Proceedings of Machine Learning Research vol 336:1–54, 2026 39th Annual Conference on Learning Theory

Testing for a Hidden Geometry in Random Graphs

Amit Silber AMITSILBER@MAIL.TAU.AC.IL
School of Electrical & Computer Engineering, Tel Aviv University

Mor Oren-Loberman ORENMOR@MAIL.TAU.AC.IL
School of Electrical & Computer Engineering, Tel Aviv University

Wasim Huleihel WASIMH@TAUEX.TAU.AC.IL
School of Electrical & Computer Engineering, Tel Aviv University

Editors: Steve Hanneke and Tor Lattimore

Abstract

In this work, we investigate the fundamental problem of detecting a faint geometric signal hidden within an otherwise random graph. We formulate this task as a hypothesis testing problem: under the null hypothesis, the observed graph is an Erdős–Rényi random graph $\mathcal{G}(n,q)$ with edge density $q \in (0,1)$; under the alternative, a high-dimensional geometric structure is clandestinely embedded. Specifically, a random geometric graph $\mathcal{G}(k,q,d)$ on $k \leq n$ vertices is planted inside $\mathcal{G}(n,q)$, where each of the k vertices corresponds to an independent random point drawn uniformly from the unit sphere $\mathbb{S}^{d-1}$, and edges are formed according to latent proximity, resulting in the same edge probability q.

Our objective is to characterize the limits of detectability of this hidden geometry, from both statistical and computational perspectives. We derive sharp information-theoretic lower bounds that characterize the regimes in which detection is fundamentally impossible, expressed explicitly in terms of the problem parameters. Complementing these impossibility results, we propose and analyze several algorithms that provably attain these limits whenever detection is feasible.

We also explore the algorithmic landscape of the problem and investigate which regimes admit efficient, polynomial-time testing procedures. As in many other structured high-dimensional inference problems, our model exhibits a pronounced *easy–hard–impossible* phase transition: there exist regimes in which detection is statistically possible yet computationally prohibitive, as well as regimes in which detection is impossible even with unbounded computational resources. As concrete evidence of this computational barrier, we show that the entire class of low-degree polynomial algorithms fails in the conjecturally hard regime, highlighting a sharp separation between statistical possibility and algorithmic feasibility.

Keywords: Random graphs, testing high-dimensional geometry, statistical-computational limits.

1. Introduction

Networks with latent structure arise throughout modern data science, from social and biological systems to communication and information networks. A large body of work models such data using random graphs endowed with an underlying geometric or feature-based structure, where vertices correspond to latent points in a metric space and edges form preferentially between nearby points. Canonical examples include random geometric graphs, random dot product graphs, and more general latent space models; see, for example, the monograph Penrose (2007) and the references therein. These models provide a principled way to capture dependencies and correlations that are absent from edge-independent baselines such as the Erdős–Rényi random graph.

In many applications, however, the latent variables themselves are unobserved, and only the graph structure is available. This leads to a fundamental statistical question: given a single observed graph, can one determine whether its edges encode an underlying geometric structure, or whether they are effectively indistinguishable from those of a purely random graph? From a probabilistic perspective, this question has been studied in the high-dimensional setting, where geometry may be lost as the ambient dimension grows. Early work showed that classical random geometric graphs become asymptotically indistinguishable from Erdős–Rényi graphs as the dimension tends to infinity Devroye et al. (2011). Subsequent results identified a sharp phase transition in the dense regime, revealing that high-dimensional geometric structure disappears precisely when the dimension exceeds a cubic threshold in the number of vertices Bubeck et al. (2016). More recent extensions have explored how noise and softened geometric dependence affect this transition; see, for instance, Liu and Rácz (2023a,b).

While these results address the detectability of *global* geometric structure, many problems of interest involve *localized* signals that are present only on a small subset of vertices. Detecting such faint, structured signals hidden inside high-dimensional noise is a recurring theme across modern statistics, probability, and theoretical computer science. In network data, this theme manifests through *planted subgraph* models, where one observes a random graph drawn under one of two hypotheses: either a featureless null model, or an alternative in which a small subset of vertices carries additional structure. These models have played a central role in clarifying the statistical limits of detection and in revealing striking computational–statistical gaps, where information-theoretic possibility does not coincide with what is achievable by efficient algorithms; see, for example, the dense-subgraph and community-detection literature Arias-Castro and Verzelen (2014); Verzelen and Arias-Castro (2015), as well as recent complexity frameworks, see, e.g., Brennan et al. (2018); Hopkins (2018); Brennan and Bresler (2020).

A planted geometric structure. In this paper, we study the problem of detecting a small hidden *geometric* structure embedded in a random graph. We formulate this task as a hypothesis testing problem. Under the null hypothesis, the observed graph is an Erdős–Rényi random graph $\mathcal{G}(n, q)$ with edge density $q \in (0, 1)$. Under the alternative, a high-dimensional random geometric graph $\mathcal{G}(k, q, d)$ on $k \leq n$ vertices is planted inside $\mathcal{G}(n, q)$. Each of the k planted vertices corresponds to an independent random point drawn uniformly from the unit sphere $\mathbb{S}^{d-1}$, and edges within the planted subgraph arise from latent proximity, resulting in *the same*[1] marginal edge probability q. All remaining edges behave as in the null model.

This formulation captures a setting in which geometry is both *localized*—only a small subset of vertices participates in the geometric structure—and *high-dimensional*. The latter aspect is particularly important: as the ambient dimension grows, geometric information becomes increasingly diffuse, and it is known that high-dimensional geometric graphs may become indistinguishable from Erdős–Rényi graphs. Understanding where this loss of geometry occurs, and how it interacts with the size of the planted subgraph, is central to our investigation.

From a modeling perspective, the proposed alternative can be viewed as a geometry-enriched analogue of classical planted dense-subgraph or community models. Rather than postulating an *ad hoc* increase in edge probability on a hidden vertex set, the planted structure here is generated

1. Allowing a distinct edge probability inside the planted set introduces an explicit density contrast, in which case detection is governed primarily by classical planted dense subgraph considerations and the ambient dimension d plays no essential role in the dense setting. We briefly discuss this regime later for comparison.

by a latent geometric mechanism that induces correlations among edges. In particular, the planted subgraph is not merely "denser" than the background—it carries a coherent geometric footprint that couples many local statistics. To our knowledge, this work is the first to provide a sharp detection theory for a planted structure whose defining feature is latent geometry.

Beyond statistical detectability, our study is also motivated by computational considerations. Many planted-structure problems in random graphs exhibit pronounced gaps between what is information-theoretically possible and what can be achieved by efficient algorithms. A key objective of this work is to determine whether such computational–statistical gaps arise in the presence of latent geometry, and to characterize the resulting phase transitions between easy, hard, and impossible regimes.

Main contributions. We investigate the fundamental limits of detecting latent geometric structure from both statistical and computational perspectives. Our first contribution is an information-theoretic characterization of detectability in terms of the parameters (n, k, d). A central technical challenge in this setting is that naive second-moment methods for the likelihood ratio can be dominated by rare tail events and therefore fail to capture the true statistical threshold. To address this issue, we develop a truncated second-moment analysis, inspired by classical ideas in minimax hypothesis testing and truncation techniques in high-dimensional statistics Ingster (1997); Verzelen and Arias-Castro (2015). This approach yields a finite second-moment regime and allows us to identify sharp conditions under which detection is information-theoretically impossible.

We complement these lower bounds by proposing and analyzing three testing procedures that achieve the statistical threshold (up to constants and logarithmic factors). The first is a *vanilla signed-triangle* test, based on counting signed triangles in the observed graph, extending ideas originally developed for detecting global high-dimensional geometry Bubeck et al. (2016). The second is a *scan signed-triangle* test, which counts signed triangles over every subset of $\binom{[n]}{k}$ vertices, corresponding to all candidate planted sets. The third is a *geometry-agnostic scan test* that ignores the latent geometric structure and instead attempts to localize the planted subset purely through density-type evidence, connecting our problem to classical planted dense-subgraph and community-detection models Arias-Castro and Verzelen (2014). Together, these tests demonstrate that latent geometry can be detected optimally using both geometry-aware and geometry-agnostic procedures, depending on the regime.

