Louigi Addario-Berry, Nicolas Broutin, Luc Devroye, and Gábor Lugosi. On combinatorial testing problems. *Annals of Statistics*, 38(5):3063–3092, 2010. URL <https://doi.org/10.1214/10-AOS817>.
- Louigi Addario-Berry, Omer Angel, Gábor Lugosi, Miklós Z. Rácz, and Tselil Schramm. The statistical

- threshold for planted matchings and spanning trees. arXiv: 2602.07669, 2026. URL <https://arxiv.org/abs/2602.07669>.
- Thomas D. Ahle. Sharp and simple bounds for the raw moments of the binomial and Poisson distributions. *Statistics & Probability Letters*, 182:109306, 2022. URL <https://doi.org/10.1016/j.spl.2021.109306>.
- Kalle Alaluusua and B. R. Vinay Kumar. Planted clique detection and recovery from the hypergraph adjacency matrix. arXiv: 2604.08691, 2026. URL <https://arxiv.org/abs/2604.08691>.
- Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures & Algorithms*, 13(3–4):457–466, 1998. URL <https://web.math.princeton.edu/~nalon/PDFS/clique3.pdf>.
- Noga Alon, Alexandr Andoni, Tali Kaufman, Kevin Matulef, Ronitt Rubinfeld, and Ning Xie. Testing k-wise and almost k-wise independence. In *Proceedings of the ACM Symposium on Theory of Computing*, 2007. URL <https://doi.org/10.1145/1250790.1250863>.
- Ery Arias-Castro and Nicolas Verzelen. Community detection in dense random networks. *Annals of Statistics*, 42(3):940–969, 2014. URL <https://arxiv.org/abs/1302.7099>.
- Ery Arias-Castro, Emmanuel J. Candès, Hannes Helgason, and Ofer Zeitouni. Searching for a trail of evidence in a maze. *Annals of Statistics*, 36(4):1726–1757, 2008. URL <https://arxiv.org/abs/math/0701668>.
- Ery Arias-Castro, Emmanuel J. Candès, and Arnaud Durand. Detection of an anomalous cluster in a network. *Annals of Statistics*, 39(1):278–304, 2011. URL <https://arxiv.org/abs/1001.3209>.
- Ery Arias-Castro, Rui M. Castro, Ervin Tanczos, and Meng Wang. Distribution-free detection of structured anomalies: Permutation and rank-based scans. *Journal of the American Statistical Association*, 113(522):789–801, 2018. URL <https://arxiv.org/abs/1508.03002>.
- Afonso S. Bandeira, Jess Banks, Dmitriy Kunisky, Cristopher Moore, and Alexander S. Wein. Spectral planting and the hardness of refuting cuts, colorability, and communities in random graphs. In *Proceedings of the Conference on Learning Theory*, 2021. URL <https://proceedings.mlr.press/v134/bandeira21a.html>.
- 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*, 2022. URL <https://arxiv.org/abs/2205.09727>.
- Jess Banks, Cristopher Moore, Roman Vershynin, Nicolas Verzelen, and Jiaming Xu. Information-theoretic bounds and phase transitions in clustering, sparse PCA, and submatrix localization. In *Proceedings of the IEEE International Symposium on Information Theory*, 2017. URL <https://doi.org/10.1109/ISIT.2017.8006706>.
- 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. URL <https://doi.org/10.1137/17M1138236>.
- Quentin Berthet and Philippe Rigollet. Complexity theoretic lower bounds for sparse principal component detection. In *Proceedings of the Conference on Learning Theory*, 2013. URL <https://proceedings.mlr.press/v30/Berthet13.html>.
- Matthew S. Brennan and Guy Bresler. Reducibility and statistical-computational gaps from secret leakage. In *Proceedings of the Conference on Learning Theory*, 2020. URL <https://proceedings.mlr.press/v125/brennan20a.html>.

- Matthew S. Brennan, Guy Bresler, and Wasim Huleihel. Reducibility and computational lower bounds for problems with planted sparse structure. In *Proceedings of the Conference on Learning Theory*, 2018. URL <https://proceedings.mlr.press/v75/brennan18a.html>.
- Matthew S. Brennan, Guy Bresler, and Wasim Huleihel. Universality of computational lower bounds for submatrix detection. In *Proceedings of the Conference on Learning Theory*, 2019. URL <https://proceedings.mlr.press/v99/brennan19a.html>.
- Rares-Darius Buhai, Jun-Ting Hsieh, Aayush Jain, and Pravesh K. Kothari. The quasi-polynomial low-degree conjecture is false. In *Proceedings of the IEEE Symposium on Foundations of Computer Science*, 2025. URL <https://arxiv.org/abs/2505.17360>.
- Cristina Butucea and Yuri I. Ingster. Detection of a sparse submatrix of a high-dimensional noisy matrix. *Bernoulli*, 19(5B):2652–2688, 2013. URL <https://arxiv.org/abs/1109.0898>.
- T. Tony Cai, Tengyuan Liang, and Alexander Rakhlin. Computational and statistical boundaries for submatrix localization in a large noisy matrix. *Annals of Statistics*, 45(4):1403–1430, 2017. URL <https://doi.org/10.1214/16-AOS1488>.
- Alexandra Carpentier, Simone Maria Giancola, Christophe Giraud, and Nicolas Verzelen. Low-degree lower bounds via almost orthonormal bases. arXiv: 2509.09353, 2025. URL <https://arxiv.org/abs/2509.09353>.
- Urmisha Chatterjee, Karissa Huang, Ritabrata Karmakar, B. R. Vinay Kumar, Gábor Lugosi, Nandan Malhotra, Anirban Mandal, and Maruf Alam Tarafdar. Detecting weighted hidden cliques. arXiv: 2506.21543, 2025. URL <https://arxiv.org/abs/2506.21543>.
- Siyu Chen, Theodor Misiakiewicz, Ilias Zadik, and Peiyuan Zhang. An optimized Franz–Parisi criterion and its equivalence with SQ lower bounds. arXiv: 2506.06259, 2025. URL <https://arxiv.org/abs/2506.06259>.
- Xuan Chen and Shuangping Li. Rigorous low-degree implications for planted subgraph detection: Noise and treewidth. arXiv: 2608.06279, 2026. URL <https://arxiv.org/abs/2608.06279>.
- 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. URL <https://jmlr.org/papers/v17/14-330.html>.
- Luca Corinzia, Paolo Penna, Wojciech Szpankowski, and Joachim M. Buhmann. Statistical and computational thresholds for the planted k -densest sub-hypergraph problem. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, 2022. URL <https://proceedings.mlr.press/v151/corinzia22a.html>.
- Marom Dadon, Wasim Huleihel, and Tamir Bendory. Detection and recovery of hidden submatrices. *IEEE Transactions on Signal and Information Processing over Networks*, 10:69–82, 2024. URL <https://doi.org/10.1109/TSIPN.2024.3352264>.
- Víctor H. de la Peña and Evarist Giné. *Decoupling*. Springer, 1999. URL <https://doi.org/10.1007/978-1-4612-0537-1>.
- Yash Deshpande and Andrea Montanari. Finding hidden cliques of size $\sqrt{N/e}$ in nearly linear time. *Foundations of Computational Mathematics*, 15(4):1069–1128, 2015a. URL <https://doi.org/10.1007/s10208-014-9215-y>.
- Yash Deshpande and Andrea Montanari. Improved sum-of-squares lower bounds for hidden clique and hidden submatrix problems. In *Proceedings of the Conference on Learning Theory*, 2015b. URL <https://proceedings.mlr.press/v40/Deshpande15.html>.

- Abhishek Dhawan, Cheng Mao, and Alexander S. Wein. Detection of dense subhypergraphs by low-degree polynomials. *Random Structures & Algorithms*, 66(1):e21279, 2025. URL <https://doi.org/10.1002/rsa.21279>.
- Jingqiu Ding, Yiding Hua, Lucas Slot, and David Steurer. Low-degree evidence for computational transition of recovery rate in stochastic block model. In *Advances in Neural Information Processing Systems*, 2025. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/a789f19f6454f858ad74bc8b7581fe36-Paper-Conference.pdf.
- David Donoho and Jiashun Jin. Higher criticism for detecting sparse heterogeneous mixtures. *Annals of Statistics*, 32(3):962–994, 2004. URL <https://doi.org/10.1214/009053604000000265>.
- Dor Elimelech and Wasim Huleihel. Detecting arbitrary planted subgraphs in random graphs. arXiv: 2503.19069, 2025. URL <https://arxiv.org/abs/2503.19069>.
- Vitaly Feldman, Elena Grigorescu, Lev Reyzin, Santosh S. Vempala, and Ying Xiao. Statistical algorithms and a lower bound for detecting planted cliques. *Journal of the ACM*, 64(2):1–37, 2017. URL <https://doi.org/10.1145/3046674>.
- Helmut Finner. A generalization of Hölder’s inequality and some probability inequalities. *Annals of Probability*, 20(4):1893–1901, 1992. URL <https://doi.org/10.1214/aop/1176989534>.
- Evarist Giné, Rafał Łatała, and Joel Zinn. Exponential and moment inequalities for U-statistics. In *High Dimensional Probability II*, pages 13–38. Birkhäuser Boston, 2000. URL https://doi.org/10.1007/978-1-4612-1358-1_2.
- Subhajit Goswami and Rajarshi Mukherjee. Sharp minimax risks and phase transitions in sparse submatrix detection. arXiv: 2605.31583, 2026. URL <https://arxiv.org/abs/2605.31583>.
- Bruce Hajek, Yihong Wu, and Jiaming Xu. Computational lower bounds for community detection on random graphs. In *Proceedings of the Conference on Learning Theory*, 2015. URL <https://proceedings.mlr.press/v40/Hajek15.html>.
- Bruce Hajek, Yihong Wu, and Jiaming Xu. Submatrix localization via message passing. *Journal of Machine Learning Research*, 18(186):1–52, 2018. URL <http://jmlr.org/papers/v18/17-297.html>.
- Shuichi Hirahara and Nobutaka Shimizu. Planted clique conjectures are equivalent. In *Proceedings of the ACM Symposium on Theory of Computing*, 2024. URL <https://doi.org/10.1145/3618260.3649751>.
- Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. URL <https://www.jstor.org/stable/2282952>.
- Justin Holmgren and Alexander S. Wein. Counterexamples to the low-degree conjecture. In *Proceedings of the Innovations in Theoretical Computer Science Conference*, 2021. URL <https://doi.org/10.4230/LIPIcs.ITCS.2021.75>.
- Samuel B. Hopkins. Statistical inference and the sum of squares method. Ph.D. thesis, Cornell University, 2018. URL <https://hdl.handle.net/1813/59618>.
- Samuel B. Hopkins and David Steurer. Efficient Bayesian estimation from few samples: Community detection and related problems. In *Proceedings of the IEEE Symposium on Foundations of Computer Science*, 2017. URL <https://arxiv.org/abs/1710.00264>.
- Samuel B. Hopkins, Pravesh K. Kothari, Aaron Potechin, Prasad Raghavendra, Tselil Schramm, and David Steurer. The power of sum-of-squares for detecting hidden structures. In *Proceedings of the IEEE Symposium on Foundations of Computer Science*, 2017. URL <https://arxiv.org/abs/1710.05017>.

- Zhe Hou and Jingcheng Liu. Recovery thresholds for hidden weighted sparse graphs. arXiv: 2606.14335, 2026. URL <https://arxiv.org/abs/2606.14335>.
- Jun-Ting Hsieh, Daniel M. Kane, Pravesh K. Kothari, Jerry Li, Sidhanth Mohanty, and Stefan Tiegel. Rigorous implications of the low-degree heuristic. arXiv: 2601.05850, 2026. URL <https://arxiv.org/abs/2601.05850>.
- Wasim Huleihel. Inferring hidden structures in random graphs. *IEEE Transactions on Signal and Information Processing over Networks*, 8:855–867, 2022. URL <https://doi.org/10.1109/TSIPN.2022.3211208>.
- Wasim Huleihel. Recovery of planted subgraphs. In *Proceedings of the Conference on Learning Theory*, 2026. URL <https://proceedings.mlr.press/v336/huleihel26a.html>.
- Mark Jerrum. Large cliques elude the Metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992. URL <https://doi.org/10.1002/rsa.3240030402>.
- Pravesh K. Kothari, Santosh S. Vempala, Alexander S. Wein, and Jeff Xu. Is planted coloring easier than planted clique? In *Proceedings of the Conference on Learning Theory*, 2023. URL <https://proceedings.mlr.press/v195/kothari23a.html>.
- Dmitriy Kunisky. Hypothesis testing with low-degree polynomials in the Morris class of exponential families. In *Proceedings of the Conference on Learning Theory*, 2021. URL <https://proceedings.mlr.press/v134/kunisky21a.html>.
- Dmitriy Kunisky. Low coordinate degree algorithms I: Universality of computational thresholds for hypothesis testing. *Annals of Statistics*, 53(2):774–801, 2025. URL <https://doi.org/10.1214/24-AOS2484>.
- Dmitriy Kunisky, Alexander S. Wein, and Afonso S. Bandeira. Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. arXiv: 1907.11636, 2019. URL <https://arxiv.org/abs/1907.11636>.
- Lucien M. Le Cam. *Locally Asymptotically Normal Families of Distributions: Certain Approximations to Families of Distributions and Their Use in the Theory of Estimation and Testing Hypotheses*. University of California Press, 1960.
- Yuetian Luo and Anru R. Zhang. Tensor clustering with planted structures: Statistical optimality and computational limits. *Annals of Statistics*, 50(1):584–613, 2022. URL <https://doi.org/10.1214/21-AOS2123>.
- Zongming Ma and Yihong Wu. Computational barriers in minimax submatrix detection. *Annals of Statistics*, 43(3):1089–1116, 2015. URL <https://doi.org/10.1214/14-AOS1300>.
- Cheng Mao, Alexander S. Wein, and Shenduo Zhang. Detection-recovery gap for planted dense cycles. In *Proceedings of the Conference on Learning Theory*, 2023. URL <https://proceedings.mlr.press/v195/mao23a.html>.
- Cheng Mao, Alexander S. Wein, and Shenduo Zhang. Information-theoretic thresholds for planted dense cycles. *IEEE Transactions on Information Theory*, 71(2):1266–1282, 2025. URL <https://doi.org/10.1109/TIT.2024.3521305>.
- Songtao Mao. The polynomial-time low-degree conjecture is false. arXiv: 2607.20318, 2026. URL <https://arxiv.org/abs/2607.20318>.
- Andrea Montanari, Daniel Reichman, and Ofer Zeitouni. On the limitation of spectral methods: From the Gaussian hidden clique problem to rank-one perturbations of Gaussian tensors. *IEEE Transactions on Information Theory*, 63(3):1572–1579, 2017. URL <https://doi.org/10.1109/TIT.2016.2637959>.

- John W. Moon. *Counting Labelled Trees*. Canadian Mathematical Congress, 1970. URL <https://books.google.com/books?id=B6dDcgAACAAJ>.
- Matvey Mosievskiy and Lev Reyzin. Multiple planted structures below $\sqrt{n}$: An SoS integrality gap and an SQ lower bound. arXiv: 2604.07278, 2026. URL <https://arxiv.org/abs/2604.07278>.
- Elchanan Mossel, Jonathan Niles-Weed, Youngtak Sohn, Nike Sun, and Ilias Zadik. Sharp thresholds in inference of planted subgraphs. *Annals of Applied Probability*, 35(1):523–563, 2025. URL <https://doi.org/10.1214/24-AAP2120>.
- Mor Oren-Loberman, Dvir Jerbi, Tamir Bendory, and Wasim Huleihel. Inhomogeneous submatrix detection. arXiv: 2603.09602, 2026. URL <https://arxiv.org/abs/2603.09602>.
- Amelia Perry, Alexander S. Wein, Afonso S. Bandeira, and Ankur Moitra. Optimality and sub-optimality of PCA i: Spiked random matrix models. *Annals of Statistics*, 46(5):2416–2451, 2018. URL <https://arxiv.org/abs/1612.07728>.
- Amelia Perry, Alexander S. Wein, and Afonso S. Bandeira. Statistical limits of spiked tensor models. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 56(1):230–264, 2020. URL <https://doi.org/10.1214/19-AIHP960>.
- Asaf Rotenberg, Wasim Huleihel, and Ofer Shayevitz. Planted bipartite graph detection. *IEEE Transactions on Information Theory*, 70(6):4319–4334, 2024. URL <https://doi.org/10.1109/TIT.2024.3382228>.
- Tselil Schramm and Alexander S. Wein. Computational barriers to estimation from low-degree polynomials. *Annals of Statistics*, 50(3):1833–1858, 2022. URL <https://arxiv.org/abs/2008.02269>.
- James Sharpnack, Akshay Krishnamurthy, and Aarti Singh. Near-optimal anomaly detection in graphs using Lovász extended scan statistic. In *Advances in Neural Information Processing Systems*, 2013. URL <https://proceedings.neurips.cc/paper/2013/hash/860320be12a1c050cd7731794e231bd3-Abstract.html>.
- James Sharpnack, Alessandro Rinaldo, and Aarti Singh. Detecting anomalous activity on networks with the graph Fourier scan statistic. *IEEE Transactions on Signal Processing*, 64(2):364–379, 2016. URL <https://doi.org/10.1109/TSP.2015.2481866>.
- Anda Skeja, Daniel Gutiérrez Espinoza, Fiona Skerman, and Alexander S. Wein. Sharp low-degree thresholds for planted-vs-planted testing. arXiv: 2606.05266, 2026. URL <https://arxiv.org/abs/2606.05266>.
- Youngtak Sohn and Alexander S. Wein. Sharp phase transitions in estimation with low-degree polynomials. In *Proceedings of the ACM Symposium on Theory of Computing*, 2025. URL <https://doi.org/10.1145/3717823.3718294>.
- Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer, 2009. URL <https://doi.org/10.1007/b13794>.
- Aad W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998. URL <https://doi.org/10.1017/CB09780511802256>.
- Nicolas Verzelen and Ery Arias-Castro. Community detection in sparse random networks. *Annals of Applied Probability*, 25(6):3465–3510, 2015. URL <https://doi.org/10.1214/14-AAP1080>.
- Alexander S. Wein. Computational complexity of statistics: New insights from low-degree polynomials. arXiv: 2506.10748, 2025. URL <https://arxiv.org/abs/2506.10748>.
- Xifan Yu, Ilias Zadik, and Peiyuan Zhang. Counting stars is constant-degree optimal for detecting any planted subgraph: Extended abstract. In *Proceedings of the Conference on Learning Theory*, 2024. URL <https://proceedings.mlr.press/v247/yu24a.html>.

Appendix

The appendix contains supporting technical results, complete proofs, and additional examples of latent structures. Appendix A collects the technical lemmas used throughout the paper, and Appendix B contains the proofs. Appendix C studies further examples that complement the problems considered in the main text. A detailed outline is provided below.

Outline

A	Technical lemmas	37
B	Proofs	39
B.1	Proof of Lemma 1	39
B.2	Proof of Proposition 2	39
B.3	Proof of Lemma 3	41
B.4	Proof of Lemma 4	42
B.5	Proof of Theorem 1	43
B.6	Proof of Corollary 1	43
B.7	Proof of Theorem 2	47
B.8	Proof of Corollary 2	57
B.9	Proof of Theorem 3	68
B.10	Proof of Corollary 3	76
B.11	Proof of Proposition 3	77
B.12	Proof of Theorem 4	83
B.13	Proof of Proposition 4	86
B.14	Proof of Theorem 5	88
B.15	Alternative proofs of computational lower bounds for multiple cliques and hypercliques	90
C	Additional examples of latent structures	98
C.1	Examples from Addario-Berry et al. (2010)	98
C.2	Further examples	102

A Technical lemmas

This section collects the moment and decoupling inequalities used in the proofs.

The following lemma provides a moment bound for sub-Poissonian random variables. Its assumptions hold for Poisson and binomial random variables, as well as for random variables dominated by them in the convex order. A recurring example in this paper is a hypergeometric random variable, which can be represented as a sum of negatively associated Bernoulli random variables.

Lemma 5 ([Ahle 2022](#), Corollary 1). *Let X be a nonnegative random variable with mean $\lambda > 0$ whose moment generating function satisfies $\mathbb{E}[\exp(tX)] \leq \exp\{\lambda(e^t - 1)\}$ for every $t > 0$. Then, for every $d > 0$,*

$$\mathbb{E}[(X/\lambda)^d] \leq \left(\frac{d/\lambda}{\log(1 + d/\lambda)} \right)^d \leq \left(1 + \frac{d}{2\lambda} \right)^d \leq \exp\left(\frac{d^2}{2\lambda} \right).$$

Finner’s inequality generalizes Hölder’s inequality by accounting for the coordinates on which each function depends in a product measure space.

Lemma 6 (Finner 1992, Theorem 2.1). Let I and J be finite index sets, and let $\emptyset \neq T_j \subseteq I$ and $p_j \geq 1$ for each $j \in J$. For each $i \in I$, define $J_i := \{j \in J : i \in T_j\}$, and assume that $\sum_{j \in J_i} p_j^{-1} = 1$. Let $(\Omega_i, \mathcal{A}_i, \mu_i)$, $i \in I$, be measure spaces, and for each $T \subseteq I$, write

$$(\Omega_T, \mathcal{A}_T, \mu_T) := \left(\prod_{i \in T} \Omega_i, \bigotimes_{i \in T} \mathcal{A}_i, \bigotimes_{i \in T} \mu_i \right).$$

Suppose that $f_j \in L_{p_j}(\Omega_{T_j}, \mathcal{A}_{T_j}, \mu_{T_j})$ for each $j \in J$, and view f_j as a function on Ω_I depending only on the coordinates in T_j . Then $\prod_{j \in J} |f_j| \in L_1(\Omega_I, \mathcal{A}_I, \mu_I)$ and

$$\int \prod_{j \in J} |f_j| d\mu_I \leq \prod_{j \in J} \left(\int |f_j|^{p_j} d\mu_{T_j} \right)^{1/p_j}.$$

The next inequality bounds moments of a U-statistic by those of its decoupled counterpart; see de la Peña and Giné (1999); Mao et al. (2025).

Lemma 7. Let $X_1, \dots, X_m$ be i.i.d. random variables taking values in a measurable space $(\mathcal{X}, \mathcal{A})$, and let $(X'_1, \dots, X'_m)$ be an independent copy of this sequence. For each ordered pair (i, j) with $i \neq j$, let $f_{ij} : \mathcal{X}^2 \rightarrow \mathbb{R}$ be a measurable kernel satisfying $\mathbb{E}[|f_{ij}(X_i, X_j)|] < \infty$. There exists an absolute constant $C > 0$ such that, for every convex nondecreasing function $\Phi : [0, \infty) \rightarrow [0, \infty)$ with $\mathbb{E}[\Phi(|f_{ij}(X_i, X_j)|)] < \infty$ for all $i \neq j$,

$$\mathbb{E}\Phi\left(\left|\sum_{\substack{i, j \in [m] \\ i \neq j}} f_{ij}(X_i, X_j)\right|\right) \leq \mathbb{E}\Phi\left(C \left|\sum_{\substack{i, j \in [m] \\ i \neq j}} f_{ij}(X_i, X'_j)\right|\right).$$

The same bound applies to sums over $i < j$ by setting $f_{ij} = 0$ for $i > j$.

The following version applies to U-statistics of any fixed order.

Lemma 8 (de la Peña and Giné 1999, Theorem 3.1.1). Let $m \geq t \geq 1$ be integers, and let $X_1, \dots, X_m$ be independent random variables taking values in a measurable space $(\mathcal{X}, \mathcal{A})$. For $j \in [t]$, let $(X_i^{(j)})_{i=1}^m$ be copies of this sequence that are mutually independent and independent of the original sequence. For each $\mathbf{i} = (i_1, \dots, i_t) \in [m]^t$, let $H_{\mathbf{i}} : \mathcal{X}^t \rightarrow \mathbb{R}$ be a measurable kernel satisfying $\mathbb{E}[|H_{\mathbf{i}}(X_{i_1}, \dots, X_{i_t})|] < \infty$, and assume that $H_{\mathbf{i}} \equiv 0$ whenever $\mathbf{i}$ has repeated indices. Then, for every convex nondecreasing function $\Phi : [0, \infty) \rightarrow [0, \infty)$ with $\mathbb{E}[\Phi(|H_{\mathbf{i}}(X_{i_1}, \dots, X_{i_t})|)] < \infty$ for all $\mathbf{i} \in [m]^t$,

$$\mathbb{E}\left[\Phi\left(\left|\sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(X_{i_1}, \dots, X_{i_t})\right|\right)\right] \leq \mathbb{E}\left[\Phi\left(C_t \left|\sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(X_{i_1}^{(1)}, \dots, X_{i_t}^{(t)})\right|\right)\right],$$

where $C_t = 2^t \prod_{r=2}^t (r^r - 1)$ depends only on t .

Finally, we record a moment bound for decoupled U-statistics with nonnegative kernels.

Lemma 9 (Giné et al. 2000, Proposition 2.1). Let $m, t \in \mathbb{N}$, $d > 1$, and let $\{X_i^{(j)} : i \in [m], j \in [t]\}$ be independent random variables. For each $\mathbf{i} = (i_1, \dots, i_t) \in [m]^t$, let $H_{\mathbf{i}}$ be a nonnegative measurable function satisfying $\mathbb{E}[H_{\mathbf{i}}(\mathbf{X}_{\mathbf{i}})^d] < \infty$, where $\mathbf{X}_{\mathbf{i}} := (X_{i_1}^{(1)}, \dots, X_{i_t}^{(t)})$. For $J \subseteq [t]$, write $\mathbf{i}_J := (i_j)_{j \in J}$, let $J^c := [t] \setminus J$, and define $\mathcal{F}_J := \sigma(X_i^{(j)} : j \in J, i \in [m])$. Let $\mathbb{E}_J$ denote expectation over all variables $X_i^{(j)}$ with $j \in J$, holding the remaining variables fixed. There exists a universal constant $C > 0$ such that

$$\mathbb{E}\left[\left(\sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(\mathbf{X}_{\mathbf{i}})\right)^d\right] \leq C^{td} \sum_{J \subseteq [t]} \left[d^{|J|d} \sum_{\mathbf{i}_J \in [m]^{|J|}} \mathbb{E}_J\left\{\left(\sum_{\mathbf{i}_{J^c} \in [m]^{t-|J|}} \mathbb{E}_{J^c}[H_{\mathbf{i}}(\mathbf{X}_{\mathbf{i}}) | \mathcal{F}_J]\right)^d\right\}\right],$$

By convention, $\sum_{\mathbf{i}_0} a = a$ and $\mathbb{E}_{\emptyset}[a] = a$.

B Proofs

B.1 Proof of Lemma 1

The variational characterization of total variation and the Bretagnolle–Huber inequality (Tsybakov, 2009, Section 2.4.1) give, for every test Δ_n ,

$$\begin{aligned} R_{\nu_n}(\Delta_n) &\geq 1 - \text{TV}(\mathbb{P}_{\nu_n}, \mathbb{Q}_n) \\ &\geq 1 - \sqrt{1 - \exp\{-D_{\text{KL}}(\mathbb{P}_{\nu_n} \parallel \mathbb{Q}_n)\}}. \end{aligned} \quad (16)$$

Jensen’s inequality under $\mathbb{P}_{\nu_n}$ yields

$$\begin{aligned} D_{\text{KL}}(\mathbb{P}_{\nu_n} \parallel \mathbb{Q}_n) &= \mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_{\nu_n}} [\log L_{\nu_n}(\mathbf{Y})] \\ &\leq \log \mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_{\nu_n}} [L_{\nu_n}(\mathbf{Y})] \\ &= \log \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}(\mathbf{Y})^2] \leq \log C. \end{aligned}$$

Substitution into Equation (16) proves the claim. An alternative proof based on classical Le Cam contiguity theory (Le Cam, 1960) is given, for example, by Montanari et al. (2017, Lemma 2).

B.2 Proof of Proposition 2

We prove the three bounds in Proposition 2 and the degree-test bound in Remark 3.

Outlier test. Recall that $t_{\text{out}} := \sqrt{2 \log(2n/\delta)}$. Under $\mathbb{Q}$, each coordinate is standard normal, so the union bound and the Gaussian tail bound give

$$\mathbb{Q}\{\Delta_{\text{out}}(\mathbf{Y}) = 1\} \leq n \exp(-t_{\text{out}}^2/2) = \frac{\delta}{2}.$$

For the type II error, fix any s in the support of ν , so that $|s| = K \geq 1$. Under $\mathbb{P}_s$, the variables $\{Y_\ell : \ell \in s\}$ are independent and distributed as $\text{N}(\mu, 1)$. Writing Φ for the standard normal distribution function, we have

$$\begin{aligned} \mathbb{P}_s\{\Delta_{\text{out}}(\mathbf{Y}) = 0\} &= \mathbb{P}_s\left\{\max_{\ell \in [n]} Y_\ell \leq t_{\text{out}}\right\} \\ &\leq \mathbb{P}_s\left\{\max_{\ell \in s} Y_\ell \leq t_{\text{out}}\right\} \\ &= \prod_{\ell \in s} \mathbb{P}_s\{Y_\ell \leq t_{\text{out}}\} \\ &= \Phi(t_{\text{out}} - \mu)^{|s|}. \end{aligned}$$

For $\mu \geq t_{\text{out}}$, the Gaussian tail bound therefore gives

$$\begin{aligned} \mathbb{P}_s\{\Delta_{\text{out}}(\mathbf{Y}) = 0\} &\leq \exp\left(-\frac{|s|(\mu - t_{\text{out}})^2}{2}\right) \\ &\leq \exp\left(-\frac{(\mu - t_{\text{out}})^2}{2}\right). \end{aligned}$$

The last inequality uses $|s| \geq 1$. This bound is at most $\delta/2$ whenever $\mu \geq t_{\text{out}} + \sqrt{2 \log(2/\delta)}$. Since it holds uniformly over s in the support of ν , averaging the type II error and adding the type I error yields $R_\nu(\Delta_{\text{out}}) \leq \delta$.

Average test. Under $\mathbb{Q}$, the average statistic $T_{\text{avg}}(\mathbf{Y})$ has distribution $\text{N}(0, n)$, while under $\mathbb{P}_s$ it has distribution $\text{N}(\mu K, n)$ for every s in the support of ν . At the threshold $t_{\text{avg}} := \mu K/2$, the Gaussian tail bound therefore gives

$$R_\nu(\Delta_{\text{avg}}) \leq 2 \exp\left(-\frac{\mu^2 K^2}{8n}\right).$$

The right-hand side is at most δ whenever $\mu \geq \sqrt{(8n/K^2) \log(2/\delta)}$, as claimed.

Scan test. For $t_{\text{scan}} := (\mu K + \sqrt{2K \log |\mathcal{C}|})/2$, the signal condition in Proposition 2 implies

$$\begin{aligned} t_{\text{scan}} &\geq \sqrt{2K \log |\mathcal{C}|} + \sqrt{2K \log(2/\delta)} \geq \sqrt{2K \log(2|\mathcal{C}|/\delta)}, \\ \mu K - t_{\text{scan}} &\geq \sqrt{2K \log(2/\delta)}. \end{aligned}$$

Under $\mathbb{Q}$, the sum over any fixed candidate $s \in \mathcal{C}$ has distribution $\mathcal{N}(0, K)$. The union bound therefore yields

$$\mathbb{Q}\{\Delta_{\text{scan}}(\mathbf{Y}) = 1\} \leq |\mathcal{C}| \exp\left(-\frac{t_{\text{scan}}^2}{2K}\right) \leq \frac{\delta}{2}.$$

For the type II error, fix any s in the support of ν . Under $\mathbb{P}_s$, the sum $\sum_{\ell \in s} Y_\ell$ has distribution $\mathcal{N}(\mu K, K)$ and is bounded above by $T_{\text{scan}}(\mathbf{Y})$. Hence,

$$\begin{aligned} \mathbb{P}_s\{\Delta_{\text{scan}}(\mathbf{Y}) = 0\} &\leq \mathbb{P}_s\left\{\sum_{\ell \in s} Y_\ell \leq t_{\text{scan}}\right\} \\ &\leq \exp\left(-\frac{(\mu K - t_{\text{scan}})^2}{2K}\right) \leq \frac{\delta}{2}. \end{aligned}$$

Averaging over s and adding the type I error gives $R_\nu(\Delta_{\text{scan}}) \leq \delta$.

Degree test. We now prove the bound in Remark 3, using the threshold $t_{\text{deg}} := \mu \mathfrak{d}_{\max}/2$. Under $\mathbb{Q}$, each weighted degree satisfies

$$T_i^{\text{deg}}(\mathbf{Y}) = \sum_{\substack{e \in [m]_2 \\ i \in e}} Y_e \sim \mathcal{N}(0, m-1),$$

because there are $m-1$ edges incident to each vertex. The union bound and the Gaussian tail bound give

$$\begin{aligned} \mathbb{Q}\{\Delta_{\text{deg}}(\mathbf{Y}) = 1\} &= \mathbb{Q}\{T_{\text{deg}}(\mathbf{Y}) > t_{\text{deg}}\} \\ &\leq \sum_{i=1}^m \mathbb{Q}\left\{T_i^{\text{deg}}(\mathbf{Y}) > \frac{\mathfrak{d}_{\max}\mu}{2}\right\} \\ &\leq m \exp\left(-\frac{\mathfrak{d}_{\max}^2\mu^2}{8(m-1)}\right). \end{aligned}$$

For the type II error, fix any s in the support of ν and choose a vertex $i_s \in [m]$ with $\mathfrak{d}_s(i_s) = \mathfrak{d}_{\max}$. Under $\mathbb{P}_s$, we have $T_{i_s}^{\text{deg}}(\mathbf{Y}) \sim \mathcal{N}(\mu \mathfrak{d}_{\max}, m-1)$. Since $T_{\text{deg}}(\mathbf{Y}) \geq T_{i_s}^{\text{deg}}(\mathbf{Y})$, we have

$$\begin{aligned} \mathbb{P}_s\{\Delta_{\text{deg}}(\mathbf{Y}) = 0\} &= \mathbb{P}_s\{T_{\text{deg}}(\mathbf{Y}) \leq t_{\text{deg}}\} \\ &\leq \mathbb{P}_s\left\{T_{i_s}^{\text{deg}}(\mathbf{Y}) \leq \frac{\mathfrak{d}_{\max}\mu}{2}\right\} \\ &\leq \exp\left(-\frac{\mathfrak{d}_{\max}^2\mu^2}{8(m-1)}\right), \end{aligned}$$

where the last inequality follows from the Gaussian tail bound. The bound holds uniformly over s in the support of ν , so averaging the type II error and adding the type I error gives

$$\begin{aligned} R_\nu(\Delta_{\text{deg}}) &\leq (m+1) \exp\left(-\frac{\mathfrak{d}_{\max}^2\mu^2}{8(m-1)}\right) \\ &\leq 2m \exp\left(-\frac{\mathfrak{d}_{\max}^2\mu^2}{8(m-1)}\right). \end{aligned}$$

Consequently, $R_\nu(\Delta_{\text{deg}}) \leq \delta$ whenever

$$\mu \geq \sqrt{\frac{8(m-1)}{\mathfrak{d}_{\max}^2} \log \frac{2m}{\delta}}.$$

B.3 Proof of Lemma 3

Let $\dot{V} := \{i \in [m] : \dot{A}_i \leq k\}$ be the vertex set of S , and define $\dot{V}'$ using an independent copy $\dot{\mathbf{A}}'$ of $\dot{\mathbf{A}}$. Then $H := |\dot{V} \cap \dot{V}'|$ has the hypergeometric distribution with parameters (m, k, k) , and $|S \cap S'| = \binom{H}{2}$. The function $x \mapsto \exp\{\mu^2 x(x-1)/2\}$ is convex, so the comparison between sampling with and without replacement in [Hoeffding \(1963\)](#) gives

$$\begin{aligned} \mathbb{E}_{S, S' \sim \nu} [\exp\{\mu^2 |S \cap S'| \}] &= \mathbb{E}_H \left[\exp \left\{ \mu^2 \binom{H}{2} \right\} \right] \\ &\leq \mathbb{E}_B \left[\exp \left\{ \mu^2 \binom{B}{2} \right\} \right], \end{aligned}$$

where $B \sim \text{Binomial}(k, k/m)$. The case $k = 1$ is trivial because the overlap is empty, so assume $k \geq 2$. Take $B_1, \dots, B_k \stackrel{\text{iid}}{\sim} \text{Bernoulli}(k/m)$ and set $B := \sum_{i=1}^k B_i$. Since $\binom{B}{2} = \sum_{i < j} B_i B_j$, we have

$$\mathbb{E}_B \left[\exp \left\{ \mu^2 \binom{B}{2} \right\} \right] = \mathbb{E} \left[\exp \left(\mu^2 \sum_{i < j} B_i B_j \right) \right]. \quad (17)$$

We apply Finner's inequality (Lemma 6) with $I = [k]$, $J = [k]_2$, and $T_{(i,j)} = \{i, j\}$. Each coordinate B_i appears in $k-1$ factors, so choosing $p_{ij} = k-1$ makes the sum of the reciprocal exponents equal to one at every coordinate. Since $B_i B_j \sim \text{Bernoulli}(k^2/m^2)$ for $i < j$, the inequality yields

$$\begin{aligned} \mathbb{E} \left[\exp \left(\mu^2 \sum_{i < j} B_i B_j \right) \right] &\leq \prod_{i < j} \mathbb{E} [\exp\{\mu^2 (k-1) B_i B_j\}]^{1/(k-1)} \\ &= \mathbb{E} [\exp\{\mu^2 (k-1) B_1 B_2\}]^{k/2} \\ &= \left(1 + \frac{k^2}{m^2} (\exp\{\mu^2 (k-1)\} - 1) \right)^{k/2} \\ &\leq \exp \left(\frac{k^3}{2m^2} (\exp(\mu^2 k) - 1) \right), \end{aligned} \quad (18)$$

where the last inequality uses $\log(1+x) \leq x$ for $x \geq 0$ and $k-1 \leq k$. Thus, the likelihood-ratio second moment is bounded whenever $\frac{k^3}{m^2} (\exp(\mu^2 k) - 1) = O(1)$. More explicitly, for any fixed constant $C > 0$, the condition

$$\mu \leq \sqrt{\frac{1}{k} \log \left(1 + C \frac{m^2}{k^3} \right)}$$

ensures that the second moment is at most $e^{C/2}$. We verify this condition in the two regimes of Lemma 3.

First, if $k^3 m^{-2} \leq m^{-\delta}$ and $\mu \leq \sqrt{\delta k^{-1} \log m}$, then

$$\frac{k^3}{2m^2} (\exp(\mu^2 k) - 1) \leq \frac{m^{-\delta}}{2} (m^\delta - 1) \leq \frac{1}{2}.$$

Consequently, the second moment is at most $e^{1/2}$.

Second, suppose that $k^3 m^{-2} = \Omega(1)$ and $\mu = O(mk^{-2})$. Choose constants $c_0, C_0 > 0$ such that $k^3/m^2 \geq c_0$ and $\mu \leq C_0 m/k^2$ for all sufficiently large m , so that $\mu^2 k \leq C_0^2/c_0$. Using $e^x - 1 \leq xe^x$ for $x \geq 0$, we obtain

$$\frac{k^3}{2m^2} (\exp(\mu^2 k) - 1) \leq \frac{\mu^2 k^4}{2m^2} \exp(\mu^2 k) \leq \frac{C_0^2}{2} \exp(C_0^2/c_0),$$

which is bounded independently of m and proves the second assertion.

B.4 Proof of Lemma 4

Fix constants $\delta \in (0, 0.1)$ and $\epsilon \in (0, \delta)$ as in the lemma, and set $D := \lfloor \log^{1+\epsilon} m \rfloor$. By Proposition 1, the low-degree likelihood-ratio second moment is

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \tilde{\nu}} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \right].$$

As in Appendix B.3, let $H := |\dot{V} \cap \dot{V}'|$, so that H has the hypergeometric distribution with parameters (m, k, k) and $|S \cap S'| = \binom{H}{2} \leq H^2/2$. For each integer $d \geq 1$, the function $x \mapsto x^{2d}$ is convex on $[0, \infty)$, so the comparison with $B \sim \text{Binomial}(k, k/m)$ gives $\mathbb{E}_H[H^{2d}] \leq \mathbb{E}_B[B^{2d}]$. Combining these bounds with Lemma 5, due to Ahle (2022, Corollary 1), yields

$$\begin{aligned} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2] &\leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{2^d d!} \left(\frac{2d}{\log(1 + 2dm/k^2)} \right)^{2d} \\ &= 1 + \sum_{d=1}^D \frac{2^d \mu^{2d}}{d!} \left(\frac{d}{\log(1 + 2dm/k^2)} \right)^{2d}. \end{aligned} \tag{19}$$

It remains to bound the series in the two regimes of the lemma.

- **Case 1.** $k^2/m \leq m^{-\delta}$.

Since $\log(1 + 2dm/k^2) \geq \log(m/k^2)$ for $d \geq 1$, the series in Equation (19) is at most $\sum_{d=1}^D T_d$, where

$$T_d := \frac{2^d \mu^{2d}}{d!} \left(\frac{d}{\log(m/k^2)} \right)^{2d}.$$

For $1 \leq d < D$, we have

$$\begin{aligned} \frac{T_{d+1}}{T_d} &= \frac{2\mu^2(d+1)}{\log^2(m/k^2)} \left(1 + \frac{1}{d} \right)^{2d} \\ &\leq \frac{2e^2 \mu^2 D}{\log^2(m/k^2)}, \end{aligned}$$

where we used $(1 + 1/d)^d \leq e$. Thus, if $\mu \leq \log(m/k^2)/(2e\sqrt{D})$, then $T_{d+1} \leq T_d/2$ and

$$\sum_{d=1}^D T_d \leq 2T_1 = \frac{4\mu^2}{\log^2(m/k^2)} \leq \frac{1}{e^2 D}.$$

The assumption $k^2/m \leq m^{-\delta}$ gives $\log(m/k^2) \geq \delta \log m$. Since $D \leq \log^{1+\epsilon} m$, the required condition holds whenever $\mu \leq C(\log m)^{(1-\epsilon)/2}$ with $0 < C \leq \delta/(2e)$. This proves the first assertion.

- **Case 2.** $k^2/m \geq \log^{1+\delta} m$.

Uniformly over $d \in [D]$, we have $0 < dm/k^2 \leq \log^{\epsilon-\delta} m \rightarrow 0$, since $\epsilon < \delta$. Using $\log(1 + 2x) \geq x$ for $0 \leq x \leq 1$, we obtain, for all sufficiently large m ,

$$\log \left(1 + \frac{2dm}{k^2} \right) \geq \frac{dm}{k^2}, \quad d \in [D].$$

Consequently, the series in Equation (19) satisfies

$$\begin{aligned} \sum_{d=1}^D \frac{2^d \mu^{2d}}{d!} \left(\frac{d}{\log(1 + 2dm/k^2)} \right)^{2d} &\leq \sum_{d=1}^D \frac{1}{d!} \left(\frac{2\mu^2 k^4}{m^2} \right)^d \\ &\leq \exp \left(\frac{2\mu^2 k^4}{m^2} \right) - 1. \end{aligned}$$

This bound is $O(1)$ whenever $\mu = O(mk^{-2})$, proving the second assertion.

By Lemma 2, the bounded second moment rules out strong separation by polynomials of degree at most D . Its interpretation as a lower bound against all polynomial-time tests is conditional on Conjecture 1.

B.5 Proof of Theorem 1

For the fixed-size planted clique problem, $n = \binom{m}{2}$, $K = \binom{k}{2}$, and $|\mathcal{C}| = \binom{m}{k}$. We combine the upper bounds in Proposition 2 with Lemmas 3 and 4 to identify the thresholds at polynomial order.

Upper bounds. Applying Proposition 2 with prior ν gives the following guarantees.

- Taking the target risk to be $\delta_m = 1/\log m$, the outlier test satisfies $R_\nu(\Delta_{\text{out}}) \rightarrow 0$ whenever

$$\mu \geq C\sqrt{\log m}$$

for a sufficiently large constant $C > 0$.

- The average-test risk bound proved in Appendix B.2 gives

$$R_\nu(\Delta_{\text{avg}}) \leq 2 \exp\left(-\frac{\mu^2 K^2}{8n}\right).$$

Since $\sqrt{n}/K \asymp m/k^2$, the risk tends to zero whenever $\mu = \omega(mk^{-2})$.

- For any fixed $c \in (0, 1)$, the scan test satisfies $R_\nu(\Delta_{\text{scan}}) \rightarrow 0$ whenever $k \leq (1 - c)m$ and

$$\mu \geq C\sqrt{\frac{1}{k} \log \frac{m}{k}}$$

for a sufficiently large constant $C > 0$ depending only on c . This follows by taking $\delta_m = 1/\log m$ and using $\log \binom{m}{k} \leq k \log(em/k)$, together with the assumption that k grows at least polynomially in m .

The outlier and average tests run in $O(m^2)$ time. In the regime $k \ll m^{2/3}$, the condition $k \leq (1 - c)m$ holds eventually, and a direct scan over all $\binom{m}{k}$ candidates requires superpolynomial time because k grows at least polynomially in m .

Statistical threshold. If $k \ll m^{2/3}$, then $k^3/m^2 \leq m^{-\delta}$ eventually for some constant $\delta \in (0, 0.1)$. Thus, Lemma 3(i) and Lemma 1 rule out strong detection whenever $\mu \ll k^{-1/2}$, while the scan test achieves strong detection whenever $\mu \gg k^{-1/2}$. If $k \gg m^{2/3}$, then Lemma 3(ii) rules out strong detection whenever $\mu \ll mk^{-2}$, while the average test achieves strong detection whenever $\mu \gg mk^{-2}$. These bounds give the two branches of μ_{stat}^* in Theorem 1.

Computational threshold. If $k \ll m^{1/2}$, then $k^2/m \leq m^{-\delta}$ eventually for some constant $\delta \in (0, 0.1)$. Choose any $\epsilon \in (0, \delta)$. By Lemma 4(i), the low-degree likelihood-ratio second moment is bounded for degrees $D_m = \log^{1+\epsilon} m$ whenever $\mu \ll 1$, while the outlier test achieves strong detection whenever $\mu \gg 1$. If $k \gg m^{1/2}$, then $k^2/m \geq \log^{1+\delta} m$ eventually for any fixed $\delta \in (0, 0.1)$, so Lemma 4(ii) bounds the low-degree second moment whenever $\mu \ll mk^{-2}$. The average test achieves strong detection whenever $\mu \gg mk^{-2}$. Under Conjecture 1, these low-degree bounds give the corresponding lower bounds against all polynomial-time tests. Since logarithmic factors are subpolynomial in m , Definition 3 yields the stated two branches of μ_{comp}^* and completes the proof.

B.6 Proof of Corollary 1

In the random-size model, $\tilde{\mathcal{C}}$ is the collection of clique edge sets of arbitrary size in the complete graph with m vertices. For the latent vector $\tilde{\mathbf{A}}$, define the vertex set $\tilde{V} := \{l \in [m] : \mathbf{1}(\tilde{A}_l \leq k) = 1\}$ of the planted clique. We define $\tilde{V}'$ similarly for a random vector $\tilde{\mathbf{A}}'$. With a slight abuse of notation, we use $\tilde{\nu}$ for both the distribution of $\tilde{V}$ (and also of $\tilde{V}'$) and the induced distribution of S on $\tilde{\mathcal{C}}$.

Conditioning and total variation. We first compare the total variation distance between the alternative distribution and the alternative induced by a conditional mixing measure. Given a mixing measure $\tilde{\nu}$, consider an event $\mathcal{E}$ and define a new mixing measure $\tilde{\nu}_{\mathcal{E}}$ as the conditional distribution of $\tilde{\nu}$ given $\mathcal{E}$, i.e., for any $s \in \tilde{\mathcal{C}}$, $\tilde{\nu}_{\mathcal{E}}(s) = \tilde{\nu}(s | \mathcal{E})$. Here and below, conditioning on $\mathcal{E}$ refers to the joint construction of the latent structure and observations, and $\tilde{\nu}(\mathcal{E})$ denotes the probability of the latent event. In the density calculation below, we use the same symbols for probability laws and their Lebesgue densities. The corresponding alternative distribution $\mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}$ satisfies

$$\mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}(\cdot) = \mathbb{E}_{S \sim \tilde{\nu}_{\mathcal{E}}}[\mathbb{P}_S(\cdot)] = \mathbb{E}_{S \sim \tilde{\nu}}[\mathbb{P}_S(\cdot) | \mathcal{E}] := \mathbb{P}_{\tilde{\nu}}(\cdot | \mathcal{E}).$$

We compare $\mathbb{P}_{\tilde{\nu}}$ and $\mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}$ in terms of total variation distance. Note that the distribution $\mathbb{P}_{\tilde{\nu}}$ can be decomposed as

$$\begin{aligned} \mathbb{P}_{\tilde{\nu}}(\cdot) &= \mathbb{P}_{\tilde{\nu}}(\cdot | \mathcal{E})\tilde{\nu}(\mathcal{E}) + \mathbb{P}_{\tilde{\nu}}(\cdot | \mathcal{E}^c)(1 - \tilde{\nu}(\mathcal{E})) \\ &= \mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}(\cdot)\tilde{\nu}(\mathcal{E}) + \mathbb{P}_{\tilde{\nu}}(\cdot | \mathcal{E}^c)\tilde{\nu}(\mathcal{E}^c). \end{aligned}$$

Therefore, we have

$$\mathbb{P}_{\tilde{\nu}}(\cdot) - \mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}(\cdot) = \tilde{\nu}(\mathcal{E}^c)(\mathbb{P}_{\tilde{\nu}}(\cdot | \mathcal{E}^c) - \mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}(\cdot)),$$

which implies that

$$\begin{aligned} \text{TV}(\mathbb{P}_{\tilde{\nu}}, \mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}) &= \frac{1}{2} \int |\mathbb{P}_{\tilde{\nu}}(\mathbf{y}) - \mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}(\mathbf{y})| d\mathbf{y} \\ &= \tilde{\nu}(\mathcal{E}^c) \cdot \frac{1}{2} \int |\mathbb{P}_{\tilde{\nu}}(\mathbf{y} | \mathcal{E}^c) - \mathbb{P}_{\tilde{\nu}_{\mathcal{E}}}(\mathbf{y})| d\mathbf{y} \\ &\leq \tilde{\nu}(\mathcal{E}^c). \end{aligned} \tag{20}$$

Upper bounds. To apply the tests introduced in Proposition 2, we need to address the randomness in the size of the clique. At a vanishing cost in total variation, we modify the mixing measure $\tilde{\nu}$ so that, under the alternative hypothesis, the number of clique vertices is at least $c \cdot k$ for some constant $c \in (0, 1)$. To be precise, consider the event defined by $\underline{\mathcal{E}} = \underline{\mathcal{E}}(c) := \{\tilde{V} \in 2^{[m]} : |\tilde{V}| \geq c \cdot k\}$, where $c \in (0, 1)$ is a constant. Define a new mixing measure $\tilde{\nu}$ as the conditional distribution of $\tilde{\nu}$ given $\underline{\mathcal{E}}$, i.e., for any $s \in \tilde{\mathcal{C}}$, $\tilde{\nu}(s) = \tilde{\nu}(s | \underline{\mathcal{E}})$. For the mixing measures $\tilde{\nu}$ and $\tilde{\nu}$, the corresponding risks $R_{\tilde{\nu}}(\Delta)$ and $R_{\tilde{\nu}}(\Delta)$ of a test Δ satisfy

$$\begin{aligned} R_{\tilde{\nu}}(\Delta) &= \mathbb{Q}\{\Delta(\mathbf{Y}) = 1\} + \mathbb{P}_{\tilde{\nu}}\{\Delta(\mathbf{Y}) = 0\} \\ &= \mathbb{Q}\{\Delta(\mathbf{Y}) = 1\} + \mathbb{P}_{\tilde{\nu}}\{\Delta(\mathbf{Y}) = 0\} + (\mathbb{P}_{\tilde{\nu}}\{\Delta(\mathbf{Y}) = 0\} - \mathbb{P}_{\tilde{\nu}}\{\Delta(\mathbf{Y}) = 0\}) \\ &\leq R_{\tilde{\nu}}(\Delta) + \text{TV}(\mathbb{P}_{\tilde{\nu}}, \mathbb{P}_{\tilde{\nu}}) \\ &\leq R_{\tilde{\nu}}(\Delta) + \tilde{\nu}(\underline{\mathcal{E}}^c), \end{aligned} \tag{21}$$

where the last step follows from Equation (20). Since $\mathbf{1}\{l \in \tilde{V}\}$ for $l \in [m]$ are i.i.d. Bernoulli(k/m) random variables and thus $|\tilde{V}| = \sum_{l=1}^m \mathbf{1}\{l \in \tilde{V}\}$ follows a binomial distribution with parameters m and k/m , by the Chernoff bound, we have

$$\tilde{\nu}(\underline{\mathcal{E}}^c) = \Pr\{|\tilde{V}| < c \cdot k\} \leq \exp\left(-\frac{(1-c)^2}{2}k\right),$$

which converges to zero as long as $k \rightarrow \infty$. Therefore, it suffices to show $R_{\tilde{\nu}}(\Delta) \rightarrow 0$ to conclude that $R_{\tilde{\nu}}(\Delta) \rightarrow 0$.

When ck and Ck specify vertex counts below, we round them down and up, respectively; the Chernoff bounds remain valid under these roundings. We apply the three tests in Proposition 2 with the modified mixing measure $\tilde{\nu}$. Note that every S in the support of $\tilde{\nu}$ contains a clique on $c \cdot k$ vertices. Thus, the theoretical guarantees of the tests we analyzed in Appendix B.5 still hold with k replaced by $c \cdot k$. More precisely, the average test uses threshold $\mu\binom{ck}{2}/2$, and the scan searches all $[ck]$ -vertex cliques. The same error bounds hold for larger planted cliques because the average statistic has mean at least $\mu\binom{ck}{2}$ and the scan includes a fully signaled candidate. Therefore, we have the following upper bounds on the critical signal strength μ^* :

- For some constant $C > 0$, the outlier test achieves $R_{\underline{\nu}}(\Delta_{\text{out}}) \rightarrow 0$ whenever $\mu \geq C\sqrt{\log m}$.
- The average test achieves $R_{\underline{\nu}}(\Delta_{\text{avg}}) \rightarrow 0$ as long as $\mu = \omega(\frac{m}{k^2})$.
- For some constant $C > 0$, the scan test achieves $R_{\underline{\nu}}(\Delta_{\text{scan}}) \rightarrow 0$ if $k \leq (1 - c_0)m$ for a fixed constant $c_0 \in (0, 1)$ and

$$\mu \geq C\sqrt{\frac{1}{k} \log \frac{m}{k}}.$$

The restriction on k for the scan bound holds in the regime $k \ll m^{2/3}$ where that bound is needed.

Statistical lower bound. We now derive a statistical lower bound on the critical signal strength μ^* . We first consider the case $k \leq \kappa m$ for some constant $\kappa \in (0, 1)$ and address the case $k > \kappa m$ separately at the end of the statistical lower-bound argument.

When $k \leq \kappa m$, we again condition the mixing measure, this time to obtain a lower bound on the risk. Consider the event $\bar{\mathcal{E}} = \bar{\mathcal{E}}(C) := \{\tilde{V} \in 2^{[m]} : |\tilde{V}| \leq C \cdot k\}$, where $C \in (1, 1/\kappa)$ is a constant. Define a new mixing measure $\tilde{\nu}$ as the conditional distribution of $\tilde{\nu}$ given $\bar{\mathcal{E}}$, i.e., for any $s \in \bar{\mathcal{C}}$, $\tilde{\nu}(s) = \tilde{\nu}(s | \bar{\mathcal{E}})$. By Chernoff's bound, we have

$$\tilde{\nu}(\bar{\mathcal{E}}^c) = \Pr\{|\tilde{V}| > C \cdot k\} \leq \exp\left(-\frac{(C-1)^2}{2C}k\right) \rightarrow 0$$

as $k \rightarrow \infty$.

For any test Δ , by Equation (20) and Equation (16), together with the argument that follows, we have

$$\begin{aligned} R_{\tilde{\nu}}(\Delta) &\geq 1 - \text{TV}(\mathbb{Q}, \mathbb{P}_{\tilde{\nu}}) \\ &\geq 1 - (\text{TV}(\mathbb{Q}, \mathbb{P}_{\tilde{\nu}}) + \text{TV}(\mathbb{P}_{\tilde{\nu}}, \mathbb{P}_{\tilde{\nu}})) \\ &\geq 1 - \sqrt{1 - \frac{1}{\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} \left[\left(\frac{d\mathbb{P}_{\tilde{\nu}}(\mathbf{Y})}{d\mathbb{Q}(\mathbf{Y})} \right)^2 \right]}} - \tilde{\nu}(\bar{\mathcal{E}}^c). \end{aligned} \tag{22}$$

Since $\tilde{\nu}(\bar{\mathcal{E}}^c) \rightarrow 0$ and $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} \left[\left(\frac{d\mathbb{P}_{\tilde{\nu}}(\mathbf{Y})}{d\mathbb{Q}(\mathbf{Y})} \right)^2 \right] > 1$, to keep $R_{\tilde{\nu}}(\Delta)$ bounded away from zero asymptotically, it suffices to establish that

$$\begin{aligned} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} \left[\left(\frac{d\mathbb{P}_{\tilde{\nu}}(\mathbf{Y})}{d\mathbb{Q}(\mathbf{Y})} \right)^2 \right] &= \mathbb{E}_{S, S' \sim \tilde{\nu}} [\exp\{\mu^2 |S \cap S'| \}] \\ &= \mathbb{E}_{\tilde{V}, \tilde{V}' \sim \tilde{\nu}} \left[\exp \left\{ \mu^2 \binom{|\tilde{V} \cap \tilde{V}'|}{2} \right\} \right] = O(1). \end{aligned}$$

We claim that, under $\tilde{\nu}$, $|\tilde{V} \cap \tilde{V}'|$ is stochastically dominated by a hypergeometric distribution with parameters (m, Ck, Ck) . The following coupling proves the claim: by construction, $\tilde{V}$ and $\tilde{V}'$ are independent samples from $\tilde{\nu}$. Add $Ck - |\tilde{V}|$ (resp. $Ck - |\tilde{V}'|$) elements to $\tilde{V}$ (resp. $\tilde{V}'$), drawn uniformly at random from $[m] \setminus \tilde{V}$ (resp. $[m] \setminus \tilde{V}'$), thus producing a new set $\dot{V}$ (resp. $\dot{V}'$) of cardinality exactly Ck . By the permutation invariance of $\tilde{\nu}$, $\dot{V}$ and $\dot{V}'$ are independent and uniformly distributed over the set $\{a \in 2^{[m]} : \sum_{l=1}^m \mathbf{1}\{l \in a\} = Ck\}$. Under this coupling, we have $|\tilde{V} \cap \tilde{V}'| \leq |\dot{V} \cap \dot{V}'|$ almost surely, which implies that

$$\Pr\{|\tilde{V} \cap \tilde{V}'| \geq t\} \leq \Pr\{|\dot{V} \cap \dot{V}'| \geq t\}, \quad \forall t \geq 0. \tag{23}$$

Therefore, we have

$$\begin{aligned} \mathbb{E}_{S, S' \sim \tilde{\nu}} [\exp\{\mu^2 |S \cap S'| \}] &= \mathbb{E}_{\tilde{V}, \tilde{V}'} \left[\exp \left\{ \mu^2 \binom{|\tilde{V} \cap \tilde{V}'|}{2} \right\} \right] \\ &\stackrel{(a)}{\leq} \mathbb{E}_{\dot{V}, \dot{V}'} \left[\exp \left\{ \mu^2 \binom{|\dot{V} \cap \dot{V}'|}{2} \right\} \right] \\ &\stackrel{(b)}{\leq} \mathbb{E}_B \left[\exp \left\{ \mu^2 \binom{B}{2} \right\} \right] \\ &\stackrel{(c)}{\leq} \exp \left(\frac{(Ck)^3}{2m^2} (\exp(\mu^2 Ck) - 1) \right), \end{aligned} \tag{24}$$

where $B \sim \text{Binomial}(Ck, Ck/m)$. Above, step (a) follows from the stochastic dominance relation in Equation (23) and step (b) from the fact that the hypergeometric distribution with parameters (m, Ck, Ck) is dominated in the convex ordering by the binomial distribution with parameters $(Ck, Ck/m)$ (Hoeffding, 1963), applied to the convex function $x \mapsto \exp\{\mu^2 x(x-1)/2\}$. Step (c) is Equation (18) with k replaced by Ck . We can therefore apply the same analysis in Appendix B.3, replacing k by Ck , to obtain the same statistical lower bound.

When $k > \kappa m$, we can directly bound the second moment without truncation. Since $|S \cap S'| \leq \binom{m}{2}$ almost surely, we have

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n}[L_{\tilde{\nu}}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \tilde{\nu}}[e^{\mu^2 |S \cap S'|}] \leq \exp\left(\mu^2 \binom{m}{2}\right).$$

Since $k > \kappa m$, the right-hand side is bounded whenever $\mu = O(m/k^2)$, yielding the same statistical lower bound as in the case $k \leq \kappa m$.

Computational lower bound. To derive a computational lower bound, we present two approaches: (i) working directly with the original mixing measure (without truncation), and (ii) working with a truncated mixing measure, as in the statistical lower bound analysis.

(i) **Original mixing measure.** The starting point is the following bound for any integer degree cutoff $D \geq 0$:

$$\begin{aligned} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2] &= \mathbb{E}_{S, S' \sim \tilde{\nu}}\left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d\right] \\ &= \mathbb{E}_{\tilde{V}, \tilde{V}' \sim \tilde{\nu}}\left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} \binom{|\tilde{V} \cap \tilde{V}'|}{2}^d\right] \\ &\leq \sum_{d=0}^D \frac{\mu^{2d}}{2^d d!} \mathbb{E}_{\tilde{V}, \tilde{V}' \sim \tilde{\nu}}[|\tilde{V} \cap \tilde{V}'|^{2d}] \\ &= 1 + \sum_{d=1}^D \frac{\mu^{2d}}{2^d d!} \mathbb{E}_{\tilde{V}, \tilde{V}' \sim \tilde{\nu}}[|\tilde{V} \cap \tilde{V}'|^{2d}]. \end{aligned} \tag{25}$$

Letting $D = D_m = \lfloor \log^{1+\epsilon} m \rfloor$ with $0 < \epsilon < \delta < 0.1$ as in Lemma 4, Conjecture 1 implies that a sufficient condition for computational indistinguishability is the boundedness of the last term above.

Observe that $\mathbb{1}\{i \in \tilde{V}\}, \mathbb{1}\{i \in \tilde{V}'\}$ are i.i.d. Bernoulli(k/m) random variables for $i \in [m]$, and thus the vertex overlap $|\tilde{V} \cap \tilde{V}'| = \sum_i \mathbb{1}\{i \in \tilde{V}\} \mathbb{1}\{i \in \tilde{V}'\}$ follows Binomial($m, (k/m)^2$), whose mean is k^2/m . Applying the sub-Poissonian moment bound (see Lemma 5) gives

$$\begin{aligned} 1 + \sum_{d=1}^D \frac{\mu^{2d}}{2^d d!} \mathbb{E}_{\tilde{V}, \tilde{V}' \sim \tilde{\nu}}[|\tilde{V} \cap \tilde{V}'|^{2d}] &\leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{2^d d!} \left(\frac{2d}{\log(1 + 2dm/k^2)}\right)^{2d} \\ &= 1 + \sum_{d=1}^D \frac{2^d \mu^{2d}}{d!} \left(\frac{d}{\log(1 + 2dm/k^2)}\right)^{2d}. \end{aligned}$$

Note that the last equation is identical to the bound for the fixed-size case (see Equation (19)). Repeating that analysis proves the bound for the original mixing measure.

(ii) **Truncated mixing measure $\tilde{\nu}$.** We now use the truncation approach from the statistical lower-bound analysis to obtain the same computational lower bound. In what follows, we assume that $k \leq \kappa m$ for some constant $\kappa \in (0, 1)$. The case $k > \kappa m$ can be handled as in the statistical lower bound analysis, using the fact that $|S \cap S'| \leq \binom{m}{2}$ almost surely and that the second moment of the low-degree likelihood ratio is bounded above by that of the full likelihood ratio.

We first compare the risks:

$$|R_{\tilde{\nu}}(\Delta) - R_{\tilde{\nu}}(\Delta)| = |\mathbb{P}_{\tilde{\nu}}\{\Delta(\mathbf{Y}) = 0\} - \mathbb{P}_{\tilde{\nu}}\{\Delta(\mathbf{Y}) = 0\}|$$

$$\begin{aligned} &\leq \text{TV}(\mathbb{P}_{\tilde{\nu}}, \mathbb{P}_{\tilde{\nu}^c}) \\ &\leq \tilde{\nu}(\tilde{\mathcal{E}}^c). \end{aligned}$$

This bound can be established in the same way as the bound in Equation (21). Using this bound, and as long as $\tilde{\nu}(\tilde{\mathcal{E}}^c) \rightarrow 0$, we immediately obtain that any polynomial-time test under the truncated mixing measure $\tilde{\nu}$ achieves strong detection if and only if it also achieves strong detection under the original mixing measure $\tilde{\nu}$. Thus, under the low-degree polynomial conjecture (Conjecture 1), it suffices to find a scaling of μ such that $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2]$ remains bounded for $D = D_m = \lfloor \log^{1+\epsilon} m \rfloor$ with ϵ as above, as this will imply that $R_{\tilde{\nu}}(\Delta)$ will not vanish for any polynomial-time test Δ and, by the equivalence above, the same will hold for $R_{\tilde{\nu}}(\Delta)$ as well.

Repeating the steps leading to Equation (25) with the truncated mixing measure $\tilde{\nu}$, we obtain

$$\begin{aligned} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2] &= \mathbb{E}_{S, S' \sim \tilde{\nu}} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \right] \\ &\leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{2^d d!} \mathbb{E}_{\tilde{V}, \tilde{V}' \sim \tilde{\nu}} [|\tilde{V} \cap \tilde{V}'|^{2d}], \end{aligned}$$

where $\tilde{V}$ and $\tilde{V}'$ are i.i.d. samples from the distribution on $2^{[m]}$ corresponding to the conditional distribution of m i.i.d. Bernoulli(k/m) given that their sum is no larger than $C \cdot k$, for some arbitrary constant $C \in (1, 1/\kappa)$. As in the coupling leading to Equation (24), enlarge $\tilde{V}$ and $\tilde{V}'$ independently to supersets $\dot{V}$ and $\dot{V}'$ of size $\lceil Ck \rceil$ by adding vertices uniformly without replacement from their respective complements in $[m]$. By permutation invariance, $\dot{V}$ and $\dot{V}'$ are independent and uniformly distributed over $\{v \subseteq [m] : |v| = \lceil Ck \rceil\}$. The coupling gives $|\tilde{V} \cap \tilde{V}'| \leq |\dot{V} \cap \dot{V}'|$ almost surely, so the former overlap is stochastically dominated by the latter, which has the hypergeometric distribution with parameters $(m, \lceil Ck \rceil, \lceil Ck \rceil)$. Now we repeat the same analysis in Appendix B.4 with k replaced by $\lceil Ck \rceil$. This substitution changes only the constants, not the asymptotic rate, and gives the same computational lower bound as above, completing the proof.

B.7 Proof of Theorem 2

In this section, we study the upper bounds provided by three tests for the multiple planted clique problem and prove Theorem 2 by establishing matching statistical and low-degree lower bounds. The computational interpretation of the latter uses Conjecture 1. A formal statement of the lower-bound results we prove is as follows:

- (Statistical lower bound) Fix any constant $\delta \in (0, 0.1)$. For either prior distribution $\dot{\nu}$ or $\tilde{\nu}$, strong detection is impossible under either of the following conditions: (i) $\gamma^2 k^3 m^{-2} \leq m^{-\delta}$ and $\mu \leq C_\delta \sqrt{k^{-1} \log m}$ for a sufficiently small constant $C_\delta > 0$; or (ii) $\gamma^2 k^3 m^{-2} = \Omega(1)$ and $\mu = O(\gamma^{-1} k^{-2} m)$.
- (Computational lower bound) Fix any constants $\delta \in (0, 0.1)$ and $\epsilon \in (0, \delta)$, and let $D_m = \lfloor \log^{1+\epsilon} m \rfloor$. For either prior distribution $\dot{\nu}$ or $\tilde{\nu}$, no polynomial of degree at most D_m achieves strong separation under either of the following conditions: (i) $\gamma k^2 m^{-1} \leq m^{-\delta}$ and $\mu \leq C(\log m)^{(1-\epsilon)/2}$ for a sufficiently small constant $C > 0$; or (ii) $\gamma k^2 m^{-1} \geq \log^{1+\delta} m$ and $\mu = O(\gamma^{-1} k^{-2} m)$.

We first establish the claimed lower bounds under the restriction $\gamma k \leq \kappa m$ for an arbitrary fixed constant $\kappa \in (0, 1)$, treating the fixed-size case and the random-size case in Appendix B.7.1 and Appendix B.7.2, respectively. We then remove this restriction by the grouping arguments below.

Throughout the following analysis, we assume that $k = \omega(\log m)$, which is automatically satisfied in the regime of interest where k is polynomial in m . The case $\gamma = 1$ follows from Theorem 1 and Corollary 1, so we assume $\gamma \geq 2$ below. When ck or Ck denotes a fixed vertex count, we round it down; this does not affect the asymptotic rates since $k \rightarrow \infty$.

B.7.1 Fixed-size case

We first consider the case where $\dot{\mathcal{C}}$ consists of all subgraphs that are the union of γ disjoint cliques of size k . In this case, we have $n = \binom{m}{2}$, $K = \gamma \cdot \binom{k}{2}$ and $|\dot{\mathcal{C}}| = \frac{1}{\gamma!} \binom{m}{k} \cdot \binom{m-k}{k} \cdots \binom{m-(\gamma-1)k}{k}$.

Upper bounds. We analyze the upper bounds on the critical signal strength μ^* obtained by the three tests in Proposition 2:

- For some constant $C > 0$, the outlier test achieves $R_{\dot{\nu}}(\Delta_{\text{out}}) \rightarrow 0$ whenever $\mu \geq C\sqrt{\log m}$.
- The average test achieves $R_{\dot{\nu}}(\Delta_{\text{avg}}) \rightarrow 0$ when $\mu = \omega(\frac{m}{\gamma k^2})$.
- Note that $|\dot{\mathcal{C}}| \leq \binom{m}{k}^\gamma$, and thus $\log |\dot{\mathcal{C}}| \leq \gamma k \log \frac{em}{k}$. Therefore, for some constant $C > 0$, the scan test achieves $R_{\dot{\nu}}(\Delta_{\text{scan}}) \rightarrow 0$ whenever

$$\mu \geq C\sqrt{\frac{1}{k} \log \frac{m}{k}}.$$

These guarantees follow by taking $\delta_m = 1/\log m$ in Proposition 2, except that for the average test we use its risk bound directly, as in Appendix B.5. Since $\gamma \geq 2$ and $\gamma k \leq m$, we have $k \leq m/2$, so $\log(em/k) \asymp \log(m/k)$.

Statistical lower bound. Fix $\kappa \in (0, 1)$, and assume that $\gamma k \leq \kappa m$. Denote the γ disjoint cliques in S as $S_1, \dots, S_\gamma$, and let $\dot{V}_1, \dots, \dot{V}_\gamma$ be their corresponding vertex membership sets. Formally, for $a \in [\gamma]$, define $\dot{V}_a := \{l \in [m] : (a-1)k < l \leq ak\}$ and let $S_a = \{(i, j) \in [m]_2 : i, j \in \dot{V}_a\}$. Then, $S = \bigcup_{a \in [\gamma]} S_a$. Let $\dot{\mathbf{A}}'$ be an independent copy of $\dot{\mathbf{A}}$, and define $S'_1, \dots, S'_\gamma, S'$ and $\dot{V}'_1, \dots, \dot{V}'_\gamma$ similarly. Then,

$$|S \cap S'| = \sum_{a,b} |S_a \cap S'_b| = \sum_{a,b} \binom{|\dot{V}_a \cap \dot{V}'_b|}{2}.$$

For brevity, write $f(x) := \exp\{\mu^2 x(x-1)/2\}$. This function is nondecreasing on the nonnegative integers and convex on $\mathbb{R}$. Then

$$\begin{aligned} \mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] &= \mathbb{E}\left[\exp\left\{\mu^2 \sum_{a=1}^{\gamma} \sum_{b=1}^{\gamma} \binom{|\dot{V}_a \cap \dot{V}'_b|}{2}\right\}\right] \\ &= \mathbb{E}\left[\prod_{a=1}^{\gamma} \prod_{b=1}^{\gamma} \exp\left\{\mu^2 \binom{|\dot{V}_a \cap \dot{V}'_b|}{2}\right\}\right] \\ &= \mathbb{E}\left[\prod_{a=1}^{\gamma} \prod_{b=1}^{\gamma} f(|\dot{V}_a \cap \dot{V}'_b|)\right] \\ &= \mathbb{E}\left[\mathbb{E}\left\{\prod_{a=1}^{\gamma} \prod_{b=1}^{\gamma} f(|\dot{V}_a \cap \dot{V}'_b|) \mid \dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}'_1, \dots, \dot{V}'_{\gamma-1}\right\}\right] \\ &= \mathbb{E}\left[\prod_{a=1}^{\gamma} \prod_{b=1}^{\gamma-1} f(|\dot{V}_a \cap \dot{V}'_b|) \mathbb{E}\left\{\prod_{a=1}^{\gamma} f(|\dot{V}_a \cap \dot{V}'_\gamma|) \mid \dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}'_1, \dots, \dot{V}'_{\gamma-1}\right\}\right]. \end{aligned} \tag{26}$$

Conditionally on $\dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}'_1, \dots, \dot{V}'_{\gamma-1}$, the random variables $\{|\dot{V}_a \cap \dot{V}'_\gamma|\}_{a=1}^{\gamma}$ are negatively associated. This is because $\dot{V}'_\gamma$ is a uniformly random subset of size k from the set $\mathcal{V}$ of $m - (\gamma-1)k$ nodes not contained in any of the first $\gamma-1$ node selections $\dot{V}'_1, \dots, \dot{V}'_{\gamma-1}$; that is, $\mathcal{V} = [m] \setminus \bigcup_{a=1}^{\gamma-1} \dot{V}'_a$. In turn, the set $\mathcal{V}$ contains up to γ disjoint (possibly empty) sets of the form $\mathcal{V} \cap \dot{V}_a$, for $a = 1, \dots, \gamma$. As a result, the random integer-valued vector

$$\left(\underbrace{|\dot{V}_1 \cap \dot{V}'_\gamma|, \dots, |\dot{V}_\gamma \cap \dot{V}'_\gamma|}_{\gamma \text{ times}}, \left|(\mathcal{V} \setminus \bigcup_{a=1}^{\gamma} \dot{V}_a) \cap \dot{V}'_\gamma\right|\right)$$

has a multivariate hypergeometric distribution with parameters k and

$$\left(\underbrace{|\mathcal{V} \cap \dot{V}_1|, \dots, |\mathcal{V} \cap \dot{V}_\gamma|}_{\gamma \text{ times}}, \left| \mathcal{V} \setminus \cup_{a=1}^\gamma \dot{V}_a \right| \right)$$

describing the outcome of sampling without replacement k out of $m - (\gamma - 1)k$ nodes, split into $\gamma + 1$ groups $\dot{V}_1 \cap \mathcal{V}, \dots, \dot{V}_\gamma \cap \mathcal{V}, \mathcal{V} \setminus \cup_{a=1}^\gamma \dot{V}_a$, with each group $\dot{V}_a \cap \mathcal{V}$, for $a = 1, \dots, \gamma$, possibly empty. Thus, as coordinates of a multivariate hypergeometric distribution, $\{|\dot{V}_a \cap \dot{V}_\gamma'|\}_{a=1}^\gamma$ are conditionally negatively associated.

Therefore,

$$\begin{aligned} & \mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-1} f(|\dot{V}_a \cap \dot{V}_b'|) \mathbb{E} \left\{ \prod_{a=1}^\gamma f(|\dot{V}_a \cap \dot{V}_\gamma'|) \mid \dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}_1', \dots, \dot{V}_{\gamma-1}' \right\} \right] \\ & \leq \mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-1} f(|\dot{V}_a \cap \dot{V}_b'|) \prod_{a=1}^\gamma \mathbb{E} \left\{ f(|\dot{V}_a \cap \dot{V}_\gamma'|) \mid \dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}_1', \dots, \dot{V}_{\gamma-1}' \right\} \right]. \end{aligned} \quad (27)$$

The same argument shows that, for each $b = 1, \dots, \gamma - 1$, the random variables $\{|\dot{V}_a \cap \dot{V}_b'|_{a=1}^\gamma$ are negatively associated conditionally on $\dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}_1', \dots, \dot{V}_{b-1}'$.

Next, for any $a, b = 1, \dots, \gamma$, it holds that

$$\mathbb{E}[f(|\dot{V}_a \cap \dot{V}_b'|) \mid \dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}_1', \dots, \dot{V}_{b-1}'] \leq \mathbb{E}[f(B)] \quad (28)$$

where $B \sim \text{Binomial}(k, p_{\gamma,k})$, and $p_{\gamma,k} := k / \{m - (\gamma - 1)k\}$. We prove this inequality below. Using it, the right-hand side of Equation (27) is bounded above by

$$\mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-1} f(|\dot{V}_a \cap \dot{V}_b'|) \prod_{a=1}^\gamma \mathbb{E}[f(B)] \right] = \mathbb{E}[f(B)]^\gamma \cdot \mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-1} f(|\dot{V}_a \cap \dot{V}_b'|) \right].$$

Plugging this bound into Equation (26) we obtain

$$\begin{aligned} & \mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] \\ & \leq \mathbb{E}[f(B)]^\gamma \cdot \mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-1} f(|\dot{V}_a \cap \dot{V}_b'|) \right] \\ & = \mathbb{E}[f(B)]^\gamma \cdot \mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-2} f(|\dot{V}_a \cap \dot{V}_b'|) \mathbb{E} \left\{ \prod_{a=1}^\gamma f(|\dot{V}_a \cap \dot{V}_{\gamma-1}'|) \mid \dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}_1', \dots, \dot{V}_{\gamma-2}' \right\} \right]. \end{aligned} \quad (29)$$

By repeating the same arguments that led to Equation (27),

$$\begin{aligned} & \mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-2} f(|\dot{V}_a \cap \dot{V}_b'|) \mathbb{E} \left\{ \prod_{a=1}^\gamma f(|\dot{V}_a \cap \dot{V}_{\gamma-1}'|) \mid \dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}_1', \dots, \dot{V}_{\gamma-2}' \right\} \right] \\ & \leq \mathbb{E}[f(B)]^\gamma \mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-2} f(|\dot{V}_a \cap \dot{V}_b'|) \right]. \end{aligned}$$

Plugging this bound into the right-hand side of Equation (29),

$$\mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] \leq \mathbb{E}[f(B)]^{2\gamma} \cdot \mathbb{E} \left[\prod_{a=1}^\gamma \prod_{b=1}^{\gamma-2} f(|\dot{V}_a \cap \dot{V}_b'|) \right].$$

Iterating over $b = \gamma - 2, \dots, 1$ finally gives

$$\mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] \leq \mathbb{E}[f(B)]^{\gamma^2} = \mathbb{E} \left[\exp \left(\mu^2 \binom{B}{2} \right) \right]^{\gamma^2}. \quad (30)$$

Proof of Equation (28). Throughout, we will condition on $\dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}'_1, \dots, \dot{V}'_{b-1}$. Then, $\dot{V}'_b$ is a uniformly random subset of size k drawn without replacement from the $M_b := m - (b-1)k$ nodes not contained in any of the previously selected sets $\dot{V}'_1, \dots, \dot{V}'_{b-1}$. Thus

$$|\dot{V}_a \cap \dot{V}'_b| \mid \dot{V}_a, \dot{V}'_1, \dots, \dot{V}'_{b-1} \sim \text{Hypergeometric}\left(M_b, \left|\dot{V}_a \setminus \bigcup_{l=1}^{b-1} \dot{V}'_l\right|, k\right).$$

Since $|\dot{V}_a \setminus \bigcup_{l=1}^{b-1} \dot{V}'_l| \leq k$ and the hypergeometric distribution is stochastically increasing in the number of successes,

$$|\dot{V}_a \cap \dot{V}'_b| \mid \dot{V}_a, \dot{V}'_1, \dots, \dot{V}'_{b-1} \preceq_{\text{st}} H,$$

where $H \sim \text{Hypergeometric}(M_b, k, k)$ almost surely. In turn, this implies that

$$\mathbb{E}[f(|\dot{V}_a \cap \dot{V}'_b|) \mid \dot{V}_a, \dot{V}'_1, \dots, \dot{V}'_{b-1}] \leq \mathbb{E}[f(H)],$$

since $f(x) = \exp(\mu^2 \binom{x}{2})$ is nondecreasing on the nonnegative integers. Furthermore, as the hypergeometric distribution is dominated by the corresponding binomial distribution in the convex ordering (Hoeffding, 1963), we have

$$\begin{aligned} \text{Hypergeometric}(M_b, k, k) &\preceq_{\text{cx}} \text{Binomial}\left(k, \frac{k}{M_b}\right) \\ &\preceq_{\text{st}} \text{Binomial}(k, p_{\gamma,k}), \end{aligned}$$

where $p_{\gamma,k} := k/(m - (\gamma-1)k)$ and the last stochastic ordering follows from the fact that $M_b \geq m - (\gamma-1)k$. Therefore, we have $\mathbb{E}[f(H)] \leq \mathbb{E}[f(B)]$, where $B \sim \text{Binomial}(k, p_{\gamma,k})$. Since f is nondecreasing on the nonnegative integers and convex, we obtain

$$\mathbb{E}[f(|\dot{V}_a \cap \dot{V}'_b|) \mid \dot{V}_1, \dots, \dot{V}_\gamma, \dot{V}'_1, \dots, \dot{V}'_{b-1}] \leq \mathbb{E}[f(B)],$$

which proves Equation (28).