Our main results are summarized in the phase diagram shown in Figure 1. Specifically, we consider an asymptotic regime in which both k and d scale polynomially with n, namely $d = \Theta(n^\alpha)$ and $k = \Theta(n^\beta)$ for some $\alpha > 0$ and $\beta \in (0, 1)$. Here, the exponent α captures the growth of the ambient dimension, while β governs the size of the planted geometric structure. It is evident that the detection problem becomes statistically more challenging as α increases or β decreases. In this regime, the (α, β) parameter space is partitioned into three distinct regions:

1. *Statistically impossible regime (gray):* detection is information-theoretically impossible when $\alpha > 2\beta \wedge 3$.

2. *Easy regime (blue):* there exists a polynomial-time algorithm for detection when $\alpha < 6\beta - 3$.

3. *Hard regime (red):* detection is information-theoretically possible when $\alpha < 2\beta$ and $\alpha > 0 \vee (6\beta - 3)$, but computationally intractable in the sense that no polynomial-time algorithm is known; moreover, the class of low-degree polynomial tests fails throughout this region.

A key conceptual and technical challenge arises in the analysis of the scan signed-triangle test. Since this procedure ranges over an exponential family of candidate vertex sets, its analysis requires exponential tail bounds for the signed-triangle statistic, rather than the first- or second-moment estimates that suffice for fixed tests. Indeed, moment-based methods constitute the main analytical tool in essentially all prior works on testing high-dimensional geometry in random graphs, including the signed-triangle analysis of Bubeck, Ding, Eldan, and Rácz Bubeck et al. (2016), subsequent extensions to noisy and softened geometric models Liu and Rácz (2023a,b), as well as related random-matrix approaches to geometric phase transitions such as anisotropic random geometric graphs and Wishart-type ensembles Eldan and Mikulincer (2020); Brennan et al. (2021).

While sharp upper-tail results for *unsigned* triangle counts in Erdős–Rényi graphs are by now classical—originating with the martingale-based approach of Kim and Vu Kim and Vu (2004) and refined through subsequent combinatorial and variational methods Janson et al. (2004); Boucheron et al. (2013); Ganguly et al. (2024)—these techniques do not extend to the signed setting, where substantial cancellations fundamentally alter the tail behavior. To overcome this obstacle, we develop new tools to characterize the relevant exponential rate, relying on strong decoupling inequalities for U-statistics de la Pena and Montgomery-Smith (1995); de la Pena and Giné (1999).

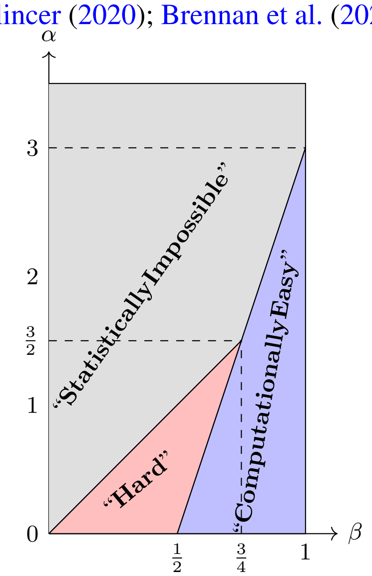

Figure 1: Phase diagram for detecting the presence of a planted random geometry subgraph.

Finally, we examine the computational landscape of the problem and provide evidence for an *easy–hard–impossible* phase transition. In particular, we show that low-degree polynomial algorithms fail in a parameter regime that is conjecturally hard, leveraging the low-degree framework for average-case hardness in high-dimensional inference Hopkins (2018); Kunisky et al. (2022). This places the detection of planted geometric structure squarely within a broader class of inference problems exhibiting sharp separations between statistical possibility and computational feasibility.

The rest of this paper is organized as follows. In Section 2, we introduce the problem setup and provide some necessary preliminaries. Section 3 presents our main results, discussions, and examples. Finally, further related work, additional notations, and all detailed formal proofs are provided in the appendix.

2. Setup and Problem Statement

We study the problem of detecting the presence of a small latent high-dimensional geometric subgraph in a random graph. Let $\mathcal{G}(n,q)$ denote the Erdős–Rényi random graph on n vertices, in which each pair of vertices is connected independently with probability q. Under the null hypothesis $\mathcal{H}_0$, the observed graph G_n is drawn from $\mathcal{G}(n,q)$.

To define the alternative hypothesis, we first recall the classical model of a random geometric graph. In this model, each vertex is associated with a point in a metric space, and an edge is present between two vertices if the distance between their corresponding labels is below a prescribed threshold. We focus on the case where the underlying metric space is the Euclidean sphere $\mathbb{S}^{d-1}$, and the latent labels are i.i.d. random vectors drawn uniformly from $\mathbb{S}^{d-1}$. We denote by $\mathcal{G}_d(n, q)$ the ensemble of such random graphs, where q is the marginal probability of an edge between any pair of vertices, and thus determines the threshold distance for connection.

Formally, $\mathcal{G}_d(n, q)$ is defined as follows. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be independent random vectors, each uniformly distributed on $\mathbb{S}^{d-1}$; we denote $\mathbf{X} \triangleq [\mathbf{x}_1, \ldots, \mathbf{x}_n]$. In $\mathcal{G}_d(n, q)$, distinct vertices i and j are connected by an edge if and only if $\langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq t_{q,d}$, where $|t_{q,d}| \leq 1$ is chosen so that $\mathbb{P}(\langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq t_{q,d}) = q$. For brevity, we define $\sigma_{ij} \triangleq \mathbb{1}\{\langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq t_{q,d}\}$, for all $i, j \in [n]$. Under the alternative hypothesis $\mathcal{H}_1$, the observed graph G_n on n vertices is generated as follows:

1. A subset $\mathcal{K}$ of k vertices is selected uniformly at random from the n vertices. Each vertex $i \in \mathcal{K}$ is associated with a latent label $\mathbf{x}_i \sim \mathsf{Unif}(\mathbb{S}^{d-1})$.
2. For any $i, j \in \mathcal{K}$, the vertices are connected by an edge if and only if $\langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq t_{q,d}$, where $t_{q,d} \in [-1, 1]$ is chosen so that $\mathbb{P}(\langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq t_{q,d}) = q$.
3. All remaining edges, namely those with at least one endpoint outside $\mathcal{K}$, are added independently with probability q.

We denote by $\mathcal{G}_d(n, k, q)$ the ensemble of random graphs generated by the above procedure. That is, $\mathcal{G}_d(n, k, q)$ consists of graphs on n vertices in which a geometric random subgraph $\mathcal{G}_d(k, p)$ is *planted* inside an Erdős–Rényi random graph $\mathcal{G}(n, q)$. The vertices in the set $\mathcal{K}$ thus form a *geometric community* whose internal connectivity is higher than that of the background graph. In this paper, we focus on the regime in which both edge probability q is a fixed constant independent of n. Our goal is to address the following *planted random geometry* (PRG) *detection problem.*

Definition 1 (PRG detection problem) *The* PRG *detection problem with parameters* (n, k, d, q)*, hereafter denoted by* $\mathsf{PRG}(n, k, d, q)$*, refers to the problem of distinguishing between the following two hypotheses:*

$$\mathcal{H}_0 : \mathsf{G}_n \sim \mathcal{G}(n, q) \qquad \text{vs.} \qquad \mathcal{H}_1 : \mathsf{G}_n \sim \mathcal{G}_d(n, k, q). \tag{1}$$

Remark 1 *We restrict attention to the setting in which the null and alternative hypotheses have identical edge marginals, so that the distinction between the two models is carried entirely by latent geometric dependencies. This restriction is deliberate. In the dense regime where edge probabilities are fixed constants, introducing a different edge probability on the planted vertex set creates an explicit density contrast. In that case, detection reduces to classical planted dense subgraph testing, and the ambient dimension does not govern the detection threshold. By contrast, when the edge marginals match, geometry becomes the sole source of signal, and the interaction between the subgraph size and the ambient dimension is fundamental. We note that when edge probabilities are allowed to vary with n, density contrast and geometry can interact nontrivially, even if they differ, but analyzing such sparse regimes is beyond the scope of this work.*

Upon observing G_n, a detection algorithm $\mathcal{A}_n(\mathsf{G}_n) \in \{0, 1\}$ for the above problem outputs a decision in $\{0, 1\}$. We define the *risk* of a detection algorithm $\mathcal{A}_n$ as the sum of its Type-I and Type-II error probabilities, namely,

$$\mathsf{R}_n(\mathcal{A}_n) = \mathbb{P}_{\mathcal{H}_0}\big(\mathcal{A}_n(\mathsf{G}_n) = 1\big) + \mathbb{P}_{\mathcal{H}_1}\big(\mathcal{A}_n(\mathsf{G}_n) = 0\big), \tag{2}$$

where $\mathbb{P}_{\mathcal{H}_0}$ and $\mathbb{P}_{\mathcal{H}_1}$ denote the probability distributions under the null and alternative hypotheses, respectively. The optimal risk is defined as $\mathsf{R}_n^\star \triangleq \inf_{\mathcal{A}_n} \mathsf{R}_n(\mathcal{A}_n)$, where the infimum is taken over all (possibly randomized) tests $\mathcal{A}_n \colon \mathsf{G}_n \mapsto \{0,1\}$. A sequence of tests $\mathcal{A}_n(\mathsf{G}_n) \in \{0,1\}$ is said to achieve *strong detection* if $\limsup_{n\to\infty} \mathsf{R}_n(\mathcal{A}_n) = 0$, and *weak detection* if $\limsup_{n\to\infty} \mathsf{R}_n(\mathcal{A}_n) < 1$. Conversely, we say that *strong detection is impossible* if $\liminf_{n\to\infty} \mathsf{R}_n^\star > 0$, and that *weak detection is impossible* if $\lim_{n\to\infty} \mathsf{R}_n^\star = 1$.