Getting back to Equation (30), by Finner's inequality (see the arguments leading to Equation (18) for details),

$$\mathbb{E}\left[\exp\left\{\mu^2 \binom{B}{2}\right\}\right] \leq \exp\left\{\frac{kp_{\gamma,k}^2}{2}(\exp(\mu^2 k) - 1)\right\}, \quad (31)$$

which implies

$$\mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] \leq \mathbb{E}\left[\exp\left\{\mu^2 \binom{B}{2}\right\}\right]^{\gamma^2} \leq \exp\left\{\frac{\gamma^2 k^3}{2(1-\kappa)^2 m^2}(\exp(\mu^2 k) - 1)\right\}, \quad (32)$$

provided that $m - (\gamma-1)k \geq (1-\kappa)m$. Thus, $\mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] = O(1)$ if $\frac{\gamma^2 k^3}{2(1-\kappa)^2 m^2}(\exp(\mu^2 k) - 1)$ is at most C for a fixed constant $C > 0$. Equivalently,

$$\mu \leq \sqrt{\frac{1}{k} \log\left(1 + C \frac{2(1-\kappa)^2 m^2}{\gamma^2 k^3}\right)}.$$

In the regime $\gamma^2 k^3 = \omega(m^2)$, we have $C \frac{(1-\kappa)^2 m^2}{\gamma^2 k^3} \leq \log(1 + C \frac{2(1-\kappa)^2 m^2}{\gamma^2 k^3})$ for all sufficiently large m . Therefore, the condition in the above equation holds as long as $\mu \leq \sqrt{C}(1-\kappa) \frac{m}{\gamma k^2}$. The same rate holds under the weaker assumption $\gamma^2 k^3/m^2 = \Omega(1)$ stated above. Indeed, choose fixed constants $c_0, C_0 > 0$ such that $\gamma^2 k^3/m^2 \geq c_0$ and $\mu \leq C_0 m/(\gamma k^2)$ eventually. Then $\mu^2 k \leq C_0^2/c_0$, and $e^x - 1 \leq x e^x$ for $x \geq 0$ gives

$$\frac{\gamma^2 k^3}{2(1-\kappa)^2 m^2}(\exp(\mu^2 k) - 1) \leq \frac{C_0^2}{2(1-\kappa)^2} \exp(C_0^2/c_0) = O(1).$$

On the other hand, in the regime $\gamma^2 k^3 = O(m^{2-\delta})$ for any $\delta > 0$, since $\log(1 + C \frac{2(1-\kappa)^2 m^2}{\gamma^2 k^3}) \geq \delta \log m + O(1)$, we have $\log(1 + C \frac{2(1-\kappa)^2 m^2}{\gamma^2 k^3}) \geq \frac{\delta}{2} \log m$ for sufficiently large m . Therefore, it suffices to have $\mu \leq \sqrt{\frac{\delta}{2k} \log m}$.

We now remove the restriction $\gamma k \leq \kappa m$ for a fixed $\kappa \in (0, 1)$ and extend the analysis to $\gamma k \leq m$. The case $\gamma = 1$ reduces to the single planted clique problem considered above, so suppose that $\gamma \geq 2$, and let $q := \lceil \frac{\gamma}{2} \rceil$. Recall that $\dot{V}_a := \{l \in [m] : (a-1)k < \dot{A}_l \leq ak\}$ and let $S_a = \{(i, j) \in [m]_2 : i, j \in \dot{V}_a\}$. Define

$$\mathcal{S}_1 := \bigcup_{r=1}^q S_r, \quad \mathcal{S}_2 := \bigcup_{r=\gamma-q+1}^{\gamma} S_r, \quad (33)$$

and define $\mathcal{S}'_1, \mathcal{S}'_2$ analogously from the independent copy $\dot{\mathbf{A}}'$. Then, observe that

$$S = \mathcal{S}_1 \cup \mathcal{S}_2, \quad S' = \mathcal{S}'_1 \cup \mathcal{S}'_2,$$

which implies that

$$|S \cap S'| \leq |\mathcal{S}_1 \cap \mathcal{S}'_1| + |\mathcal{S}_1 \cap \mathcal{S}'_2| + |\mathcal{S}_2 \cap \mathcal{S}'_1| + |\mathcal{S}_2 \cap \mathcal{S}'_2|. \quad (34)$$

By the permutation invariance of $\dot{\mathbf{A}}$, each of $\mathcal{S}_1$ and $\mathcal{S}_2$ is marginally distributed as a fixed-size multiple clique model consisting of q cliques of size exactly k , and similarly for $\mathcal{S}'_1$ and $\mathcal{S}'_2$. Moreover, since $\dot{\mathbf{A}}$ and $\dot{\mathbf{A}}'$ are independent, the cliques in $\mathcal{S}_1$ and $\mathcal{S}_2$ are independent of those in $\mathcal{S}'_1$ and $\mathcal{S}'_2$. Therefore, for $i, j \in \{1, 2\}$, the random variables $|\mathcal{S}_i \cap \mathcal{S}'_j|$ have the same distribution as the overlap between two independent multiple clique models.

Based on this observation, we first analyze the resulting statistical lower bound. Observe that

$$\begin{aligned} \mathbb{E}[\exp(\mu^2 |S \cap S'|)] &\stackrel{(a)}{\leq} \mathbb{E}\left[\exp\left(\mu^2 \left\{\sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j|\right\}\right)\right] \\ &= \mathbb{E}\left[\prod_{i,j=1}^2 \exp(\mu^2 |\mathcal{S}_i \cap \mathcal{S}'_j|)\right], \\ &\stackrel{(b)}{\leq} \prod_{i,j=1}^2 \left(\mathbb{E}[\exp(4\mu^2 |\mathcal{S}_i \cap \mathcal{S}'_j|)]\right)^{1/4} \\ &\stackrel{(c)}{=} \mathbb{E}_{\mathcal{S}_1, \mathcal{S}'_1 \sim \dot{\nu}_{q,k}}[\exp(4\mu^2 |\mathcal{S}_1 \cap \mathcal{S}'_1|)], \end{aligned} \quad (35)$$

where the inequality (a) follows from Equation (34), the inequality (b) follows from Hölder's inequality, and the equality (c) follows from the fact that all the terms have the same distribution. We note that $\frac{\gamma}{2} \leq q = \lceil \frac{\gamma}{2} \rceil \leq \frac{2}{3}\gamma$, and thus $qk \leq \frac{2}{3}\gamma k \leq \frac{2}{3}m$. Therefore, we take $\kappa = 2/3$ and apply Equation (32) with (γ, k, μ) replaced by $(q, k, 2\mu)$ to obtain

$$\begin{aligned} \mathbb{E}_{\mathcal{S}_1, \mathcal{S}'_1 \sim \dot{\nu}_{q,k}}[\exp\{4\mu^2 |\mathcal{S}_1 \cap \mathcal{S}'_1|\}] &\leq \exp\left\{\frac{q^2 k^3}{2(1-\kappa)^2 m^2} (\exp(4\mu^2 k) - 1)\right\} \\ &= \exp\left\{\frac{9q^2 k^3}{2m^2} (\exp(4\mu^2 k) - 1)\right\} \\ &\leq \exp\left\{\frac{2\gamma^2 k^3}{m^2} (\exp(4\mu^2 k) - 1)\right\}, \end{aligned}$$

where the last inequality follows from $q \leq \frac{2}{3}\gamma$. The remaining technical steps follow the analysis subsequent to Equation (31), and this completes the proof.

Computational lower bound. From Proposition 1, the second moment of the low-degree likelihood ratio can be expressed as

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\hat{\nu}}^{\leq D}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu}\left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d\right]. \quad (36)$$

We reduce boundedness of this quantity to boundedness of its counterpart in the random-size model, with mixing measure $\tilde{\nu}$ instead of $\dot{\nu}$.

Specifically, recall that the latent structure S for multiple planted cliques is defined as

$$S := \bigcup_{r=1}^{\gamma} \left\{ (i, j) \in [m]_2 : (r-1)k < A_i, A_j \leq rk \right\},$$

where $\mathbf{A} = (A_1, \dots, A_m)$ is either $\dot{\mathbf{A}}$ for the fixed-size case or $\tilde{\mathbf{A}}$ for the random-size case, and let us denote the corresponding distributions of S given γ and k as $\dot{\nu} = \dot{\nu}_{\gamma, k}$ and $\tilde{\nu} = \tilde{\nu}_{\gamma, k}$, respectively. Fix an arbitrary constant $C \in (1, 1/\kappa)$ so that $\gamma(Ck) \leq C\kappa m < m$, and consider $\tilde{\nu}_{\gamma, Ck}$, which corresponds to the random-size case with the expected clique size scaled by C . Define an event $\underline{\mathcal{E}} = \{|\tilde{V}_a| \geq k, \forall a \in [\gamma]\}$, and observe that the union bound and Chernoff's inequality give

$$\tilde{\nu}_{\gamma, Ck}(\underline{\mathcal{E}}) \geq 1 - \gamma \exp\left(-\frac{(C-1)^2 k}{2C}\right) = 1 - o(1),$$

provided that $k = \omega(\log m)$. Let us define $\underline{\mathcal{E}}' = \{|\tilde{V}'_a| \geq k, \forall a \in [\gamma]\}$ analogously for the independent copy. Writing $\Pr$ for probability under the joint law of these two independent latent samples, for sufficiently large m , we have $\Pr(\underline{\mathcal{E}} \cap \underline{\mathcal{E}}') = \tilde{\nu}_{\gamma, Ck}(\underline{\mathcal{E}})^2 \geq \frac{1}{2}$. Then, for such m the following bound holds:

$$\begin{aligned} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}_{\gamma, Ck}}^{\leq D}(\mathbf{Y})^2] &= \mathbb{E}_{S, S' \sim \tilde{\nu}_{\gamma, Ck}} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \right] \\ &= \Pr(\underline{\mathcal{E}} \cap \underline{\mathcal{E}}') \mathbb{E}_{S, S' \sim \tilde{\nu}_{\gamma, Ck}} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \middle| \underline{\mathcal{E}} \cap \underline{\mathcal{E}}' \right] \\ &\quad + \Pr(\underline{\mathcal{E}}^c \cup \underline{\mathcal{E}}'^c) \mathbb{E}_{S, S' \sim \tilde{\nu}_{\gamma, Ck}} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \middle| \underline{\mathcal{E}}^c \cup \underline{\mathcal{E}}'^c \right] \\ &\geq \frac{1}{2} \mathbb{E}_{S, S' \sim \tilde{\nu}_{\gamma, Ck}} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \middle| \underline{\mathcal{E}} \cap \underline{\mathcal{E}}' \right] \\ &\geq \frac{1}{2} \mathbb{E}_{S, S' \sim \dot{\nu}_{\gamma, k}} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \right] \\ &= \frac{1}{2} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\dot{\nu}_{\gamma, k}}^{\leq D}(\mathbf{Y})^2], \end{aligned} \tag{37}$$

where the last inequality follows from the fact that the distribution of $|S \cap S'|$ conditioned on $\underline{\mathcal{E}} \cap \underline{\mathcal{E}}'$ under $\tilde{\nu}_{\gamma, Ck}$ stochastically dominates the distribution of $|S \cap S'|$ under $\dot{\nu}_{\gamma, k}$, since on the event $\underline{\mathcal{E}} \cap \underline{\mathcal{E}}'$, each clique in the random-size model has size at least k , which is the size of the cliques in the fixed-size model. To construct the coupling, select a uniform k -subset of each clique, independently in the two copies, conditional on these events. By vertex-permutation invariance, the resulting disjoint vertex sets have the fixed-size distribution, and their edge sets are contained in those of the random-size model. In summary, we obtain the following

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}_{\gamma, k}}^{\leq D}(\mathbf{Y})^2] \leq 2 \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}_{\gamma, Ck}}^{\leq D}(\mathbf{Y})^2],$$

and thus if we can show that $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}_{\gamma, Ck}}^{\leq D}(\mathbf{Y})^2]$ is $O(1)$ for $D = D_m$ as specified above, then we can conclude that $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}_{\gamma, k}}^{\leq D}(\mathbf{Y})^2]$ is also $O(1)$ for the same D . Therefore, the computational lower bound established for the random-size case in Appendix B.7.2 can be applied directly to the fixed-size case through the above comparison. Strictly speaking, the comparison requires a bound for the random-size model with mixing measure $\tilde{\nu}_{\gamma, Ck}$, whereas the analysis on the random-size case states the result for $\tilde{\nu}_{\gamma, k}$. This distinction is immaterial for the asymptotic rate: as can be seen from the resulting bound, replacing k by Ck changes the sufficient condition on μ only by a constant factor.

We now consider the full feasible region $\gamma k \leq m$, removing the restriction $\gamma k \leq \kappa m$. Recall the definition of $\mathcal{S}_1$ and $\mathcal{S}_2$ from Equation (33) and the bound $|S \cap S'| \leq |\mathcal{S}_1 \cap \mathcal{S}'_1| + |\mathcal{S}_1 \cap \mathcal{S}'_2| + |\mathcal{S}_2 \cap \mathcal{S}'_1| + |\mathcal{S}_2 \cap \mathcal{S}'_2|$. Then, observe that

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2] &= 1 + \mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \right] \\
&\leq 1 + \mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} \left\{ \sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j| \right\}^d \right] \\
&\stackrel{(a)}{\leq} 1 + \mathbb{E} \left[\sum_{d=1}^D \frac{4^{d-1} \mu^{2d}}{d!} \sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j|^d \right] \\
&= 1 + \mathbb{E}_{\mathcal{S}_1, \mathcal{S}'_1 \sim \tilde{\nu}_{q,k}} \left[\sum_{d=1}^D \frac{(2\mu)^{2d}}{d!} |\mathcal{S}_1 \cap \mathcal{S}'_1|^d \right] \\
&\stackrel{(b)}{\leq} 1 + 2\mathbb{E}_{\mathcal{S}_1, \mathcal{S}'_1 \sim \tilde{\nu}_{q,Ck}} \left[\sum_{d=1}^D \frac{(2\mu)^{2d}}{d!} |\mathcal{S}_1 \cap \mathcal{S}'_1|^d \right]
\end{aligned} \tag{38}$$

where (a) follows from the fact that $(\sum_{i=1}^4 x_i)^d \leq 4^{d-1} \sum_{i=1}^4 x_i^d$ for $x_i \geq 0$, and (b) follows from the same comparison of nonnegative moments as in Equation (37) since $qk \leq \frac{2}{3}m$ and thus the argument holds with $\kappa = \frac{2}{3}$ and $C \in (1, 1/\kappa)$. Observe that the expectation on the right-hand side has the same form as the nonconstant part of the low-degree likelihood-ratio second moment for the random-size multiple-clique model with q cliques, clique-size parameter Ck , and signal strength 2μ . Note that $q(Ck) < m$, and $\frac{q(Ck)^2}{m} \asymp \frac{\gamma k^2}{m}$ since $\frac{C^2}{2} \frac{\gamma k^2}{m} \leq \frac{q(Ck)^2}{m} \leq \frac{2C^2}{3} \frac{\gamma k^2}{m}$. Therefore, the random-size computational lower-bound analysis in Appendix B.7.2 applies with (γ, k, μ) replaced by $(q, Ck, 2\mu)$, and these substitutions affect the resulting conditions only by constant factors. To absorb these fixed factors in the regime conditions, replace δ by any $\delta' \in (\epsilon, \delta)$ and decrease the signal constant if necessary.

B.7.2 Random-size case

We now consider the model with random clique sizes. More specifically, each vertex $v \in [m]$ is independently assigned to at most one of the γ cliques with equal probability k/m , so that the expected size of each clique is k . For $a \in [\gamma]$, define $\tilde{V}_a := \{v \in [m] : (a-1)k < \tilde{A}_v \leq ak\}$ and let $S_a = \{(i, j) \in [m]_2 : i, j \in \tilde{V}_a\}$. Let $\tilde{\mathbf{A}}'$ be an independent copy of $\tilde{\mathbf{A}}$, and define $S'_1, \dots, S'_\gamma$ and $\tilde{V}'_1, \dots, \tilde{V}'_\gamma$ analogously. Let $S := \bigcup_{a=1}^\gamma S_a$, $S' := \bigcup_{b=1}^\gamma S'_b$.

Upper bounds. As in the upper bound analysis in the random-size clique problem (see Appendix B.6), we consider a high-probability event $\underline{\mathcal{E}}$ on which the size of each clique is at least ck for some constant $c \in (0, 1)$ and control the risk of the tests conditioned on this event. Specifically, consider the event

$$\underline{\mathcal{E}} := \{|\tilde{V}_a| \geq ck, \forall a \in [\gamma]\}, \tag{39}$$

and let $\tilde{\nu}$ denote the corresponding conditional distribution, $\tilde{\nu}(\cdot) := \tilde{\nu}(\cdot | \underline{\mathcal{E}})$. Since each $|\tilde{V}_a|$ is marginally distributed as $\text{Binomial}(m, k/m)$, the union bound and Chernoff's inequality give

$$\tilde{\nu}(\underline{\mathcal{E}}^c) \leq \gamma \exp \left(- \frac{(1-c)^2}{2} k \right) = o(1),$$

provided that $\log \gamma = o(k)$, which is satisfied in the polynomial regime considered here. As shown in Equation (21), the risk difference between the original and conditional distributions vanishes:

$$R_{\tilde{\nu}}(\Delta) \leq R_{\tilde{\nu}}(\Delta) + \tilde{\nu}(\underline{\mathcal{E}}^c),$$

which implies that we have $R_{\tilde{\nu}}(\Delta) \rightarrow 0$ as long as $R_{\tilde{\nu}}(\Delta) \rightarrow 0$.

Observe that under the event $\underline{\mathcal{E}}$, each S contains γ disjoint cliques of size at least ck , and thus the tests for the fixed-size case can be applied by targeting the cliques of size ck . In particular, the average test uses

threshold $\mu\gamma(\lfloor \frac{ck}{2} \rfloor)/2$, and the scan searches over unions of γ disjoint $\lfloor ck \rfloor$ -vertex cliques. Larger planted cliques only increase the mean of the average statistic, and the scan still includes a fully signaled candidate. Replacing k with ck only affects the constant factors in the upper bound, and hence we obtain the same upper bound as in Appendix B.7.1.

Statistical lower bound. We first assume that $\gamma k \leq \kappa m$ for some constant $\kappa \in (0, 1)$, and we will remove this assumption later. We follow the truncation argument detailed in Appendix B.6, which we sketch below for the reader's convenience. Fix a constant $C \in (1, 1/\kappa)$ and define the event

$$\bar{\mathcal{E}} := \{|\tilde{V}_a| \leq Ck, \forall a \in [\gamma]\}. \quad (40)$$

Let $\bar{\nu}$ denote the conditional distribution of $\tilde{\nu}$ given $\bar{\mathcal{E}}$. Since each $|\tilde{V}_a|$ is marginally distributed as $\text{Binomial}(m, k/m)$, the union bound and Chernoff's inequality give

$$\bar{\nu}(\bar{\mathcal{E}}^c) \leq \gamma \exp\left(-\frac{(C-1)^2}{2C}k\right) = o(1),$$

where the last equality holds when $\log \gamma = o(k)$, as in the polynomial regime considered here. Therefore, by the same argument as in Equation (22), for any test Δ ,

$$R_{\bar{\nu}}(\Delta) \geq 1 - \sqrt{1 - \frac{1}{\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\bar{\nu}}(\mathbf{Y})^2]}} - \bar{\nu}(\bar{\mathcal{E}}^c),$$

and hence it suffices to show that $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\bar{\nu}}(\mathbf{Y})^2] = O(1)$. To prove this, we compare the truncated random-size model with the fixed-size model. See Appendix B.6 for detailed arguments. Under $\bar{\nu}$, the sets $\tilde{V}_1, \dots, \tilde{V}_\gamma$ are disjoint and satisfy $|\tilde{V}_a| \leq Ck$ for all $a \in [\gamma]$. Since $C\gamma k \leq C\kappa m < m$, we can enlarge each $\tilde{V}_a$ to a set $\tilde{V}_a \supseteq \tilde{V}_a$ of size exactly Ck by adding vertices uniformly at random from the unassigned vertices, and construct $\tilde{V}'_1, \dots, \tilde{V}'_\gamma$ independently in the same way from $\tilde{V}'_1, \dots, \tilde{V}'_\gamma$. By permutation invariance of $\bar{\nu}$ and of the filling procedure, $(\tilde{V}_1, \dots, \tilde{V}_\gamma)$ and $(\tilde{V}'_1, \dots, \tilde{V}'_\gamma)$ are independent and each has the same distribution as the vertex sets in the fixed-size multiple-clique model with clique size Ck .

Let $\dot{S}$ and $\dot{S}'$ denote the corresponding unions of fixed-size cliques. Since $\tilde{V}_a \subseteq \tilde{V}_a$ and $\tilde{V}'_a \subseteq \tilde{V}'_a$ for every $a \in [\gamma]$, we have $S \subseteq \dot{S}$ and $S' \subseteq \dot{S}'$, and therefore $|S \cap S'| \leq |\dot{S} \cap \dot{S}'|$ almost surely. It follows that

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\bar{\nu}}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \bar{\nu}}[\exp\{\mu^2 |S \cap S'|\}] \leq \mathbb{E}_{\dot{S}, \dot{S}'}[\exp\{\mu^2 |\dot{S} \cap \dot{S}'|\}] \quad (41)$$

$$\leq \exp\left\{\frac{C^3 \gamma^2 k^3}{2(1 - C\kappa)^2 m^2} (\exp(C\mu^2 k) - 1)\right\}, \quad (42)$$

where the last inequality follows by repeating the fixed-size analysis above with k replaced by Ck , and κ replaced by $C\kappa$. Since $C \in (1, 1/\kappa)$ is a fixed constant, this replacement only changes constant factors. Therefore, repeating the two cases in the fixed-size analysis yields the same statistical lower bound as in Theorem 2.

We next remove the restriction $\gamma k \leq \kappa m$ and consider the general case $\gamma k \leq m$. We use the grouping argument from Appendix B.7.1. Let $q := \lceil \frac{\gamma}{2} \rceil$ and define

$$\mathcal{S}_1 := \bigcup_{r=1}^q \mathcal{S}_r, \quad \mathcal{S}_2 := \bigcup_{r=\gamma-q+1}^{\gamma} \mathcal{S}_r,$$

and define $\mathcal{S}'_1, \mathcal{S}'_2$ analogously from the independent copy $\tilde{\mathbf{A}}'$. We apply the filling procedure used in Equation (41). Since $q \leq \frac{2\gamma}{3}$, it follows that $qCk \leq \frac{2C}{3}\gamma k \leq \frac{2C}{3}m$. Therefore, taking $\kappa = 2/3$, we have $qk \leq \kappa m$, and for any $C \in (1, 1/\kappa)$, $qCk \leq C\kappa m < m$. Thus, each of the q cliques in $\mathcal{S}_1$ and $\mathcal{S}'_1$ can be enlarged to a clique of size Ck , exactly as in the argument leading to Equation (41). Let us denote the corresponding unions of such cliques by $\dot{\mathcal{S}}_1$ and $\dot{\mathcal{S}}'_1$, respectively. Observe that $\mathcal{S}_1 \subseteq \dot{\mathcal{S}}_1$ and $\mathcal{S}'_1 \subseteq \dot{\mathcal{S}}'_1$. Moreover, by permutation invariance of

the conditional distribution and of the filling procedure, $\dot{S}_1$ and $\dot{S}'_1$ are independent and identically distributed with law $\dot{\nu}_{q,Ck}$. Following the analysis in Equation (35), we obtain

$$\begin{aligned}\mathbb{E}_{S,S' \sim \bar{\nu}_{\gamma,k}}[\exp\{\mu^2 |S \cap S'|\}] &\leq \mathbb{E}_{\dot{S}_1, \dot{S}'_1 \sim \dot{\nu}_{q,Ck}}[\exp(4\mu^2 |\dot{S}_1 \cap \dot{S}'_1|)] \\ &\leq \exp\left\{\frac{C^3 q^2 k^3}{2(1-2C/3)^2 m^2} (\exp(4C\mu^2 k) - 1)\right\} \\ &\leq \exp\left\{\frac{2C^3 \gamma^2 k^3}{9(1-2C/3)^2 m^2} (\exp(4C\mu^2 k) - 1)\right\},\end{aligned}$$

where the second inequality is obtained by applying the fixed-size multiple-clique bound in Equation (32) with $(\gamma, k, \mu) \mapsto (q, Ck, 2\mu)$, and the last inequality follows from $q \leq \frac{2\gamma}{3}$, which gives $q^2/2 \leq 2\gamma^2/9$. Thus, the resulting bound on μ differs from the previous bound by a constant factor.

Computational lower bound. Our objective is to bound

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} = \mathbb{E}_{S,S' \sim \bar{\nu}}\left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d\right],$$

and we focus on bounding $\mathbb{E}_{S,S' \sim \bar{\nu}}[|S \cap S'|^d]$ for each $d \in [D]$. Our approach is to bound these moments using an auxiliary clique model whose overlap moments are no smaller and are easier to analyze. For each vertex $v \in [m]$, define

$$W_v = \begin{cases} e_{(a,b)} & \text{if } v \in \tilde{V}_a \cap \tilde{V}'_b \text{ for some } a, b \in [\gamma], \\ 0 & \text{otherwise,} \end{cases} \quad (43)$$

where $e_{(a,b)} \in \{0,1\}^{\gamma \times \gamma}$ is the basis matrix with a 1 in the (a,b) -th entry and 0 elsewhere. Then $W_1, \dots, W_m$ are independent, and

$$W_v = \begin{cases} e_{(a,b)} & \text{with probability } \rho, \quad a, b \in [\gamma], \\ 0 & \text{with probability } 1 - \gamma^2 \rho. \end{cases}$$

where

$$\rho := k^2/m^2.$$

Observe that $|S \cap S'|$ can be expressed as

$$\begin{aligned}|S \cap S'| &= \sum_{a,b=1}^{\gamma} |S_a \cap S'_b| \\ &= \sum_{a,b=1}^{\gamma} \sum_{(i,j) \in [m]_2} \mathbf{1}\{i, j \in \tilde{V}_a \cap \tilde{V}'_b\} \\ &= \sum_{(i,j) \in [m]_2} \sum_{a,b=1}^{\gamma} \mathbf{1}\{i, j \in \tilde{V}_a \cap \tilde{V}'_b\} \\ &= \sum_{(i,j) \in [m]_2} \sum_{a,b=1}^{\gamma} \mathbf{1}\{W_i = e_{(a,b)}\} \mathbf{1}\{W_j = e_{(a,b)}\} \\ &= \sum_{(i,j) \in [m]_2} \mathbf{1}\{W_i = W_j \neq 0\}.\end{aligned} \quad (44)$$

An intuitive interpretation of the above expression is that $|S \cap S'|$ can be seen as the number of edges of another random γ^2 -multiple clique model; specifically, consider a model where each vertex $v \in [m]$ is assigned a label $e_{(a,b)}$ when $W_v = e_{(a,b)}$, with equal probability ρ (or 0 with probability $1 - \gamma^2 \rho$), and an edge is formed between two vertices having the same nonzero label. Note that the number of labels is γ^2 .

For notational convenience, write the overlap as $Z := |S \cap S'|$. Then, our approach is to bound $\mathbb{E}[Z^d]$ by comparing it with the d -th moment of an auxiliary random variable Z_* , which is defined as follows. Let $B_1, \dots, B_m \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$, where p is a parameter to be determined below, and let $B := \sum_{i=1}^m B_i \sim \text{Binomial}(m, p)$. Then, we set

$$Z_* := \sum_{(i,j) \in [m]_2} B_i B_j = \binom{B}{2}.$$

Then, B can be seen as the size of a random single clique in which each of the m vertices belongs to the clique independently and with probability p , and Z_* is the number of edges in the resulting random clique. We seek to find a suitable value for p such that $\mathbb{E}[Z^d] \leq \mathbb{E}[Z_*^d]$ for all $d \in [D]$, while matching the first moments. We begin with the case of $d = 1$ and compare the first moments of Z and Z_* . Letting $I_{(i,j)} := \mathbb{1}\{W_i = W_j \neq 0\}$, we have

$$\mathbb{E}[Z] = \sum_{(i,j) \in [m]_2} \mathbb{E}[I_{(i,j)}] = \sum_{(i,j) \in [m]_2} \Pr(W_i = W_j \neq 0) = \sum_{(i,j) \in [m]_2} \gamma^2 \rho^2 = \frac{m(m-1)}{2} \gamma^2 \rho^2.$$

On the other hand, denoting $J_{(i,j)} := \mathbb{1}\{B_i = B_j = 1\}$ for all $i \neq j$, the first moment of Z_* is given by

$$\mathbb{E}[Z_*] = \sum_{(i,j) \in [m]_2} \mathbb{E}[J_{(i,j)}] = \sum_{(i,j) \in [m]_2} \Pr(B_i = B_j = 1) = \sum_{(i,j) \in [m]_2} p^2 = \frac{m(m-1)}{2} p^2.$$

This motivates us to choose $p = \gamma\rho = \gamma k^2/m^2$ so that $\mathbb{E}[Z] = \mathbb{E}[Z_*]$. This is a valid Bernoulli parameter because $\gamma k \leq m$ implies $p \leq k/m \leq 1$. Next, we will show that this choice of p indeed guarantees that $\mathbb{E}[Z^d] \leq \mathbb{E}[Z_*^d]$ for all $d \geq 1$.

For any integer $d \geq 2$, write $(\mathbf{i}, \mathbf{j}) := ((i_1, j_1), \dots, (i_d, j_d)) \in ([m]_2)^d$, and observe that

$$\mathbb{E}[Z^d] = \sum_{(\mathbf{i}, \mathbf{j}) \in ([m]_2)^d} \mathbb{E}\left[\prod_{r=1}^d I_{(i_r, j_r)}\right] = \sum_{(\mathbf{i}, \mathbf{j}) \in ([m]_2)^d} \Pr\left(\prod_{r=1}^d I_{(i_r, j_r)} = 1\right).$$

Similarly, the d -th moment of Z_* can be expressed as

$$\mathbb{E}[Z_*^d] = \sum_{(\mathbf{i}, \mathbf{j}) \in ([m]_2)^d} \Pr\left(\prod_{r=1}^d J_{(i_r, j_r)} = 1\right).$$

For each $(\mathbf{i}, \mathbf{j}) \in ([m]_2)^d$, let $v = v(\mathbf{i}, \mathbf{j})$ be the number of distinct vertices in $(\mathbf{i}, \mathbf{j})$, and let $c = c(\mathbf{i}, \mathbf{j})$ be the number of connected components in the graph formed by the edges $(i_1, j_1), \dots, (i_d, j_d)$. Then, for $\prod_{r=1}^d I_{(i_r, j_r)}$ to be equal to 1, all vertices in each connected component must have the same nonzero label $e_{(a,b)}$ for some $a, b \in [\gamma]$. Since there are γ^2 possible labels, γ^{2c} is the number of ways to assign a nonzero label to each of the c connected components: each connected component requires its vertices to be associated with some pair (a, b) of cliques, for a total of $(\gamma^2)^c$ possibilities. Then, enumerate the connected components in any arbitrary way and let v_h be the number of vertices in the h th connected component, so $v = \sum_{h=1}^c v_h$. For any given assignment of pairs to connected components (as we saw, there are γ^{2c} such assignments), the probability that all the vertices in the h th connected component are of the same given type, say (a_h, b_h) , is ρ^{v_h} , yielding $\rho^{\sum_h v_h} = \rho^v$ for the overall probability that each connected component has vertices of the same given types (by independence). Therefore, we have

$$\Pr\left(\prod_{r=1}^d I_{(i_r, j_r)} = 1\right) = \gamma^{2c} \rho^v. \quad (45)$$

On the other hand, for the auxiliary model, $\prod_{r=1}^d J_{(i_r, j_r)}$ is 1 if and only if $B_i = 1$ for every vertex appearing in the selected edges, an event of probability

$$\Pr\left(\prod_{r=1}^d J_{(i_r, j_r)} = 1\right) = p^v = \gamma^v \rho^v.$$

As every connected component contains at least two vertices, $v \geq 2c$. This implies that

$$\Pr\left(\prod_{r=1}^d I_{(i_r, j_r)} = 1\right) = \gamma^{2c} \rho^v \leq \gamma^v \rho^v = \Pr\left(\prod_{r=1}^d J_{(i_r, j_r)} = 1\right),$$

which results in

$$\mathbb{E}[Z^d] = \sum_{(\mathbf{i}, \mathbf{j}) \in ([m]_2)^d} \Pr\left(\prod_{r=1}^d I_{(i_r, j_r)} = 1\right) \leq \sum_{(\mathbf{i}, \mathbf{j}) \in ([m]_2)^d} \Pr\left(\prod_{r=1}^d J_{(i_r, j_r)} = 1\right) = \mathbb{E}[Z_*^d].$$

In summary, we have shown that $\mathbb{E}[|S \cap S'|^d] \leq \mathbb{E}[Z_*^d]$ for all $d \geq 1$, which allows us to work with the sum $\sum_{d=0}^D \frac{\mu^{2d}}{d!} \mathbb{E}[Z_*^d]$. Note that $Z_* = \binom{B}{2} \leq B^2/2$, and $B \sim \text{Binomial}(m, p)$ with $p = \gamma k^2/m^2$. Following the same argument as in Appendix B.4, we apply the polynomial moment bound for binomial random variables (Ahle, 2022, Corollary 1) given in Lemma 5 to obtain

$$\sum_{d=0}^D \frac{\mu^{2d}}{2^d d!} \mathbb{E}[B^{2d}] \leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{2^d d!} \left(\frac{2d}{\log(1 + 2dm/(\gamma k^2))} \right)^{2d}. \quad (46)$$

Note that the only difference from Equation (19) is that the denominator inside the logarithm is now γk^2 instead of k^2 . Therefore, the rest of the proof can be completed by replacing k^2 with γk^2 in the analysis following Equation (19). We remark that the argument applies throughout the feasible region where $\gamma k \leq m$. Combining the upper and lower bounds proves the statistical thresholds in Theorem 2 at polynomial order, as defined in Definition 3. The computational thresholds follow in the same way under Conjecture 1.

B.8 Proof of Corollary 2

We extend the multiple-clique analysis to the hypergraph setting. We first state the statistical and low-degree lower bounds to be established; the computational interpretation of the latter uses Conjecture 1:

- (Statistical lower bound) Fix any constant $\delta \in (0, 0.1)$. For either prior distribution $\dot{\nu}$ or $\tilde{\nu}$, strong detection is impossible under either of the following conditions: (i) $\gamma^2 k^{t+1} m^{-t} \leq m^{-\delta}$ and $\mu \leq C_\delta \sqrt{k^{-(t-1)} \log m}$ for a sufficiently small constant $C_\delta > 0$; or (ii) $\gamma^2 k^{t+1} m^{-t} = \Omega(1)$ and $\mu = O(\gamma^{-1} k^{-t} m^{t/2})$.
- (Computational lower bound) Fix any constants $\delta \in (0, 0.1)$ and $\epsilon \in (0, \delta)$, and let $D_m = \lfloor \log^{1+\epsilon} m \rfloor$. For either prior distribution $\dot{\nu}$ or $\tilde{\nu}$, no polynomial of degree at most D_m achieves strong separation under either of the following conditions: (i) $\gamma^{2/t} k^2 m^{-1} \leq m^{-\delta}$ and $\mu \leq C(\log m)^{\{1-(t-1)\epsilon\}/2}$ for a sufficiently small constant $C > 0$; or (ii) $\gamma^{2/t} k^2 m^{-1} \geq \log^{1+\delta} m$ and $\mu = O(\gamma^{-1} k^{-t} m^{t/2})$.

Throughout, $t \geq 2$ is fixed and k has positive polynomial order in m , as in Corollary 2; in particular, $k = \omega(\log m)$ and $k \geq t$ eventually. Constants may depend on t , and we take $D = D_m$ in the low-degree bounds. Noninteger fixed vertex counts are rounded down, as in Appendix B.7. In both the fixed-size and random-size cases, the proof consists of three pieces: the upper bound, the statistical lower bound, and the computational lower bound. For ease of analysis, we assume that $\gamma k \leq \kappa m$ for some fixed constant $\kappa \in (0, 1)$ if needed. However, this can be relaxed to the general case $\gamma k \leq m$, following the same argument as in Appendix B.7.

Remark 7 (Extension to $\gamma k \leq m$). A key idea is that the multiple hyperclique S with $\gamma k \leq m$ can be expressed as a union of two groups of multiple hypercliques, say $\mathcal{S}_1$ and $\mathcal{S}_2$, where each group consists of $q := \lceil \gamma/2 \rceil$ hypercliques. Then, the overlap $|S \cap S'|$ can be bounded by $\sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j|$. Since $q = \lceil \gamma/2 \rceil \leq 2\gamma/3$ for $\gamma \geq 2$ and thus $qk \leq 2m/3$, we can interpret $|\mathcal{S}_i \cap \mathcal{S}'_j|$ as the overlap between two independent multiple hypercliques satisfying the restricted occupancy condition $qk \leq \kappa m$ with $\kappa = 2/3$, and apply the lower-bound results established for this regime. See Appendix B.7 for more details. Specifically, Hölder's inequality controls the full second moment, while $(x_1 + \dots + x_4)^d \leq 4^{d-1} \sum_{i=1}^4 x_i^d$ controls the low-degree moments; both comparisons only rescale μ by a constant factor. For $\gamma = 1$ and $k > \kappa m$, use $|S \cap S'| \leq \binom{m}{t}$ directly: if $\mu = O(m^{t/2}/k^t)$, then $\mu^2 \binom{m}{t} = O(1)$, which bounds both the full and low-degree second moments.

B.8.1 Fixed-size case

We consider the case where $\dot{\mathcal{C}}$ consists of all subhypergraphs that are the union of γ disjoint hypercliques of size k and dimension t . In this case, we have $n = \binom{m}{t}$, $K = \gamma \binom{k}{t}$ and $|\dot{\mathcal{C}}| = \frac{1}{\gamma!} \binom{m}{k} \cdot \binom{m-k}{k} \cdots \binom{m-(\gamma-1)k}{k}$. Define $\dot{V}_a := \{v \in [m] : (a-1)k < \dot{A}_v \leq ak\}$ and let $S_a = \{\mathbf{i} = (i_1, \dots, i_t) \in [m]_t : i_1, \dots, i_t \in \dot{V}_a\}$. Let $\dot{\mathbf{A}}'$ be an independent copy of $\dot{\mathbf{A}}$, and define $S'_1, \dots, S'_\gamma$ and $\dot{V}'_1, \dots, \dot{V}'_\gamma$ analogously. Finally, set $S := \bigcup_{a=1}^\gamma S_a$ and $S' := \bigcup_{b=1}^\gamma S'_b$.

Upper bounds. The upper bounds on the critical signal strength μ^* can be obtained by applying the three tests in Proposition 2:

- For some constant $C > 0$, the outlier test achieves $R_{\dot{\nu}}(\Delta_{\text{out}}) \rightarrow 0$ whenever $\mu \geq C\sqrt{\log m}$.
- The average test achieves $R_{\dot{\nu}}(\Delta_{\text{avg}}) \rightarrow 0$ when $\mu = \omega(\frac{m^{t/2}}{\gamma k^t})$.
- Note that $|\dot{\mathcal{C}}| \leq \binom{m}{k}^\gamma$, and thus $\log |\dot{\mathcal{C}}| \leq \gamma k \log \frac{em}{k}$. Therefore, for some constant $C > 0$, the scan test achieves $R_{\dot{\nu}}(\Delta_{\text{scan}}) \rightarrow 0$ provided that $k \leq (1 - c_0)m$ for a fixed $c_0 \in (0, 1)$ and

$$\mu \geq C \sqrt{\frac{1}{k^{t-1}} \log \frac{m}{k}}.$$

The restriction on k holds in the regime $\gamma^2 k^{t+1} \ll m^t$ where the scan bound is needed. As in Appendix B.7.1, take $\delta_m = 1/\log m$ for the outlier and scan tests, and use the average-test risk bound directly.