The algorithms we consider are either unconstrained—allowing for computationally expensive procedures—or restricted to run in polynomial time, corresponding to computationally efficient algorithms. Unconstrained algorithms are typically used to establish information-theoretic limits and to show the tightness of lower bounds. An algorithm is said to be polynomial-time if its running time is bounded by $\mathsf{poly}(n)$, where n denotes the size of the input. As discussed in the introduction, our goal is to characterize necessary and sufficient conditions under which detecting the planted geometric subgraph is possible or impossible, both in the absence of computational constraints and under polynomial-time restrictions.

3. Main Results

In this section, we present our main results. We begin with information-theoretic lower bounds establishing regimes in which detecting the planted geometric subgraph is impossible, regardless of computational constraints. We then present algorithmic upper bounds demonstrating that these limits are achievable, and conclude with computational lower bounds identifying regimes where detection is statistically possible but computationally hard.

Statistical lower bounds. We start with information-theoretic lower bounds for the $\mathsf{PRG}(n,k,d,q)$ detection problem. By the characterization of optimal hypothesis testing in terms of total variation distance, the optimal risk satisfies (see, e.g., (Tsybakov, 2009, Theorem 2.2))

$$\mathsf{R}^* = 1 - d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1}). \tag{3}$$

The following theorem establishes that when the ambient dimension grows sufficiently fast relative to the size of the planted geometric subgraph, no test can reliably detect the presence of geometry.

Theorem 2 (Impossibility of strong detection) *Consider the detection problem defined in* (1). *If the following conditions hold simultaneously:*

$$\frac{d}{k^2 \log^2 k} \to \infty \quad \text{and} \quad \frac{dn^3}{k^6 \log^2 k} \to \infty, \tag{4}$$

then $d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1}) < 1$*, i.e., strong detection is impossible.*

Moreover, the same proof technique shows that if the condition $\frac{dn^3}{k^6 \log^2 k} \to \infty$ is replaced by $\frac{dn^3}{k^6 f(k)} \to \infty$ for some function $f(k) = \omega(\log^2 k)$, then $d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1}) \to 0$, implying that weak detection is impossible.

It is instructive to compare Theorem 2 with previously known results in special cases. In the vanilla setting where $k = n$, the model reduces to testing for global high-dimensional geometry, and our lower bound essentially recovers the known phase transition at dimension $d \asymp n^3$ established in Bubeck et al. (2016), up to a $\log^2 k$ factor. We note that if the edge probability inside the planted set were taken to be $p \neq q$, then detection would be information-theoretically impossible whenever $k = O(\log n)$, regardless of the value of d. This impossibility is not driven by geometry, but instead reflects a fundamental limitation already present in classical planted clique and planted dense subgraph models: when the planted subgraph is too small, even a purely combinatorial increase in edge density cannot be reliably detected. Thus, only when $p = q$, geometry is the sole distinguishing feature, and the interplay between dimension and subgraph size becomes central.

Proof sketch of Theorem 2. We outline a sequence of reductions that control $d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1})$. Let $\mathcal{L} \triangleq \frac{\mathrm{d}\mathbb{P}_{\mathcal{H}_1}}{\mathrm{d}\mathbb{P}_{\mathcal{H}_0}}$ denote the likelihood ratio. As is well-known, to lower bound $d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1})$ it suffices to upper bound $\mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2] = 1 + \chi^2(\mathbb{P}_{\mathcal{H}_1} \| \mathbb{P}_{\mathcal{H}_0})$; if $\mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2] = 1 + o(1)$, then weak detection is impossible, and if $\mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2] = O(1)$, then strong detection is impossible.

Step 1: Truncation and stability. A direct attempt to control $\mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2]$ fails in the geometric setting, because rare tail events under $\mathcal{H}_1$ can dominate the second moment even when $d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1})$ remains bounded away from 1. To circumvent this, we instead analyze the second moment conditioned on events that are typical under $\mathcal{H}_1$. This refined approach, originally proposed by Ingster (1997); Butucea and Ingster (2013), is based on controlling the first and second moments of a truncated likelihood ratio Arias-Castro and Verzelen (2014); Wu et al. (2023). Specifically, given an event $\mathcal{F}$, measurable with respect to the observed and latent variables $(\mathsf{G}, \mathcal{K}, (\mathbf{x}_i)_{i=1}^n)$, and such that $\mathbb{P}_{\mathcal{H}_1}(\mathcal{F}) = 1 - o(1)$, we define the planted model conditioned on $\mathcal{F}$ as

$$\tilde{\mathbb{P}}_{\mathcal{H}_1}(\mathsf{G}, \mathcal{K}, \mathbf{X}) \triangleq \frac{\mathbb{P}_{\mathcal{H}_1}(\mathsf{G}, \mathcal{K}, \mathbf{X}) \mathbb{1}\left\{(\mathsf{G}, \mathcal{K}, (\mathbf{x}_i)_{i=1}^n) \in \mathcal{F}\right\}}{\mathbb{P}_{\mathcal{H}_1}(\mathcal{F})},$$

and the truncated likelihood ratio $\tilde{\mathcal{L}}(\mathsf{G}) \triangleq \mathbb{E}_{\mathcal{K},\mathbf{X}} \tilde{\mathbb{P}}_{\mathcal{H}_1}(\mathsf{G}, \mathcal{K}, \mathbf{X}) / \mathbb{P}_{\mathcal{H}_0}(\mathsf{G})$. Now, by triangle and data processing inequalities, it is evident that $d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1}) \leq d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \tilde{\mathbb{P}}_{\mathcal{H}_1}) + \mathbb{P}(\mathcal{F}^c)$, thus it suffice to control the *truncated* second moment, i.e., $\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2]$. We truncate by restricting the latent matrix (Gaussian Orthogonal Ensemble (GOE) under $\mathcal{H}_0$) to a high-probability spectral event that simultaneously controls its Frobenius and operator norms, thereby suppressing rare high-energy fluctuations that would otherwise cause the second moment to diverge. This event holds with overwhelming probability under $\mathcal{H}_1$, by standard spectral concentration for the corresponding Wishart-type latent matrix (and its GOE approximation).

Step 2: Overlap decomposition. Let $\mathcal{K}, \bar{\mathcal{K}} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}\left(\binom{[n]}{k}\right)$ be independent planted sets and define $U \triangleq \mathcal{K} \cap \bar{\mathcal{K}}$ and $Z \triangleq |U|$. A Fubini expansion combined with the conditional independence structure under $\mathcal{H}_1$ shows that the entire second-moment computation localizes onto the overlap U and yields the following useful formula:

$$\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2] \leq 1 + (1 + o(1)) \cdot \mathbb{E}_Z \left[\chi^2 \left(\tilde{\mathcal{G}}_d(Z, p) \| \mathcal{G}(Z, q)\right) \mathbb{1}_{Z \geq 2}\right],$$

where $Z \sim \mathsf{Hypergeometric}(n, k, k)$ and $\tilde{\mathcal{G}}_d(z, p)$ is the induced truncated geometric random graph distributions where latent vectors are constrained to the set $\mathcal{F}$. The rest of the proof is therefore a uniform control of the non-planted χ^2 term, followed by averaging over the random overlap Z.

Step 3: Random-matrix pushforward and a divergence chain. Following the multi-step comparison framework of Bubeck et al. (2016), we compare the truncated alternative and null graph measures by interpolating through a sequence of intermediate distributions. As in Bubeck et al. (2016), we represent the z-vertex graph measures as pushforwards of random matrix ensembles under thresholding maps: the truncated alternative $\tilde{\mathbb{P}}_1^{(z)} \equiv \tilde{\mathcal{G}}_d(Z,p)$ arises from a (truncated) Wishart matrix $\mathbf{W}(z,d) = \mathbf{X}\mathbf{X}^\top$, while the null $\mathbb{P}_0^{(z)} \equiv \mathcal{G}(Z,q)$ corresponds to a (rescaled) GOE matrix $\mathbf{M}(z,d)$. To bridge the gap between $\tilde{\mathbb{P}}_1^{(z)}$ and $\mathbb{P}_0^{(z)}$, we introduce intermediate measures $\tilde{\mathbb{P}}_0^{(z)}$ and $\tilde{\mathbb{Q}}^{(z)}$, which interpolate between the thresholded Wishart-based and GOE-based constructions before and after normalization. A decoupling lemma for D_m-divergences gives the factorization

$$1 + \chi^2(\tilde{\mathcal{G}}_d(Z,q) \| \mathcal{G}(Z,q)) \leq \sqrt{1 + D_4(\tilde{\mathbb{P}}_1^{(z)} \| \tilde{\mathbb{P}}_0^{(z)})} \sqrt[4]{1 + D_6(\tilde{\mathbb{P}}_0^{(z)} \| \tilde{\mathbb{Q}}^{(z)})} \sqrt[4]{1 + D_5(\tilde{\mathbb{Q}}^{(z)} \| \mathbb{P}_0^{(z)})},$$

where $D_m(\mathbb{P}\|\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}[(\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q})^m] - 1$, for $m \in \mathbb{N}$. Bounds on each factor above follow from perturbative density comparisons of the corresponding matrix ensembles. Crucially, while Bubeck et al. (2016) controls total variation distance—where rare catastrophic spectral events can be discarded—such events cannot be ignored when bounding D_α divergences, as they may dominate the expectation. Our truncation is therefore essential: it removes these catastrophic events while retaining overwhelming probability under both hypotheses, ensuring that all moments above remain finite and controlled. The above comparison yields the following uniform estimate. Assume $d \gg k^2 \log^2 k$. Then for all $2 \leq z \leq k$,

$$1 + \chi^2\big(\tilde{\mathcal{G}}_d(z,p) \| \mathcal{G}(z,q)\big) \leq \exp\left(C \frac{z^3 \log^2 k}{d}\right), \tag{5}$$

for a constant $C > 0$. Finally, under the scaling of Theorem 2, the hypergeometric overlap term is shown to be controlled, ensuring the desired second-moment behavior.

Upper bounds. We now turn to algorithmic upper bounds. Specifically, we propose three detection algorithms and analyze their associated risks. To this end, we first introduce some notation. Let $\mathbf{A}$ denote the adjacency matrix of the observed graph G_n; we suppress the dependence of $\mathbf{A}$ on G_n, as the underlying graph will always be clear from the context. For distinct vertices $i, j, \ell \in [n]$, define

$$\bar{\mathbf{A}}_{i,j} \triangleq \mathbf{A}_{i,j} - \mathbb{E}\left[\mathbf{A}_{i,j}\right], \qquad \mathsf{T}_{\mathsf{G}_n}(i,j,\ell) \triangleq \bar{\mathbf{A}}_{i,j}\bar{\mathbf{A}}_{i,\ell}\bar{\mathbf{A}}_{j,\ell}. \tag{6}$$

We consider the following statistics:

$$\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n) \triangleq \sum_{\{i,j,\ell\} \subset \binom{[n]}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell), \tag{7}$$

$$\mathsf{T}_{\mathsf{scan}}(\mathsf{G}_n) \triangleq \max_{\mathcal{S} \subset [n] : |\mathcal{S}| = k} \sum_{\{i,j,\ell\} \subset \binom{\mathcal{S}}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell). \tag{8}$$

The statistic $\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n)$ in (7) counts the total number of signed triangles in G_n. The scan statistic $\mathsf{T}_{\mathsf{scan}}(\mathsf{G}_n)$ in (8) enumerates all k-vertex induced subgraphs and selects the one with the largest signed-triangle count. Based on these statistics, we define the following detection algorithms:

$$\mathcal{A}_{\mathsf{triangle}}(\mathsf{G}_n) \triangleq \mathbb{1}\left\{\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n) \geq \tau_{\mathsf{triangle}}\right\}, \tag{9}$$

$$\mathcal{A}_{\mathsf{scan}}(\mathsf{G}_n) \triangleq \mathbb{1}\left\{\mathsf{T}_{\mathsf{scan}}(\mathsf{G}_n) \geq \tau_{\mathsf{scan}}\right\}, \tag{10}$$

where $\tau_{\mathsf{triangle}}, \tau_{\mathsf{scan}} \in \mathbb{R}_+$ are thresholds specified below.

We briefly motivate this choice of statistics. Consider the conditional likelihood ratio $\mathcal{L}_{n|\mathcal{K}} \triangleq \mathbb{P}_{\mathcal{H}_1|\mathcal{K}}/\mathbb{P}_{\mathcal{H}_0}$. Expanding $\mathcal{L}_{n|\mathcal{K}}$ in the orthonormal basis associated with the entries of $\mathcal{G}(n,q)$, the expansion up to degree three takes the form

$$\mathcal{L}_{n|\mathcal{K}} \approx 1 + \frac{1}{\sqrt{d}} \sum_{\{i,j,\ell\}\subset\binom{\mathcal{K}}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell). \tag{11}$$

Thus, signed triangles constitute the lowest-degree nontrivial term in the likelihood-ratio expansion, explaining why triangle-based statistics arise naturally. Since the likelihood-ratio test is optimal by the Neyman–Pearson lemma, it is reasonable to expect low-degree proxies of the likelihood ratio to be powerful as well. Our algorithmic guarantees are summarized in the following theorem.

Theorem 3 (Strong detection upper bounds) *Consider the detection problem in (1), and the signed triangle and triangle scan tests in (9), and (10), respectively, with* $\tau_{\mathsf{scan}} = \tau_{\mathsf{triangle}} \triangleq \binom{k}{3}\frac{\mathsf{C}_q}{\sqrt{d}}$*, for a constant* C_q *depending only on* q*.*

1. *If* $dn^3/k^6 \to 0$*, then* $\mathsf{R}(\mathcal{A}_{\mathsf{triangle}}) \to 0$*, as* $k,n\to\infty$*.*
2. *If* $k^2/d > C\log^2 n$*, for some constant* $C>0$*, then* $\mathsf{R}(\mathcal{A}_{\mathsf{scan}}) \to 0$*, as* $k,n\to\infty$*.*

Theorem 3 complements the information-theoretic lower bound in Theorem 2 up to polylogarithmic factors. In particular, Theorem 2 together with Theorem 3 identifies the phase transition for optimal detection (up to the stated logarithmic terms). In the case where the edge probability inside the planted set is taken to be $p \neq q$, the optimal test scans over all k-vertex subgraphs and selects the one with the largest number of edges. We show in Appendix G that strong detection is achievable whenever $k = \Omega(\log n)$.

We briefly remark on a key technical difficulty underlying Theorem 3, which is the analysis of the signed triangle scan test. Unlike the vanilla signed triangle statistic, the scan test ranges over an exponential family of candidate vertex sets, which necessitates control of exponential tail probabilities rather than moment bounds. While first- and second-moment analyses suffice for fixed tests and have been the primary tools in prior work on testing high-dimensional geometry in random graphs, they are insufficient in the scan setting due to the competing combinatorial complexity.