Statistical lower bound. Assume that $\gamma k \leq \kappa m$ for some fixed constant $\kappa \in (0, 1)$. We can write

$$\begin{aligned} \mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] &= \mathbb{E}\left[\exp\left\{\mu^2 \sum_{a,b=1}^\gamma \binom{|\dot{V}_a \cap \dot{V}'_b|}{t}\right\}\right] \\ &= \mathbb{E}\left[\prod_{a,b=1}^\gamma f(|\dot{V}_a \cap \dot{V}'_b|)\right], \end{aligned} \tag{47}$$

where we now set $f(x) := \exp\{\mu^2 \binom{x}{t}\}$ for nonnegative integers x , and extend f linearly between consecutive integers. The sequence $\binom{x}{t}$ is nondecreasing and discretely convex, so exponentiation and linear interpolation give a nondecreasing convex function on $[0, \infty)$. This extension leaves all expectations below unchanged, since their arguments are integer-valued. The conditional negative-association and convex-order arguments from Equation (26) through Equation (30) therefore apply and give

$$\mathbb{E}\left[\prod_{a,b=1}^\gamma f(|\dot{V}_a \cap \dot{V}'_b|)\right] \leq \mathbb{E}[f(B)]^{\gamma^2}, \tag{48}$$

where $B = \sum_{i=1}^k B_i$, with $B_1, \dots, B_k$ i.i.d. Bernoulli random variables with parameter $p_{\gamma,k} := k/(m - (\gamma-1)k)$. Then $B \sim \text{Binomial}(k, p_{\gamma,k})$ and

$$\mathbb{E}[f(B)] = \mathbb{E}\left[\exp\left\{\mu^2 \binom{B}{t}\right\}\right] = \mathbb{E}\left[\exp\left\{\mu^2 \sum_{\mathbf{i} \in [k]_t} B_{\mathbf{i}}\right\}\right] = \mathbb{E}\left[\prod_{\mathbf{i} \in [k]_t} \exp(\mu^2 B_{\mathbf{i}})\right]$$

where, for any $\mathbf{i} = (i_1, \dots, i_t) \in [k]_t$, we set $B_{\mathbf{i}} = B_{i_1} \cdots B_{i_t}$, and the second equality above follows from the easily verifiable identity

$$\binom{B}{t} = \sum_{\mathbf{i} \in [k]_t} B_{\mathbf{i}}. \tag{49}$$

Next we apply Finner's inequality as given in Lemma 6. In the notation of that lemma, take $I = [k]$, $J = [k]_t$, and $T_{\mathbf{i}} = \{i_1, \dots, i_t\}$ for each $\mathbf{i} = (i_1, \dots, i_t) \in [k]_t$. Thus, there are $\binom{k}{t}$ factors, each depending

on t coordinates, and we index the corresponding exponents by p_i . In particular, since each B_i appears in $\binom{k-1}{t-1}$ terms in the summation in Equation (49), we choose $p_i = \binom{k-1}{t-1}$ for all $i \in [k]_t$. The reciprocal exponents therefore sum to one at each coordinate, as required by Finner's inequality. Then, Finner's inequality (Lemma 6) yields

$$\begin{aligned}
\mathbb{E}_B \left[\exp \left(\mu^2 \sum_{i \in [k]_t} B_i \right) \right] &\leq \prod_{i \in [k]_t} \mathbb{E}_B [\exp(\mu^2 B_i)^{p_i}]^{\frac{1}{p_i}} \\
&= \prod_{i \in [k]_t} \mathbb{E}_B [\exp(\mu^2 B_i)^{\binom{k-1}{t-1}}]^{\frac{1}{\binom{k-1}{t-1}}} \\
&= \mathbb{E}_B \left[\exp \left\{ \mu^2 \binom{k-1}{t-1} B_i \right\} \right]^{k/t} \\
&\leq \mathbb{E}_B \left[\exp \left\{ \mu^2 k^{t-1} B_i \right\} \right]^{k/t} \\
&= \left\{ 1 + p_{\gamma,k}^t (\exp(\mu^2 k^{t-1}) - 1) \right\}^{k/t} \quad (\because B_i \sim \text{Bernoulli}(p_{\gamma,k}^t)) \\
&\stackrel{(a)}{\leq} \exp \left\{ \frac{k p_{\gamma,k}^t}{t} (\exp(\mu^2 k^{t-1}) - 1) \right\} \\
&\stackrel{(b)}{\leq} \exp \left\{ \frac{k^{t+1}}{t(1-\kappa)^t m^t} (\exp(\mu^2 k^{t-1}) - 1) \right\}
\end{aligned} \tag{50}$$

where (a) follows from the fact that $1 + x \leq e^x$ for all $x \in \mathbb{R}$, and (b) follows from the fact that $p_{\gamma,k} = k/(m - (\gamma - 1)k)$ and $m - (\gamma - 1)k \geq (1 - \kappa)m$. Therefore,

$$\mathbb{E}_{S, S' \sim \hat{\nu}} [\exp\{\mu^2 |S \cap S'|_J\}] \leq \exp \left\{ \frac{\gamma^2 k^{t+1}}{t(1-\kappa)^t m^t} (\exp(\mu^2 k^{t-1}) - 1) \right\}, \tag{51}$$

A sufficient condition for $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\hat{\nu}}(\mathbf{Y})^2] \leq e^C$, where $C > 0$ is fixed, is

$$\mu \leq \sqrt{\frac{1}{k^{t-1}} \log \left(1 + C \frac{t(1-\kappa)^t m^t}{\gamma^2 k^{t+1}} \right)}.$$

On one hand, if $\gamma^2 k^{t+1} = \omega(m^t)$, we have $C \frac{(1-\kappa)^t m^t}{\gamma^2 k^{t+1}} \leq \log(1 + C \frac{t(1-\kappa)^t m^t}{\gamma^2 k^{t+1}})$ for all sufficiently large m . Therefore, the condition in the above equation holds when $\mu \leq \sqrt{C} (1-\kappa)^{t/2} \gamma^{-1} k^{-t} m^{t/2}$. The same rate holds under the stated assumption $\gamma^2 k^{t+1}/m^t = \Omega(1)$. Indeed, if $\gamma^2 k^{t+1}/m^t \geq c_1 > 0$ and $\mu \leq C_1 m^{t/2}/(\gamma k^t)$ eventually, then $\mu^2 k^{t-1} \leq C_1^2/c_1$. Using $e^x - 1 \leq x e^x$ for $x \geq 0$, we obtain

$$\frac{\gamma^2 k^{t+1}}{t(1-\kappa)^t m^t} (\exp(\mu^2 k^{t-1}) - 1) \leq \frac{C_1^2}{t(1-\kappa)^t} \exp(C_1^2/c_1) = O(1).$$

On the other hand, in the regime $\gamma^2 k^{t+1} = O(m^{t-\delta})$ for any $\delta > 0$, since $\log(1 + C \frac{t(1-\kappa)^t m^t}{\gamma^2 k^{t+1}}) \geq \delta \log m + O(1)$, we have $\log(1 + C \frac{t(1-\kappa)^t m^t}{\gamma^2 k^{t+1}}) \geq \frac{\delta}{2} \log m$ for sufficiently large m . Therefore, it suffices to have $\mu \leq \sqrt{\frac{\delta}{2k^{t-1}} \log m}$.

Computational lower bound. The computational lower bound for the fixed-size case follows directly from the random-size result by the same stochastic dominance argument as in Appendix B.7. Fix $c \in (0, 1 - \kappa)$ and set $k' = k/(1 - c)$, so that $\gamma k' \leq \kappa m/(1 - c) < m$. Specifically, conditioning both independent copies of the random-size model on the high-probability event that every hyperclique has size at least $(1 - c)k'$ yields

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\hat{\nu}_{\gamma, (1-c)k'}}^{\leq D}(\mathbf{Y})^2] \leq 2 \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\hat{\nu}_{\gamma, k'}}^{\leq D}(\mathbf{Y})^2].$$

The joint conditioning event has probability at least $1 - 2\gamma \exp(-c^2 k'/2) = 1 - o(1)$, and hence at least $1/2$ eventually. Selecting a uniform k -subset of each hyperclique in each copy gives a coupling with the fixed-size

model, with containment of the corresponding hyperedge sets. This choice of k' transfers the boundedness of the LDLR second moment from the random-size case, which we cover in Appendix B.8.2, to the fixed-size case, and thus gives the desired computational lower bound. As in Appendix B.7.1, fixed multiplicative factors in the regime conditions can be absorbed by replacing δ by any $\delta' \in (\epsilon, \delta)$ and decreasing the signal constant if necessary.

B.8.2 Random-size case

We define $\tilde{V}_a := \{v \in [m] : (a-1)k < \tilde{A}_v \leq ak\}$ and $S_a := \{\mathbf{i} = (i_1, \dots, i_t) \in [m]_t : i_1, \dots, i_t \in \tilde{V}_a\}$. Also, let $\tilde{\mathbf{A}}'$ be an independent copy of $\tilde{\mathbf{A}}$, and define $S'_1, \dots, S'_\gamma$ and $\tilde{V}'_1, \dots, \tilde{V}'_\gamma$ analogously. Finally, we denote $S := \bigcup_{a=1}^\gamma S_a$ and $S' := \bigcup_{b=1}^\gamma S'_b$.

Upper bounds. Note that the same conditioning approach as in Appendix B.7.2 can be directly applied. In particular, we consider the event that each hyperclique has size at least ck for some constant $c \in (0, 1)$, and the risk difference between the original and conditional models is at most the probability of the complementary event. By the union bound and Chernoff's inequality, this probability is at most $\gamma \exp\{-(1-c)^2 k/2\} = o(1)$. Then, we can derive the upper bounds by targeting the hypercliques of size ck , which differs from the fixed-size case only by constant factors. The average test uses threshold $\mu\gamma \binom{ck}{t}/2$, and the scan searches over unions of γ disjoint $\lfloor ck \rfloor$ -vertex hypercliques.

Statistical lower bound. The truncation argument in Appendix B.7.2 applies without modification to the present hyperclique setting. Indeed, the argument from Equation (40) to Equation (41) only uses the sizes and inclusions of the underlying vertex sets, and does not depend on the hyperedge dimension t . In particular, conditioning on the event $\bar{\mathcal{E}} := \{|\tilde{V}_a| \leq Ck, \forall a \in [\gamma]\}$ for a fixed constant $C \in (1, 1/\kappa)$, we may enlarge each $\tilde{V}_a$ to a set $\dot{V}_a \supseteq \tilde{V}_a$ of size exactly Ck . Note that this is indeed feasible, since $C\gamma k \leq C\kappa m < m$. Here, $\bar{\nu}$ denotes the conditional prior, and $\bar{\nu}(\bar{\mathcal{E}}^c) \leq \gamma \exp\{-(C-1)^2 k/(2C)\} = o(1)$, so the truncation has vanishing cost in total variation. The resulting union of random-size hypercliques S is then contained in the corresponding union of fixed-size hypercliques $\dot{S}$, and the same construction for an independent copy gives $|S \cap S'| \leq |\dot{S} \cap \dot{S}'|$ almost surely. Therefore,

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\bar{\nu}}(\mathbf{Y})^2] \leq \mathbb{E}_{\dot{S}, \dot{S}'}[\exp \mu^2 |\dot{S} \cap \dot{S}'|] \leq \exp \left\{ \frac{C^{t+1} \gamma^2 k^{t+1}}{t(1-C\kappa)^t m^t} (\exp(C^{t-1} \mu^2 k^{t-1}) - 1) \right\},$$

where the inequality follows from the fixed-size analysis (see Equation (51)) with k replaced by Ck , and κ replaced by $C\kappa$. Since $C \in (1, 1/\kappa)$ is a fixed constant, this replacement only changes constant factors. Hence, repeating the same two cases as in the fixed-size analysis yields the statistical lower bound stated in Corollary 2.

Computational lower bound. We take a similar approach as in Appendix B.7.2 to bound the LDLR second moment. In particular, we will show that the moments of the overlap $|S \cap S'|$ can be tightly bounded by the moments of an auxiliary random variable Z_* , which corresponds to the number of hyperedges in a random single hyperclique model. Recall the definition of W_v in Equation (43):

$$W_v = \begin{cases} e_{(a,b)} & \text{if } v \in \tilde{V}_a \cap \tilde{V}'_b \text{ for some } a, b \in [\gamma], \\ 0 & \text{otherwise,} \end{cases} \quad (52)$$

where $\Pr(W_v = e_{(a,b)}) = \rho = k^2/m^2$ for all $a, b \in [\gamma]$ and $\Pr(W_v = 0) = 1 - \gamma^2\rho$. Denote $\mathbf{i} := (i_1, \dots, i_t) \in [m]_t$. Then, analogous to Equation (44), we can express the overlap $|S \cap S'|$ as

$$\begin{aligned}
|S \cap S'| &= \sum_{a,b=1}^{\gamma} |S_a \cap S'_b| \\
&= \sum_{a,b=1}^{\gamma} \sum_{\mathbf{i} \in [m]_t} \mathbb{1}\{i_1, \dots, i_t \in \tilde{V}_a \cap \tilde{V}'_b\} \\
&= \sum_{\mathbf{i} \in [m]_t} \sum_{a,b=1}^{\gamma} \prod_{l=1}^t \mathbb{1}\{W_{i_l} = e_{(a,b)}\} \\
&= \sum_{\mathbf{i} \in [m]_t} \mathbb{1}\{W_{i_1} = \dots = W_{i_t} \neq 0\} \\
&=: Z,
\end{aligned} \tag{53}$$

where Z can be interpreted as the number of hyperedges of another random γ^2 -multiple hyperclique model. Let $B_1, \dots, B_m \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$, where p is the probability parameter we will specify later. Define $B := \sum_{i=1}^m B_i \sim \text{Binomial}(m, p)$, and let $Z_* := \sum_{\mathbf{i} \in [m]_t} \prod_{l=1}^t B_{i_l} = \binom{B}{t}$. Let us denote $I_{\mathbf{i}} := \mathbb{1}\{W_{i_1} = \dots = W_{i_t} \neq 0\}$ and $J_{\mathbf{i}} := \prod_{l=1}^t B_{i_l}$. Then, we have

$$\mathbb{E}[Z] = \sum_{\mathbf{i} \in [m]_t} \mathbb{E}[I_{\mathbf{i}}] = \sum_{\mathbf{i} \in [m]_t} \Pr(W_{i_1} = \dots = W_{i_t} \neq 0) = \sum_{\mathbf{i} \in [m]_t} \gamma^2 \rho^t = \binom{m}{t} \gamma^2 \rho^t,$$

and

$$\mathbb{E}[Z_*] = \sum_{\mathbf{i} \in [m]_t} \mathbb{E}[J_{\mathbf{i}}] = \sum_{\mathbf{i} \in [m]_t} \Pr(B_{i_1} = \dots = B_{i_t} = 1) = \sum_{\mathbf{i} \in [m]_t} p^t = \binom{m}{t} p^t.$$

Choose $p = \gamma^{2/t} \rho$ so that $\mathbb{E}[Z] = \mathbb{E}[Z_*]$. This is a valid Bernoulli parameter since $t \geq 3$ and $\gamma k \leq m$ imply $p \leq \gamma \rho \leq k/m \leq 1$. Next, we claim that this choice of p satisfies

$$\mathbb{E}[Z^d] = \sum_{\mathbf{I} \in ([m]_t)^d} \Pr\left(\prod_{r=1}^d I_{\mathbf{i}_r} = 1\right) \leq \sum_{\mathbf{I} \in ([m]_t)^d} \Pr\left(\prod_{r=1}^d J_{\mathbf{i}_r} = 1\right) = \mathbb{E}[Z_*^d]$$

for all $d \geq 1$, where $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_d)$. We define $v = v(\mathbf{I})$ as the number of distinct vertices in $\mathbf{I}$, and $c = c(\mathbf{I})$ as the number of connected components in the hypergraph formed by the hyperedges $\mathbf{i}_1, \dots, \mathbf{i}_d$. More specifically, we call two hyperedges $\mathbf{i}_r$ and $\mathbf{i}_s$ adjacent if they share at least one vertex, and we define a connected component as a maximal collection of hyperedges linked by chains of adjacent hyperedges. Then, for $\prod_{r=1}^d I_{\mathbf{i}_r}$ to be equal to 1, all vertices in each connected component must have the same nonzero label $e_{(a,b)}$ for some $a, b \in [\gamma]$. Since there are γ^2 possible labels, we have

$$\Pr\left(\prod_{r=1}^d I_{\mathbf{i}_r} = 1\right) = \gamma^{2c} \rho^v.$$

See Equation (45) and the arguments following it for a detailed explanation of the above equation. On the other hand, for the auxiliary model, $\prod_{r=1}^d J_{\mathbf{i}_r}$ is 1 if and only if $B_i = 1$ for every vertex in the selected hyperedges, an event of probability

$$\Pr\left(\prod_{r=1}^d J_{\mathbf{i}_r} = 1\right) = p^v = \gamma^{2v/t} \rho^v.$$

Observe that every connected component contains at least one t -uniform hyperedge and therefore at least t vertices. This implies that $v \geq tc$, and thus

$$\Pr\left(\prod_{r=1}^d I_{\mathbf{i}_r} = 1\right) = \gamma^{2c} \rho^v \leq \gamma^{2v/t} \rho^v = \Pr\left(\prod_{r=1}^d J_{\mathbf{i}_r} = 1\right),$$

which proves

$$\mathbb{E}[Z^d] = \sum_{\mathbf{I} \in ([m]_t)^d} \Pr\left(\prod_{r=1}^d I_{i_r} = 1\right) \leq \sum_{\mathbf{I} \in ([m]_t)^d} \Pr\left(\prod_{r=1}^d J_{i_r} = 1\right) = \mathbb{E}[Z_*^d].$$

Consequently, we have $\mathbb{E}[|S \cap S'|^d] \leq \mathbb{E}[Z_*^d]$ for all $d \geq 1$, and thus we can bound the LDLR second moment by

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \tilde{\nu}} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \right] \leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} \mathbb{E}[Z_*^d].$$

Recall that $Z_* = \binom{B}{t} \leq B^t/t!$, and $B \sim \text{Binomial}(m, p)$ with $p = \gamma^{2/t} k^2/m^2$. Since t is a fixed constant, the problem reduces to bounding the td th moments of a binomial random variable, which can be done by applying the polynomial moment bound for binomial random variables (Ahle, 2022, Corollary 1) given in Lemma 5. The resulting bound is

$$\begin{aligned} 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} \mathbb{E}[Z_*^d] &\leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{(t!)^d d!} \mathbb{E}[B^{td}] \\ &\leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{(t!)^d d!} \left(\frac{td}{\log(1 + tdm/(\gamma^{2/t} k^2))} \right)^{td}. \end{aligned} \quad (54)$$

Let $\eta := \gamma^{2/t} k^2$, and split the analysis into two cases:

- **Case 1:** $\eta = O(m^{1-\delta})$ for some constant $\delta > 0$.

Note that

$$\begin{aligned} \sum_{d=1}^D \frac{\mu^{2d}}{(t!)^d d!} \left(\frac{td}{\log(1 + tdm/\eta)} \right)^{td} &\leq \sum_{d=1}^D \frac{t^{td} \mu^{2d}}{(t!)^d d!} \left(\frac{d}{\log(m/\eta)} \right)^{td} \\ &=: \sum_{d=1}^D T_d, \end{aligned}$$

and this is bounded as long as (i) $T_1 = O(1)$ and (ii) $T_{d+1}/T_d \leq 1/2$ for all $d \in [D-1]$. The first condition is equivalent to $\mu = O(\log^{t/2}(m/\eta))$. For the second condition, observe that the ratio between consecutive terms is

$$\begin{aligned} \frac{T_{d+1}}{T_d} &= \frac{t^{t(d+1)} \mu^{2(d+1)}}{(t!)^{d+1} (d+1)!} \left(\frac{d+1}{\log(m/\eta)} \right)^{t(d+1)} \cdot \frac{(t!)^d d!}{t^d \mu^{2d}} \left(\frac{\log(m/\eta)}{d} \right)^{td} \\ &= \frac{t^t \mu^2}{(t!)(d+1)} \left(\frac{d+1}{d} \right)^{td} \left(\frac{d+1}{\log(m/\eta)} \right)^t \\ &\leq \frac{t^t}{t!} \mu^2 e^t \frac{D^{t-1}}{\log^t(m/\eta)}, \end{aligned}$$

where the last inequality follows from the fact that $(1 + 1/d)^d \leq e$ for all $d \geq 1$. This ratio is at most $1/2$ whenever

$$\mu \leq \sqrt{\frac{t!}{2t^t e^t} \frac{\log^{t/2}(m/\eta)}{D^{(t-1)/2}}}.$$

Observe that the second condition is more stringent than the first one for any $D = \Omega(1)$. Also, since $\eta = O(m^{1-\delta})$ implies $\log(m/\eta) = \Theta(\log m)$, the stated choice $D = D_m$ makes the preceding condition hold whenever $\mu \leq C(\log m)^{\{1-(t-1)\epsilon\}/2}$ for a sufficiently small constant $C > 0$ depending on δ and t . Indeed, $\log(m/\eta) \geq (\delta/2) \log m$ eventually and $D_m \leq \log^{1+\epsilon} m$. In particular, if $\mu \leq \log^{1/2-\xi} m$ for some constant $\xi > 0$, any $\epsilon \in (0, \min\{\delta, 2\xi/(t-1)\})$ makes the second condition hold eventually.

- **Case 2:** $\eta = \Omega(m \log^{1+\delta} m)$ for some constant $\delta > 0$. Recall that $D = D_m = \lfloor \log^{1+\epsilon} m \rfloor$ with $0 < \epsilon < \delta$. Then, there exists $M > 0$ such that for all $m \geq M$ and $d \in [D]$, we have $tdm/\eta < 1$. Since $\log(1+tx) \geq x$, for all $0 < x < 1$ and $t \geq 2$, we have that, for all such $m \geq M$,

$$\log \left(1 + t \frac{dm}{\eta} \right) \geq \frac{dm}{\eta}.$$

Using this bound,

$$\begin{aligned} \sum_{d=1}^D \frac{\mu^{2d}}{(t!)^d d!} \left(\frac{td}{\log(1 + tdm/\eta)} \right)^{td} &\leq \sum_{d=1}^D \frac{t^{td} \mu^{2d}}{(t!)^d d!} \left(\frac{\eta}{m} \right)^{td} \\ &= \sum_{d=1}^D \frac{1}{d!} \left(\frac{t^t \mu^2 \eta^t}{t! m^t} \right)^d \\ &\leq \exp \left(\frac{t^t \mu^2 \eta^t}{t! m^t} \right). \end{aligned}$$

In turn, a sufficient condition for this series to remain bounded is

$$\mu = O \left(\frac{m^{t/2}}{\eta^{t/2}} \right) = O \left(\frac{m^{t/2}}{\gamma k^t} \right).$$

The upper and lower bounds yield the thresholds in Corollary 2 at polynomial order, in the sense of Definition 3. The computational conclusion is conditional on Conjecture 1.

B.8.3 Proof of multiple p -partite planted hypercliques problem

In this section, we formally define the multiple p -partite planted hypercliques problem (Remark 4) and verify that the statistical and computational thresholds are the same as those of the multiple planted hypercliques problem.

We consider a p -partite hypergraph with p disjoint vertex classes, each containing m vertices, so that the set of possible hyperedges is indexed by $[m]^p$. Let $\dot{\mathbf{A}}^{(1)}, \dots, \dot{\mathbf{A}}^{(p)}$ be independent random vectors, where each $\dot{\mathbf{A}}^{(j)} = (\dot{A}_1^{(j)}, \dots, \dot{A}_m^{(j)})$ is a uniformly random permutation of $[m]$ in the fixed-size case. For each $a \in [\gamma]$ and $j \in [p]$, define the vertex set

$$\dot{V}_a^{(j)} := \{v \in [m] : (a-1)k < \dot{A}_v^{(j)} \leq ak\}.$$

The edge set of the a th planted p -partite hyperclique is defined by

$$S_a := \dot{V}_a^{(1)} \times \dots \times \dot{V}_a^{(p)} = \{\mathbf{i} = (i_1, \dots, i_p) \in [m]^p : i_j \in \dot{V}_a^{(j)} \text{ for all } j \in [p]\}.$$

The latent structure is then given by

$$S := \bigcup_{a=1}^{\gamma} S_a.$$

For the random-size case, we replace $\dot{\mathbf{A}}^{(1)}, \dots, \dot{\mathbf{A}}^{(p)}$ with $\tilde{\mathbf{A}}^{(1)}, \dots, \tilde{\mathbf{A}}^{(p)}$, where each $\tilde{\mathbf{A}}^{(j)} = (\tilde{A}_1^{(j)}, \dots, \tilde{A}_m^{(j)})$ is a vector of i.i.d. uniform random variables on the interval $[0, m]$. We denote the induced distributions of S in the fixed-size and random-size cases by $\dot{\nu}$ and $\tilde{\nu}$, respectively. In the fixed-size case, each $\dot{V}_a^{(j)}$ has cardinality k , and hence $|S| = \gamma k^p$.

Since the proof closely follows that of the original hyperclique problem (Appendix B.8), rather than repeating the entire argument from scratch, we focus on verifying that the key arguments remain valid under the present p -partite extension. In the following, we will focus on the statistical lower bound for the fixed-size case and the computational lower bound for the random-size case, as the other cases can be handled analogously using truncation and stochastic dominance arguments similar to those in Appendix B.8.

Statistical lower bound for the fixed-size case. Let S' be an independent copy of S , and define $\dot{V}_a'^{(j)}$ and S'_a analogously. The overlap between S and S' is given by

$$|S \cap S'| = \sum_{a,b=1}^{\gamma} \prod_{j=1}^p |\dot{V}_a^{(j)} \cap \dot{V}_b'^{(j)}|.$$

Our starting point is to confirm the following inequality, which is the counterpart of Equation (30) and Equation (48):

$$\mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] \leq \mathbb{E}[F(B_1, \dots, B_p)]^{\gamma^2}, \quad (55)$$

where $F(x_1, \dots, x_p) := \exp(\mu^2 \prod_{j=1}^p x_j)$ and $B_1, \dots, B_p$ are independent binomial random variables with parameters k and $q_{\gamma,k} := k/(m - (\gamma - 1)k)$. Throughout the proof, we assume that $\gamma k \leq \kappa m$ for some constant $\kappa \in (0, 1)$, as the extension to the case $\gamma k = \Theta(m)$ follows from the covering argument in Appendix B.8.

To prove this inequality, we follow the argument in Equation (26) and the subsequent analysis. Write $X_{ab}^{(j)} := |\dot{V}_a^{(j)} \cap \dot{V}_b'^{(j)}|$ for $a, b \in [\gamma]$ and $j \in [p]$ for notational convenience.

The full array of these overlap counts need not be negatively associated, as the following calculation shows. Fix $j \in [p]$ and suppose that $\gamma \geq 2$, so that $2k \leq m$. Conditional on the unprimed vertex sets, any two vertices $u \in \dot{V}_1^{(j)}$ and $v \in \dot{V}_2^{(j)}$ are distinct, and uniform sampling of the primed disjoint blocks gives

$$\Pr\{u \in \dot{V}_1'^{(j)}, v \in \dot{V}_2'^{(j)} \mid \dot{\mathbf{A}}^{(j)}\} = \frac{k}{m} \frac{k}{m-1}.$$

Summing over these vertices gives conditional moments that do not depend on the unprimed sets, so unconditionally

$$\mathbb{E}[X_{11}^{(j)}] = \mathbb{E}[X_{22}^{(j)}] = \frac{k^2}{m}, \quad \mathbb{E}[X_{11}^{(j)} X_{22}^{(j)}] = \frac{k^4}{m(m-1)}.$$

Consequently,

$$\text{Cov}(X_{11}^{(j)}, X_{22}^{(j)}) = \frac{k^4}{m(m-1)} - \frac{k^4}{m^2} = \frac{k^4}{m^2(m-1)} > 0. \quad (56)$$

This positive covariance rules out negative association of the full array, but does not affect the conditional negative association used below.

For each $b \in [\gamma]$, let $\mathcal{F}_{b-1}$ denote the sigma-field generated by $\{\dot{V}_a^{(j)} : a \in [\gamma], j \in [p]\}$ and $\{\dot{V}_l'^{(j)} : l \in [b-1], j \in [p]\}$. Conditionally on $\mathcal{F}_{b-1}$, each $\dot{V}_b'^{(j)}$ is a uniformly random subset of size k drawn without replacement from the remaining $M_b := m - (b-1)k$ vertices. Since $\dot{V}_1^{(j)}, \dots, \dot{V}_{\gamma}^{(j)}$ are disjoint, the random variables $\{X_{ab}^{(j)}\}_{a=1}^{\gamma}$ are coordinate marginals of a multivariate hypergeometric distribution and hence are negatively associated. Moreover, the pairs of permutations used in different partitions are independent, and $\mathcal{F}_{b-1}$ records the revealed vertex sets separately within each partition. Consequently, conditional on $\mathcal{F}_{b-1}$, the p random vectors $\mathbf{X}^{(j)} := (X_{1b}^{(j)}, \dots, X_{\gamma b}^{(j)})$ for $j \in [p]$ are independent, while the coordinates within each vector are negatively associated. Based on these properties, for each $b \in [\gamma]$, we obtain

$$\begin{aligned} & \mathbb{E}\left[\prod_{a=1}^{\gamma} F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}\right] \\ & \stackrel{(a)}{=} \mathbb{E}\left[\mathbb{E}\left\{\prod_{a=1}^{\gamma} F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}\right\} \mid \mathcal{F}_{b-1}\right] \\ & \stackrel{(b)}{\leq} \mathbb{E}\left[\prod_{a=1}^{\gamma} \mathbb{E}[F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}] \mid \mathcal{F}_{b-1}\right] \\ & \stackrel{(c)}{=} \mathbb{E}\left[\mathbb{E}\left\{\prod_{a=1}^{\gamma} \mathbb{E}[F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}] \mid \mathcal{F}_{b-1}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-2)}\right\} \mid \mathcal{F}_{b-1}\right] \\ & \stackrel{(d)}{\leq} \mathbb{E}\left[\prod_{a=1}^{\gamma} \mathbb{E}[F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-2)}] \mid \mathcal{F}_{b-1}\right] \end{aligned}$$

$$\leq \dots \leq \prod_{a=1}^{\gamma} \mathbb{E}[F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}],$$

where each step is justified as follows. (a) and (c) are direct applications of the law of iterated expectations. For (b), note that the function F given $\mathcal{F}_{b-1}$ and $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}$ is $F(X_{ab}^{(1)}, \dots, X_{ab}^{(p-1)}, x) = \exp(\mu^2 x \prod_{j=1}^{p-1} X_{ab}^{(j)})$, which is a nonnegative nondecreasing function of x . Then, due to the conditional independence of $\mathbf{X}^{(p)}$ from $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}$ given $\mathcal{F}_{b-1}$, additional conditioning on $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}$ does not change the law of $\mathbf{X}^{(p)}$, and hence the negative association of $\{X_{ab}^{(p)}\}_{a=1}^{\gamma}$ gives the corresponding inequality. For (d), the argument is analogous to that for (b), except that we need to verify that taking the conditional expectation of F over its p th coordinate preserves the nonnegative and nondecreasing properties in its $(p-1)$ th coordinate. Indeed, note that

$$\begin{aligned} & \mathbb{E}[F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}] \\ &= \sum_{t=0}^k \exp\left(X_{ab}^{(p-1)} \mu^2 \prod_{j=1}^{p-2} X_{ab}^{(j)} \cdot t\right) \Pr(X_{ab}^{(p)} = t \mid \mathcal{F}_{b-1}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}) \\ &= \sum_{t=0}^k \exp\left(X_{ab}^{(p-1)} \mu^2 \prod_{j=1}^{p-2} X_{ab}^{(j)} \cdot t\right) \Pr(X_{ab}^{(p)} = t \mid \mathcal{F}_{b-1}), \end{aligned}$$

where the last equality follows from the conditional independence of $\mathbf{X}^{(p)}$ from $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-1)}$ given $\mathcal{F}_{b-1}$. Observe that $\mu^2 \prod_{j=1}^{p-2} X_{ab}^{(j)}$ is a fixed nonnegative constant when $\mathcal{F}_{b-1}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p-2)}$ are given. Therefore, the last expression is a nonnegative linear combination of nondecreasing functions of $X_{ab}^{(p-1)}$, and hence is itself a nonnegative nondecreasing function of $X_{ab}^{(p-1)}$. This verifies the required property for (d).

Next, for any $a, b \in [\gamma]$ and $j \in [p]$, from the proof of Equation (28) we obtain: for every nondecreasing convex function φ ,

$$\mathbb{E}[\varphi(X_{ab}^{(j)}) \mid \mathcal{F}_{b-1}] \leq \mathbb{E}[\varphi(B_j)]$$

almost surely. We take $B_1, \dots, B_p$ to be mutually independent and independent of all the vertex selections. For fixed nonnegative values of the other coordinates, F is an exponential function with a nonnegative coefficient in the remaining coordinate, and hence is nondecreasing and convex in that coordinate. By conditional independence across partitions, the preceding comparison can therefore be applied successively to each coordinate, yielding

$$\begin{aligned} & \mathbb{E}[F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}] \\ & \leq \mathbb{E}[F(B_1, X_{ab}^{(2)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}] \\ & \leq \dots \leq \mathbb{E}[F(B_1, \dots, B_p) \mid \mathcal{F}_{b-1}] = \mathbb{E}[F(B_1, \dots, B_p)]. \end{aligned}$$

Combining this bound with the negative association argument, we conclude that, for every $b \in [\gamma]$,

$$\mathbb{E}\left[\prod_{a=1}^{\gamma} F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}\right] \leq \prod_{a=1}^{\gamma} \mathbb{E}[F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}] \leq \mathbb{E}[F(B_1, \dots, B_p)]^{\gamma}.$$

Therefore, it follows that

$$\begin{aligned} & \mathbb{E}\left[\prod_{a=1}^{\gamma} \prod_{l=1}^b F(X_{al}^{(1)}, \dots, X_{al}^{(p)})\right] \\ &= \mathbb{E}\left[\prod_{a=1}^{\gamma} \prod_{l=1}^{b-1} F(X_{al}^{(1)}, \dots, X_{al}^{(p)}) \mathbb{E}\left[\prod_{a=1}^{\gamma} F(X_{ab}^{(1)}, \dots, X_{ab}^{(p)}) \mid \mathcal{F}_{b-1}\right]\right] \\ & \leq \mathbb{E}[F(B_1, \dots, B_p)]^{\gamma} \mathbb{E}\left[\prod_{a=1}^{\gamma} \prod_{l=1}^{b-1} F(X_{al}^{(1)}, \dots, X_{al}^{(p)})\right]. \end{aligned}$$

Iterating this inequality over $b = \gamma, \dots, 1$, with the empty product interpreted as 1, gives

$$\mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] \leq \mathbb{E}[F(B_1, \dots, B_p)]^{\gamma^2},$$

which proves the desired inequality.

Now, it remains to check whether applying Finner's inequality to the right-hand side of the above inequality yields a bound equivalent to that in Equation (50). Write $B_j = \sum_{i=1}^k Z_i^{(j)}$ for $j \in [p]$, where $\{Z_i^{(j)} : i \in [k], j \in [p]\}$ are i.i.d. Bernoulli($q_{\gamma,k}$) random variables. For each $\mathbf{i} = (i_1, \dots, i_p) \in [k]^p$, define $Z_{\mathbf{i}} := \prod_{j=1}^p Z_{i_j}^{(j)}$. Then,

$$\mathbb{E}[F(B_1, \dots, B_p)] = \mathbb{E}\left[\exp\left\{\mu^2 \prod_{j=1}^p B_j\right\}\right] = \mathbb{E}\left[\exp\left\{\mu^2 \sum_{\mathbf{i} \in [k]^p} Z_{\mathbf{i}}\right\}\right].$$

Since each $Z_i^{(j)}$ appears in exactly k^{p-1} terms in the above summation, we choose the exponents $p_{\mathbf{i}} = k^{p-1}$ for all $\mathbf{i} \in [k]^p$. Applying Finner's inequality (Lemma 6) gives

$$\begin{aligned} \mathbb{E}[F(B_1, \dots, B_p)] &\leq \prod_{\mathbf{i} \in [k]^p} \mathbb{E}[\exp\{\mu^2 k^{p-1} Z_{\mathbf{i}}\}]^{\frac{1}{k^{p-1}}} \\ &= \mathbb{E}[\exp\{\mu^2 k^{p-1} Z_{\mathbf{i}}\}]^k \\ &= \{1 + q_{\gamma,k}^p (\exp(\mu^2 k^{p-1}) - 1)\}^k \\ &\leq \exp\{k q_{\gamma,k}^p (\exp(\mu^2 k^{p-1}) - 1)\}, \end{aligned}$$

where the second equality follows from $Z_{\mathbf{i}} \sim \text{Bernoulli}(q_{\gamma,k}^p)$ and the last inequality follows from $1 + x \leq e^x$. Combining this bound with the preceding inequality and recalling that $q_{\gamma,k} = k/(m - (\gamma - 1)k)$, we obtain

$$\begin{aligned} \mathbb{E}[\exp\{\mu^2 |S \cap S'|\}] &\leq \mathbb{E}[F(B_1, \dots, B_p)]^{\gamma^2} \\ &\leq \exp\left\{\frac{\gamma^2 k^{p+1}}{(1 - \kappa)^p m^p} (\exp(\mu^2 k^{p-1}) - 1)\right\}, \end{aligned} \tag{57}$$

provided that $\gamma k \leq \kappa m$ for some fixed $\kappa \in (0, 1)$. Comparing this bound with Equation (50), the resulting statistical lower bound for the multiple p -partite planted hypercliques problem equivalent to that of the multiple planted hypercliques problem can be obtained by following the remaining steps in Appendix B.8.

Computational lower bound for the random-size case. We take an approach similar to that in Appendix B.8.2. In particular, we interpret the overlap $|S \cap S'|$ as the number of hyperedges in a random γ^2 -multiple p -partite hyperclique model and compare its moments with those of an auxiliary single p -partite hyperclique. Throughout this argument, $p \geq 2$ is fixed and $\gamma k \leq m$.

Let S' be an independent copy of S , and write $\tilde{V}_a^{(j)}$ and $\tilde{V}_b'^{(j)}$ for the corresponding vertex membership sets. For each $j \in [p]$ and $v \in [m]$, define

$$W_v^{(j)} := \begin{cases} e_{(a,b)}, & \text{if } v \in \tilde{V}_a^{(j)} \cap \tilde{V}_b'^{(j)} \text{ for some } a, b \in [\gamma], \\ 0, & \text{otherwise.} \end{cases}$$

Since the vertex assignments are independent across both vertices and partitions in the random-size model, the random variables $\{W_v^{(j)} : v \in [m], j \in [p]\}$ are independent, with

$$\Pr(W_v^{(j)} = e_{(a,b)}) = \rho := \frac{k^2}{m^2}, \quad \Pr(W_v^{(j)} = 0) = 1 - \gamma^2 \rho.$$

For a hyperedge $\mathbf{i} = (i_1, \dots, i_p) \in [m]^p$, let $I_{\mathbf{i}} := \mathbb{1}\{W_{i_1}^{(1)} = \dots = W_{i_p}^{(p)} \neq 0\}$. Then,

$$|S \cap S'| = \sum_{a,b=1}^{\gamma} \sum_{\mathbf{i} \in [m]^p} \prod_{j=1}^p \mathbb{1}\{i_j \in \tilde{V}_a^{(j)} \cap \tilde{V}_b'^{(j)}\}$$

$$\begin{aligned}
&= \sum_{\mathbf{i} \in [m]^p} \mathbb{1}\{W_{i_1}^{(1)} = \dots = W_{i_p}^{(p)} \neq 0\} \\
&= \sum_{\mathbf{i} \in [m]^p} I_{\mathbf{i}} \\
&=: Z.
\end{aligned}$$

Thus, Z can be seen as the number of hyperedges of another random γ^2 -multiple p -partite hyperclique model. For the auxiliary model, let $\{U_v^{(j)} : v \in [m], j \in [p]\}$ be i.i.d. Bernoulli(q_*) random variables, independent of all the vertex selections, where $q_* := \gamma^{2/p}\rho$. Define

$$B_j := \sum_{v=1}^m U_v^{(j)}, \quad J_{\mathbf{i}} := \prod_{j=1}^p U_{i_j}^{(j)}, \quad Z_* := \sum_{\mathbf{i} \in [m]^p} J_{\mathbf{i}} = \prod_{j=1}^p B_j.$$

Then $B_1, \dots, B_p$ are independent Binomial(m, q_*) random variables, and Z_* is the number of hyperedges in the resulting random single p -partite hyperclique.