Proof sketch of Theorem 3. Let

$$\mathsf{T}_{\mathcal{S}}(\mathsf{G}) \triangleq \sum_{\{i,j,\ell\}\subset\binom{\mathcal{S}}{3}} \bar{\mathbf{A}}_{ij}\bar{\mathbf{A}}_{i\ell}\bar{\mathbf{A}}_{j\ell}, \qquad \bar{\mathbf{A}}_{ij} \triangleq \mathbf{A}ij - q, \tag{12}$$

and then the scan statistic is $\mathsf{T}_{\mathsf{scan}}(\mathsf{G}) \triangleq \max_{|\mathcal{S}|=k} \mathsf{T}_{\mathcal{S}}(\mathsf{G})$. As for the Type-II error, under $\mathcal{H}_1$, the planted set $\mathcal{K}$ achieves the maximum, so

$$\mathbb{P}_{\mathcal{H}_1}\left(\mathsf{T}_{\mathsf{scan}} < \tau_{\mathsf{scan}}\right) \leq \mathbb{P}_{\mathcal{H}_1}\left(\mathsf{T}_{\mathcal{K}} < \tau_{\mathsf{scan}}\right). \tag{13}$$

Moreover, $\mathbb{E}_{\mathcal{H}_1}[\mathsf{T}_{\mathcal{K}}] \geq \binom{k}{3}\frac{C_q}{\sqrt{d}}$. A Chebyshev bound yields

$$\mathbb{P}_{\mathcal{H}_1}\left(\mathsf{T}_{\mathcal{K}} < \tau_{\mathsf{scan}}\right) \leq \frac{\mathsf{Var}_{\mathcal{H}_1}(\mathsf{T}_{\mathcal{K}})}{(\mathbb{E}_{\mathcal{H}_1}[\mathsf{T}_{\mathcal{K}}] - \tau_{\mathsf{scan}})^2} \leq C\frac{dk^3+k^4}{k^6}, \tag{14}$$

where the second inequality is from (Bubeck et al., 2016, Eq. (29)). Thus, we obtain Type-II converging to zero, provided $d = o(k^3)$ and $k \to \infty$. The more challenging part is the Type-I error. By the union bound over all k-subsets $\mathcal{S} \subset [n]$,

$$\mathbb{P}_{\mathcal{H}_0}\left(\mathsf{T}_{\mathsf{scan}} \geq \tau_{\mathsf{scan}}\right) \leq \binom{n}{k} \cdot \mathbb{P}_{\mathcal{H}_0}\left(\mathsf{T}_{\mathcal{K}_0} \geq \tau_{\mathsf{scan}}\right), \tag{15}$$

where $\mathcal{K}_0$ is any fixed set of size k (by symmetry). Thus, it suffices to establish an upper bound on the tail probability of the signed triangle count, namely,

$$\mathbb{P}_{\mathcal{H}_0}\left(\mathsf{T}_{\mathcal{K}_0} \geq \tau_{\mathsf{scan}}\right) \leq \; ? \quad \text{for} \quad \tau_{\mathsf{scan}} \gtrsim \frac{k^3}{\sqrt{d}}. \tag{16}$$

While sharp upper-tail results for *unsigned* triangle counts in Erdős–Rényi graphs are by now classical—originating with the martingale-based approach of Kim and Vu Kim and Vu (2004) and refined through subsequent combinatorial and variational methods Janson et al. (2004); Boucheron et al. (2013)—these techniques do not seem to extend to the signed setting, where substantial cancellations fundamentally alter the tail behavior. To overcome this obstacle, we decouple the cubic U-statistic and reduce its Laplace transform to that of a quadratic form. Concretely, write

$$\mathsf{T}_{\mathcal{K}} = \frac{1}{3} \sum_{e=\{i,j\} \subset \binom{\mathcal{K}}{2}} \bar{\mathbf{A}}_{ij} W_{ij}, \qquad W_{ij} \triangleq \sum_{\ell \in \mathcal{K} \setminus \{i,j\}} \bar{\mathbf{A}}_{i\ell} \bar{\mathbf{A}}_{j\ell}. \tag{17}$$

A decoupling inequality for canonical U-statistics de la Pena and Montgomery-Smith (1995); de la Pena and Giné (1999) implies that for all $\theta \in \mathbb{R}$,

$$\mathbb{E} \exp(\theta \mathsf{T}_{\mathcal{K}}) \leq \mathbb{E} \exp\left(c_0 \theta^2 \sum_{e \subset \mathcal{K}} W_e^2\right). \tag{18}$$

Next, a concentration argument shows $\sum_{e \subset \binom{\mathcal{K}}{2}} W_e^2 = O(k^3 \log n)$ with probability $1 - e^{-\Omega(k \log n)}$. Combining these bounds with Markov's inequality and optimizing in θ gives

$$\mathbb{P}_{\mathcal{H}_0}(\mathsf{T}_{\mathcal{K}} \geq t) \leq \exp\left(-c \frac{t^2}{k^3 \log n}\right) + \exp(-c' k \log n). \tag{19}$$

Plugging $t = \tau_{\mathsf{scan}} \asymp k^3/\sqrt{d}$ yields

$$\mathbb{P}_{\mathcal{H}_0}(\mathsf{T}_{\mathsf{scan}} \geq \tau_{\mathsf{scan}}) \leq \exp\left(k \log \frac{en}{k} - c \frac{k^3}{d \log n}\right) + \exp\left(-k(c'-1)\log n\right), \tag{20}$$

which vanishes when $k^2 \gg d \log^2 n$ and by taking $c' > 1$ (we refer to the proof for precise details). Combining the Type-I and Type-II bounds yields the stated guarantee for the scan test.

Computational lower bounds. We begin with a brief overview of the low-degree polynomial method. The central idea of this approach is to use low-degree multivariate polynomials in the entries of the observed data as surrogates for computationally efficient procedures. This heuristic is motivated by the observation that many known polynomial-time algorithms can be approximated, in an appropriate sense, by polynomials of bounded degree. The foundations of this methodology originate from a sequence of works in the sum-of-squares optimization literature Barak et al. (2016); Hopkins (2018); Hopkins and Steurer (2017); Hopkins et al. (2017).

We follow the notation and formalism introduced in Hopkins (2018); Kunisky et al. (2022). Let $\mathbb{P}_{\mathcal{H}_0}$ be a probability distribution on $\Omega_n = \{0,1\}^{\binom{n}{2}}$. This distribution induces an inner product on measurable functions $f, g : \Omega_n \to \mathbb{R}$ given by $\langle f, g\rangle_{\mathcal{H}_0} \triangleq \mathbb{E}_{\mathcal{H}_0}[f(\mathsf{G})g(\mathsf{G})]$, with associated norm $\|f\|_{\mathcal{H}_0} = \langle f, f\rangle_{\mathcal{H}_0}^{1/2}$. We denote by $L^2(\mathbb{P}_{\mathcal{H}_0})$ the Hilbert space of functions with finite $\|\cdot\|_{\mathcal{H}_0}$ norm, equipped with this inner product.

In the absence of computational constraints, the Neyman–Pearson lemma implies that the likelihood ratio test achieves the optimal tradeoff between Type-I and Type-II error probabilities. Equivalently, the likelihood ratio is the optimal distinguisher between $\mathbb{P}_{\mathcal{H}_0}$ and $\mathbb{P}_{\mathcal{H}_1}$ in the $L^2(\mathbb{P}_{\mathcal{H}_0})$ sense. Writing $\mathcal{L}_n \triangleq \mathbb{P}_{\mathcal{H}_1}/\mathbb{P}_{\mathcal{H}_0}$ for the likelihood ratio, standard second-moment arguments show that if $\|\mathcal{L}_n\|_{\mathcal{H}_0}^2$ remains bounded as $n \to \infty$, then $\mathbb{P}_{\mathcal{H}_1}$ is contiguous to $\mathbb{P}_{\mathcal{H}_0}$. In this case, the two distributions are statistically indistinguishable, in the sense that no test can simultaneously drive both error probabilities to zero.

The low-degree method asks whether a similar conclusion holds when one restricts attention to low-degree polynomials. To formalize this, let $\mathcal{V}_{n,\leq D} \subset L^2(\mathbb{P}_{\mathcal{H}_0})$ denote the subspace of polynomial functions $\Omega_n \to \mathbb{R}$ of total degree at most D. Let $\mathcal{P}_{\leq D} : L^2(\mathbb{P}_{\mathcal{H}_0}) \to \mathcal{V}_{n,\leq D}$ be the corresponding orthogonal projection operator. The *D-low-degree likelihood ratio* is defined as $\mathcal{L}_{n,\leq D} \triangleq \mathcal{P}_{\leq D}\mathcal{L}_n$, that is, the orthogonal projection of the likelihood ratio onto the space of degree-D polynomials with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}_0}$. Since the full likelihood ratio is the optimal L^2 distinguisher, the projected likelihood ratio plays an analogous optimal role among all degree-D polynomials. The following lemma formalizes this property (see e.g., Hopkins and Steurer (2017); Hopkins et al. (2017); Kunisky et al. (2022)).

Lemma 4 (Optimality of the low-degree likelihood ratio) *Consider the optimization problem*

$$\max \mathbb{E}_{\mathcal{H}_1}\left[f(\mathsf{G})\right] \quad \text{s.t.} \quad \mathbb{E}_{\mathcal{H}_0}\left[f^2(\mathsf{G})\right] = 1, f \in \mathcal{V}_{n,\leq D}, \tag{21}$$

The unique maximizer is $f^\star = \mathcal{L}_{n,\leq D}/\left\|\mathcal{L}_{n,\leq D}\right\|_{\mathcal{H}_0}$, and the optimal value equals $\left\|\mathcal{L}_{n,\leq D}\right\|_{\mathcal{H}_0}$.

In the computationally unbounded setting, boundedness of $\|\mathcal{L}_n\|_{\mathcal{H}_0}$ implies statistical indistinguishability between $\mathbb{P}_{\mathcal{H}_0}$ and $\mathbb{P}_{\mathcal{H}_1}$. The low-degree method asserts that an analogous principle governs computational limits, with $\mathcal{L}_{n,\leq D}$ playing the role of the likelihood ratio when attention is restricted to efficiently computable tests. This intuition is captured by the following informal version of the low-degree conjecture (see Hopkins (2018); Hopkins and Steurer (2017); Hopkins et al. (2017) and (Kunisky et al., 2022, Conj. 1.16)).

Conjecture 5 (Low-degree conjecture (informal)) *If there exist $\epsilon > 0$ and $D = D(n) \geq (\log n)^{1+\epsilon}$ such that $\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0}$ remains bounded as $n \to \infty$, then no polynomial-time algorithm can distinguish $\mathbb{P}_{\mathcal{H}_0}$ from $\mathbb{P}_{\mathcal{H}_1}$ (i.e., achieve strong detection).*

In what follows, we use Conjecture 5 to provide evidence for the statistical–computational gap observed above. We start with the following result.

Theorem 6 (Statistical-computational gap) *Consider the problem in (1). Suppose there exists $\varepsilon > 0$ such that: 1) $k \leq n^{1/2-\varepsilon}$ or 2) $\frac{k^6}{n^3 d}\log^3 d \leq n^{-\varepsilon}$ and $k = \Omega(\sqrt{n})$, for all large n. Then, there exists $C = C(\varepsilon)$ such that if $D \leq C\log n/\log\log n$, one has $\left\|\mathcal{L}_{n,\leq D}\right\|_{\mathcal{H}_0} = O(1)$. On the other hand, if $\frac{k^6}{n^3 d} = \omega(1)$, then $\left\|\mathcal{L}_{n,\leq D}\right\|_{\mathcal{H}_0} = \omega(1)$.*

Theorem 6 establishes boundedness of the low-degree likelihood ratio up to $D \leq \frac{\log n}{\log \log n}$, which is near-logarithmic and thus provides evidence for computational hardness within this framework. This interpretation is consistent with analogous results in related planted problems (e.g., Bangachev and Bresler (2024), Elimelech and Huleihel (2025), Wein (2025)).

Together with Conjecture 5, Theorem 6 suggests that, if low-degree polynomials are taken as a proxy for efficient computation, then no polynomial-time algorithm can distinguish the null and alternative hypotheses in the regime $1 \vee \frac{k^6}{n^3} \ll d \ll k^2$. Equivalently, low-degree polynomials fail whenever either the problem is planted-clique hard, or the geometric signal is sufficiently weak that even the signed-triangle statistic is ineffective.

These predictions align precisely with the statistical–computational tradeoffs identified above. A more explicit characterization of the computational barrier, exhibiting its dependence on D, can be extracted from the proof of Theorem 6; for clarity of exposition, we have chosen to present the simplified formulation stated above.

A key ingredient in the proof of Theorem 6 is a recent powerful result (Bangachev and Bresler, 2024, Thm. 1.1), which bounds centered subgraph moments (equivalently, Fourier coefficients) of spherical random geometric graphs by showing they decay polynomially in d. This decay, together with the elementary observation that tree-like subgraphs contribute nothing to the low-degree likelihood, ensures that only cyclic subgraphs contribute to the low-degree likelihood; among these their total contribution can be controlled when $\frac{k^6}{n^3} \ll d$.

Proof sketch of Theorem 6. We apply the low-degree framework (see, e.g., Hopkins (2018); Kunisky et al. (2022); Wein (2025)). Under $\mathcal{H}_0$, the Fourier characters $\chi_\alpha(\mathsf{G}) = \prod_{e\in\alpha} \frac{G_e - q}{\sqrt{q(1-q)}}$ form an orthonormal basis, so Parseval gives

$$\|\mathcal{L}_{n,\leq D}\|^2_{\mathcal{H}_0} = 1 + \sum_{1\leq|\alpha|\leq D} \left(\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})]\right)^2. \tag{22}$$

In our model, a coefficient is nonzero only if all edges of α lie within the planted set, and it equals

$$\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})] = \frac{(k)_{v(\alpha)}}{(n)_{v(\alpha)}} \mathbb{E}\left[\prod_{e\in\alpha} Z_e\right], \tag{23}$$

where $Z_e \triangleq \frac{\sigma_e - q}{\sqrt{q(1-q)}}$, for $e \in \binom{[n]}{2}$. In the regime $k \leq n^{1/2-\varepsilon}$, by bounding $|Z_e| \leq \sqrt{(1-q)/q}$ we effectively dominate the geometric model by planted clique model. It is well-known (see, e.g., Hopkins (2018)) that for the planted clique model, the low-degree second moment is bounded, i.e., $\|\mathcal{L}_{n,\leq D}\|^2_{\mathcal{H}_0} = O(1)$, if and only if $k \leq n^{1/2-\epsilon}$, uniformly over all d. Consider now the regime $k = \Omega(\sqrt{n})$. If the graph induced by α has a leaf, the centered moment vanishes by conditioning on the leaf vertex; hence only cyclic patterns (minimum degree ≥ 2) contribute. For such patterns, the Fourier decay theorem of Bangachev–Bresler (Bangachev and Bresler, 2024, Thm. 1.1) yields

$$\left|\mathbb{E}\left[\prod_{e\in\alpha} Z_e\right]\right| \leq A^{|\alpha|} \left(\frac{B|\alpha||v(\alpha)|\mathsf{polylog}(d)}{\sqrt{d}}\right)^{\Omega(|v(\alpha)|)}, \tag{24}$$

where $A, B > 0$ are some constants, and $|v(\alpha)|$ is the number of vertices in the graph induced by α. Summing these bounds over all $|\alpha| \leq D$ via a crude combinatorial count gives

$$\|\mathcal{L}_{n,\leq D}\|^2_{\mathcal{H}_0} \lesssim 1 + (CD)^D \beta_n \quad \text{with} \quad \beta_n \triangleq \frac{k^6}{n^3 d} \log^3 d. \tag{25}$$

Hence, if $\beta_n \leq n^{-\varepsilon}$ with $k = \Omega(\sqrt{n})$, then choosing $D \leq c \frac{\log n}{\log\log n}$ implies $\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0} = O(1)$. Conversely, the upper bound is obtained by isolating the contribution of triangles: when $d \ll k^6/n^3$, the triangle moment alone causes the low-degree norm to diverge.

Acknowledgments

This work is supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1734/21).

References

Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2009.

Ery Arias-Castro and Nicolas Verzelen. Community detection in dense random networks. *The Annals of Statistics*, 42(3):940–969, 2014.

Kiril Bangachev and Guy Bresler. On the fourier coefficients of high-dimensional random geometric graphs. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC 2024)*, pages 549–560, Vancouver, BC, Canada, 2024. ACM. doi: 10.1145/3596830.3602707. Also available as arXiv:2402.12589.

B. Barak, S. B. Hopkins, J. Kelner, P. Kothari, A. Moitra, and A. Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 428–437, 2016.

Gianmarco Bet, Kay Bogerd, Rui M. Castro, and Remco van der Hofstad. Detecting a botnet in a network. *Mathematics and Statistics Learning*, 3(3–4):315–343, 2020.

Jinho Bok, Shuangping Li, and Sophie H. Yu. Detection of local geometry in random graphs: Information-theoretic and computational limits. *arXiv preprint arXiv:2603.24545*, 2026.

Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, Oxford, 2013. ISBN 9780199535255.