Next, we show that $\mathbb{E}[Z^d] \leq \mathbb{E}[Z_*^d]$ for every integer $d \geq 1$. For an ordered tuple $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_d) \in ([m]^p)^d$, let $v = v(\mathbf{I})$ be the number of distinct vertices appearing in the hyperedges $\mathbf{i}_1, \dots, \mathbf{i}_d$ and let $c = c(\mathbf{I})$ be the number of connected components in the hypergraph formed by them. For $\prod_{r=1}^d I_{\mathbf{i}_r}$ to equal 1, all vertices in each connected component must share the same nonzero label. There are γ^2 possible labels for each component, and different components may use the same label. By independence of the vertex labels, it follows that

$$\Pr\left(\prod_{r=1}^d I_{\mathbf{i}_r} = 1\right) = \gamma^{2c}\rho^v.$$

On the other hand, $\prod_{r=1}^d J_{\mathbf{i}_r} = 1$ precisely when all the v distinct vertices are selected in the auxiliary model. Hence,

$$\Pr\left(\prod_{r=1}^d J_{\mathbf{i}_r} = 1\right) = q_*^v = \gamma^{2v/p}\rho^v.$$

Every connected component contains at least one p -partite hyperedge and therefore at least p vertices. Thus, $v \geq pc$, and consequently

$$\Pr\left(\prod_{r=1}^d I_{\mathbf{i}_r} = 1\right) = \gamma^{2c}\rho^v \leq \gamma^{2v/p}\rho^v = \Pr\left(\prod_{r=1}^d J_{\mathbf{i}_r} = 1\right).$$

Summing this comparison over all $\mathbf{I} \in ([m]^p)^d$, we obtain

$$\begin{aligned}
\mathbb{E}[Z^d] &= \sum_{\mathbf{I} \in ([m]^p)^d} \Pr\left(\prod_{r=1}^d I_{\mathbf{i}_r} = 1\right) \\
&\leq \sum_{\mathbf{I} \in ([m]^p)^d} \Pr\left(\prod_{r=1}^d J_{\mathbf{i}_r} = 1\right) = \mathbb{E}[Z_*^d].
\end{aligned} \tag{58}$$

We now apply this moment comparison to the low-degree likelihood ratio. By Proposition 1,

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2] &= 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} \mathbb{E}[Z^d] \\
&\leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} \mathbb{E}[Z_*^d] = 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} \prod_{j=1}^p \mathbb{E}[(B_j)^d],
\end{aligned}$$

where the last equality follows from the independence of $B_1, \dots, B_p$. Applying Lemma 5 gives $\mathbb{E}[(B_j)^d] \leq (d/\log(1 + dm/\gamma^{2/p}k^2))^d$, and consequently,

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2] \leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} \left(\frac{d}{\log(1 + dm/\gamma^{2/p}k^2)} \right)^{pd}.$$

The remaining technical steps to derive the computational lower bound from this inequality are identical to those in Equation (54) and subsequent analysis, and hence we omit the details.

B.9 Proof of Theorem 3

We first state the statistical and low-degree lower bounds to be established; the computational interpretation of the latter uses Conjecture 1.

- (Statistical lower bound) Fix any constant $\delta \in (0, 0.1)$. For either prior distribution $\dot{\nu}$ or $\tilde{\nu}$, strong detection is impossible under either of the following conditions: (i) $\ell^2 \Delta m^{-2} \leq m^{-\delta}$ and $\mu \leq C_\delta \sqrt{\Delta^{-1} \log m}$ for a sufficiently small constant $C_\delta > 0$; or (ii) $\ell^2 \Delta m^{-2} = \Omega(1)$ and $\mu = O(\ell^{-1} \Delta^{-1} m)$.
- (Computational lower bound) Fix any constants $\delta \in (0, 0.1)$ and $\epsilon \in (0, \delta)$, and let $D_m = \lfloor \log^{1+\epsilon} m \rfloor$. For either prior distribution $\dot{\nu}$ or $\tilde{\nu}$, no polynomial of degree at most D_m achieves strong separation under either of the following conditions: (i) $\ell \Delta m^{-1} \leq m^{-\delta}$ and $\mu \leq C(\log m)^{(1-\epsilon)/2}$ for a sufficiently small constant $C > 0$; or (ii) $\ell \Delta m^{-1} \geq \log^{1+\delta} m$ and $\mu = O(\ell^{-1} \Delta^{-1} m)$.

We assume that $\Delta = \omega(\log m)$ throughout the proof, which is satisfied in the polynomial regime considered here. We retain the standing assumption $2\Delta < \ell \leq m$ and take $D = D_m$ in the low-degree bounds.

Recall that our convention $[m]_2 = (i, j) : 1 \leq i < j \leq m$ represents each undirected edge by listing its endpoints in increasing order to avoid double-counting. In the proofs below, with a slight abuse of notation, let us write (i_a, i_b) to denote the undirected edge joining vertices i_a and i_b regardless of their ordering, instead of the formal representation $(\min\{i_a, i_b\}, \max\{i_a, i_b\})$.

B.9.1 Fixed-size case

We consider the case where $\dot{\mathcal{C}}$ consists of all fixed-size thick cycles of length ℓ and thickness Δ . In this case, we have $n = \binom{m}{2}$, $K = \ell \Delta$.

Upper bounds. The upper bounds on the critical signal strength μ^* follow from the tests listed in Proposition 2:

- For some constant $C > 0$, the outlier test achieves $R_{\dot{\nu}}(\Delta_{\text{out}}) \rightarrow 0$ whenever $\mu \geq C\sqrt{\log m}$.
- The average test achieves $R_{\dot{\nu}}(\Delta_{\text{avg}}) \rightarrow 0$ when $\mu = \omega(\frac{m}{\ell \Delta})$.
- For the scan test, instead of detecting the entire thick cycle, we target a clique of size $\Delta + 1$, which is a subgraph of the thick cycle. We apply the scan test with parameters $|\dot{\mathcal{C}}_{\text{clique}}| := \binom{m}{\Delta+1}$ and $K_{\text{clique}} := \binom{\Delta+1}{2}$. Then, for some constant $C > 0$, the scan test achieves $R_{\dot{\nu}}(\Delta_{\text{scan}}) \rightarrow 0$ when

$$\mu \geq C \sqrt{\frac{1}{\Delta} \log \frac{m}{\Delta}}.$$

For the outlier and clique-scan guarantees, take $\delta_m = 1/\log m$ in Proposition 2; for the average test, use the risk bound from Appendix B.2 directly.

Argument for lower bounds. For both lower bounds, we cover the thick cycle by a carefully chosen collection of overlapping cliques. Controlling their overlap allows us to apply the multiple-clique analysis with parameters $(\gamma, k) = (\ell/(2\Delta), 2\Delta)$.

In detail, for the explicit construction below, assume that $q := \ell/(2\Delta)$ is an integer. We will extend the analysis to the case where ℓ is not divisible by 2Δ at the end of this subsection. Given $\dot{A}_1, \dots, \dot{A}_m$, sort the vertices in $[m]$ as $i_1, \dots, i_m$ according to the values of $\dot{A}_1, \dots, \dot{A}_m$ in ascending order. Recalling the definition of the cycle distance $d_{\text{cyc}}^{(\ell)}(a, b) := \min\{|a - b|, \ell - |a - b|\}$, the thick cycle S is composed of edges whose endpoints have cyclic distance at most Δ along the base cycle $(i_1, i_2), \dots, (i_{\ell-1}, i_\ell), (i_\ell, i_1)$ with vertices $i_1, \dots, i_\ell$. Now consider two subsets of the vertices, both indexed by $r \in [q]$:

$$V_{1,r} = \{i_{2\Delta(r-1)+1}, \dots, i_{2\Delta r}\} \text{ for all } r \in [q],$$

and

$$V_{2,r} = \{i_{2\Delta(r-1)+\Delta+1}, \dots, i_{2\Delta r+\Delta}\} \text{ for all } r \in [q],$$

where $i_{\ell+s} := i_s$ for $s \in [\ell]$; that is, the indices are interpreted modulo ℓ . To each vertex set $V_{1,r}$ and $V_{2,r}$ associate a corresponding clique $S_{1,r} = \{(i_a, i_b) \in [m]_2 : i_a, i_b \in V_{1,r}\}$ and $S_{2,r} = \{(i_a, i_b) \in [m]_2 : i_a, i_b \in V_{2,r}\}$, respectively, for all $r \in [q]$. We then form a collection of cliques $\mathcal{S}$, which we will refer to as the patchwork of S , as

$$\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2,$$

where

$$\mathcal{S}_1 = \cup_{r=1}^q S_{1,r} \quad \text{and} \quad \mathcal{S}_2 = \cup_{r=1}^q S_{2,r}.$$

By construction, $\mathcal{S}_1$ and $\mathcal{S}_2$ each consist of q disjoint cliques of size 2Δ and the patchwork $\mathcal{S}$ covers the thick cycle S . Indeed, consider an edge $(i_a, i_b) \in S$. Orient its shorter cyclic arc clockwise, exchanging the endpoint names if necessary and interpreting their indices cyclically. If i_a and i_b are in the same vertex set, say $V_{1,r}$ for some $r \in [q]$, then (i_a, i_b) is covered by the corresponding clique. Otherwise, i_a and i_b straddle the boundary of two consecutive vertex sets, say $V_{1,r}$ and $V_{1,r+1}$ for some $r \in [q]$ (where $V_{1,q+1} := V_{1,1}$). Since $d_{\text{cyc}}^{(\ell)}(a, b) \leq \Delta$, the two vertices are at most Δ positions apart along the base cycle. Therefore, this forces i_a to be in the last Δ vertices of $V_{1,r}$ and i_b to be in the first Δ vertices of $V_{1,r+1}$. In this case, both vertices i_a and i_b are contained in the vertex set $V_{2,r}$, and thus (i_a, i_b) is covered by the corresponding clique $S_{2,r}$.

Now, let S and S' be independent copies, each uniformly distributed over the set of length- ℓ , Δ -thick cycles as defined in Equation (10), with corresponding patchworks $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ and $\mathcal{S}' = \mathcal{S}'_1 \cup \mathcal{S}'_2$, respectively. Then, the overlap size $|S \cap S'|$ can be bounded as

$$\begin{aligned} |S \cap S'| &\leq |(\mathcal{S}_1 \cup \mathcal{S}_2) \cap (\mathcal{S}'_1 \cup \mathcal{S}'_2)| \\ &\leq |\mathcal{S}_1 \cap \mathcal{S}'_1| + |\mathcal{S}_1 \cap \mathcal{S}'_2| + |\mathcal{S}_2 \cap \mathcal{S}'_1| + |\mathcal{S}_2 \cap \mathcal{S}'_2|. \end{aligned} \tag{59}$$

We remark that $\mathcal{S}_1$ and $\mathcal{S}_2$ are not independent, and neither are $\mathcal{S}'_1$ and $\mathcal{S}'_2$. Instead, each of them is marginally distributed as a uniform distribution over the set of all unions of q disjoint cliques of size 2Δ , and $\mathcal{S}_1$ and $\mathcal{S}_2$ are independent of $\mathcal{S}'_1$ and $\mathcal{S}'_2$. Based on this observation, we analyze the statistical and computational thresholds of the thick cycle problem by bounding $|\mathcal{S}_i \cap \mathcal{S}'_j|$ for $i, j \in \{1, 2\}$ using the results from the multiple planted cliques problem.

Statistical lower bound. We have that

$$\begin{aligned}
\mathbb{E}[\exp(\mu^2 |S \cap S'|)] &\stackrel{(a)}{\leq} \mathbb{E}\left[\exp\left(\mu^2 \left\{\sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j|\right\}\right)\right] \\
&= \mathbb{E}\left[\prod_{i,j=1}^2 \exp(\mu^2 |\mathcal{S}_i \cap \mathcal{S}'_j|)\right], \\
&\stackrel{(b)}{\leq} \prod_{i,j=1}^2 \left(\mathbb{E}[\exp(4\mu^2 |\mathcal{S}_i \cap \mathcal{S}'_j|)]\right)^{1/4} \\
&\stackrel{(c)}{=} \mathbb{E}[\exp(4\mu^2 |\mathcal{S}_1 \cap \mathcal{S}'_1|)],
\end{aligned} \tag{60}$$

where (a) follows from Equation (59), (b) from Hölder's inequality, and (c) from the fact that the overlaps $|\mathcal{S}_i \cap \mathcal{S}'_j|$ are identically distributed for all $i, j \in \{1, 2\}$. We leverage our analysis of the multiple cliques problem. For ease of analysis, let us assume that $\ell \leq \kappa m$ for some constant $\kappa \in (0, 1)$. We will relax this restriction later. Then, we refer to Equation (30) and Equation (31) and obtain, with $B \sim \text{Binomial}(2\Delta, p_{q,2\Delta})$ where $p_{q,2\Delta} := 2\Delta/(m - (q-1)2\Delta)$,

$$\begin{aligned}
\mathbb{E}[\exp(\mu^2 |S \cap S'|)] &\leq \mathbb{E}\left[\exp\left\{4\mu^2 \binom{B}{2}\right\}\right]^{q^2} \\
&\leq \exp\left\{\frac{q^2(2\Delta)^3}{2(1-\kappa)^2 m^2} (\exp((2\mu)^2(2\Delta)) - 1)\right\} \\
&\leq \exp\left\{\frac{\ell^2 \Delta}{(1-\kappa)^2 m^2} (\exp(8\mu^2 \Delta) - 1)\right\},
\end{aligned} \tag{61}$$

provided that $m - (q-1)2\Delta \geq (1-\kappa)m$. Therefore, $\mathbb{E}[\exp(\mu^2 |S \cap S'|)] \leq e^C$ for some $C > 0$, as long as

$$\mu \leq \sqrt{\frac{1}{8\Delta} \log\left(1 + C \frac{(1-\kappa)^2 m^2}{\ell^2 \Delta}\right)}.$$

In the regime $\ell^2 \Delta = \omega(m^2)$, we have $C \frac{(1-\kappa)^2 m^2}{2\ell^2 \Delta} \leq \log(1 + C \frac{(1-\kappa)^2 m^2}{\ell^2 \Delta})$ for all sufficiently large m . Therefore, the condition in the above equation holds as long as $\mu \leq \frac{\sqrt{C}}{4} \frac{(1-\kappa)m}{\ell \Delta}$. The same rate holds under the stated assumption $\ell^2 \Delta/m^2 = \Omega(1)$. Indeed, if $\ell^2 \Delta/m^2 \geq c_0 > 0$ and $\mu \leq C_0 m/(\ell \Delta)$ eventually, then $8\mu^2 \Delta \leq 8C_0^2/c_0$. Using $e^x - 1 \leq xe^x$ for $x \geq 0$ gives

$$\frac{\ell^2 \Delta}{(1-\kappa)^2 m^2} (\exp(8\mu^2 \Delta) - 1) \leq \frac{8C_0^2}{(1-\kappa)^2} \exp(8C_0^2/c_0) = O(1).$$

On the other hand, in the regime $\ell^2 \Delta = O(m^{2-\delta})$ for any $\delta > 0$, since $\log(1 + C \frac{(1-\kappa)^2 m^2}{\ell^2 \Delta}) \geq \delta \log m + O(1)$, we have $\log(1 + C \frac{(1-\kappa)^2 m^2}{\ell^2 \Delta}) \geq \frac{\delta}{2} \log m$ for sufficiently large m . Therefore, it suffices to have $\mu \leq \sqrt{\frac{\delta}{16\Delta} \log m}$.

For the general case $\ell \leq m$, we apply the grouping argument on $\mathcal{S}_1$ and $\mathcal{S}'_1$, as in Appendix B.7.1. When $q = 1$, $\mathcal{S}_1$ consists of a single clique of size 2Δ , so the statistical and computational moment bounds follow directly from the corresponding single-clique analyses. Hence, we assume $q \geq 2$ in the grouping argument below. In particular, note that $\mathcal{S}_1$ consists of q cliques of size 2Δ , i.e., $\mathcal{S}_1 = \bigcup_{r=1}^q \mathcal{S}_{1,r}$. Let $h = \lceil q/2 \rceil$. Define

$$\mathcal{T}_1 := \bigcup_{r=1}^h \mathcal{S}_{1,r}, \quad \mathcal{T}_2 := \bigcup_{r=q-h+1}^q \mathcal{S}_{1,r}, \tag{62}$$

and define $\mathcal{T}'_1$ and $\mathcal{T}'_2$ similarly. This construction is analogous to that in Equation (33). Since $h \leq 2q/3$, each grouped model satisfies $h(2\Delta) \leq 2m/3$. Therefore, following the subsequent analysis in Equation (34) and

Equation (35), we obtain

$$\begin{aligned}
\mathbb{E}_{\mathcal{S}_1, \mathcal{S}'_1 \sim \dot{\nu}_{q, 2\Delta}} [\exp(4\mu^2 |\mathcal{S}_1 \cap \mathcal{S}'_1|)] &\leq \mathbb{E}_{\mathcal{T}_1, \mathcal{T}'_1 \sim \dot{\nu}_{h, 2\Delta}} [\exp(16\mu^2 |\mathcal{T}_1 \cap \mathcal{T}'_1|)] \\
&\leq \exp \left\{ \frac{2q^2(2\Delta)^3}{m^2} (\exp(16\mu^2(2\Delta)) - 1) \right\} \\
&= \exp \left\{ \frac{4\ell^2\Delta}{m^2} (\exp(32\mu^2\Delta) - 1) \right\}.
\end{aligned} \tag{63}$$

The resulting bounds on μ are equivalent to those obtained in the previous analysis up to constant factors, and establish the statistical lower bound when 2Δ divides ℓ .

Computational lower bound. Observe that

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\dot{\nu}}^{\leq D}(\mathbf{Y})^2] &= 1 + \mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} |\mathcal{S} \cap \mathcal{S}'|^d \right] \\
&\stackrel{(a)}{\leq} 1 + \mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} \left\{ \sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j| \right\}^d \right] \\
&\stackrel{(b)}{\leq} 1 + \mathbb{E} \left[\sum_{d=1}^D \frac{4^{d-1} \mu^{2d}}{d!} \sum_{i,j=1}^2 |\mathcal{S}_i \cap \mathcal{S}'_j|^d \right] \\
&\stackrel{(c)}{=} 1 + \mathbb{E} \left[\sum_{d=1}^D \frac{4^d \mu^{2d}}{d!} |\mathcal{S}_1 \cap \mathcal{S}'_1|^d \right],
\end{aligned} \tag{64}$$

where (a) follows from Equation (59), (b) follows from the fact that $(\sum_{i=1}^4 x_i)^d \leq 4^{d-1} \sum_{i=1}^4 x_i^d$ for $x_i \geq 0$, and (c) follows from the fact that the overlaps $|\mathcal{S}_i \cap \mathcal{S}'_j|$ are identically distributed for all $i, j \in \{1, 2\}$. Suppose that $\ell \leq \kappa m$ for some constant $\kappa \in (0, 1)$, and notice that the last term can be viewed as the second moment of the low-degree likelihood ratio of the fixed-size multiple cliques problem, which is Equation (36), with 2μ instead of μ , and with parameters $k = 2\Delta$ and $\gamma = q = \ell/(2\Delta)$. As shown in Equation (37), the fixed-size bound transfers from the random-size model after inflating the clique-size parameter by a fixed factor. Thus, up to a fixed multiplicative constant, the analysis reduces to a quantity of the form on the right-hand side of Equation (46), with $k = C(2\Delta)$, $\gamma = q = \ell/(2\Delta)$, and signal strength 2μ , where $C \in (1, 1/\kappa)$ is fixed. These substitutions change only constant factors in the sufficient signal conditions, and the analysis in Appendix B.7.2 applies. We omit the details.

Similar to the statistical lower bound, for the general case $\ell \leq m$, we apply the grouping argument in Equation (62). Following Equation (38) and the subsequent analysis therein, we obtain a bound of the same order as in the preceding analysis.

Handling the case where ℓ is not divisible by 2Δ . Now, we extend the above analysis to the case where ℓ is not divisible by 2Δ . Let $t := \ell - 2\Delta q$ be the remainder, where $q := \lfloor \ell/(2\Delta) \rfloor$ is the quotient. If we construct $\mathcal{S}_1$ and $\mathcal{S}_2$ as before, their union may fail to cover some edges of $\mathcal{S}$ incident to the last t vertices. Our main idea is to augment the patchwork by introducing two additional coverings, $\mathcal{S}_3$ and $\mathcal{S}_4$, constructed analogously to $\mathcal{S}_1$ and $\mathcal{S}_2$ but rotated to cover the remaining edges, so that

$$\mathcal{S} \subseteq \mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4.$$

Specifically, for each $r \in [q]$, recall the definition of $V_{1,r}$ and $V_{2,r}$ and additionally define $V_{3,r}$ and $V_{4,r}$ as

$$\begin{aligned}
V_{1,r} &= \{i_{2\Delta(r-1)+1}, \dots, i_{2\Delta r}\}, \\
V_{2,r} &= \{i_{2\Delta(r-1)+\Delta+1}, \dots, i_{2\Delta r+\Delta}\}, \\
V_{3,r} &= \{i_{2\Delta q+2\Delta(r-1)+1}, \dots, i_{2\Delta q+2\Delta r}\},
\end{aligned}$$

$$V_{4,r} = \{i_{2\Delta q+2\Delta(r-1)+\Delta+1}, \dots, i_{2\Delta q+2\Delta r+\Delta}\},$$

where the indices are interpreted modulo ℓ , as before. For $a \in [4]$ and $r \in [q]$, define the corresponding clique

$$S_{a,r} := \{(u, v) \in [m]_2 : u, v \in V_{a,r}\},$$

and let $\mathcal{S}_a := \bigcup_{r=1}^q S_{a,r}$ for $a \in [4]$. By construction, each $\mathcal{S}_a$, $a \in [4]$, consists of q disjoint cliques of size 2Δ . Also, we claim that $\mathcal{S}$ covers the thick cycle S . To see this, consider an arbitrary edge $(i_a, i_b) \in S$. Without loss of generality, assume that i_a is the starting vertex, in the sense that the clockwise cyclic distance from a to b is at most Δ . If a lies among the first $2\Delta q$ vertices, following the same argument as before, (i_a, i_b) is covered by $\mathcal{S}_1 \cup \mathcal{S}_2$, as in the previous analysis. Otherwise, if a lies between $2\Delta q + 1$ and ℓ , then i_a should be in either $V_{3,1}$ or $V_{4,1}$, and again following the same argument as before, (i_a, i_b) is covered by $\mathcal{S}_3 \cup \mathcal{S}_4$. Thus, the resulting patchwork $\mathcal{S} = \bigcup_{a=1}^4 \mathcal{S}_a$ covers the thick cycle S .

Now, let S' be an independent copy of S , with corresponding patchwork $\mathcal{S}' = \bigcup_{a=1}^4 \mathcal{S}'_a$. Then, we have

$$|S \cap S'| \leq \sum_{a,b=1}^4 |\mathcal{S}_a \cap \mathcal{S}'_b|.$$

Observe that each $\mathcal{S}_a$ is marginally distributed as a uniform union of q disjoint cliques of size 2Δ , and $\mathcal{S}_a$ and $\mathcal{S}'_b$ are independent for all $a, b \in [4]$. Therefore, the overlaps $|\mathcal{S}_a \cap \mathcal{S}'_b|$ are identically distributed for all $a, b \in [4]$.

For the statistical lower bound, following the same argument as before, we obtain

$$\mathbb{E} \left[\exp \left(\mu^2 \sum_{a,b=1}^4 |\mathcal{S}_a \cap \mathcal{S}'_b| \right) \right] \stackrel{(a)}{\leq} \prod_{a,b=1}^4 \left(\mathbb{E} [\exp (16\mu^2 |\mathcal{S}_a \cap \mathcal{S}'_b|)] \right)^{1/16} \stackrel{(b)}{=} \mathbb{E} [\exp (16\mu^2 |\mathcal{S}_1 \cap \mathcal{S}'_1|)],$$

where (a) follows from Hölder's inequality and (b) from the fact that the overlaps $|\mathcal{S}_a \cap \mathcal{S}'_b|$ are identically distributed for all pairs $(a, b) \in [4]^2$. Similarly, for the computational lower bound, we have

$$\mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} \left\{ \sum_{a,b=1}^4 |\mathcal{S}_a \cap \mathcal{S}'_b|^d \right\} \right] \stackrel{(a)}{\leq} \mathbb{E} \left[\sum_{d=1}^D \frac{16^{d-1} \mu^{2d}}{d!} \sum_{a,b=1}^4 |\mathcal{S}_a \cap \mathcal{S}'_b|^d \right] \stackrel{(b)}{=} \mathbb{E} \left[\sum_{d=1}^D \frac{16^d \mu^{2d}}{d!} |\mathcal{S}_1 \cap \mathcal{S}'_1|^d \right],$$

where (a) follows by applying the inequality $(\sum_{i=1}^n x_i)^d \leq n^{d-1} \sum_{i=1}^n x_i^d$ for $x_i \geq 0$, $n \in \mathbb{N}$ and $d \geq 1$, and (b) follows from the identical distributions of the overlaps.

The preceding statistical and computational lower-bound analyses following Equation (60) and Equation (64) can then be applied to the new bounds, with only a change in the constant factors. This completes the extension to general ℓ , which may not be divisible by 2Δ .

B.9.2 Random-size case

Upper bounds. We introduce an event on which the number of active vertices in the random-size thick cycle is at least $c_1 \ell$, and each active vertex has at least $c_2 \Delta$ active neighbors in each direction, where $c_1, c_2 \in (0, 1)$ are fixed constants. Formally, we let $V := \{i \in [m] : \tilde{A}_i \in (0, \ell]\}$ be the set of active vertices; then the number of active vertices satisfies $|V| \sim \text{Binomial}(m, \ell/m)$. For each $i \in V$, we define

$$N_i^+ := \sum_{j \neq i} \mathbf{1}\{\tilde{A}_j \in (0, \ell], d_{\text{cyc}}^{(\ell),+}(\tilde{A}_i, \tilde{A}_j) \leq \Delta\},$$

$$N_i^- := \sum_{j \neq i} \mathbf{1}\{\tilde{A}_j \in (0, \ell], d_{\text{cyc}}^{(\ell),-}(\tilde{A}_i, \tilde{A}_j) \leq \Delta\},$$

where $d_{\text{cyc}}^{(\ell),+}(a, b) := b - a \pmod{\ell}$ and $d_{\text{cyc}}^{(\ell),-}(a, b) := a - b \pmod{\ell}$ are the clockwise and counterclockwise distances on a circle of circumference ℓ between a and b , respectively. We define an event

$$\mathcal{E} := \left\{ |V| \geq c_1 \ell, \min_{i \in V} N_i^+ \geq c_2 \Delta, \min_{i \in V} N_i^- \geq c_2 \Delta \right\},$$

where $1 > c_1, c_2 > 0$ are constants. Then, observe that the Chernoff bound guarantees the existence of a constant $c_3 > 0$ such that

$$\Pr(|V| \geq c_1 \ell) \geq 1 - \exp(-c_3 \ell),$$

Conditional on $i \in V$ and its latent location, each of N_i^+ and N_i^- has distribution $\text{Binomial}(m-1, \Delta/m)$. Hence, for some fixed $c_4 > 0$ and all sufficiently large m ,

$$\Pr(N_i^+ \geq c_2 \Delta \mid i \in V) \geq 1 - \exp(-c_4 \Delta), \quad \Pr(N_i^- \geq c_2 \Delta \mid i \in V) \geq 1 - \exp(-c_4 \Delta).$$

We apply the union bound to obtain

$$\begin{aligned} \Pr(\exists i \in V : N_i^+ < c_2 \Delta) &\leq \sum_{i=1}^m \Pr(i \in V, N_i^+ < c_2 \Delta) \\ &\leq m \frac{\ell}{m} \exp(-c_4 \Delta) = \ell \exp(-c_4 \Delta), \end{aligned}$$

and similarly, $\Pr(\exists i \in V : N_i^- < c_2 \Delta) \leq \ell \exp(-c_4 \Delta)$. Therefore, we apply the union bound again to obtain

$$\Pr(\mathcal{E}^c) \leq \exp(-c_3 \ell) + 2\ell \exp(-c_4 \Delta) = o(1),$$

where the last equality follows from $\Delta = \omega(\log m)$. Now, for an event $\mathcal{E}$ defined in terms of the auxiliary latent vector $\tilde{\mathbf{A}}$, define the induced conditional distribution of S by

$$\tilde{\nu}(B) := \frac{\Pr(S(\tilde{\mathbf{A}}) \in B, \mathcal{E})}{\Pr(\mathcal{E})}, \quad B \subseteq \mathcal{C}.$$

The total variation distance between $\tilde{\nu}$ and $\tilde{\nu} := \tilde{\nu}(\cdot \mid \mathcal{E})$ vanishes, as does the risk difference $|R_{\tilde{\nu}}(\Delta) - R_{\tilde{\nu}}(\Delta)|$ for any test Δ . To apply the tests, set $h := \lfloor c_2 \Delta \rfloor$. On $\mathcal{E}$, each active vertex has at least h neighbors in each of two disjoint oriented arcs, so $|S| \geq |V|h \geq c_1 \ell h$. Also, any active vertex and the next h active vertices clockwise form a clique: they all lie in a clockwise arc of length at most Δ . Thus, the average test with threshold $\mu c_1 \ell h/2$, the scan over $(h+1)$ -vertex cliques, and the outlier test give the same upper bounds as in the fixed-size case, up to constant factors.

Truncation argument for lower bounds. For the lower bounds, we reduce the analysis of the random-size thick cycle case to the fixed-size thick cycle case via a truncation argument similar to that in Appendix B.6. Specifically, let V be the set of active vertices in the random-size thick cycle, i.e., $V := \{i \in [m] : \tilde{A}_i \in (0, \ell]\}$. Then, observe that $|V| \sim \text{Binomial}(m, \ell/m)$, and thus $\mathbb{E}[|V|] = \ell$. Then, for any constant $C_1 > 1$, the Chernoff bound guarantees the existence of a constant $c_1 > 0$ such that

$$\Pr(|V| \geq C_1 \ell) \leq \exp(-c_1 \ell) = o(1), \tag{65}$$

where c_1 only depends on C_1 . Note that if ℓ is large, then $C_1 \ell$ may exceed m . Therefore, for later use, let $r_{\max} := \min\{\lfloor C_1 \ell \rfloor, m\}$.

For each $i \in V$, let N_i be the number of other active vertices within cyclic distance Δ of i , that is,

$$N_i := \sum_{j \neq i} \mathbb{1}\{\tilde{A}_j \in (0, \ell], d_{\text{cyc}}^{(\ell)}(\tilde{A}_i, \tilde{A}_j) \leq \Delta\}, \quad i \in V,$$

which is the degree of vertex i in the thick cycle S . Since the total length of the two oriented arcs within distance Δ from i is 2Δ , we have $N_i \mid \{i \in V\} \sim \text{Binomial}(m-1, 2\Delta/m)$, and thus $\mathbb{E}[N_i \mid i \in V] = (m-1)2\Delta/m$. By the Chernoff bound, there exists a constant $c_2 > 0$ such that

$$\Pr(N_i \geq 2C_2 \Delta \mid i \in V) \leq \exp(-c_2 \Delta),$$

where $C_2 > 1$ is a constant that can be chosen arbitrarily. Taking a union bound over the m possible vertex labels gives

$$\Pr\left(\max_{i \in V} N_i \geq 2C_2 \Delta\right) \leq m \exp(-c_2 \Delta) = o(1), \tag{66}$$

where the last equality follows from $\Delta = \omega(\log m)$. Define the event

$$\mathcal{E} := \left\{ 1 \leq |V| \leq r_{\max}, \max_{i \in V} N_i \leq 2C_2\Delta \right\}.$$

The additional nonemptiness condition costs at most $\Pr(|V| = 0) = (1 - \ell/m)^m \leq e^{-\ell} = o(1)$. Then, $\Pr(\mathcal{E}) = 1 - o(1)$ by Equation (65) and Equation (66). Define the conditional distribution of S by

$$\tilde{\nu}(B) := \frac{\Pr(S(\tilde{\mathbf{A}}) \in B, \mathcal{E})}{\Pr(\mathcal{E})}, \quad B \subseteq \mathcal{C}.$$

Following the same truncation argument as in Appendix B.6, the lower-bound analysis under the truncated mixing measure $\tilde{\nu}$ is sufficient to establish the same lower bounds for the original mixing measure $\tilde{\nu}$.

We further refine the event $\mathcal{E}$ by conditioning on the number of active vertices $|V| = r$ for $r \in [r_{\max}]$. All sums and suprema over these indices below are understood to include only conditional laws with positive mixing weight. For $r \in [r_{\max}]$, consider the event $\mathcal{E}_r := \{|V| = r, \max_{i \in V} N_i \leq 2C_2\Delta\}$ and let $\tilde{\nu}_r(\cdot) := \tilde{\nu}(\cdot | \mathcal{E}_r)$ be the corresponding conditional mixing measure given $\mathcal{E}_r$. Also, let $S_r \sim \tilde{\nu}_r$. Given a conditional latent realization with edge set s_r , denote the active vertices as $v_1, \dots, v_r \in [m]$, arranged according to their latent locations on a circle of circumference ℓ ; that is, $0 < \tilde{A}_{v_1} < \dots < \tilde{A}_{v_r} \leq \ell$. Write $b_* := \lfloor 2C_2\Delta \rfloor$ for the integer covering bandwidth. As a covering of s_r , we consider the cyclic band graph

$$\dot{s}_r := \{(v_a, v_b) \in [m]_2 : d_{\text{cyc}}^{(r)}(a, b) \leq 2C_2\Delta\},$$

where $d_{\text{cyc}}^{(r)}(a, b) := \min\{|a - b|, r - |a - b|\}$ is the cyclic distance on a circle of circumference r . We claim that $s_r \subseteq \dot{s}_r$. Indeed, for any edge $(v_a, v_b) \in s_r$, we have $d_{\text{cyc}}^{(\ell)}(\tilde{A}_{v_a}, \tilde{A}_{v_b}) \leq \Delta$ by definition. Choose one of the two oriented arcs connecting $\tilde{A}_{v_a}$ and $\tilde{A}_{v_b}$ whose length is at most Δ . Without loss of generality, assume that this arc from $\tilde{A}_{v_a}$ to $\tilde{A}_{v_b}$ is in the clockwise direction. If R is the number of cyclic rank steps along this oriented arc, then $d_{\text{cyc}}^{(r)}(a, b) \leq R$; the opposite arc may have fewer rank steps. Then, every active vertex along this arc after v_a including v_b must be within distance Δ from v_a , and thus $R \leq N_{v_a}$. Under the event $\mathcal{E}_r$, we have $N_{v_a} \leq 2C_2\Delta$, and thus $d_{\text{cyc}}^{(r)}(a, b) \leq 2C_2\Delta$.

If $r > 2b_* + 1$, this cover is a fixed-size thick cycle with length r and thickness b_* ; if $r \leq 2b_* + 1$, it is the complete graph on the r active vertices. Applying the construction to the conditional latent variables defines the random cover $\dot{S}_r$. Then, we have $S_r \subseteq \dot{S}_r$ almost surely. It remains to verify that this cover has the uniform fixed-size law (or the uniform r -vertex clique law in the latter case). To see this, it suffices to observe that conditionally on the event $\mathcal{E}_r$, the ordered active vertices $v_1, \dots, v_r$ are uniformly distributed over all possible ordered r -tuples of distinct vertices in $[m]$. Consider any two ordered r -tuples of distinct vertices $(i_1, \dots, i_r)$ and $(i'_1, \dots, i'_r)$ in $[m]$. Then, there exists a permutation π on $[m]$ such that $\pi(i_j) = i'_j$ for all $j \in [r]$. Since the latent variables $\tilde{A}_1, \dots, \tilde{A}_m$ are exchangeable and the event $\mathcal{E}_r$ is invariant under vertex labeling, relabeling by π gives

$$\Pr((V_1, \dots, V_r) = (i_1, \dots, i_r)) = \Pr((V_1, \dots, V_r) = (i'_1, \dots, i'_r)),$$

where $V_1, \dots, V_r$ are the random indices defined by the ordered active vertices under the conditional latent law corresponding to $S_r \sim \tilde{\nu}_r$. Lastly, observe that $\dot{S}_r$ is a deterministic function of $(V_1, \dots, V_r)$, so it has the claimed uniform law.

Statistical lower bound. Let $S_r, S'_r \sim \tilde{\nu}_r$ be independent, and let $\dot{S}_r, \dot{S}'_r$ be their corresponding covering fixed-size thick cycles. Then, $|S_r \cap S'_r| \leq |\dot{S}_r \cap \dot{S}'_r|$ almost surely, which implies

$$\mathbb{E}[\exp(\mu^2 |S_r \cap S'_r|)] \leq \mathbb{E}[\exp(\mu^2 |\dot{S}_r \cap \dot{S}'_r|)].$$

We next aggregate the above bound over $r \in [r_{\max}]$. Observe that the truncated mixing measure $\tilde{\nu}$ can be expressed as a mixture of the conditional mixing measures $\tilde{\nu}_r$, i.e., $\tilde{\nu}(\cdot) = \sum_{r=1}^{r_{\max}} \Pr(|V| = r | \mathcal{E}) \tilde{\nu}_r(\cdot)$. Let $S, S' \sim \tilde{\nu}$ be independent, and let $S_r, S'_r \sim \tilde{\nu}_r$ be independent for each $r \in [r_{\max}]$. Then,

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\tilde{\nu}}(\mathbf{Y})^2] = \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} \left[\left(\sum_{r=1}^{r_{\max}} \Pr(|V| = r | \mathcal{E}) L_{\tilde{\nu}_r}(\mathbf{Y}) \right)^2 \right]$$

$$\begin{aligned}
&= \sum_{r=1}^{r_{\max}} \sum_{r'=1}^{r_{\max}} \Pr(|V| = r \mid \mathcal{E}) \Pr(|V| = r' \mid \mathcal{E}) \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\bar{v}_r}(\mathbf{Y}) L_{\bar{v}_{r'}}(\mathbf{Y})] \\
&\stackrel{(\dagger)}{\leq} \sum_{r=1}^{r_{\max}} \sum_{r'=1}^{r_{\max}} \Pr(|V| = r \mid \mathcal{E}) \Pr(|V| = r' \mid \mathcal{E}) \sqrt{\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\bar{v}_r}(\mathbf{Y})^2] \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\bar{v}_{r'}}(\mathbf{Y})^2]} \\
&\leq \sum_{r=1}^{r_{\max}} \sum_{r'=1}^{r_{\max}} \Pr(|V| = r \mid \mathcal{E}) \Pr(|V| = r' \mid \mathcal{E}) \sup_{r \in [r_{\max}]} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\bar{v}_r}(\mathbf{Y})^2] \\
&= \sup_{r \in [r_{\max}]} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\bar{v}_r}(\mathbf{Y})^2] \\
&= \sup_{r \in [r_{\max}]} \mathbb{E}[\exp(\mu^2 |S_r \cap S'_r|)] \\
&\leq \sup_{r \in [r_{\max}]} \mathbb{E}[\exp(\mu^2 |\dot{S}_r \cap \dot{S}'_r|)]. \tag{67}
\end{aligned}$$

where $(\dagger)$ follows from the Cauchy–Schwarz inequality. Therefore, $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\bar{v}}(\mathbf{Y})^2]$ is bounded as long as $\mathbb{E}[\exp(\mu^2 |\dot{S}_r \cap \dot{S}'_r|)]$ is uniformly bounded in $r \in [r_{\max}]$. For $r > 2b_* + 1$, apply the four-component extension of Equation (63) with length r and thickness $b_* \leq 2C_2\Delta$. The Hölder factor increases from 4 to 16, so the exponential constant 32 in that bound becomes 128. Here $q = \lfloor r/(2b_*) \rfloor$ satisfies $2b_*q \leq r$; the single-clique case $q = 1$ is bounded by the same expression. For $2 \leq r \leq 2b_* + 1$, the single-clique bound in Equation (18) instead gives an upper bound $\exp\{r^3(\exp(\mu^2 r) - 1)/(2m^2)\}$, which is no larger than the bound below since $r \leq 4C_2\Delta + 1 \leq 5C_2\Delta$. For $r = 1$, the overlap is zero. Thus, in each case we obtain the uniform estimate

$$\mathbb{E}[\exp(\mu^2 |\dot{S}_r \cap \dot{S}'_r|)] \leq \exp\left(\frac{8C_2r^2\Delta}{m^2} (\exp(256\mu^2 C_2\Delta) - 1)\right).$$

Since the displayed upper bound is increasing in r , it is bounded above by its value at $r = r_{\max}$. Therefore, we have

$$\sup_{r \in [r_{\max}]} \mathbb{E}[\exp(\mu^2 |\dot{S}_r \cap \dot{S}'_r|)] \leq \exp\left(\frac{8C_1^2 C_2 \ell^2 \Delta}{m^2} (\exp(256C_2\mu^2 \Delta) - 1)\right),$$

since $r_{\max} \leq C_1\ell$. The analysis of this bound only differs from the analysis we conducted in the fixed-size case by constant factors. Thus, without any divisibility assumption, the statistical lower bound for the random-size thick cycle problem is identical to the fixed-size one.