Matthew Brennan and Guy Bresler. Reducibility and statistical-computational gaps from secret leakage. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125, pages 648–847, 09–12 Jul 2020.

Matthew Brennan, Guy Bresler, and Wasim Huleihel. Reducibility and computational lower bounds for problems with planted sparse structure. In *Proceedings of the 31st Conference On Learning Theory*, volume 75, pages 48–166, 06–09 Jul 2018.

Matthew Brennan, Guy Bresler, and Dheeraj Nagaraj. Phase transitions for detecting latent geometry in random graphs. *Probability Theory and Related Fields*, 178(3):1215–1289, 2020.

Matthew Brennan, Guy Bresler, and Brice Huang. De finetti–style results for wishart matrices: Combinatorial structure and phase transitions. *arXiv preprint*, 2021. URL https://arxiv.org/abs/2103.14011.

Matthew S. Brennan, Guy Bresler, and Brice Huang. Threshold for detecting high dimensional geometry in anisotropic random geometric graphs. *Random Structures & Algorithms*, 64(1): 125–137, 2024.

Sébastien Bubeck, Jian Ding, Ronen Eldan, and Miklós Z Rácz. Testing for high-dimensional geometry in random graphs. *Random Structures & Algorithms*, 49(3):503–532, 2016.

Cristina Butucea and Yuri I Ingster. Detection of a sparse submatrix of a high-dimensional noisy matrix. *Bernoulli*, 19(5B):2652–2688, 2013.

Victor H. de la Pena and E. Giné. *Decoupling: From Dependence to Independence*. Probability and its Applications. Springer, New York, 1999.

Victor H. de la Pena and S. J. Montgomery-Smith. Decoupling Inequalities for the Tail Probabilities of Multivariate U-Statistics. *The Annals of Probability*, 23(2):806 – 816, 1995.

Carl P. Dettmann and Orestis Georgiou. Random geometric graphs with general connection functions. *Phys. Rev. E*, 93:032313, Mar 2016.

Luc Devroye, András György, Gábor Lugosi, and Frederic Udina. High-dimensional random geometric graphs and their clique number. *Electronic Journal of Probability*, 16:2481–2508, 2011. doi: 10.1214/EJP.v16-967. URL https://doi.org/10.1214/EJP.v16-967.

Ronen Eldan and Danny Mikulincer. Information and dimensionality of anisotropic random geometric graphs. *Annals of Probability*, 48(4):1980–2015, 2020. doi: 10.1214/19-AOP1406.

Dor Elimelech and Wasim Huleihel. Detecting arbitrary planted subgraphs in random graphs. In *Proceedings of the Thirty Eighth Conference on Learning Theory*, volume 291 of *Proceedings of Machine Learning Research*, pages 1691–1798. PMLR, 2025.

Shirshendu Ganguly, Elisabeth Hiesmayr, and Kyeongsik Nam. Upper tail behavior of the number of triangles in random graphs with constant average degree. *Combinatorica*, 44:699–740, 2024. doi: 10.1007/s00493-024-00086-3.

Valentin Gutev. On real-valued functions of lipschitz type. *Expositiones Mathematicae*, 43(5): 125701, 2025.

S. B. Hopkins and D. Steurer. Efficient bayesian estimation from few samples: Community detection and related problems. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 379–390, 2017.

Samuel B. Hopkins. *Statistical Inference and the Sum of Squares Method*. PhD thesis, Cornell University, 2018.

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. *Proceedings of the fifty-eighth IEEE Foundations of Computer Science (FOCS)*, pages 720–731, 2017.

Wasim Huleihel. Inferring hidden structures in random graphs. *IEEE Transactions on Signal and Information Processing over Networks*, 8:855–867, 2022. doi: 10.1109/TSIPN.2022.3211208.

Yuri I. Ingster. Some problems of hypothesis testing leading to infinitely divisible distributions. *Mathematical Methods of Statistics*, 6(1):47–69, 1997.

Svante Janson, Krzysztof Oleszkiewicz, and Andrzej Ruciński. Upper tails for subgraph counts in random graphs. *Israel Journal of Mathematics*, 142:61–92, 2004. doi: 10.1007/BF02771528.

Jeong Han Kim and Van H Vu. Divide and conquer martingales and the number of triangles in a random graph. *Random Structures & Algorithms*, 24(2):166–174, 2004.

Dmitriy Kunisky, Alexander S. Wein, and Afonso S. Bandeira. Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. In *Mathematical Analysis, its Applications and Computation*, volume 385 of *Springer Proceedings in Mathematics & Statistics*, pages 1–50. Springer, 2022. URL https://arxiv.org/abs/1907.11636. ISAAC Congress 2019.

Suqi Liu and Miklós Z. Rácz. Phase transition in noisy high-dimensional random geometric graphs. *Electronic Journal of Statistics*, 17(2):3512–3574, 2023a. doi: 10.1214/23-EJS2162. URL https://doi.org/10.1214/23-EJS2162.

Suqi Liu and Miklós Z. Rácz. A probabilistic view of latent space graphs and phase transitions. *Bernoulli*, 29(3):2417–2441, 2023b. doi: 10.3150/22-BEJ1547. URL https://doi.org/10.3150/22-BEJ1547.

Srinivasan Parthasarathy, David Sivakoff, Minghao Tian, and Yusu Wang. A quest to unravel the metric structure behind perturbed networks. In *33rd International Symposium on Computational Geometry (SoCG 2017)*, volume 77 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 53:1–53:15. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2017. doi: 10.4230/LIPIcs.SoCG.2017.53.

Mathew Penrose. *Random Geometric Graphs*. Oxford University Press, Oxford, 2007.

Yury Polyanskiy and Yihong Wu. *Information Theory: From Coding to Learning*. Cambridge University Press, 2024.

Igal Sason. Bounds on f-divergences and related distances. *CCIT Report*, 859(859), 2014.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, 2009. doi: 10.1007/b13794.

Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press, Cambridge, 2018. ISBN 9781108415194.

Nicolas Verzelen and Ery Arias-Castro. Community detection in sparse random networks. *The Annals of Applied Probability*, 25(6):3465–3510, 2015.

Alexander S. Wein. Computational complexity of statistics: New insights from low-degree polynomials. *arXiv preprint arXiv:2506.10748*, Jun 2025. doi: 10.48550/arXiv.2506.10748. URL https://arxiv.org/abs/2506.10748.

Michael Wilsher, Carl P. Dettmann, and Ayalvadi Ganesh. Connectivity in one-dimensional soft random geometric graphs. *Physical Review E*, 102(6):062312, 2020. doi: 10.1103/PhysRevE.102.062312.

Yihong Wu, Jiaming Xu, and Sophie H. Yu. Testing correlation of unlabeled random graphs. *Annals of Applied Probability*, 33(4):2519–2558, 2023.

Appendix A. Related Work

The study of random geometric graphs in high dimensions traces back to early work showing, via a multivariate central limit theorem, that geometric graphs become asymptotically indistinguishable from Erdős–Rényi graphs as the dimension grows Devroye et al. (2011). This phenomenon was later sharpened by work that identified a precise phase transition in the dense regime, demonstrating that geometry is lost when the dimension exceeds a cubic threshold in the number of vertices Bubeck et al. (2016). These results also uncovered deep connections between geometric graph models and classical random matrix ensembles, showing that a matching Wishart-to-GOE transition occurs at the same scale (see also Eldan and Mikulincer (2020); Brennan et al. (2024)).

Extensions of these results have explored broader distributional assumptions and anisotropic settings, including information-theoretic limits for detecting geometry when the latent distribution is anisotropic Eldan and Mikulincer (2020). In the sparse regime, where the edge probability vanishes with the graph size, it was conjectured that geometry should be lost at significantly lower dimensions Bubeck et al. (2016). Progress toward this conjecture, breaking the cubic barrier, was subsequently obtained Brennan et al. (2020), with further refinements appearing in recent work on partial and masked Wishart ensembles Brennan et al. (2021). This remains an active and rapidly developing area. More recently, a unified framework for noisy high-dimensional geometric graphs that interpolate between purely geometric and Erdős–Rényi models was introduced in Liu and Rácz (2023a,b). These results quantify how the strength of geometric dependence interacts with dimensionality to determine detectability. Our work complements this line by focusing on a *localized* setting, where geometric structure is present only on a planted subset of vertices.