Computational lower bound. Note that a decomposition analogous to that in Equation (67) can be applied to the low-degree likelihood ratio, which yields

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\bar{v}}^{\leq D}(\mathbf{Y})^2] &\leq \sup_{r \in [r_{\max}]} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} [L_{\bar{v}_r}^{\leq D}(\mathbf{Y})^2] \\
&= 1 + \sup_{r \in [r_{\max}]} \mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} |S_r \cap S'_r|^d \right] \\
&\leq 1 + \sup_{r \in [r_{\max}]} \mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} |\dot{S}_r \cap \dot{S}'_r|^d \right].
\end{aligned}$$

We apply the four-component extension of Equation (64) from the fixed-size analysis, allowing every conditional length r , to obtain

$$\sup_{r \in [r_{\max}]} \mathbb{E} \left[\sum_{d=1}^D \frac{\mu^{2d}}{d!} |\dot{S}_r \cap \dot{S}'_r|^d \right] \leq \sup_{r \in [r_{\max}]} \mathbb{E} \left[\sum_{d=1}^D \frac{16^d \mu^{2d}}{d!} |\mathcal{S}_{1,r} \cap \mathcal{S}'_{1,r}|^d \right],$$

where $\mathcal{S}_{1,r}$ and $\mathcal{S}'_{1,r}$ are the first patchwork components of $\dot{S}_r$ and $\dot{S}'_r$, respectively. For $r \leq 2b_* + 1$, take these components to be the covering cliques themselves, so the displayed inequality still holds. For the

four-component construction with $r > 2b_* + 1$, each component is a union of $q = \lfloor r/(2b_*) \rfloor$ disjoint cliques of size $k = 2b_*$, and hence has the fixed-size multiple-clique law with parameters $(\gamma, k) = (q, 2b_*)$. For fixed k , coupling these models using the first q blocks of a common random permutation shows that the overlap moments are nondecreasing in the integer q . Also, $qk^2 \leq 2b_*r \leq 4C_1C_2\ell\Delta$, so the multiple-clique moment bound is controlled uniformly by the same binomial series as in the fixed-size analysis, with its overlap parameter changed only by a constant factor. For the covering cliques when $r \leq 2b_* + 1$, use Equation (19); their corresponding parameter satisfies $r^2 = O(\Delta^2) = O(\ell\Delta)$, giving the same uniform control. The four-component construction therefore gives the same computational lower-bound rates for all conditional lengths r , without any divisibility assumption.

B.10 Proof of Corollary 3

In general, the formal upper and lower bounds we aim to establish are the same as those for the thick cycle problem; see Appendix B.9. In this section, we show how the thick path problem reduces to the thick cycle problem. As in Appendix B.9, assume $2\Delta < \ell \leq m$ and $\Delta = \omega(\log m)$ throughout.

Upper bounds. We note that the analysis of the upper bounds for the thick path problem is nearly identical to that of the thick cycle problem in Appendix B.9, with minor modifications due to the absence of wrap-around edges. We address such modifications in the following.

For the fixed-size case, the outlier and clique-scan tests apply without change. For the average test, the number of edges is $\ell\Delta - \frac{\Delta(\Delta+1)}{2}$ instead of $\ell\Delta$. Under $2\Delta < \ell$, this count lies between $\ell\Delta/2$ and $\ell\Delta$, so the same upper-bound rate follows.

For the random-size case, we modify the truncation event to handle the vertices close to the endpoints of the interval $(0, \ell]$. Specifically, we modify the definition of N_i^+ and N_i^- for each $i \in V$ as follows:

$$N_i^+ := \sum_{j \neq i} \mathbf{1}\{\tilde{A}_j \in (0, \ell], 0 \leq d_{\text{lin}}^{(\ell),+}(\tilde{A}_i, \tilde{A}_j) \leq \Delta\},$$

$$N_i^- := \sum_{j \neq i} \mathbf{1}\{\tilde{A}_j \in (0, \ell], 0 \leq d_{\text{lin}}^{(\ell),-}(\tilde{A}_i, \tilde{A}_j) \leq \Delta\},$$

where $d_{\text{lin}}^{(\ell),+}(a, b) := b - a$ and $d_{\text{lin}}^{(\ell),-}(a, b) := a - b$ are the signed displacements on the interval $(0, \ell]$ between a and b in the right and left directions, respectively. Also, to handle the vertices close to the boundary, define

$$V^+ := \{i \in V : \tilde{A}_i \leq \ell - \Delta\}, \quad V^- := \{i \in V : \tilde{A}_i \geq \Delta\}.$$

Consider the truncation event

$$\underline{\mathcal{E}} := \left\{ |V| \geq c_1\ell, \min_{i \in V^+} N_i^+ \geq c_2\Delta, \min_{i \in V^-} N_i^- \geq c_2\Delta \right\},$$

where $0 < c_2 < c_1 < 1$ are fixed constants, and let $\tilde{\nu} := \tilde{\nu}(\cdot | \underline{\mathcal{E}})$. We interpret a minimum over an empty set as $+\infty$. We claim that under this event, S contains a thick path with parameters $\lfloor c_1\ell \rfloor$ and $\lfloor c_2\Delta \rfloor$. To see this, let $0 < \tilde{A}_{v_1} < \dots < \tilde{A}_{v_{|V|}} \leq \ell$ be the ordered latent locations of the active vertices for $S \sim \tilde{\nu}$. Write $h := \lfloor c_2\Delta \rfloor$, and $r := \lfloor c_1\ell \rfloor$. Since $c_2 < c_1$ and $2\Delta < \ell$, we have $r \geq 2h$ and $h \rightarrow \infty$. Then, observe that $\dot{S} := \{(v_a, v_b) : 1 \leq a < b \leq r, b - a \leq h\}$ is a thick path with parameters r and h , and it suffices to show that $\dot{S} \subseteq S$ almost surely. For $a < b$ satisfying $b - a \leq h$:

- **Case 1:** $\tilde{A}_{v_a} \leq \ell - \Delta$. In this case, the event $\underline{\mathcal{E}}$ guarantees that there are at least h other active vertices in $(\tilde{A}_{v_a}, \tilde{A}_{v_a} + \Delta]$, and thus $\tilde{A}_{v_b} \leq \tilde{A}_{v_a} + \Delta$, provided that $b - a \leq h$. Therefore, $(v_a, v_b) \in S$.
- **Case 2:** $\tilde{A}_{v_b} \geq \Delta$. Similar to Case 1, there are at least h other active vertices in $[\tilde{A}_{v_b} - \Delta, \tilde{A}_{v_b})$, and thus $\tilde{A}_{v_a} \geq \tilde{A}_{v_b} - \Delta$, provided that $b - a \leq h$. Therefore, $(v_a, v_b) \in S$.

- **Case 3:** $\tilde{A}_{v_a} > \ell - \Delta$ and $\tilde{A}_{v_b} < \Delta$. Note that this case can only occur if $\Delta > \ell/2$. In this case, $\tilde{A}_{v_b} - \tilde{A}_{v_a} < \Delta$, and thus $(v_a, v_b) \in S$. Thus, this case does not arise under the standing assumption $2\Delta < \ell$.

Therefore, we have $\dot{S} \subseteq S$ almost surely, and the upper bound analysis for detecting $\dot{S}$ is identical to that of the fixed-size thick path problem up to constant factors. Indeed, $\dot{S}$ has $rh - h(h+1)/2 \geq rh/2$ edges and contains an $(h+1)$ -vertex clique, with $r = \Theta(\ell)$ and $h = \Theta(\Delta)$. Under the original latent law, conditional on $i \in V^+$ and its latent location, N_i^+ has distribution $\text{Binomial}(m-1, \Delta/m)$; the analogous conditional law holds for N_i^- when $i \in V^-$. Chernoff bounds and a union bound therefore give, for some fixed $c_3, c_4 > 0$,

$$\Pr(\mathcal{E}^c) \leq \exp(-c_3\ell) + 2m \exp(-c_4\Delta) = o(1).$$

The comparison of risks using total variation in Appendix B.6 then transfers these upper bounds from the truncated prior to $\tilde{\nu}$.

Lower bounds. The lower bound follows from a coupling with the thick cycle problem. For either choice of the auxiliary vector $\dot{\mathbf{A}}$ or $\tilde{\mathbf{A}}$, a thick path of length ℓ

$$S_{\text{path}} := \{(i, j) \in [m]_2 : A_i, A_j \in (0, \ell], d_{\text{lin}}^{(\ell)}(A_i, A_j) \leq \Delta\}$$

is a subgraph of the corresponding thick cycle of length ℓ

$$S_{\text{cycle}} := \{(i, j) \in [m]_2 : A_i, A_j \in (0, \ell], d_{\text{cyc}}^{(\ell)}(A_i, A_j) \leq \Delta\}.$$

Indeed, recall that $d_{\text{lin}}^{(\ell)}(a, b) := |a - b|$ and $d_{\text{cyc}}^{(\ell)}(a, b) := \min\{|a - b|, \ell - |a - b|\}$. Therefore, for any $i, j \in [m]$, if $d_{\text{lin}}^{(\ell)}(A_i, A_j) \leq \Delta$, then $d_{\text{cyc}}^{(\ell)}(A_i, A_j) \leq \Delta$, which in turn guarantees that $S_{\text{path}} \subseteq S_{\text{cycle}}$. Using an independent copy of the same latent vector to define the primed paths and cycles, we obtain

$$|S_{\text{path}} \cap S'_{\text{path}}| \leq |S_{\text{cycle}} \cap S'_{\text{cycle}}|$$

almost surely. The full and low-degree likelihood-ratio second moments are expectations of increasing functions of this overlap, so the moment bounds in Appendix B.9 transfer to thick paths. In the random-size case, the same coupling remains valid when both latent copies are conditioned on the high-probability event used for the thick-cycle lower bounds. Thus, the statistical and computational lower bounds established for the thick cycle problem in Theorem 3 imply the same bounds for the thick path problem.

B.11 Proof of Proposition 3

Throughout this section, we use the notation

$$m_0 := \psi'(\theta_0), \quad m_1 := \psi'(\theta_1), \quad v_0 := \psi''(\theta_0) > 0, \\ d_\theta := m_1 - m_0 > 0, \quad \tau_n := d_\theta / \sqrt{v_0}.$$

Let $\delta = \delta_n \in (0, 1)$, and retain the assumption $|S| = K \geq 1$ almost surely under $S \sim \nu$. Recall the local regularity condition in Definition 4: we say that the NEF models P_{θ_0} and P_{θ_1} are locally regular at scale r_n if $[\theta_j - r_n, \theta_j + r_n] \subseteq \Theta$, $j \in \{0, 1\}$, and there exists a constant $C \geq 1$ such that

$$\sup_{\substack{j \in \{0, 1\} \\ |u| \leq r_n}} \frac{\psi''(\theta_j + u)}{\psi''(\theta_0)} \leq C,$$

for sufficiently large n . Note that C is no less than 1, since the supremum includes $j = 0$ and $u = 0$. In the following lemma, we show that the centered sum of independent samples from the NEF models satisfies a local sub-Gaussian tail bound under the local regularity condition.

Lemma 10 (Local concentration for NEF models). *Suppose that the NEF models P_{θ_0} and P_{θ_1} are locally regular at scale r_n with constant $C \geq 1$. For $N = N_n \geq 1$, let $Y_1, \dots, Y_N$ be independent samples such that $Y_i \sim P_{\theta^{(i)}}$ for $\theta^{(i)} \in \{\theta_0, \theta_1\}$. Then, for all sufficiently large n and every $0 \leq x \leq CNv_0r_n$, the centered sum satisfies the following tail bounds:*

$$\begin{aligned} \Pr \left\{ \sum_{i=1}^N (Y_i - \psi'(\theta^{(i)})) \geq x \right\} &\leq \exp \left(-\frac{x^2}{2CNv_0} \right), \\ \Pr \left\{ \sum_{i=1}^N (Y_i - \psi'(\theta^{(i)})) \leq -x \right\} &\leq \exp \left(-\frac{x^2}{2CNv_0} \right). \end{aligned}$$

Proof. Note that the NEF models allow the following equality for the cumulant generating function:

$$\log (\mathbb{E}_{Y \sim P_\theta} [\exp \{t[Y - \psi'(\theta)]\}]) = \psi(\theta + t) - \psi(\theta) - t\psi'(\theta), \quad \theta, \theta + t \in \Theta,$$

which can also be rewritten as the remainder term in the Taylor expansion:

$$\psi(\theta + t) - \psi(\theta) - t\psi'(\theta) = t^2 \int_0^1 (1-s)\psi''(\theta + st)ds.$$

Under the local regularity condition in Definition 4, we have an upper bound on the curvature of the log-partition function in the neighborhood of θ_0 and θ_1 :

$$\sup_{\substack{\theta \in \{\theta_0, \theta_1\} \\ |u| \leq r_n}} \psi''(\theta + u) \leq Cv_0.$$

For $\theta \in \{\theta_0, \theta_1\}$, the centered variable $Y - \psi'(\theta)$ is sub-exponential with parameters $(Cv_0, 1/r_n)$ in the following sense: for $|t| \leq r_n$, its log moment generating function satisfies

$$\begin{aligned} \log (\mathbb{E}_{Y \sim P_\theta} [\exp \{t[Y - \psi'(\theta)]\}]) &= t^2 \int_0^1 (1-s)\psi''(\theta + st)ds \\ &\leq \frac{Cv_0t^2}{2}. \end{aligned}$$

By the independence of $Y_1, \dots, Y_N$, the sum $\sum_{i=1}^N (Y_i - \psi'(\theta^{(i)}))$ is sub-exponential with parameters $(CNv_0, 1/r_n)$; its log moment generating function is bounded by $CNv_0t^2/2$ for $|t| \leq r_n$. For $0 \leq x \leq CNv_0r_n$, choose $t = x/(CNv_0) \leq r_n$ in the Chernoff bound to obtain

$$\Pr \left\{ \sum_{i=1}^N (Y_i - \psi'(\theta^{(i)})) \geq x \right\} \leq \exp \left(-tx + \frac{CNv_0t^2}{2} \right) = \exp \left(-\frac{x^2}{2CNv_0} \right).$$

Applying the same argument to the negative of the centered sum gives the lower-tail bound. $\square$

Thus, the proof procedure for the average, scan, and degree tests consists of the following two steps: (i) we specify the scale r_n for the local regularity of the NEF models, and (ii) we check that each target deviation lies in the range $x \leq CNv_0r_n$ of Lemma 10 and is no larger than the corresponding gap between the mean and the test threshold. The gap itself may exceed the local concentration range; in that case, we use event inclusion to apply the lemma at the smaller target deviation.

B.11.1 Average test

For the average test, define

$$r_{\text{avg},n} := \sqrt{\frac{2}{nv_0} \log \frac{2}{\delta}},$$

and assume that the NEF models are locally regular at scale $r_{\text{avg},n}$ with constant C_{avg} . Then, the average test Δ_{avg} achieves $R_\nu(\Delta_{\text{avg}}) \leq \delta$ if

$$\tau_n \geq \sqrt{\frac{8C_{\text{avg}}n}{K^2} \log \frac{2}{\delta}}.$$

To prove the claim, we control the Type I and Type II errors of the average test separately by $\delta/2$. Recall that the average test statistic and the threshold for the average test are defined as follows:

$$T_{\text{avg}}(\mathbf{Y}) := \sum_{i=1}^n Y_i, \quad t_{\text{avg}} := nm_0 + \frac{Kd_\theta}{2}.$$

Type I error. Under the null hypothesis H_0 , we have $Y_i \sim P_{\theta_0}$ independently for $i = 1, \dots, n$, which implies that the test statistic $\sum_{i=1}^n Y_i$ has mean nm_0 . Therefore,

$$\mathbb{Q}\{T_{\text{avg}}(\mathbf{Y}) > t_{\text{avg}}\} = \Pr_{Y_i \sim P_{\theta_0}} \left\{ \sum_{i=1}^n (Y_i - m_0) > \frac{Kd_\theta}{2} \right\}.$$

If the local regularity condition holds for the sequence $r_{\text{avg},n}$ with a constant $C_{\text{avg}} \geq 1$, then for any $0 \leq x \leq C_{\text{avg}}nv_0r_{\text{avg},n}$, the local sub-Gaussian concentration inequality in Lemma 10 guarantees

$$\Pr_{Y_i \sim P_{\theta_0}} \left\{ \sum_{i=1}^n (Y_i - m_0) \geq x \right\} \leq \exp \left(-\frac{x^2}{2C_{\text{avg}}nv_0} \right).$$

To make the right-hand side less than or equal to $\delta/2$, define a target deviation $x_\delta := \sqrt{2C_{\text{avg}}nv_0 \log(2/\delta)}$, so that $\exp(-x_\delta^2/(2C_{\text{avg}}nv_0)) = \delta/2$. Then, if

$$\frac{Kd_\theta}{2} \geq x_\delta,$$

we have

$$\left\{ \sum_{i=1}^n (Y_i - m_0) > \frac{Kd_\theta}{2} \right\} \subseteq \left\{ \sum_{i=1}^n (Y_i - m_0) > x_\delta \right\},$$

which implies that the Type I error is bounded by $\delta/2$. The condition $\frac{Kd_\theta}{2} \geq x_\delta$ is equivalent to

$$\tau_n \geq \sqrt{\frac{8C_{\text{avg}}n}{K^2} \log \frac{2}{\delta}}.$$

Finally, it remains to check that the target deviation x_δ satisfies the condition $x_\delta \leq C_{\text{avg}}nv_0r_{\text{avg},n}$ in Lemma 10, which holds as long as

$$r_{\text{avg},n} \geq \sqrt{\frac{2}{C_{\text{avg}}nv_0} \log \frac{2}{\delta}}.$$

Since $C_{\text{avg}} \geq 1$, by the definition of $r_{\text{avg},n} = \sqrt{2/(nv_0 \log(2/\delta))}$, the condition is satisfied.

Type II error. Fix any s in the support of ν , and define $\theta_s^{(i)} := \theta_1$ for $i \in s$ and $\theta_s^{(i)} := \theta_0$ for $i \notin s$. Under $\mathbb{P}_s$, we have $\mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_s} [\sum_{i=1}^n Y_i] = \sum_{i=1}^n \psi'(\theta_s^{(i)}) = (n-K)m_0 + Km_1 = nm_0 + Kd_\theta$. Therefore, the Type II error for this fixed s satisfies

$$\mathbb{P}_s\{T_{\text{avg}}(\mathbf{Y}) \leq t_{\text{avg}}\} = \Pr_{Y_i \sim P_{\theta_s^{(i)}}} \left\{ \sum_{i=1}^n (Y_i - \psi'(\theta_s^{(i)})) \leq -\frac{Kd_\theta}{2} \right\}.$$

We repeat the same argument as in the Type I error case, and the only difference is that we apply the local sub-Gaussian concentration inequality in Lemma 10 to the lower tail. Consequently, the Type II error is also bounded by $\delta/2$ under the same condition

$$\tau_n \geq \sqrt{\frac{8C_{\text{avg}}n}{K^2} \log \frac{2}{\delta}}.$$

Because the bound is uniform over all s in the support of ν ,

$$\mathbb{E}_{S \sim \nu} [\mathbb{P}_S \{\Delta_{\text{avg}}(\mathbf{Y}) = 0\}] \leq \frac{\delta}{2}.$$

Combining the Type I and Type II error bounds, we conclude that the average test achieves $R_\nu(\Delta_{\text{avg}}) \leq \delta$ under the stated condition on τ_n .

B.11.2 Scan test

For the scan test, define

$$r_{\text{scan},n} := \sqrt{\frac{2}{Kv_0} \log \frac{2|\mathcal{C}|}{\delta}},$$

and assume that the NEF models are locally regular at scale $r_{\text{scan},n}$ with constant C_{scan} . Consider the scan statistic and the threshold defined as

$$T_{\text{scan}}(\mathbf{Y}) = \max_{s \in \mathcal{C}} \sum_{l \in s} Y_l, \quad t_{\text{scan}} = Km_0 + \frac{Kd_\theta + \sqrt{2KC_{\text{scan}}v_0 \log |\mathcal{C}|}}{2}.$$

Then, the scan test $\Delta_{\text{scan}} := \mathbb{1}\{T_{\text{scan}}(\mathbf{Y}) > t_{\text{scan}}\}$ achieves $R_\nu(\Delta_{\text{scan}}) \leq \delta$ if

$$\tau_n \geq \sqrt{\frac{2C_{\text{scan}}}{K}} \left(2\sqrt{\log \frac{2}{\delta}} + \sqrt{\log |\mathcal{C}|} \right).$$

We control the Type I and Type II errors separately as follows.

Type I error. Note that for any fixed $s \in \mathcal{C}$, the sum $\sum_{l \in s} Y_l$ has mean Km_0 under the null hypothesis H_0 . Therefore, by the union bound, we obtain

$$\begin{aligned} \mathbb{Q}\{T_{\text{scan}}(\mathbf{Y}) > t_{\text{scan}}\} &\leq \sum_{s \in \mathcal{C}} \mathbb{Q}\left\{ \sum_{l \in s} Y_l > t_{\text{scan}} \right\} \\ &= \sum_{s \in \mathcal{C}} \Pr_{Y_l \sim P_{\theta_0}} \left\{ \sum_{l \in s} (Y_l - m_0) > \frac{Kd_\theta + \sqrt{2KC_{\text{scan}}v_0 \log |\mathcal{C}|}}{2} \right\} \\ &= |\mathcal{C}| \cdot \Pr_{Y_l \sim P_{\theta_0}} \left\{ \sum_{l \in s} (Y_l - m_0) > \frac{Kd_\theta + \sqrt{2KC_{\text{scan}}v_0 \log |\mathcal{C}|}}{2} \right\}. \end{aligned}$$

In the last line, s is any fixed member of $\mathcal{C}$; all candidate sums have the same distribution under $\mathbb{Q}$ because each candidate has size K . No independence across candidates is needed for the union bound. If the local regularity condition holds for the sequence $r_{\text{scan},n}$ with a constant $C_{\text{scan}} \geq 1$, then for any $0 \leq x \leq C_{\text{scan}}Kv_0r_{\text{scan},n}$, the local sub-Gaussian concentration inequality in Lemma 10 guarantees

$$\Pr_{Y_l \sim P_{\theta_0}} \left\{ \sum_{l \in s} (Y_l - m_0) > x \right\} \leq \exp \left(-\frac{x^2}{2C_{\text{scan}}Kv_0} \right).$$

We aim to make the right-hand side less than or equal to $\delta/(2|\mathcal{C}|)$, and thus define a target deviation $x_{\delta,\text{I}} := \sqrt{2C_{\text{scan}}Kv_0 \log(2|\mathcal{C}|/\delta)}$. Then, it remains to verify two conditions.

First, the target deviation $x_{\delta,\text{I}}$ should satisfy the condition $x_{\delta,\text{I}} \leq C_{\text{scan}}Kv_0r_{\text{scan},n}$, equivalently, we require

$$r_{\text{scan},n} \geq \sqrt{\frac{2}{C_{\text{scan}}Kv_0} \log \frac{2|\mathcal{C}|}{\delta}}.$$

By the definition of $r_{\text{scan},n} = \sqrt{2/(Kv_0) \log(2|\mathcal{C}|/\delta)}$, the condition is satisfied since $C_{\text{scan}} \geq 1$.

Second, we require

$$\frac{Kd_\theta + \sqrt{2KC_{\text{scan}}v_0 \log |\mathcal{C}|}}{2} \geq x_{\delta, \text{I}}, \quad (68)$$

so that we can bound the Type I error as

$$\begin{aligned} \mathbb{Q}\{T_{\text{scan}}(\mathbf{Y}) > t_{\text{scan}}\} &\leq |\mathcal{C}| \cdot \Pr_{Y_l \sim P_{\theta_0}} \left\{ \sum_{l \in s} (Y_l - m_0) > \frac{Kd_\theta + \sqrt{2KC_{\text{scan}}v_0 \log |\mathcal{C}|}}{2} \right\} \\ &\leq |\mathcal{C}| \cdot \Pr_{Y_l \sim P_{\theta_0}} \left\{ \sum_{l \in s} (Y_l - m_0) > x_{\delta, \text{I}} \right\} \\ &\leq |\mathcal{C}| \cdot \frac{\delta}{2|\mathcal{C}|} = \frac{\delta}{2}. \end{aligned}$$

Substituting the definition of $x_{\delta, \text{I}}$ into Equation (68) gives the equivalent condition

$$Kd_\theta \geq \sqrt{2KC_{\text{scan}}v_0} \left(2\sqrt{\log \frac{2|\mathcal{C}|}{\delta}} - \sqrt{\log |\mathcal{C}|} \right),$$

and since $d_\theta/\sqrt{v_0} = \tau_n$, this holds as long as

$$\tau_n \geq \sqrt{\frac{2C_{\text{scan}}}{K}} \left(2\sqrt{\log \frac{2|\mathcal{C}|}{\delta}} - \sqrt{\log |\mathcal{C}|} \right). \quad (69)$$

Type II error. Fix any s^* in the support of ν . Under $\mathbb{P}_{s^*}$, the mean of $\sum_{l \in s^*} Y_l$ is $Km_1 = Km_0 + Kd_\theta$. Therefore, the Type II error for this fixed s^* satisfies

$$\begin{aligned} \mathbb{P}_{s^*}\{T_{\text{scan}}(\mathbf{Y}) \leq t_{\text{scan}}\} &\leq \mathbb{P}_{s^*}\left\{ \sum_{l \in s^*} Y_l \leq t_{\text{scan}} \right\} \\ &= \Pr_{Y_l \sim P_{\theta_1}} \left\{ \sum_{l \in s^*} (Y_l - m_1) \leq -\frac{Kd_\theta - \sqrt{2KC_{\text{scan}}v_0 \log |\mathcal{C}|}}{2} \right\}. \end{aligned}$$

Again, if the local regularity condition holds for the sequence $r_{\text{scan}, n}$ with a constant $C_{\text{scan}} \geq 1$, then for any $0 \leq x \leq C_{\text{scan}}Kv_0r_{\text{scan}, n}$, the local sub-Gaussian concentration inequality in Lemma 10 guarantees

$$\Pr_{Y_l \sim P_{\theta_1}} \left\{ \sum_{l \in s^*} (Y_l - m_1) \leq -x \right\} \leq \exp \left(-\frac{x^2}{2C_{\text{scan}}Kv_0} \right).$$

To control the Type II error by $\delta/2$, the local sub-Gaussian concentration inequality suggests the target deviation $x_{\delta, \text{II}} := \sqrt{2C_{\text{scan}}Kv_0 \log(2/\delta)}$, so that $\exp(-x_{\delta, \text{II}}^2/(2C_{\text{scan}}Kv_0)) = \delta/2$, and it remains to check the following two conditions.

First, we verify that the target deviation $x_{\delta, \text{II}}$ lies in the local concentration regime, i.e., $x_{\delta, \text{II}} \leq C_{\text{scan}}Kv_0r_{\text{scan}, n}$, which is equivalent to

$$r_{\text{scan}, n} \geq \sqrt{\frac{2}{C_{\text{scan}}Kv_0} \log \frac{2}{\delta}}.$$

By the definition of $r_{\text{scan}, n} = \sqrt{2/(Kv_0 \log(2|\mathcal{C}|/\delta))}$, the condition is satisfied since $C_{\text{scan}} \geq 1$ and $|\mathcal{C}| \geq 1$, hence the local sub-Gaussian concentration lemma applies. Also, it remains to show that

$$\frac{Kd_\theta - \sqrt{2KC_{\text{scan}}v_0 \log |\mathcal{C}|}}{2} \geq x_{\delta, \text{II}},$$

so that the Type II error is bounded by $\delta/2$. Note that it suffices to have

$$Kd_\theta \geq \sqrt{2KC_{\text{scan}}v_0} \left(2\sqrt{\log \frac{2}{\delta}} + \sqrt{\log |\mathcal{C}|} \right),$$

which is equivalent to

$$\tau_n \geq \sqrt{\frac{2C_{\text{scan}}}{K}} \left(2\sqrt{\log \frac{2}{\delta}} + \sqrt{\log |\mathcal{C}|} \right). \quad (70)$$

Therefore, under this condition, the Type II error is at most $\delta/2$ uniformly over all s^* in the support of ν , and hence

$$\mathbb{E}_{S \sim \nu} [\mathbb{P}_S \{\Delta_{\text{scan}}(\mathbf{Y}) = 0\}] \leq \frac{\delta}{2}.$$

Finally, we combine the Type I and Type II error controls; observe that the signal strength condition for the Type II error control in Equation (70) is stronger than the condition for the Type I error control in Equation (69), since

$$\begin{aligned} 2\sqrt{\log \frac{2}{\delta}} + \sqrt{\log |\mathcal{C}|} &= 2 \left(\sqrt{\log \frac{2}{\delta}} + \sqrt{\log |\mathcal{C}|} \right) - \sqrt{\log |\mathcal{C}|} \\ &\geq 2\sqrt{\log |\mathcal{C}| + \log \frac{2}{\delta}} - \sqrt{\log |\mathcal{C}|}, \end{aligned}$$

where the last inequality follows from the fact that $\sqrt{a} + \sqrt{b} \geq \sqrt{a+b}$ for all $a, b \geq 0$. Therefore, if the condition in Equation (70) is satisfied, both Type I and Type II errors are controlled by $\delta/2$, and we conclude that the scan test achieves $R_\nu(\Delta_{\text{scan}}) \leq \delta$ under the stated condition on τ_n and the local regularity condition.

B.11.3 Degree test

For the degree test, recall the scale and threshold from Proposition 3:

$$\begin{aligned} r_{\text{deg}} &:= \sqrt{\frac{2}{(m-1)v_0} \log \frac{2m}{\delta}}, \\ t_{\text{deg}} &:= (m-1)m_0 + \frac{\mathfrak{d}_{\max} d_\theta}{2}. \end{aligned}$$

We assume the signal-strength condition in that proposition and retain its assumption that every s in the support of ν has maximum signal degree $\mathfrak{d}_{\max} \geq 1$. Recall that, under the local regularity condition at scale r_{deg} with constant C_{deg} , Lemma 10 guarantees that for any sum of $m-1$ independent observations from P_{θ_0} and P_{θ_1} and every $0 \leq x \leq C_{\text{deg}}(m-1)v_0 r_{\text{deg}}$,

$$\Pr \left\{ \sum_{j=1}^{m-1} (Y_j - \mathbb{E}[Y_j]) \geq x \right\} \leq \exp \left(- \frac{x^2}{2C_{\text{deg}}(m-1)v_0} \right),$$

and the analogous bound holds for the lower tail.

Type I error. Under the null hypothesis, for every vertex $i \in [m]$, the weighted degree $T_i^{\text{deg}}(\mathbf{Y})$ is a sum of $m-1$ independent observations from P_{θ_0} and thus has mean $(m-1)m_0$. Hence, by the union bound,

$$\mathbb{Q} \{ T_{\text{deg}}(\mathbf{Y}) > t_{\text{deg}} \} \leq \sum_{i=1}^m \mathbb{Q} \left\{ T_i^{\text{deg}}(\mathbf{Y}) - (m-1)m_0 > \frac{\mathfrak{d}_{\max} d_\theta}{2} \right\}.$$

Define the target deviation

$$x_{\delta, \text{I}} := \sqrt{2C_{\text{deg}}(m-1)v_0 \log \frac{2m}{\delta}}.$$

By the definition of r_{deg} and the fact that $C_{\text{deg}} \geq 1$, we have $x_{\delta, \text{I}} \leq C_{\text{deg}}(m-1)v_0 r_{\text{deg}}$, so the local concentration bound (Lemma 10) applies. Moreover, the assumed lower bound on τ_n is equivalent to

$$\frac{\mathfrak{d}_{\max} d_\theta}{2} \geq x_{\delta, \text{I}}.$$

Therefore,

$$\begin{aligned} \sum_{i=1}^m \mathbb{Q} \left\{ T_i^{\deg}(\mathbf{Y}) - (m-1)m_0 > \frac{\mathfrak{d}_{\max} d_\theta}{2} \right\} &\leq \sum_{i=1}^m \mathbb{Q} \{ T_i^{\deg}(\mathbf{Y}) - (m-1)m_0 > x_{\delta, \text{I}} \} \\ &\leq m \exp \left(- \frac{x_{\delta, \text{I}}^2}{2C_{\deg}(m-1)v_0} \right) = \frac{\delta}{2}. \end{aligned}$$

Type II error. Fix any s in the support of ν . By assumption, there exists a vertex $i_s \in [m]$ such that $\mathfrak{d}_s(i_s) = \mathfrak{d}_{\max}$. Under $\mathbb{P}_s$, the weighted degree of this vertex has mean

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_s} [T_{i_s}^{\deg}(\mathbf{Y})] = (m-1)m_0 + \mathfrak{d}_{\max} d_\theta.$$

Since $T_{\deg}(\mathbf{Y}) \geq T_{i_s}^{\deg}(\mathbf{Y})$, the Type II error for this fixed s satisfies

$$\begin{aligned} \mathbb{P}_s \{ T_{\deg}(\mathbf{Y}) \leq t_{\deg} \} &\leq \mathbb{P}_s \{ T_{i_s}^{\deg}(\mathbf{Y}) \leq t_{\deg} \} \\ &\leq \mathbb{P}_s \left\{ T_{i_s}^{\deg}(\mathbf{Y}) - \mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_s} [T_{i_s}^{\deg}(\mathbf{Y})] \leq - \frac{\mathfrak{d}_{\max} d_\theta}{2} \right\}. \end{aligned}$$

Define

$$x_{\delta, \text{II}} := \sqrt{2C_{\deg}(m-1)v_0 \log \frac{2}{\delta}}.$$

Since $\log(2/\delta) \leq \log(2m/\delta)$, we have

$$x_{\delta, \text{II}} \leq x_{\delta, \text{I}} \leq C_{\deg}(m-1)v_0 r_{\deg},$$

and the assumed lower bound on τ_n also implies $\mathfrak{d}_{\max} d_\theta / 2 \geq x_{\delta, \text{II}}$. Therefore, by the lower-tail bound in Lemma 10,

$$\begin{aligned} \mathbb{P}_s \left\{ T_{i_s}^{\deg}(\mathbf{Y}) - \mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_s} [T_{i_s}^{\deg}(\mathbf{Y})] \leq - \frac{\mathfrak{d}_{\max} d_\theta}{2} \right\} &\leq \mathbb{P}_s \left\{ T_{i_s}^{\deg}(\mathbf{Y}) - \mathbb{E}_{\mathbf{Y} \sim \mathbb{P}_s} [T_{i_s}^{\deg}(\mathbf{Y})] \leq -x_{\delta, \text{II}} \right\} \\ &\leq \exp \left(- \frac{x_{\delta, \text{II}}^2}{2C_{\deg}(m-1)v_0} \right) = \frac{\delta}{2}. \end{aligned}$$

The bound holds uniformly over every s in the support of ν , and hence

$$\mathbb{E}_{S \sim \nu} [\mathbb{P}_S \{ \Delta_{\deg}(\mathbf{Y}) = 0 \}] \leq \frac{\delta}{2}.$$

Combining the Type I and Type II error bounds yields $R_\nu(\Delta_{\deg}) \leq \delta$.

B.12 Proof of Theorem 4

Second moment of the likelihood ratio. We first derive the exact second-moment identity for the likelihood ratio. Fix one of the two models and suppress the superscript $M \in \{A, B\}$ for notational convenience. Write $\chi^2 = \chi^2(\theta_0, \theta_1) = \chi_n^2$ and $D = D_n$ for brevity. Write

$$L_{\theta_0, \theta_1}(y) = \frac{dP_{\theta_1}(y)}{dP_{\theta_0}(y)} = \exp \{ (\theta_1 - \theta_0)y - (\psi(\theta_1) - \psi(\theta_0)) \},$$

for the one-coordinate likelihood ratio. The assumption $\psi(2\theta_1 - \theta_0) < \infty$ ensures that this ratio belongs to $L^2(P_{\theta_0})$. Then, for any fixed $s \in \mathcal{C}_n$, the likelihood ratio of the n -dimensional observation $\mathbf{Y} = (Y_1, \dots, Y_n)$ is $L_{s,n}(\mathbf{y}) = \prod_{l \in s} L_{\theta_0, \theta_1}(y_l)$. Since the alternative is a mixture over $S \sim \nu_n$, the overall likelihood ratio is

$$L_{\nu_n}(\mathbf{y}) = \mathbb{E}_{S \sim \nu_n} [L_{S,n}(\mathbf{y})].$$

Let S and S' be independent copies with law ν_n . The second moment of the likelihood ratio under the null is therefore:

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}(\mathbf{Y})^2] &= \mathbb{E}_{S, S' \sim \nu_n} \left[\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{S,n}(\mathbf{Y}) L_{S',n}(\mathbf{Y})] \right] \\
&= \mathbb{E}_{S, S' \sim \nu_n} \left[\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} \left[\prod_{l \in S \cap S'} L_{\theta_0, \theta_1}(Y_l)^2 \prod_{i \in S \Delta S'} L_{\theta_0, \theta_1}(Y_i) \right] \right] \\
&= \mathbb{E}_{S, S' \sim \nu_n} \left[\prod_{l \in S \cap S'} \mathbb{E}_{Y_l \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y_l)^2] \prod_{i \in S \Delta S'} \mathbb{E}_{Y_i \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y_i)] \right] \\
&= \mathbb{E}_{S, S' \sim \nu_n} \left[(1 + \chi^2)^{|S \cap S'|} \right],
\end{aligned}$$

as $\mathbb{E}_{Y \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y)] = 1$, and $\chi^2 = \mathbb{E}_{Y \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y)^2] - 1$. Therefore, for two models with the same prior ν_n , if $\chi_A^2 = \chi_B^2$, then

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n^A} [L_{\nu_n}^A(\mathbf{Y})^2] = \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n^B} [L_{\nu_n}^B(\mathbf{Y})^2].$$

Second moment of the low-degree likelihood ratio. We now study the low-degree likelihood ratio. It suffices to consider integers $D \geq 1$, since $D = 0$ gives $L_{\nu_n}^{\leq 0} \equiv 1$. Again we suppress the superscript $M \in \{A, B\}$ for notational convenience. Let $\{\phi_d\}_{d \geq 0}$ be orthonormal polynomials in $L^2(P_{\theta_0})$, with $\phi_0 \equiv 1$ and $\phi_1(y) = (y - \psi'(\theta_0))/\sqrt{\psi''(\theta_0)}$. Note that if P_{θ_0} has finite support, then the sequence $\{\phi_d\}_{d \geq 0}$ terminates. For each available degree, define

$$r_d(\theta_0, \theta_1) := \mathbb{E}_{Y \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y) \phi_d(Y)], \quad d \geq 1.$$

In the finite-support case, extend r_d by zero beyond the last available degree. For an admissible multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_{\geq 0}^n$, write

$$\Phi_\alpha(\mathbf{y}) = \prod_{l=1}^n \phi_{\alpha_l}(y_l), \quad |\alpha| = \sum_{l=1}^n \alpha_l,$$

and $\text{supp}(\alpha) = \{l : \alpha_l > 0\}$. In the finite-support case, retain only indices for which every ϕ_{α_l} exists, and set coefficients at all other indices to zero in the sums below. Then, the set of tensor products $\{\Phi_\alpha : |\alpha| \leq D\}$ forms an orthonormal basis for the space of polynomials of degree at most D in $L^2(\mathbb{Q}_n)$. Let

$$\begin{aligned}
R_\alpha &:= \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}(\mathbf{Y}) \Phi_\alpha(\mathbf{Y})] \\
&= \mathbb{E}_{S \sim \nu_n} \left[\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{S,n}(\mathbf{Y}) \Phi_\alpha(\mathbf{Y})] \right] \\
&= \mathbb{E}_{S \sim \nu_n} \left[\prod_{l=1}^n \mathbb{E}_{Y \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y)^{\mathbb{1}\{l \in S\}} \phi_{\alpha_l}(Y)] \right].
\end{aligned}$$

We now focus on the inner expectation. For each $l \in [n]$, there are only three possible cases: (i) if $\alpha_l = 0$, then $\mathbb{E}_{Y \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y)^{\mathbb{1}\{l \in S\}} \phi_{\alpha_l}(Y)] = 1$; (ii) if $\alpha_l \geq 1$ and $l \in S$, then the factor equals $\mathbb{E}_{Y \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y) \phi_{\alpha_l}(Y)] = r_{\alpha_l}(\theta_0, \theta_1)$; (iii) if $\alpha_l \geq 1$ and $l \notin S$, then $\mathbb{E}_{Y \sim P_{\theta_0}} [\phi_{\alpha_l}(Y)] = 0$. Therefore, the entire product is zero unless $\text{supp}(\alpha) \subseteq S$. Consequently,

$$\begin{aligned}
R_\alpha &= \mathbb{E}_{S \sim \nu_n} \left[\mathbb{1}\{\text{supp}(\alpha) \subseteq S\} \prod_{l \in \text{supp}(\alpha)} r_{\alpha_l}(\theta_0, \theta_1) \right] \\
&= \Pr_{S \sim \nu_n} [\text{supp}(\alpha) \subseteq S] \prod_{l \in \text{supp}(\alpha)} r_{\alpha_l}(\theta_0, \theta_1).
\end{aligned}$$

By the same logic,

$$R_\alpha^2 = \Pr_{S, S' \sim \nu_n} [\text{supp}(\alpha) \subseteq S \cap S'] \prod_{l \in \text{supp}(\alpha)} r_{\alpha_l}(\theta_0, \theta_1)^2.$$

Now, recall that the low-degree likelihood ratio is

$$L_{\nu_n}^{\leq D}(\mathbf{Y}) = \sum_{|\alpha| \leq D} R_\alpha \Phi_\alpha(\mathbf{Y}).$$

Therefore, Parseval's identity gives

$$\begin{aligned} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D}(\mathbf{Y})^2] &= \sum_{|\alpha| \leq D} R_\alpha^2 \\ &= \sum_{|\alpha| \leq D} \Pr_{S, S' \sim \nu_n} [\text{supp}(\alpha) \subseteq S \cap S'] \prod_{l \in \text{supp}(\alpha)} r_{\alpha_l}(\theta_0, \theta_1)^2 \\ &= \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{\substack{\alpha: \text{supp}(\alpha) \subseteq S \cap S' \\ |\alpha| \leq D}} \prod_{l \in \text{supp}(\alpha)} r_{\alpha_l}(\theta_0, \theta_1)^2 \right]. \end{aligned}$$

We group this sum according to the support size $d = |\text{supp}(\alpha)|$, which ranges from 0 to D , and obtain

$$\begin{aligned} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D}(\mathbf{Y})^2] &= \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^D \sum_{\substack{\alpha: \text{supp}(\alpha) \subseteq S \cap S' \\ |\text{supp}(\alpha)|=d \\ |\alpha| \leq D}} \prod_{l \in \text{supp}(\alpha)} r_{\alpha_l}(\theta_0, \theta_1)^2 \right] \\ &= \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^D \sum_{\substack{T \subseteq S \cap S' \\ |T|=d}} \sum_{\substack{\alpha: \text{supp}(\alpha)=T \\ |\alpha| \leq D}} \prod_{l \in T} r_{\alpha_l}(\theta_0, \theta_1)^2 \right] \\ &= \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D \wedge |S \cap S'|} \binom{|S \cap S'|}{d} \sum_{\substack{\alpha_1, \dots, \alpha_d \geq 1 \\ \alpha_1 + \dots + \alpha_d \leq D}} \prod_{i=1}^d r_{\alpha_i}(\theta_0, \theta_1)^2 \right], \end{aligned} \quad (71)$$

The $d = 0$ term corresponds to the empty support; its empty product and inner sum are both 1.

Now, we bound this expression from above and below. For $1 \leq d \leq D$, the case $\alpha_1 = \dots = \alpha_d = 1$ is always included in the sum, and thus we have

$$\sum_{\substack{\alpha_1, \alpha_2, \dots, \alpha_d \geq 1 \\ \alpha_1 + \dots + \alpha_d \leq D}} \prod_{i=1}^d r_{\alpha_i}(\theta_0, \theta_1)^2 \geq r_1(\theta_0, \theta_1)^{2d}, \quad (72)$$

which gives the lower bound

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D}(\mathbf{Y})^2] \geq \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D \wedge |S \cap S'|} \binom{|S \cap S'|}{d} r_1(\theta_0, \theta_1)^{2d} \right].$$

On the other hand, for each $d \geq 1$,

$$\begin{aligned} \sum_{\substack{\alpha_1, \dots, \alpha_d \geq 1 \\ \alpha_1 + \dots + \alpha_d \leq D}} \prod_{i=1}^d r_{\alpha_i}(\theta_0, \theta_1)^2 &\leq \sum_{\alpha_1, \dots, \alpha_d \geq 1} \prod_{i=1}^d r_{\alpha_i}(\theta_0, \theta_1)^2 \\ &= \left(\sum_{\alpha \geq 1} r_\alpha(\theta_0, \theta_1)^2 \right)^d \\ &\leq \|L_{\theta_0, \theta_1} - 1\|_{L^2(P_{\theta_0})}^{2d} \\ &= \chi^2(\theta_0, \theta_1)^d. \end{aligned}$$

The norm bound is Bessel's inequality applied to $L_{\theta_0, \theta_1} - 1$ and the orthonormal polynomials of positive degree. The above inequality, when combined with Equation (72), yields

$$\chi^2(\theta_0, \theta_1)^d \geq r_1(\theta_0, \theta_1)^{2d}, \quad (73)$$

and, when combined with Equation (71),

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D}(\mathbf{Y})^2] \leq \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D \wedge |S \cap S'|} \binom{|S \cap S'|}{d} \chi^2(\theta_0, \theta_1)^d \right].$$

Thus, we have shown that

$$\begin{aligned} \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D \wedge |S \cap S'|} \binom{|S \cap S'|}{d} r_1(\theta_0, \theta_1)^{2d} \right] &\leq \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n} [L_{\nu_n}^{\leq D}(\mathbf{Y})^2] \\ &\leq \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D \wedge |S \cap S'|} \binom{|S \cap S'|}{d} \chi^2(\theta_0, \theta_1)^d \right]. \end{aligned} \quad (74)$$

Now we compare $r_1(\theta_0, \theta_1)^2$ and $\chi^2(\theta_0, \theta_1)$. Note that $r_1(\theta_0, \theta_1)^2 = \tau_n^2$, since

$$\begin{aligned} r_1(\theta_0, \theta_1) &= \mathbb{E}_{Y \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y) \phi_1(Y)] \\ &= \mathbb{E}_{Y \sim P_{\theta_1}} [\phi_1(Y)] \\ &= \frac{\mathbb{E}_{Y \sim P_{\theta_1}} [Y] - \mathbb{E}_{Y \sim P_{\theta_0}} [Y]}{\sqrt{\text{Var}_{Y \sim P_{\theta_0}} [Y]}} \\ &= \frac{\psi'(\theta_1) - \psi'(\theta_0)}{\psi''(\theta_0)^{1/2}}. \end{aligned}$$

Since the NEF is nondegenerate and $\theta_0 \neq \theta_1$, we have $\tau_n \neq 0$. By Equation (73),

$$\frac{\chi^2(\theta_0, \theta_1)}{\tau_n^2} \geq 1, \quad \log \frac{\chi^2(\theta_0, \theta_1)}{\tau_n^2} \geq 0.$$

Under the assumption $D \log(\chi^2(\theta_0, \theta_1)/\tau_n^2) = O(1)$, for $0 \leq d \leq D$ we have

$$\begin{aligned} r_1(\theta_0, \theta_1)^{2d} &= \exp\{-d \log(\chi^2(\theta_0, \theta_1)/\tau_n^2)\} \chi^2(\theta_0, \theta_1)^d \\ &\geq \exp\{-D \log(\chi^2(\theta_0, \theta_1)/\tau_n^2)\} \chi^2(\theta_0, \theta_1)^d \\ &\geq e^{-C} \chi^2(\theta_0, \theta_1)^d \end{aligned}$$

for some fixed $C > 0$, uniformly over $0 \leq d \leq D$ and all sufficiently large n . Therefore, we obtain

$$\mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D \wedge |S \cap S'|} \binom{|S \cap S'|}{d} r_1(\theta_0, \theta_1)^{2d} \right] \geq e^{-C} \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D \wedge |S \cap S'|} \binom{|S \cap S'|}{d} \chi^2(\theta_0, \theta_1)^d \right],$$

which implies that the two bounds in Equation (74) are equivalent up to a constant factor. The comparison constants depend only on an eventual upper bound for $D \log(\chi^2/\tau_n^2)$. This completes the proof.

B.13 Proof of Proposition 4

In this proof, we parse the condition $D \log(\chi^2(\theta_0, \theta_1)/\tau_n^2) = O(1)$ and discuss its implications for the fixed-null setting. For brevity, write $D = D_n$, $h = h_n$, and $\chi^2 = \chi^2(\theta_0, \theta_1) = \chi_n^2$. We first fix $\theta_0 \in \Theta$ and let $\theta_1 = \theta_0 + h$, where $h \rightarrow 0$ with $h \neq 0$. Define $v_0 := \psi^{(2)}(\theta_0) > 0$ and $c_3 := \psi^{(3)}(\theta_0)$. Because θ_0 is an interior point, the

segments $\{\theta_0 + uh : u \in [0, 2]\}$ lie in a fixed compact subset of Θ for all sufficiently small $|h|$. In particular, $\theta_0 + 2h \in \Theta$, so the chi-square divergence below is finite. Then, we have

$$\begin{aligned}
1 + \chi^2 &= \mathbb{E}_{Y \sim P_{\theta_0}} [L_{\theta_0, \theta_1}(Y)^2] \\
&= \mathbb{E}_{Y \sim P_{\theta_0}} [\exp\{2(\theta_1 - \theta_0)Y - 2(\psi(\theta_1) - \psi(\theta_0))\}] \\
&= \exp\{-2(\psi(\theta_1) - \psi(\theta_0))\} \mathbb{E}_{Y \sim P_{\theta_0}} [\exp\{2(\theta_1 - \theta_0)Y\}] \\
&\stackrel{(\dagger)}{=} \exp\{-2(\psi(\theta_1) - \psi(\theta_0))\} \exp\{\psi(2\theta_1 - \theta_0) - \psi(\theta_0)\} \\
&= \exp\{\psi(2\theta_1 - \theta_0) - 2\psi(\theta_1) + \psi(\theta_0)\} \\
&= \exp\{\psi(\theta_0 + 2h) - 2\psi(\theta_0 + h) + \psi(\theta_0)\},
\end{aligned}$$

where $(\dagger)$ follows from $\psi(\theta_0 + t) - \psi(\theta_0) = \log \mathbb{E}_{Y \sim P_{\theta_0}} [\exp(tY)]$ at $t = 2h$. By Taylor expansion with Lagrange remainder, there exist $u_1, u_2 \in [0, 1]$ such that

$$\begin{aligned}
\psi(\theta_0 + 2h) &= \psi(\theta_0) + 2h\psi^{(1)}(\theta_0) + 4h^2 \frac{\psi^{(2)}(\theta_0)}{2} + 8h^3 \frac{\psi^{(3)}(\theta_0)}{6} \\
&\quad + 16h^4 \frac{\psi^{(4)}(\theta_0 + 2u_1h)}{24}
\end{aligned}$$

and

$$2\psi(\theta_0 + h) = 2\psi(\theta_0) + 2h\psi^{(1)}(\theta_0) + 2h^2 \frac{\psi^{(2)}(\theta_0)}{2} + 2h^3 \frac{\psi^{(3)}(\theta_0)}{6} + 2h^4 \frac{\psi^{(4)}(\theta_0 + u_2h)}{24}.$$

Using the above identities, we obtain

$$\psi(\theta_0 + 2h) - 2\psi(\theta_0 + h) + \psi(\theta_0) = v_0h^2 + c_3h^3 + O(h^4),$$

where the remainder is controlled by $\sup_{u \in [0, 2]} |\psi^{(4)}(\theta_0 + uh)|$ and is uniform for sufficiently small $|h|$. Thus, as $h \rightarrow 0$,

$$\chi^2 = \exp\{v_0h^2 + c_3h^3 + O(h^4)\} - 1,$$

and therefore, using the fact that $e^x - 1 = x + O(x^2)$ for $x \rightarrow 0$,

$$\chi^2 = v_0h^2 + c_3h^3 + O(h^4) = \Theta(h^2), \tag{75}$$

for $h \rightarrow 0$. As for $\tau_n = \frac{\psi'(\theta_1) - \psi'(\theta_0)}{\psi''(\theta_0)^{1/2}}$, the Taylor expansion of $\psi^{(1)}(\theta_1)$ around θ_0 yields, for some $u_3 \in [0, 1]$,

$$\begin{aligned}
\tau_n^2 &= \frac{(v_0h + c_3h^2/2 + \psi^{(4)}(\theta_0 + u_3h)h^3/6)^2}{v_0} \\
&= \frac{v_0^2h^2 + v_0c_3h^3 + O(h^4)}{v_0} \\
&= v_0h^2 + c_3h^3 + O(h^4) = \Theta(h^2).
\end{aligned} \tag{76}$$

For $|h| \leq 1$, the final remainder $O(h^4)$ can be bounded in absolute value by h^4 times

$$c_4 + \frac{\max\{c_3^2, c_4^2, |c_3|c_4\}}{v_0},$$

with $c_4 := \sup_{u \in [0, 2]} |\psi^{(4)}(\theta_0 + uh)|$, which is uniformly bounded for small $|h|$. Therefore, as $h \rightarrow 0$,

$$\frac{\chi^2}{\tau_n^2} = \frac{v_0h^2 + c_3h^3 + O(h^4)}{v_0h^2 + c_3h^3 + O(h^4)} = \frac{1 + \frac{c_3}{v_0}h + O(h^2)}{1 + \frac{c_3}{v_0}h + O(h^2)} = 1 + O(h^2),$$

which implies that, as $h \rightarrow 0$,

$$\log \frac{\chi^2}{\tau_n^2} = \log(1 + O(h^2)) = O(h^2).$$

The last expression, combined with Equation (75) and Equation (76), gives

$$\log \frac{\chi_n^2}{\tau_n^2} = O(\chi^2) = O(\tau_n^2),$$

as $h \rightarrow 0$. Consequently, $D_n h_n^2 = O(1)$ implies

$$D_n \log \frac{\chi_n^2}{\tau_n^2} = O(D_n h_n^2) = O(1).$$

Moreover, Equations (75) and (76) show that h_n^2 , χ_n^2 , and τ_n^2 are comparable up to fixed positive constants, so the three sufficient degree conditions in Proposition 4 are equivalent.

The fixed-null assumption can be relaxed when the Taylor expansions above hold uniformly. In particular, let $\theta_{0,n}$ vary with n , set $\theta_{1,n} = \theta_{0,n} + h_n$ with $h_n \rightarrow 0$, and suppose that the segments $\{\theta_{0,n} + u h_n : u \in [0, 2]\}$ are contained in a fixed compact subset of $\text{int}(\Theta)$. Then $\psi''(\theta_{0,n})$ is bounded away from zero and the derivatives of ψ appearing in the expansions are uniformly bounded, so the preceding calculation gives

$$\frac{\chi_n^2}{\tau_n^2} = 1 + O(h_n^2)$$

with an implicit constant independent of n . Consequently, the conclusion of Proposition 4 continues to hold for such varying null parameters. Without this condition on the null parameters, uniform derivative bounds or a positive lower bound on $\psi''(\theta_{0,n})$ may fail, so the extension requires model-specific verification, as in the sparse Bernoulli regimes discussed in Section 4.3.1.

B.14 Proof of Theorem 5

Fix an arbitrary positive sequence $h_n^B \rightarrow 0$ for which the effective signal strength τ_n^B vanishes polynomially. We first construct a chi-square-matched sequence $h_n^A \rightarrow 0$. For $M \in \{A, B\}$ and sufficiently small $h \geq 0$, define

$$g_M(h) := \chi^2(P_{c_M+h}^M \| P_{c_M}^M) = \exp \{ \psi^M(c_M + 2h) - 2\psi^M(c_M + h) + \psi^M(c_M) \} - 1.$$

This function is continuous with $g_M(0) = 0$, and for sufficiently small $h > 0$,

$$\frac{d}{dh} \log(1 + g_M(h)) = 2 \{ (\psi^M)'(c_M + 2h) - (\psi^M)'(c_M + h) \} > 0,$$

where the strict inequality follows from $(\psi^M)'' > 0$. Also, the expansion in Equation (75) gives $g_M(h) = (\psi^M)''(c_M) h^2 (1 + O(h))$. Hence, for all sufficiently large n , continuity and strict monotonicity give a unique sufficiently small $h_n^A > 0$ such that $g_A(h_n^A) = g_B(h_n^B)$, and $h_n^A \rightarrow 0$.

For $M \in \{A, B\}$, write

$$\chi_{M,n}^2 := \chi^2(P_{\theta_1^M}^M \| P_{\theta_0^M}^M).$$

By the local alternative expansion in Proposition 4,

$$\frac{\chi_{M,n}^2}{(\tau_n^M)^2} = 1 + O((h_n^M)^2) = 1 + o(1).$$

The two models have matching chi-square divergences, $\chi_{A,n}^2 = \chi_{B,n}^2 = \chi_n^2$, and both effective signal strengths are positive because $h_n^M > 0$ and $(\psi^M)'' > 0$. Therefore,

$$\frac{\tau_n^A}{\tau_n^B} \rightarrow 1.$$

In particular, the matched effective signal strength τ_n^A also vanishes polynomially, so the assumptions for model A apply to this sequence. Hence, multiplication by a $1 + o(1)$ factor preserves the relations of polynomial separation $\ll$ and $\gg$. In particular, for any positive sequence a_n ,

$$\tau_n^A \ll a_n \iff \tau_n^B \ll a_n, \quad \tau_n^A \gg a_n \iff \tau_n^B \gg a_n.$$

We first consider the statistical lower bound. By Theorem 4, the second moment of the likelihood ratio under model M is

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n^M} [L_{\nu_n}^M(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu_n} [(1 + \chi_n^2)^{|S \cap S'|}].$$

Thus, under chi-square matching, the likelihood-ratio second moments of models A and B are identical. Therefore, if $\tau_n^B \ll \tau_{\text{stat}}^{A,*}$, then $\tau_n^A \ll \tau_{\text{stat}}^{A,*}$, and boundedness of the second moment for model A implies the same boundedness for model B . Hence, by Lemma 1, the statistical lower bound at the scale $\tau_{\text{stat}}^{A,*}$ transfers to model B .

Next, consider the computational lower bound. Suppose that $\tau_n^B \ll \tau_{\text{comp}}^{A,*}$. Then $\tau_n^A \ll \tau_{\text{comp}}^{A,*}$ as well. By the standing assumption that the effective signal strengths vanish polynomially and that D_n is polylogarithmic,

$$D_n(\tau_n^M)^2 = o(1), \quad M \in \{A, B\}.$$

It follows from Proposition 4 that

$$D_n \log \left(\frac{\chi_n^2}{(\tau_n^M)^2} \right) = O(1), \quad M \in \{A, B\}.$$

Therefore, the low-degree part of Theorem 4 applies to both models and gives

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}_n^M} [L_{\nu_n}^{M, \leq D_n}(\mathbf{Y})^2] \asymp \mathbb{E}_{S, S' \sim \nu_n} \left[\sum_{d=0}^{D_n \wedge |S \cap S'|} \binom{|S \cap S'|}{d} (\chi_n^2)^d \right], \quad M \in \{A, B\}.$$

The expression on the right-hand side is common to the two models. Thus, boundedness of the degree- D_n likelihood-ratio second moment for model A implies boundedness for model B . By Lemma 2, this rules out strong separation by degree- D_n polynomials under model B . Under Conjecture 1, with the same degree requirement as for model A , the computational lower bound at the scale $\tau_{\text{comp}}^{A,*}$ therefore transfers to model B .

It remains to establish the corresponding upper bounds. Because each null parameter is a fixed interior point, the neighborhoods $[\theta_j^M - r_n, \theta_j^M + r_n]$, $j \in \{0, 1\}$, are contained in Θ_M for all sufficiently large n whenever $r_n \rightarrow 0$. Since both models are in the local alternative regime, Proposition 5 gives, for every sequence $r_n \rightarrow 0$,

$$\sup_{\substack{j \in \{0, 1\} \\ |u| \leq r_n}} \frac{(\psi^M)''(\theta_j^M + u)}{(\psi^M)''(\theta_0^M)} = 1 + o(1), \quad M \in \{A, B\}.$$

Hence, the local regularity conditions required for the average, degree, and scan tests are satisfied with a constant arbitrarily close to one, provided that the corresponding radius converges to zero. For the average test and, when applicable, the degree test, this holds by choosing the target risk $\delta_n \rightarrow 0$ sufficiently slowly, so that $\log(2/\delta_n) = o(n)$ and $\log(2m/\delta_n) = o(m)$, respectively. For the scan test, since $(\psi^M)''(\theta_0^M)$ is a fixed positive constant,

$$r_{\text{scan}}^2 \asymp \frac{\log |\mathcal{C}| + \log(1/\delta_n)}{K},$$

where the radius is evaluated separately for each model. Whenever the scan-test sufficient condition gives strong detection in model A , the assumption $\tau_n^A \rightarrow 0$ forces $K \rightarrow \infty$. Thus, the assumption $\log |\mathcal{C}| = o(K)$ allows us to choose $\delta_n \rightarrow 0$ sufficiently slowly so that $r_{\text{scan}} \rightarrow 0$ for both models.

By Proposition 3 and Remark 3, the sufficient conditions for these tests depend on the NEF model only through its effective signal strength, up to the local regularity constant. Therefore, the corresponding average, degree, or scan test that yields the matching upper bound for model A yields an upper bound of the same order for model B . Because the latent-structure class and the test statistic are unchanged, this transfer also preserves the computational complexity of the test. Consequently, strong detection in model B is achievable whenever $\tau_n^B \gg \tau_{\text{stat}}^{A,*}$, and polynomial-time strong detection is achievable whenever $\tau_n^B \gg \tau_{\text{comp}}^{A,*}$.

Since the starting admissible signal sequence in model B was arbitrary, combining the transferred lower and upper bounds shows that $\tau_{\text{stat}}^{A,*}$ and $\tau_{\text{comp}}^{A,*}$ are also valid statistical and computational thresholds, respectively, for model B , in the sense of Definition 5. This proves the claim.

B.15 Alternative proofs of computational lower bounds for multiple cliques and hypercliques

In this subsection, we provide alternative proofs of the computational lower bounds for the random-size multiple clique problem (Theorem 2) and the random-size multiple planted hypercliques problem (Corollary 2). We follow the approach of Mao et al. (2025), based on U -statistics and their decoupled counterparts (de la Peña and Giné, 1999; Giné et al., 2000). We note that the resulting bounds match those obtained from the original proofs in Appendices B.7.2 and B.8.2, up to logarithmic factors.

B.15.1 Multiple random-size cliques

Recall the definition in Equation (43):

$$W_v = \begin{cases} e_{(a,b)} & \text{if } v \in \tilde{V}_a \cap \tilde{V}_b' \text{ for some } a, b \in [\gamma], \\ 0 & \text{otherwise,} \end{cases} \quad (77)$$

and note that $\Pr(W_v = e_{(a,b)}) = \rho = k^2/m^2$ for all $a, b \in [\gamma]$. The labels $W_1, \dots, W_m$ are i.i.d., with $\Pr(W_v = 0) = 1 - \gamma^2\rho$; the feasibility condition $\gamma k \leq m$ ensures that $\gamma^2\rho \leq 1$. Define the kernel function $h : (\{0, 1\}^{\gamma \times \gamma})^2 \rightarrow \mathbb{R}$ by $h(w_1, w_2) := \sum_{a,b=1}^{\gamma} \mathbb{1}\{w_1 = e_{(a,b)}\} \mathbb{1}\{w_2 = e_{(a,b)}\}$. From Equation (44), the overlap size $|S \cap S'|$ can be expressed as

$$|S \cap S'| = \sum_{(i,j) \in [m]_2} h(W_i, W_j). \quad (78)$$

Hence, this can be seen as a U -statistic of order 2 with kernel h . We bound the overlap size $|S \cap S'|$ using its decoupled version. Consider an independent copy $W'_1, \dots, W'_m$ of $W_1, \dots, W_m$, and the kernel

$$H_{i,j}(w_1, w_2) := \mathbb{1}\{i \neq j\} h(w_1, w_2), \quad (i, j) \in [m] \times [m].$$

Note that $H_{i,j}$ is zero whenever $i = j$, and is equal to $h(w_1, w_2)$ otherwise. Therefore, since h is symmetric, we obtain

$$\begin{aligned} |S \cap S'| &= \sum_{(i,j) \in [m]_2} h(W_i, W_j) \\ &= \frac{1}{2} \sum_{(i,j) \in [m]^2} H_{i,j}(W_i, W_j) \\ &=: U(W_1, \dots, W_m). \end{aligned}$$

Define the decoupled version of $U(W_1, \dots, W_m)$ as

$$U'(W_1, \dots, W_m; W'_1, \dots, W'_m) = \frac{1}{2} \sum_{(i,j) \in [m] \times [m]} H_{i,j}(W_i, W'_j).$$

Then, by Lemma 7, there exists a constant $c > 0$ such that

$$\mathbb{E} \left[\left(\sum_{(i,j) \in [m]^2} H_{i,j}(W_i, W_j) \right)^d \right] \leq c^d \mathbb{E} \left[\left(\sum_{(i,j) \in [m]^2} H_{i,j}(W_i, W'_j) \right)^d \right],$$

for every integer $d \geq 1$, which implies that

$$\mathbb{E}[|S \cap S'|^d] \leq c'^d \mathbb{E} \left[\left(\sum_{(i,j) \in [m]^2} H_{i,j}(W_i, W'_j) \right)^d \right], \quad (79)$$

for $c' = c/2$.

For notational convenience, define

$$\lambda := \gamma^2 \left(\frac{k^2}{m} \right)^2.$$

Then, when $d = 1$,

$$\begin{aligned} \mathbb{E} \left[\sum_{(i,j) \in [m]^2} H_{i,j}(W_i, W'_j) \right] &= \sum_{(i,j) \in [m]^2} \mathbb{E} [H_{i,j}(W_i, W'_j)] \\ &= \sum_{i \neq j} \mathbb{E} [H_{i,j}(W_i, W'_j)] \\ &= \sum_{i \neq j} \Pr (H_{i,j}(W_i, W'_j) = 1) \\ &= \sum_{i \neq j} \sum_{a,b=1}^{\gamma} \Pr (W_i = e_{(a,b)}) \Pr (W'_j = e_{(a,b)}) \\ &= \sum_{i \neq j} \sum_{a,b=1}^{\gamma} \rho^2 \\ &\leq m^2 \gamma^2 \rho^2 \\ &= \lambda. \end{aligned} \tag{80}$$

For $d \geq 2$, Lemma 9 yields

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{(i,j) \in [m]^2} H_{i,j}(W_i, W'_j) \right)^d \right] &\leq C_0^{2d} \sum_{J \subseteq [2]} d^{|J|d} \sum_{\mathbf{i}_J} \mathbb{E}_J \left[\left(\sum_{\mathbf{i}_{J^c}} \mathbb{E}_{J^c} [H_{i_1, i_2}(W_{i_1}, W'_{i_2}) \mid \mathcal{F}_J] \right)^d \right] \\ &=: C_0^{2d} \sum_{J \subseteq [2]} Q_J, \end{aligned}$$

where $C_0 > 0$ is a universal constant, $\mathbf{i}_J := (i_l : l \in J)$, and $\mathbf{i}_{J^c} := (i_l : l \in J^c)$. Here, $\mathcal{F}_J$ and $\mathbb{E}_J$ are defined as in Lemma 9, with $(W_i)_{i=1}^m$ and $(W'_i)_{i=1}^m$ as the two independent coordinate blocks. We bound the four terms Q_J for $J = \emptyset, \{1\}, \{2\}, \{1, 2\}$ separately. When $J = \emptyset$,

$$\begin{aligned} Q_{\emptyset} &= \left(\sum_{i=j} \mathbb{E} [H_{i,j}(W_i, W'_j)] + \sum_{i \neq j} \mathbb{E} [H_{i,j}(W_i, W'_j)] \right)^d \\ &= \left(\sum_{i \neq j} \mathbb{E} [H_{i,j}(W_i, W'_j)] \right)^d \\ &\stackrel{(a)}{=} \left(\sum_{i \neq j} \Pr (H_{i,j}(W_i, W'_j) = 1) \right)^d \\ &= \left(\sum_{i \neq j} \sum_{a,b=1}^{\gamma} \Pr (W_i = e_{(a,b)}) \Pr (W'_j = e_{(a,b)}) \right)^d \\ &= \left(\sum_{i \neq j} \sum_{a,b=1}^{\gamma} \rho^2 \right)^d \\ &\leq (m^2 \gamma^2 \rho^2)^d \\ &= \lambda^d, \end{aligned} \tag{81}$$

where (a) follows from the fact that $H_{i,j}$ takes values in $\{0, 1\}$.

Next, consider $J = \{1\}$. For ease of notation, for $i \in [m]$, let $Q_{\{1\}}(i) := \sum_{j \in [m]} \mathbb{E} [H_{i,j}(W_i, W'_j) \mid W_i]$. Then,

$$Q_{\{1\}} = d^d \sum_{i \in [m]} \mathbb{E} \left[\left(\sum_{j \in [m]} \mathbb{E} [H_{i,j}(W_i, W'_j) \mid W_i] \right)^d \right]$$

$$= d^d \sum_{i \in [m]} \mathbb{E} [Q_{\{1\}}(i)^d].$$

If $W_i = 0$, then $Q_{\{1\}}(i) = 0$. On the other hand, if $W_i = e_{(a,b)}$ for some $a, b \in [\gamma]$, then $\Pr(W'_j = e_{(a,b)}) = \rho$ for every $j \neq i$. Therefore

$$Q_{\{1\}}(i) \leq m\rho \mathbb{1}\{W_i \neq 0\}.$$

As $\Pr(W_i \neq 0) = \gamma^2 \rho$, we obtain

$$\begin{aligned} Q_{\{1\}} &= d^d \sum_{i \in [m]} \mathbb{E} [Q_{\{1\}}(i)^d] \\ &\leq d^d \sum_{i=1}^m \mathbb{E} [m^d \rho^d \mathbb{1}\{W_i \neq 0\}] \\ &= d^d \sum_{i=1}^m (m^d \rho^d) \Pr(W_i \neq 0) \\ &= d^d m^{d+1} \gamma^2 \rho^{d+1} \\ &= d^d \lambda \left(\frac{k^2}{m} \right)^{d-1}. \end{aligned}$$

By symmetry, the same bound holds for $J = \{2\}$. Therefore, the total contribution from $J = \{1\}$ and $J = \{2\}$ is at most

$$Q_{\{1\}} + Q_{\{2\}} \leq 2d^d \lambda \left(\frac{k^2}{m} \right)^{d-1}. \quad (82)$$

Finally, consider $J = \{1, 2\}$. Since $H_{i,j}$ takes values in $\{0, 1\}$, we have

$$H_{i,j}(W_i, W'_j)^d = H_{i,j}(W_i, W'_j).$$

Thus,

$$\begin{aligned} Q_{\{1,2\}} &= d^{2d} \sum_{(i,j) \in [m]^2} \mathbb{E} [H_{i,j}(W_i, W'_j)^d] \\ &= d^{2d} \sum_{(i,j) \in [m]^2} \mathbb{E} [H_{i,j}(W_i, W'_j)] \\ &= d^{2d} \sum_{(i,j) \in [m]^2} \Pr(H_{i,j}(W_i, W'_j) = 1) \\ &= d^{2d} \sum_{i \neq j} \Pr(H_{i,j}(W_i, W'_j) = 1) \\ &= d^{2d} \sum_{i \neq j} \sum_{a,b=1}^{\gamma} \Pr(W_i = e_{(a,b)}) \Pr(W'_j = e_{(a,b)}) \\ &\leq d^{2d} m^2 \gamma^2 \rho^2 \\ &= d^{2d} \lambda. \end{aligned} \quad (83)$$

Combining the bounds Equations (81) to (83), we obtain, for any integer $d \geq 2$,

$$\mathbb{E} \left[\left(\sum_{(i,j) \in [m]^2} H_{i,j}(W_i, W'_j) \right)^d \right] \leq C_0^{2d} \left[\lambda^d + 2d^d \lambda \left(\frac{k^2}{m} \right)^{d-1} + d^{2d} \lambda \right],$$

for some universal constant $C_0 > 0$.

Combining the above bound with Equation (80), which deals with the case of $d = 1$, and with Equation (79), and taking a sufficiently large universal constant $C \geq 1$ to absorb the constants from decoupling and the moment bounds, we arrive at the bound

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} \left[L_{\tilde{\nu}}^{\leq D}(\mathbf{Y})^2 \right] \leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} C^d \left[\lambda^d + 2d^d \lambda \left(\frac{k^2}{m} \right)^{d-1} + d^{2d} \lambda \right].$$

We now handle the terms on the right-hand side of the last expression. For the first term, we have

$$\sum_{d=1}^D \frac{\mu^{2d}}{d!} C^d \lambda^d \leq \exp(C\mu^2 \lambda) - 1,$$

which is $O(1)$ whenever $\mu = O(m/(\gamma k^2))$. Next, we bound the quantity

$$\sum_{d=1}^D \frac{\mu^{2d}}{d!} C^d \left[2d^d \lambda \left(\frac{k^2}{m} \right)^{d-1} + d^{2d} \lambda \right].$$

We distinguish two cases.

Case 1: $\lambda \leq 1$. Since $\lambda = \gamma^2(k^2/m)^2$ and $\gamma \geq 1$, we have $k^2/m \leq 1$. Therefore, for every $d \geq 1$,

$$2d^d \lambda \left(\frac{k^2}{m} \right)^{d-1} + d^{2d} \lambda \leq 3\lambda d^{2d}.$$

Using $d! \geq (d/e)^d$, we obtain

$$\sum_{d=1}^D \frac{\mu^{2d}}{d!} C^d \left[2d^d \lambda \left(\frac{k^2}{m} \right)^{d-1} + d^{2d} \lambda \right] \leq 3\lambda \sum_{d=1}^D (eC\mu^2 d)^d \leq 3\lambda \sum_{d=1}^D (eC\mu^2 D)^d,$$

which is of order $O(1)$ whenever $\mu^2 D \leq c_1$ for a sufficiently small constant $c_1 > 0$.

Case 2: $\lambda > 1$. We first note that, for every $d \geq 1$,

$$\lambda \left(\frac{k^2}{m} \right)^{d-1} \leq \lambda^d. \tag{84}$$

When $k^2/m \leq 1$, this bound is straightforward. On the other hand, if $k^2/m > 1$, then

$$\lambda \left(\frac{k^2}{m} \right)^{d-1} = \gamma^2 \left(\frac{k^2}{m} \right)^{d+1} \leq \gamma^{2d} \left(\frac{k^2}{m} \right)^{2d} = \lambda^d,$$

as claimed, where we used the facts that $d+1 \leq 2d$ and $\gamma \geq 1$.

Using Equation (84) and since $\lambda > 1$,

$$2d^d \lambda \left(\frac{k^2}{m} \right)^{d-1} + d^{2d} \lambda \leq 3d^{2d} \lambda^d.$$

It follows that

$$\begin{aligned} \sum_{d=1}^D \frac{\mu^{2d}}{d!} C^d \left[2d^d \lambda \left(\frac{k^2}{m} \right)^{d-1} + d^{2d} \lambda \right] &\leq 3 \sum_{d=1}^D \frac{(C\mu^2 \lambda)^d d^{2d}}{d!} \\ &\leq 3 \sum_{d=1}^D (eC\mu^2 \lambda d)^d \\ &\leq 3 \sum_{d=1}^D (eC\mu^2 \lambda D)^d, \end{aligned}$$

where in the second inequality we use again the bound $d! \geq (d/e)^d$. The last quantity is $O(1)$ whenever $\mu^2 \lambda D \leq c_2$ for a sufficiently small constant $c_2 > 0$.

A sufficient condition for both cases is

$$\mu \leq \frac{c}{\sqrt{D}} \left(\frac{m}{\gamma k^2} \wedge 1 \right),$$

where $c = \min\{\sqrt{c_1}, \sqrt{c_2}\}$. For $D \geq 1$, this condition also gives $\mu^2 \lambda = O(1)$, so the degree- D likelihood-ratio second moment is $O(1)$. By Lemma 2, no degree- D polynomial strongly separates the two hypotheses in this regime.

B.15.2 Multiple random-size hypercliques

Throughout this part, $t \geq 3$ is fixed, and constants may depend on t but not on m, γ, k, d , or D . We recall Equation (43) and define the kernel function $h : (\{0, 1\}^{\gamma \times \gamma})^t \rightarrow \mathbb{R}$ by

$$h(w_1, \dots, w_t) := \sum_{a,b=1}^{\gamma} \prod_{l=1}^t \mathbb{1}\{w_l = e_{(a,b)}\}.$$

Following the analysis in Equation (78), the overlap size $|S \cap S'|$ can be expressed as

$$|S \cap S'| = \sum_{\mathbf{i} \in [m]_t} h(W_{i_1}, \dots, W_{i_t}), \quad (85)$$

which can be seen as a U -statistic of order t with kernel h . For the decoupled version, let us define the kernel

$$H_{\mathbf{i}}(w_1, \dots, w_t) := \mathbb{1}\{i_1, \dots, i_t \text{ are pairwise distinct}\} h(w_1, \dots, w_t),$$

for each $\mathbf{i} \in [m]^t$. Note that $H_{\mathbf{i}}$ is zero whenever any two indices coincide, and is equal to $h(w_1, \dots, w_t)$ otherwise. Therefore, since h is symmetric, we obtain

$$\begin{aligned} |S \cap S'| &= \sum_{\mathbf{i} \in [m]_t} h(W_{i_1}, \dots, W_{i_t}) \\ &= \frac{1}{t!} \sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(W_{i_1}, \dots, W_{i_t}) \\ &=: U(W_1, \dots, W_m). \end{aligned}$$

Let the decoupled version of U be defined as

$$U'(W_1^{(1)}, \dots, W_m^{(1)}; \dots; W_1^{(t)}, \dots, W_m^{(t)}) = \frac{1}{t!} \sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(W_{i_1}^{(1)}, \dots, W_{i_t}^{(t)}),$$

where the t sequences $(W_i^{(j)})_{i=1}^m$, $j \in [t]$, are mutually independent copies of $(W_i)_{i=1}^m$ and are independent of the original sequence. Then, by Lemma 8, there exists a constant $c > 0$ depending only on t such that

$$\mathbb{E} \left[\left(\sum_{\mathbf{i} \in [m]_t} H_{\mathbf{i}}(W_{i_1}, \dots, W_{i_t}) \right)^d \right] \leq c^d \mathbb{E} \left[\left(\sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(W_{i_1}^{(1)}, \dots, W_{i_t}^{(t)}) \right)^d \right],$$

for every integer $d \geq 1$, which implies that

$$\mathbb{E}[|S \cap S'|^d] \leq c'^d \mathbb{E} \left[\left(\sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) \right)^d \right], \quad (86)$$

for some constant $c' > 0$ depending only on t and $\mathbf{W}_i := (W_{i_1}^{(1)}, \dots, W_{i_t}^{(t)})$. For $d = 1$, the decoupled sum has expectation at most $m^t \gamma^2 \rho^t$. For integers $d \geq 2$, we apply Lemma 9 to bound the d -th moment:

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) \right)^d \right] &\leq C^{td} \sum_{J \subseteq [t]} \left[d^{|J|d} \sum_{\mathbf{i}_J} \mathbb{E}_J \left\{ \left(\sum_{\mathbf{i}_{J^c} \in [m]^{t-|J|}} \mathbb{E}_{J^c} [H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) \mid \mathcal{F}_J] \right)^d \right\} \right] \\ &= C^{td} \sum_{r=0}^t \sum_{\substack{J \subseteq [t] \\ |J|=r}} \left[d^{rd} \sum_{\mathbf{i}_J \in [m]^r} \mathbb{E}_J \left\{ \left(\sum_{\mathbf{i}_{J^c} \in [m]^{t-r}} \mathbb{E}_{J^c} [H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) \mid \mathcal{F}_J] \right)^d \right\} \right], \end{aligned}$$

where $\mathcal{F}_J = \sigma(W_i^{(j)} : j \in J, i \in [m])$, J^c is the complement of J in $[t]$, $\sum_{\mathbf{i}_\emptyset} a = a$ and $\mathbb{E}_\emptyset[a] = a$ by convention. Let

$$\begin{aligned} Q_J(\mathbf{i}_J) &:= \sum_{\mathbf{i}_{J^c} \in [m]^{t-r}} \mathbb{E}_{J^c} [H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) \mid \mathcal{F}_J] \\ &= \sum_{\mathbf{i}_{J^c} \in [m]^{t-r}} \Pr_{J^c} (H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) = 1 \mid \mathcal{F}_J), \end{aligned}$$

where $r = |J|$. For $J \neq \emptyset$, let $\mathcal{E}_{\mathbf{i}_J}$ be the event that the r variables $(W_{i_j}^{(j)})_{j \in J}$ have the same nonzero label, i.e.,

$$\mathcal{E}_{\mathbf{i}_J} = \{ \exists (a, b) \in [\gamma]^2 \text{ such that } \forall j \in J, W_{i_j}^{(j)} = e_{(a,b)} \}.$$

Then, under the event $\mathcal{E}_{\mathbf{i}_J}$, the only way to have $H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) = 1$ is to have $W_{i_j}^{(j)} = e_{(a,b)}$ for all $j \in J^c$ as well, where (a, b) is the common label determined by the variables indexed by J . By independence, these remaining labels match with probability ρ^{t-r} , where $\rho := k^2/m^2$; repeated vertex indices can only make $H_{\mathbf{i}}$ vanish. On the other hand, if $\mathcal{E}_{\mathbf{i}_J}$ does not hold, we have $H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) = 0$ almost surely. Therefore, we have

$$Q_J(\mathbf{i}_J) = \sum_{\mathbf{i}_{J^c} \in [m]^{t-r}} \Pr_{J^c} (H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) = 1 \mid \mathcal{F}_J) \leq m^{t-r} \rho^{t-r} \mathbf{1}\{\mathcal{E}_{\mathbf{i}_J}\}.$$

Also note that $\Pr(\mathcal{E}_{\mathbf{i}_J}) = \gamma^2 \rho^r$ for $r \geq 1$. Therefore,

$$\begin{aligned} &\sum_{\substack{J \subseteq [t] \\ |J|=r}} \sum_{\mathbf{i}_J \in [m]^r} \mathbb{E}_J \left\{ \left(\sum_{\mathbf{i}_{J^c} \in [m]^{t-r}} \mathbb{E}_{J^c} [H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) \mid \mathcal{F}_J] \right)^d \right\} \\ &= \sum_{\substack{J \subseteq [t] \\ |J|=r}} \sum_{\mathbf{i}_J \in [m]^r} \mathbb{E}_J [Q_J(\mathbf{i}_J)^d] \\ &\leq \sum_{\substack{J \subseteq [t] \\ |J|=r}} \sum_{\mathbf{i}_J \in [m]^r} m^{(t-r)d} \rho^{(t-r)d} \Pr(\mathcal{E}_{\mathbf{i}_J}) \\ &\leq \binom{t}{r} m^r (m\rho)^{(t-r)d} \gamma^2 \rho^r \\ &= \binom{t}{r} \gamma^2 (m\rho)^{(t-r)d+r}, \end{aligned}$$

for $r \geq 1$. For $J = \emptyset$, we have

$$Q_\emptyset^d = \left(\sum_{\mathbf{i} \in [m]^t} \Pr (H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) = 1) \right)^d \leq (m^t \cdot \gamma^2 \rho^t)^d.$$

Summarizing the above analysis, we obtain

$$\mathbb{E} \left[\left(\sum_{\mathbf{i} \in [m]^t} H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) \right)^d \right] \leq C^{td} \sum_{r=0}^t \sum_{\substack{J \subseteq [t] \\ |J|=r}} \left[d^{rd} \sum_{\mathbf{i}_J \in [m]^r} \mathbb{E}_J \left\{ \left(\sum_{\mathbf{i}_{J^c} \in [m]^{t-r}} \mathbb{E}_{J^c} [H_{\mathbf{i}}(\mathbf{W}_{\mathbf{i}}) \mid \mathcal{F}_J] \right)^d \right\} \right]$$

$$\begin{aligned}
&\leq C^{td} \left[(m^t \cdot \gamma^2 \rho^t)^d + \sum_{r=1}^t \binom{t}{r} d^{rd} \gamma^2 (m\rho)^{(t-r)d+r} \right] \\
&\leq C^{td} \left[\gamma^{2d} \left(\frac{k^2}{m} \right)^{td} + \gamma^2 \sum_{r=1}^t \binom{t}{r} d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)d+r} \right] \\
&= C^{td} \left[\lambda^d + \lambda \sum_{r=1}^t \binom{t}{r} d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} \right] \\
&\leq C_1^{td} \left[\lambda^d + \lambda \sum_{r=1}^t d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} \right],
\end{aligned}$$

where $\lambda = \gamma^2(k^2/m)^t$ and $C_1 = 2C$, and the last inequality follows from the fact that $\binom{t}{r} \leq 2^t$ for all $r \in [t]$. The same bound holds for $d = 1$ after enlarging C_1 if necessary, by the first-moment bound above. Therefore, combining this bound with Equation (86) and further enlarging C_1 to absorb the decoupling constant, we have

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}} &\leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} \left[\lambda^d + \lambda \sum_{r=1}^t d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} \right] \\
&\leq 1 + \sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} \lambda^d + \sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} \lambda \sum_{r=1}^t d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)},
\end{aligned}$$

and it is sufficient to show

$$\text{(I)} : \sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} \lambda^d = O(1), \quad \text{(II)} : \sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} \lambda \sum_{r=1}^t d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} = O(1).$$

For (I), note that

$$\sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} \lambda^d \leq \exp(\mu^2 C_1^t \lambda) - 1,$$

and thus it suffices to show $\mu^2 C_1^t \lambda = O(1)$, which is equivalent to $\mu = O(\frac{m^{t/2}}{\gamma k^t})$. For (II), we split into two cases.

- **Case 1:** $\lambda \leq 1$. In this case, observe that

$$\frac{k^2}{m} \leq \lambda^{1/t} \leq 1,$$

since $\gamma \geq 1$. Therefore, for every $r \in [t]$ and $d \in \mathbb{N}$, we have

$$\lambda \sum_{r=1}^t d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} \leq \lambda \sum_{r=1}^t d^{td} \leq t d^{td},$$

as $\lambda \leq 1$ and $d^{rd} \leq d^{td}$. Therefore, note that (II) can be upper bounded by

$$\begin{aligned}
\sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} \lambda \sum_{r=1}^t d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} &\leq \sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} t d^{td} \\
&\leq t \sum_{d=1}^D \mu^{2d} e^d C_1^{td} d^{(t-1)d} \quad (\because d! \geq (d/e)^d) \\
&\leq t \sum_{d=1}^D \mu^{2d} e^d C_1^{td} D^{(t-1)d}
\end{aligned}$$

$$= t \sum_{d=1}^D (e\mu^2 C_1^t D^{t-1})^d,$$

which is $O(1)$ when $\mu^2 D^{t-1} \leq c_1$, for a sufficiently small constant $1 > c_1 > 0$.

- **Case 2:** $\lambda > 1$. If $\frac{k^2}{m} \leq 1$, we follow the same argument as in Case 1 to obtain

$$\lambda d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} \leq \lambda d^{rd} \leq \lambda d^{td} \leq \lambda^d d^{td},$$

as $\lambda > 1$. On the other hand, if $\frac{k^2}{m} > 1$, then observe that for every $r \in [t]$ and $d \in \mathbb{N}$,

$$(t-r)d + r \leq td,$$

and thus

$$\lambda d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} = d^{rd} \gamma^2 \left(\frac{k^2}{m} \right)^{(t-r)d+r} \leq d^{td} \gamma^2 \left(\frac{k^2}{m} \right)^{td}.$$

Since

$$\gamma^2 \left(\frac{k^2}{m} \right)^{td} \leq \gamma^{2d} \left(\frac{k^2}{m} \right)^{td} = \lambda^d$$

we have

$$\lambda d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} \leq d^{td} \lambda^d.$$

Therefore, in either case, we have

$$\lambda \sum_{r=1}^t d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} \leq t \lambda^d d^{td},$$

and thus (II) can be upper bounded by

$$\begin{aligned} \sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} \lambda \sum_{r=1}^t d^{rd} \left(\frac{k^2}{m} \right)^{(t-r)(d-1)} &\leq \sum_{d=1}^D \frac{\mu^{2d}}{d!} C_1^{td} t \lambda^d d^{td} \\ &\leq \sum_{d=1}^D \mu^{2d} e^d C_1^{td} t \lambda^d d^{(t-1)d} \\ &\leq t \sum_{d=1}^D (e\mu^2 C_1^t \lambda D^{t-1})^d, \end{aligned}$$

which is $O(1)$ when $\mu^2 \lambda D^{t-1} \leq c_2$, for a sufficiently small constant $c_2 > 0$.

Combining the above two cases, we conclude that a sufficient condition is

$$\mu \leq \frac{c}{D^{(t-1)/2}} \left(\frac{m^{t/2}}{\gamma k^t} \wedge 1 \right),$$

for $c = \min\{\sqrt{c_1}, \sqrt{c_2}\}$. For $D \geq 1$, this condition also bounds term (I), so the degree- D likelihood-ratio second moment is $O(1)$ and Lemma 2 rules out strong separation by degree- D polynomials. For fixed t and polylogarithmic D , the factors $D^{-1/2}$ and $D^{-(t-1)/2}$ in the two bounds are logarithmic losses. Thus, when D satisfies the degree requirement in Conjecture 1, these estimates recover the computational lower bounds in Theorem 2 and Corollary 2, up to logarithmic factors, under that conjecture.

C Additional examples of latent structures

C.1 Examples from [Addario-Berry et al. \(2010\)](#)

In this subsection, we collect the results from other combinatorial testing examples discussed in [Addario-Berry et al. \(2010\)](#) for completeness. First, note that the second moment of the low-degree likelihood ratio is always bounded above by that of the full likelihood ratio, i.e.,

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\nu}^{\leq D}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu} \left[\sum_{d=0}^D \frac{\mu^{2d}}{d!} |S \cap S'|^d \right] \leq \mathbb{E}_{S, S' \sim \nu} [\exp(\mu^2 |S \cap S'|)] = \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\nu}(\mathbf{Y})^2]. \quad (87)$$

This implies that a statistical lower bound obtained by controlling the second moment of the full likelihood ratio also yields a computational lower bound through the low-degree framework. Here, the lower bound against polynomial-time tests is unconditional, since statistical impossibility already rules out every test. This observation is particularly useful for problems in which there is no statistical-computational gap. Indeed, for the problems considered in this subsection, the statistical lower bounds established by [Addario-Berry et al. \(2010\)](#) are matched by upper bounds achieved by polynomial-time tests, so no separate analysis of computational lower bounds is necessary.

Accordingly, our goal in this subsection is not to establish new lower-bound results for these problems. Rather, we aim to complement the analysis of [Addario-Berry et al. \(2010\)](#), which relies on structural properties of the underlying set classes, such as symmetry and negative association, by providing alternative proofs based on more explicit characterizations of the overlap distribution $|S \cap S'|$ or problem-specific combinatorial arguments.

Throughout this subsection, we assume that the prior ν is uniform over $\mathcal{C}$ and that $|s| = K$ for all $s \in \mathcal{C}$. Also, we denote $N = |\mathcal{C}|$. In each overlap calculation, S and S' are independent draws from ν , and $C > 1$ is a fixed second-moment bound independent of the ambient dimension. For each problem, we aim to derive a sufficient condition on μ under which $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\nu}(\mathbf{Y})^2]$ remains bounded by a constant, thereby establishing the impossibility of strong detection via Lemma 1. We also identify the test that achieves a matching upper bound.

C.1.1 Disjoint sets

Consider the case where $\mathcal{C}$ is the collection of N disjoint sets of size K . In this case, $|S \cap S'|$ is a random variable that takes the value K with probability $1/N$ and 0 otherwise. Therefore, we have

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\nu}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu}[\exp(\mu^2 |S \cap S'|)] = \frac{1}{N} \exp(\mu^2 K) + \frac{N-1}{N},$$

and thus, if

$$\mu \leq \sqrt{\frac{1}{K} \log(1 + N(C-1))},$$

then $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\nu}(\mathbf{Y})^2]$ is bounded by C . When $N \rightarrow \infty$, the matching upper bound (up to a constant factor) can be obtained by the scan test in Proposition 2, which runs in $O(NK)$ time, hence $O(n)$ time, because the candidate sets are disjoint. If N remains bounded, the corresponding condition for strong detection is $\mu\sqrt{K} \rightarrow \infty$.

C.1.2 K -subsets

We consider the case where $\mathcal{C}$ consists of all K -subsets of $[n]$. In this case, the overlap size $|S \cap S'|$ between two randomly chosen K -sets S and S' from $\mathcal{C}$ follows a hypergeometric distribution with parameters (n, K, K) . By the convex-order comparison with a binomial random variable ([Hoeffding, 1963](#), Theorem 4), we have

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_{\nu}(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu}[\exp(\mu^2 |S \cap S'|)]$$

$$\begin{aligned}
&\leq \mathbb{E}[\exp(\mu^2 B)] \\
&= \left(1 + \frac{K}{n}(e^{\mu^2} - 1)\right)^K \\
&\leq \exp\left(\frac{K^2}{n}(e^{\mu^2} - 1)\right),
\end{aligned}$$

where $B \sim \text{Binomial}(K, K/n)$. Therefore, if

$$\mu \leq \sqrt{\log\left(1 + \frac{n}{K^2} \log C\right)},$$

then $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2]$ is bounded by C . We highlight two regimes. (i) If $K^2 = O(n^{1-\delta})$ for some fixed $\delta \in (0, 0.1)$, then $\log(1 + \frac{n}{K^2} \log C) \geq \delta \log n + O(1)$, so $\log(1 + \frac{n}{K^2} \log C) \geq \frac{\delta}{2} \log n$ for sufficiently large n . Therefore, it suffices to have $\mu \leq \sqrt{\frac{\delta}{2} \log n}$. A matching upper bound up to a constant factor is achieved by the outlier test in Proposition 2, which runs in $O(n)$ time. (ii) If $K^2 = \omega(n)$, then $\frac{n}{2K^2} \log C \leq \log(1 + \frac{n}{K^2} \log C)$ for sufficiently large n . Thus, the sufficient condition is $\mu \leq \sqrt{\frac{1}{2} \frac{n}{K^2} \log C}$. A matching signal scale is achieved by the average test in Proposition 2, which runs in $O(n)$ time and achieves strong detection whenever $\mu K / \sqrt{n} \rightarrow \infty$.

C.1.3 Perfect matchings

We consider the case where $\mathcal{C}$ consists of all perfect matchings in a complete bipartite graph $K_{m,m}$. We have $n = m^2$, $N = m!$, and $K = m$. As pointed out in [Addario-Berry et al. \(2010\)](#), the overlap size $|S \cap S'|$ between two randomly chosen perfect matchings S and S' from $\mathcal{C}$ follows the distribution of the number of fixed points in a random permutation of size m ; let us denote this random variable by Z , i.e., $Z := \sum_{i=1}^m \mathbf{1}\{\pi(i) = i\}$, where π is a uniformly random permutation of $[m]$. Then, we have

$$\Pr(Z = r) = \frac{\binom{m}{r} D_{m-r}}{m!}, \quad r = 0, 1, \dots, m,$$

where D_{m-r} is the number of derangements of size $m-r$, with $D_0 = 1$. Hence,

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2] &= \mathbb{E}_{S, S' \sim \nu}[\exp(\mu^2 |S \cap S'|)] \\
&= \mathbb{E}_Z[\exp(\mu^2 Z)] \\
&= \sum_{r=0}^m \exp(\mu^2 r) \frac{\binom{m}{r} D_{m-r}}{m!} \\
&= \sum_{r=0}^m \exp(\mu^2 r) \frac{1}{r!} \sum_{j=0}^{m-r} \frac{(-1)^j}{j!} \\
&= \sum_{\substack{r, j \geq 0 \\ r+j \leq m}} \frac{\exp(\mu^2 r) (-1)^j}{r! j!} \\
&= \sum_{k=0}^m \frac{1}{k!} \sum_{r=0}^k \binom{k}{r} (\exp(\mu^2))^r (-1)^{k-r} \quad \text{where } k := r + j \\
&= \sum_{k=0}^m \frac{1}{k!} (\exp(\mu^2) - 1)^k \\
&\leq \exp(\exp(\mu^2) - 1).
\end{aligned}$$

Therefore, if

$$\mu \leq \sqrt{\log(1 + \log C)},$$

then $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2]$ is bounded by C . Thus, every bounded sequence $\mu = O(1)$ precludes strong detection, whereas the average test in Proposition 2 achieves strong detection whenever $\mu \rightarrow \infty$, since $K/\sqrt{n} = 1$.

Negative association of matching indicators. [Addario-Berry et al. \(2010, Section 4.3\)](#) establish the same statistical lower bound for perfect matchings by invoking Proposition 3.4, which requires negative association of the edge-inclusion indicators along a fixed reference matching. As the calculation below shows, these indicators are not negatively associated when $m \geq 2$. The exact overlap calculation above nevertheless yields the stated lower bound without this assumption. For completeness, we recall the definition of negative association and give a simple counterexample.

Definition 6 (Negative association ([Addario-Berry et al., 2010](#))). *A collection of random variables $Y_1, \dots, Y_n$ is said to be negatively associated if for any pair of disjoint subsets $A_1, A_2 \subset [n]$ and any two coordinatewise nondecreasing functions $f : \mathbb{R}^{|A_1|} \rightarrow \mathbb{R}$ and $g : \mathbb{R}^{|A_2|} \rightarrow \mathbb{R}$,*

$$\mathbb{E}[f(Y_i : i \in A_1)g(Y_j : j \in A_2)] \leq \mathbb{E}[f(Y_i : i \in A_1)]\mathbb{E}[g(Y_j : j \in A_2)].$$

For $m \geq 2$, fix the reference matching $S' = \{(1, 1), (2, 2), \dots, (m, m)\}$ by symmetry, and let S be uniformly distributed over $\mathcal{C}$. For $(i, j) \in S'$, define the edge-inclusion indicator $I_{i,j} := \mathbb{1}\{(i, j) \in S\}$. Consider $I_{1,1}$ and $I_{2,2}$, and take $f(x) = g(x) = x$. Then, we have

$$\mathbb{E}[I_{1,1}I_{2,2}] = \Pr((1, 1) \in S, (2, 2) \in S) = \frac{(m-2)!}{m!} = \frac{1}{m(m-1)},$$

and

$$\mathbb{E}[I_{1,1}]\mathbb{E}[I_{2,2}] = \Pr((1, 1) \in S) \Pr((2, 2) \in S) = \frac{(m-1)!}{m!} \cdot \frac{(m-1)!}{m!} = \frac{1}{m^2}.$$

Since $1/(m(m-1)) > 1/m^2$ for all $m \geq 2$, this shows that $I_{1,1}$ and $I_{2,2}$ are not negatively associated.

C.1.4 Stars

Consider a network of m nodes, where a *star* is a subgraph of the complete graph consisting of all $K = m - 1$ edges incident to a fixed node. Consider the set $\mathcal{C}$ of all stars in the complete graph. For $m \geq 3$, we have $n = \binom{m}{2}$ and $N = m$. In this case, the overlap size $|S \cap S'|$ has the distribution

$$|S \cap S'| = \begin{cases} m-1, & \text{with probability } \frac{1}{m}, \quad \text{the centers coincide} \\ 1, & \text{with probability } \frac{m-1}{m}, \quad \text{otherwise.} \end{cases}$$

Therefore, we have

$$\begin{aligned} \mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2] &= \mathbb{E}_{S, S' \sim \nu}[\exp(\mu^2 |S \cap S'|)] \\ &= \frac{1}{m} \exp(\mu^2(m-1)) + \frac{m-1}{m} \exp(\mu^2) \\ &= \frac{1}{m} (\exp(\mu^2(m-1)) - 1) + \frac{m-1}{m} (\exp(\mu^2) - 1) + 1 \\ &\leq \frac{2}{m} (\exp(\mu^2(m-1)) - 1) + 1, \end{aligned}$$

where the inequality follows from $e^{(m-1)\mu^2} - 1 \geq (m-1)(e^{\mu^2} - 1)$ by the convexity of $f(x) = e^x$. Thus, if

$$\mu \leq \sqrt{\frac{1}{m-1} \log \left(1 + \frac{m(C-1)}{2} \right)},$$

then $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2]$ is bounded by C . Note that the scan test achieves the matching upper bound; all star sums can be accumulated in one pass over the n observations, so this scan takes $O(n)$ time.

C.1.5 Spanning trees

Consider a network of m nodes, where a *spanning tree* is a subgraph of the complete graph which is a tree and connects all m nodes. We have $n = \binom{m}{2}$, $N = m^{m-2}$, and $K = m - 1$. Note that while [Addario-Berry et al. \(2010\)](#) provide two different ways to obtain a statistical lower bound in this case, we provide an alternative approach focusing on combinatorial properties of spanning trees. A key step is to interpret each realization of $S \cap S'$ such that $|S \cap S'| = r$ as a forest with m nodes and r edges, including isolated vertices. For example, when $m = 6$ and $r = 3$, one possible realization of $S \cap S'$ is $\{(1, 2), (2, 3), (4, 5)\}$, which can be interpreted as a forest F with $c = 3$ components whose vertex sets are $\{\{1, 2, 3\}, \{4, 5\}, \{6\}\}$. We identify each such forest F with its edge set; its vertex set is always $[m]$. Write $F \subset A$ for the event that all edges of F belong to the edge set A ; for example, $F \subset S \cap S'$ means that all edges in the forest F are included in both S and S' . Write $r(F)$ and $c(F)$ for the numbers of edges and components of F , respectively; then $c(F) = m - r(F)$. We introduce a useful lemma to count the number of spanning trees containing a given forest:

Lemma 11 ([Moon 1970](#), Theorem 6.1). *Let F be a forest on vertex set $[m]$ with $c(F)$ components. If we denote the number of nodes in each component as $m_1, \dots, m_{c(F)}$, then the number of spanning trees containing F is given by*

$$m^{c(F)-2} \prod_{i=1}^{c(F)} m_i.$$

Therefore, given a forest F , the probability that independent random spanning trees S and S' both contain F is given by

$$\begin{aligned} \Pr(F \subset S \cap S') &= \Pr(F \subset S) \Pr(F \subset S') \\ &= \left(\frac{m^{c(F)-2} \prod_{i=1}^{c(F)} m_i}{m^{m-2}} \right)^2 \\ &\stackrel{(a)}{=} \left(\frac{\prod_{i=1}^{c(F)} m_i}{m^{r(F)}} \right)^2 \\ &\stackrel{(b)}{\leq} \left(\frac{2^{\sum_i (m_i-1)}}{m^{r(F)}} \right)^2 \\ &= \left(\frac{4}{m^2} \right)^{r(F)}, \end{aligned}$$

where (a) follows from the relationship $c(F) = m - r(F)$, and (b) follows from the fact that $m_i \leq 2^{m_i-1}$ for all $m_i \geq 1$. Now, denoting $Z := |S \cap S'|$, observe that, for integers $0 \leq R \leq m - 1$, the R -th falling factorial of Z can be expressed as

$$(Z)_R = Z(Z-1)\dots(Z-R+1) = \sum_{I_R \in [m]_{2,R}} \mathbf{1}\{(i_1, j_1), \dots, (i_R, j_R) \in S \cap S'\},$$

where we denote an ordered R -tuple of edges as $I_R := ((i_1, j_1), \dots, (i_R, j_R))$, and the collection of ordered R -tuples of distinct edges as $[m]_{2,R} := \{I_R : (i_a, j_a) \in [m]_2, (i_a, j_a) \neq (i_b, j_b) \text{ for all } 1 \leq a \neq b \leq R\}$. We use the conventions $(Z)_0 = 1$ and $[m]_{2,0} = \{()\}$. Then, the R -th factorial moment can be expressed as

$$\mathbb{E}[(Z)_R] = \sum_{I_R \in [m]_{2,R}} \Pr((i_1, j_1), \dots, (i_R, j_R) \in S \cap S').$$

Now, observe that if the edges $(i_1, j_1), \dots, (i_R, j_R)$ contain a cycle, then there is no spanning tree containing all those edges, and thus $\Pr((i_1, j_1), \dots, (i_R, j_R) \in S \cap S') = 0$. Therefore, the only nonzero contributions to the above summation come from those I_R that form a forest with R edges. Also, note that we can bound the size of the collection as $|[m]_{2,R}| \leq n(n-1)\dots(n-R+1) \leq n^R$, where $n = \binom{m}{2}$.

$$\mathbb{E}[(Z)_R] = \sum_{I_R \in [m]_{2,R}} \Pr((i_1, j_1), \dots, (i_R, j_R) \in S \cap S')$$

$$\begin{aligned}
&\leq n^R \left(\frac{4}{m^2} \right)^R \\
&= \left(\frac{2(m-1)}{m} \right)^R \\
&\leq 2^R.
\end{aligned}$$

For $t \geq 0$, we have

$$\exp(tZ) = (1 + (e^t - 1))^Z = \sum_{R=0}^Z \frac{(e^t - 1)^R}{R!} (Z)_R.$$

Hence, we rewrite the moment generating function of Z using the factorial moments as

$$\mathbb{E}[\exp(tZ)] = \sum_{R=0}^{m-1} \frac{(e^t - 1)^R}{R!} \mathbb{E}[(Z)_R] \leq \sum_{R=0}^{m-1} \frac{(e^t - 1)^R}{R!} 2^R \leq \sum_{R=0}^{\infty} \frac{(2(e^t - 1))^R}{R!} = \exp(2(e^t - 1)).$$

Therefore, we have

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu}[\exp(\mu^2 |S \cap S'|)] \leq \exp(2(e^{\mu^2} - 1)).$$

Thus, if

$$\mu \leq \sqrt{\log \left(1 + \frac{\log C}{2} \right)},$$

then $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2]$ is bounded by C . The matching signal scale is achieved by the average test, which strongly detects whenever $\mu \rightarrow \infty$, since $K/\sqrt{n} = (m-1)/\sqrt{\binom{m}{2}} \rightarrow \sqrt{2}$.

C.2 Further examples

As in the preceding subsection, let ν be uniform over $\mathcal{C}$, let S, S' be independent draws from ν , and fix a second-moment bound $C > 1$ independent of the ambient dimension.

C.2.1 k -vertex stars

We consider a variant of the star graph in Appendix C.1.4, consisting of k vertices and hence $K = k-1$ edges, where $3 \leq k \leq m-1$. In this case, we have $n = \binom{m}{2}$ and $N = \binom{m}{k} \cdot k$. Note that the overlap size $|S \cap S'|$ between two randomly chosen k -vertex stars S and S' from $\mathcal{C}$ has the following conditional distributions:

$$|S \cap S'| = \begin{cases} A \sim \text{Hypergeometric}(m-1, k-1, k-1), & \text{if the centers coincide} \\ B \sim \text{Bernoulli}(\frac{k-1}{m-1}), & \text{if the center of } S' \text{ is a leaf of } S \\ 0, & \text{otherwise.} \end{cases}$$

These three cases have probabilities $1/m$, $(k-1)/m$, and $(m-k)/m$, respectively. Therefore, denoting $p = \frac{k-1}{m-1}$, we have

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2] &= \mathbb{E}_{S, S' \sim \nu}[\exp(\mu^2 |S \cap S'|)] \\
&= \frac{1}{m} \mathbb{E}[\exp(\mu^2 A)] + \frac{k-1}{m} \mathbb{E}[\exp(\mu^2 B)] + \frac{m-k}{m} \cdot 1 \\
&= \frac{1}{m} \mathbb{E}[\exp(\mu^2 A)] + \frac{k-1}{m} [p\{\exp(\mu^2) - 1\} + 1] + \frac{m-k}{m} \\
&= \frac{1}{m} \mathbb{E}[\exp(\mu^2 A)] + \frac{m-1}{m} + \frac{(k-1)^2}{m(m-1)} \{\exp(\mu^2) - 1\} \\
&\stackrel{(a)}{\leq} \frac{1}{m} [p\{\exp(\mu^2) - 1\} + 1]^{k-1} + \frac{m-1}{m} + \frac{(k-1)^2}{m(m-1)} \{\exp(\mu^2) - 1\}
\end{aligned}$$

$$\begin{aligned}
&= 1 + \frac{1}{m} \left([p\{\exp(\mu^2) - 1\} + 1]^{k-1} - 1 \right) + \frac{(k-1)^2}{m(m-1)} \{ \exp(\mu^2) - 1 \} \\
&= 1 + \frac{1}{m} \left(\exp \left\{ (k-1) \log [p\{\exp(\mu^2) - 1\} + 1] \right\} - 1 \right) + \frac{(k-1)^2}{m(m-1)} \{ \exp(\mu^2) - 1 \} \\
&\stackrel{(b)}{\leq} 1 + \frac{1}{m} \left(\exp [(k-1)p\{\exp(\mu^2) - 1\}] - 1 \right) + \frac{(k-1)^2}{m(m-1)} \{ \exp(\mu^2) - 1 \} \\
&= 1 + \frac{1}{m} \left(\exp \left[\frac{(k-1)^2}{m-1} \{ \exp(\mu^2) - 1 \} \right] - 1 \right) + \frac{(k-1)^2}{m(m-1)} \{ \exp(\mu^2) - 1 \} \\
&= 1 + \frac{1}{m} (e^z - 1 + z), \quad \left(z := \frac{(k-1)^2}{m-1} \{ \exp(\mu^2) - 1 \} \right),
\end{aligned}$$

where (a) uses the convex-order comparison of a hypergeometric random variable with a binomial random variable (Hoeffding, 1963, Theorem 4), and (b) uses $\log(1+x) \leq x$. Then, since $z \leq e^z - 1$, we have

$$\begin{aligned}
\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2] &\leq 1 + \frac{1}{m} (e^z - 1 + z) \\
&\leq 1 + \frac{2}{m} (e^z - 1) \\
&= 1 + \frac{2}{m} \left(\exp \left[\frac{(k-1)^2}{m-1} \{ \exp(\mu^2) - 1 \} \right] - 1 \right).
\end{aligned}$$

Therefore, $\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2]$ is bounded by C if

$$\mu \leq \sqrt{\log \left\{ 1 + \frac{m-1}{(k-1)^2} \log \left(1 + \frac{m(C-1)}{2} \right) \right\}}.$$

We further simplify this sufficient condition by considering two regimes. (i) Suppose that $k^2 = O(m^{1-\delta})$ for some fixed $\delta \in (0, 0.1)$. Since

$$\log \left(1 + \frac{m-1}{(k-1)^2} \log \left(1 + \frac{m(C-1)}{2} \right) \right) \geq \delta \log m + O(1),$$

the quantity on the left is at least $\frac{\delta}{2} \log m$ for all sufficiently large m . Hence, $\mu \leq \sqrt{\frac{\delta}{2} \log m}$ is sufficient to ensure that the second moment remains bounded by C . This lower bound matches, up to a constant factor, the upper bound achieved by the outlier test, which runs in $O(n)$ time. (ii) Suppose that $k^2 = \omega(m \log m)$. In this regime,

$$\frac{m-1}{(k-1)^2} \log \left(1 + \frac{m(C-1)}{2} \right) = o(1).$$

Applying $\log(1+x) \geq x/2$ for $0 \leq x \leq 1$, we obtain

$$\log \left(1 + \frac{m-1}{(k-1)^2} \log \left(1 + \frac{m(C-1)}{2} \right) \right) \geq \frac{(m-1) \log m}{4(k-1)^2}$$

for all sufficiently large m . Consequently, it suffices to require $\mu \leq \frac{\sqrt{(m-1) \log m}}{2(k-1)}$. The resulting lower bound matches, up to a constant factor, the upper bound attained by the degree test in Remark 3, which can be implemented in $O(n)$ time.

C.2.2 Paths and cycles of length k

We first consider the path case. For integers $1 \leq k \leq m-1$, possibly depending on m , we consider the case where $\mathcal{C}$ consists of all paths of length k in a complete graph with m nodes. We have $n = \binom{m}{2}$ and $N = m(m-1) \cdots (m-k)/2$, with $K = k$. We take an approach similar to that in Appendix C.1.5. Note that each realization of $S \cap S'$ is a set of disjoint paths with a total of $r \leq k$ edges. We denote the corresponding

forest on $[m]$ by F , and let F_p be the subforest obtained by discarding all isolated vertices. Also, let $v(F_p)$ and $c(F_p)$ be the numbers of nodes and components in F_p , respectively. For example, when $m = 5$ and $k = 3$, one possible realization of $S \cap S'$ is $\{(1, 2), (2, 3)\}$, which can be interpreted as a forest with $c(F) = 3$ components whose vertex sets are $\{\{1, 2, 3\}, \{4\}, \{5\}\}$; F_p is the path 1–2–3, with $c(F_p) = 1$ and $v(F_p) = 3$. Then, note that we have $r(F) = r(F_p) = v(F_p) - c(F_p)$, and we can count the number of paths of length k containing a given F_p with $v(F_p) \leq k + 1$ as

$$\binom{m - v(F_p)}{k + 1 - v(F_p)} \cdot 2^{c(F_p)} (k + 1 - r(F_p))! / 2,$$

where the first term accounts for the number of ways to choose the remaining nodes, the second term accounts for the number of directions for each component, the last term accounts for the number of ways to arrange all the nodes and components, and the division by 2 accounts for the fact that a path can be traversed in two directions. If $v(F_p) > k + 1$, no path of length k can contain F_p . Therefore, given a feasible F_p , the probability that a uniformly random path S of length k contains F_p is bounded by

$$\begin{aligned} \Pr(F_p \subset S) &= \frac{\binom{m - v(F_p)}{k + 1 - v(F_p)} \cdot 2^{c(F_p)} (k + 1 - r(F_p))! / 2}{m(m - 1) \cdots (m - k) / 2} \\ &= \frac{(m - v(F_p))!}{(k + 1 - v(F_p))! (m - k - 1)!} \cdot \frac{(m - k - 1)!}{m!} \cdot 2^{c(F_p)} (k + 1 - r(F_p))! \\ &= \frac{2^{c(F_p)}}{(m)_{v(F_p)}} \cdot \frac{(k + 1 - r(F_p))!}{(k + 1 - v(F_p))!} \\ &\leq \frac{2^{c(F_p)}}{(m)_{v(F_p)}} \cdot (k + 1)^{c(F_p)}. \end{aligned}$$

Set $Z := |S \cap S'|$. For $1 \leq R \leq k$, let $c_R := \min\{R, k + 1 - R\}$; a feasible forest with R edges has at most c_R nontrivial components. In the sums below, F_p ranges over edge sets of disjoint nontrivial paths with $v(F_p) \leq k + 1$, and $v = R + c$ whenever $r(F_p) = R$ and $c(F_p) = c$. Now, considering the R -th falling factorial of Z , we have

$$\begin{aligned} \mathbb{E}[(Z)_R] &= \sum_{I_R \in [m]_{2,R}} \Pr((i_1, j_1), \dots, (i_R, j_R) \in S \cap S') \\ &= \sum_{I_R \in [m]_{2,R}} \Pr((i_1, j_1), \dots, (i_R, j_R) \in S)^2 \\ &= R! \sum_{r(F_p)=R} \Pr(F_p \subset S)^2 \\ &\leq R! \sum_{c=1}^{c_R} \sum_{\substack{r(F_p)=R \\ c(F_p)=c}} \Pr(F_p \subset S)^2 \\ &\leq R! \sum_{c=1}^{c_R} \left(\frac{2^c}{(m)_v} \cdot (k + 1)^c \right)^2 \cdot |\{F_p : r(F_p) = R, c(F_p) = c\}|. \end{aligned}$$

We count the size of the collection $\{F_p : r(F_p) = R, c(F_p) = c\}$ for $1 \leq c \leq c_R$ as follows. We first choose and arrange $v = R + c$ nodes from m nodes, which can be done in $(m)_v$ ways. Then, we partition those v nodes into c components containing at least one edge, which can be done in $\binom{R-1}{c-1}$ ways. Also, we divide by 2^c to account for the fact that each component can be traversed in two directions, and thus we only count one of them. Last, we divide by $c!$ to account for the indistinguishability of components. Therefore, we have

$$|\{F_p : r(F_p) = R, c(F_p) = c\}| = \frac{(m)_v}{2^c c!} \binom{R-1}{c-1}.$$

Then, observe that

$$\begin{aligned}
\mathbb{E}[(Z)_R] &\leq R! \sum_{c=1}^{c_R} \frac{2^{2c}}{(m)_v^2} \cdot (k+1)^{2c} \cdot \frac{(m)_v}{2^c c!} \binom{R-1}{c-1} \\
&\leq R! \sum_{c=1}^{c_R} \frac{2^c (k+1)^{2c}}{(m)_v c!} \binom{R-1}{c-1} \\
&\stackrel{(\dagger)}{\leq} R! \sum_{c=1}^{c_R} \frac{2^c (k+1)^{2c}}{(m/e)^v c!} \binom{R-1}{c-1} \\
&\leq R! \sum_{c=1}^R \frac{2^c e^v (k+1)^{2c}}{m^v c!} \binom{R-1}{c-1},
\end{aligned}$$

where $(\dagger)$ uses $(m)_v \geq (m!)^{v/m} \geq (m/e)^v$ for $v \leq k+1 \leq m$. The last line extends the nonnegative sum to all $1 \leq c \leq R$, with $v = R+c$ in each summand. Now, for $t \geq 0$ with $e^t - 1 < m/e$, we bound the moment generating function of Z as follows:

$$\begin{aligned}
\mathbb{E}[\exp(tZ)] &= \sum_{R=0}^k \frac{(e^t - 1)^R}{R!} \mathbb{E}[(Z)_R] \\
&\leq 1 + \sum_{R=1}^k (e^t - 1)^R \sum_{c=1}^R \frac{2^c e^v (k+1)^{2c}}{m^v c!} \binom{R-1}{c-1} \\
&= 1 + \sum_{R=1}^k \left(\frac{e(e^t - 1)}{m} \right)^R \sum_{c=1}^R \left(\frac{2e(k+1)^2}{m} \right)^c \frac{1}{c!} \binom{R-1}{c-1} \\
&= 1 + \sum_{c=1}^k \underbrace{\left(\frac{2e(k+1)^2}{m} \right)^c}_{:=x} \frac{1}{c!} \sum_{R=c}^k \binom{R-1}{c-1} \underbrace{\left(\frac{e(e^t - 1)}{m} \right)^R}_{:=y} \\
&\leq 1 + \sum_{c=1}^k x^c \frac{1}{c!} \sum_{R=c}^{\infty} \binom{R-1}{c-1} y^R \\
&= 1 + \sum_{c=1}^k x^c \frac{1}{c!} \sum_{R'=0}^{\infty} \binom{R'+c-1}{c-1} y^{R'+c} \quad \text{where } R' := R - c \\
&= 1 + \sum_{c=1}^k x^c \frac{y^c}{c!} \sum_{R'=0}^{\infty} \binom{R'+c-1}{c-1} y^{R'} \\
&\stackrel{(\dagger)}{=} 1 + \sum_{c=1}^k x^c \frac{1}{c!} \cdot \frac{y^c}{(1-y)^c} \\
&\leq \exp\left(\frac{xy}{1-y}\right) \\
&= \exp\left(\frac{2e^2(k+1)^2(e^t - 1)}{m\{m - e(e^t - 1)\}}\right),
\end{aligned}$$

The infinite-series bound and the negative-binomial identity in $(\dagger)$ require $0 \leq y < 1$. For $t \geq 0$, this is equivalent to $e^t - 1 < m/e$, and hence to $e^{\mu^2} - 1 < m/e$ when $t = \mu^2$. We verify below that the final sufficient condition guarantees this restriction. Therefore, we have

$$\mathbb{E}_{\mathbf{Y} \sim \mathbb{Q}}[L_\nu(\mathbf{Y})^2] = \mathbb{E}_{S, S' \sim \nu}[\exp(\mu^2 |S \cap S'|)] \leq \exp\left(\frac{2e^2(k+1)^2(e^{\mu^2} - 1)}{m\{m - e(e^{\mu^2} - 1)\}}\right) = \exp\left(\frac{2e^2(k+1)^2 z}{m(m - ez)}\right),$$

where we denote $z := e^{\mu^2} - 1$. Note that, for $0 \leq z < m/e$, the right-hand side is bounded by C if and only if

$$z \leq \frac{m^2 \log C}{2e^2(k+1)^2 + em \log C}, \tag{88}$$

which is equivalent to

$$\mu \leq \sqrt{\log \left(1 + \frac{m^2 \log C}{2e^2(k+1)^2 + em \log C} \right)}.$$

Observe that the condition $e^{\mu^2} - 1 < m/e$ is automatically satisfied when Equation (88) holds.

We now examine this sufficient condition in two regimes. (i) Assume that $k = O(m^{1-\delta})$ for some fixed $\delta \in (0, 0.1)$. Since

$$2e^2(k+1)^2 + em \log C = O(m^{2-2\delta}),$$

it follows that

$$\log \left(1 + \frac{m^2 \log C}{2e^2(k+1)^2 + em \log C} \right) \geq 2\delta \log m + O(1).$$

Hence, for sufficiently large m , the condition $\mu \leq \sqrt{\delta \log m}$ guarantees that the second moment is bounded by C . A matching upper bound up to a constant factor is obtained by the outlier test in Proposition 2, which runs in $O(n)$ time. (ii) Assume that $k = \Theta(m)$. Since $k \leq m - 1$, we obtain

$$\frac{m^2 \log C}{2e^2(k+1)^2 + em \log C} \geq \frac{\log C}{2e^2 + e \log C/m} \geq \frac{\log C}{4e^2}$$

for all sufficiently large m . Therefore, it is sufficient to have

$$\mu \leq \sqrt{\log \left(1 + \frac{\log C}{4e^2} \right)}.$$

This establishes a constant-order statistical lower bound, which matches the signal scale achieved by the average test, since the test achieves strong detection whenever $\mu = \omega(m/k) = \omega(1)$.

We now consider cycles of length $3 \leq k \leq m$. The cycle case can be handled using a grouping argument similar to those employed in the proofs of our main results, such as the multiple planted cliques problem (see Appendix B.7). The main idea is to cover a cycle by two paths of length $\lceil k/2 \rceil$ and analyze their overlaps instead of directly studying the overlap between cycles. Given a uniformly random cycle S , choose a starting edge and an orientation uniformly at random, and denote the successive edges in the chosen direction by $S_1, \dots, S_k$. Let $q := \lceil k/2 \rceil$, as in Equation (33), and set $\mathcal{S}_1 := \{S_1, \dots, S_q\}$ and $\mathcal{S}_2 := \{S_{k-q+1}, \dots, S_k\}$. The two paths are edge-disjoint when k is even and share one edge when k is odd. Define $\mathcal{S}'_1$ and $\mathcal{S}'_2$ analogously for an independent cycle S' , and observe that

$$S = \mathcal{S}_1 \cup \mathcal{S}_2, \quad S' = \mathcal{S}'_1 \cup \mathcal{S}'_2.$$

Following the same approach as in Equation (34) and the subsequent analysis therein, and applying Hölder's inequality to the four pairwise overlaps, we obtain

$$\mathbb{E}[\exp(\mu^2 |S \cap S'|)] \leq \mathbb{E}[\exp(4\mu^2 |\mathcal{S}_1 \cap \mathcal{S}'_1|)].$$

Note that each of $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}'_1$, and $\mathcal{S}'_2$ is marginally distributed as a uniformly random path of length q , and the blocks from S are independent of those from S' . Then, applying the second moment bound established for the path case with path length q and signal strength 2μ , we obtain a sufficient condition on μ that differs only by a constant factor. Finally, depending on the regime, the matching upper bounds are achieved by the same tests as in the path case.