Beyond hard-threshold geometric graphs, a substantial literature studies *soft* random geometric graphs, in which the probability of an edge depends smoothly on latent distances or inner products. Such models arise naturally in wireless communication, social networks, and biological systems. A foundational probabilistic treatment of geometric and soft geometric graphs, including connectivity properties in fixed dimensions, was developed in Penrose (2007). Subsequent works have analyzed connectivity and phase transitions from both probabilistic and statistical physics perspectives; see, for example, Dettmann and Georgiou (2016) and related studies of one-dimensional and perturbed geometric networks Parthasarathy et al. (2017); Wilsher et al. (2020).

Detecting hidden structure in random graphs has also been extensively studied through planted subgraph models, including planted dense subgraphs and community detection. Sharp detection thresholds and computational barriers in dense regimes were characterized in Arias-Castro and Verzelen (2014); Verzelen and Arias-Castro (2015). More general formulations of planted subgraph detection, including arbitrary planted patterns, were studied in Huleihel (2022); Elimelech and Huleihel (2025), among others. These works typically posit an explicit increase in edge probability on a hidden vertex set, with edges remaining conditionally independent, a modeling choice that stands in contrast to the geometric mechanism considered here. In contrast, the alternative considered in the present paper is generated by a latent geometric mechanism, which induces structured dependencies among edges. This places our work at the intersection of planted subgraph detection and geometric random graph theory and, to our knowledge, provides the first sharp detection theory for a planted structure whose defining feature is high-dimensional geometry.

Concurrent and independent work by Bok et al. (2026) studies the same geometric detection problem as in our paper and derives similar information-theoretic and computational limits. They also analyze the sparse regime where the edge density vanishes polynomially with the size of the graph, establishing corresponding statistical lower and upper bounds.

Finally, we mention Bet et al. (2020). This paper studies the problem of detecting a botnet hidden within a larger network. The authors model the network as a graph and assume that a botnet appears as a group of nodes with slightly different connectivity patterns compared to normal users. More specifically, the hypothesis testing problem considered in their paper, is formulated with a geometric random graph as the null model, while under the alternative a "small" Erdős–Rényi subgraph is planted within the geometric random graph. In contrast, our setting reverses this perspective: the null hypothesis is an Erdős–Rényi graph, and under the alternative we plant a geometric random graph. As a result, the underlying models, as well as the corresponding results and techniques, differ substantially.

Appendix B. Notation

Given a probability distribution $\mathbb{P}$, we write $\mathbb{P}^{\otimes n}$ for the law of the n-dimensional random vector $(X_1, X_2, \ldots, X_n)$, where the coordinates are independent and identically distributed according to $\mathbb{P}$. Likewise, $\mathbb{P}^{\otimes m\times n}$ denotes the distribution on $\mathbb{R}^{m\times n}$ whose entries are i.i.d. with common distribution $\mathbb{P}$. For a finite or measurable set $\mathcal{X}$, $\mathsf{Unif}[\mathcal{X}]$ denotes the uniform distribution on $\mathcal{X}$. The notation $X \perp\!\!\!\perp Y$ indicates that the random variables X and Y are independent.

The spectral (operator) norm of a symmetric matrix $\mathbf{A}$ is written as $\|\mathbf{A}\|_{\mathsf{op}}$, and $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. For an $n \times n$ matrix $\mathbf{A}$ and a set $S \subseteq \{1, 2, \ldots, n\}$, we use $\mathbf{A}|_S$ and $\mathbf{A}[S]$ to denote the submatrix obtained by restricting $\mathbf{A}$ to the rows and columns indexed by S.

We write $\mathcal{N}(\mu, \sigma^2)$ for a univariate normal random variable with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}_{\geq 0}$. Let Φ denote the cumulative distribution function of a standard normal random variable, given by $\Phi(x) = \int_{-\infty}^{x} e^{-t^2/2}\mathrm{d}t$, and $\bar{\Phi}$ its complement, i.e $\bar{\Phi}(x) = 1 - \Phi(x)$. For probability measures $\mathbb{P}$ and $\mathbb{Q}$, we define the total variation distance, and χ^2-divergence, respectively, by $d_{\mathsf{TV}}(\mathbb{P}, \mathbb{Q}) = \frac{1}{2}\int |\mathrm{d}\mathbb{P} - \mathrm{d}\mathbb{Q}|$, $\chi^2(\mathbb{P}||\mathbb{Q}) = \int \frac{(\mathrm{d}\mathbb{P} - \mathrm{d}\mathbb{Q})^2}{\mathrm{d}\mathbb{Q}}$. We denote by $\mathsf{Bern}(p)$ and $\mathsf{Binomial}(n, p)$ the Bernoulli and Binomial distributions with parameters p and (n, p), respectively, and by $\mathsf{Hypergeometric}(n, k, m)$ the Hypergeometric distribution with parameters (n, k, m).

Throughout, we use standard asymptotic notation. For two positive sequences $\{a_n\}$ and $\{b_n\}$, we write $a_n = O(b_n)$ if there exists a constant C such that $a_n \leq C b_n$ for all n, and $a_n = \Omega(b_n)$ if $b_n = O(a_n)$. We write $a_n = \Theta(b_n)$ if both $a_n = O(b_n)$ and $a_n = \Omega(b_n)$ hold. Moreover, $a_n = o(b_n)$ (equivalently, $b_n = \omega(a_n)$) if $a_n/b_n \to 0$ as $n \to \infty$. We write $a_n \ll b_n$ to indicate that a_n is polynomially smaller than b_n in n, that is, $\liminf_{n\to\infty} \log a_n < \liminf_{n\to\infty} \log b_n$. For $n \in \mathbb{N}$, we let $[n] = \{1, 2, \ldots, n\}$. For real numbers a and b, we define $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$. Unless otherwise specified, C denotes a generic constant independent of the problem parameters and may vary from line to line. For integers n and m we denote by $(n)_m$ the falling factorial, i.e., $(n)_m = n(n-1)(n-2)\cdots(n-m+1)$. Finally, for a set $S \subseteq \mathbb{R}$, $\mathbb{1}\{S\}$

denotes its indicator function; we also write $\mathbb{1}_S(x)$ with the same meaning, namely, $\mathbb{1}_S(x) = 1$ if $x \in S$ and 0 otherwise.

Appendix C. Proof of Theorem 2

In this subsection we prove Theorem 2. Our aim is to derive a lower bound on the optimal risk, thereby precluding the possibility of successful detection. Define the likelihood ratio as

$$\mathcal{L}(\mathbf{A}) \triangleq \frac{\mathrm{d}\mathbb{P}_{\mathcal{H}_1}}{\mathrm{d}\mathbb{P}_{\mathcal{H}_0}}(\mathbf{A}), \tag{26}$$

namely, the Radon–Nikodym derivative of $\mathbb{P}_{\mathcal{H}_1}$ relative to the measure $\mathbb{P}_{\mathcal{H}_0}$, and we recall that $\mathbf{A}$ denotes the graph adjacency matrix. It is classical (see, e.g., (Tsybakov, 2009, Theorem 2.2)) that the test minimizing the risk is the likelihood ratio test

$$\mathcal{A}^*(\mathbf{A}) \triangleq \mathbb{1}\{\mathcal{L}(\mathbf{A}) \geq 1\}, \tag{27}$$

and that the corresponding optimal risk satisfies

$$\mathsf{R}^* = 1 - d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1}). \tag{28}$$

Recalling that the chi-square divergence admits the representation $\chi^2(\mathbb{P}_{\mathcal{H}_1} \| \mathbb{P}_{\mathcal{H}_0}) = \mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2] - 1$, it follows from standard inequalities relating total variation and chi-square divergences (see, e.g., (Tsybakov, 2009, Sec. 2), (Sason, 2014, Prop. 3)) that

$$\chi^2(\mathbb{P}_{\mathcal{H}_1} \| \mathbb{P}_{\mathcal{H}_0}) \geq \max\left\{ \frac{1}{2(1 - d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1}))} - 1, \left(2 d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1})\right)^2 \right\}. \tag{29}$$

Consequently, the optimal risk admits the lower bound

$$\mathsf{R}^* \geq \max\left\{ 1 - \frac{1}{2}\sqrt{\chi^2(\mathbb{P}_{\mathcal{H}_1} \| \mathbb{P}_{\mathcal{H}_0})}, \frac{1}{2\left(1 + \chi^2(\mathbb{P}_{\mathcal{H}_1} \| \mathbb{P}_{\mathcal{H}_0})\right)} \right\}. \tag{30}$$

In particular, the optimal risk remains bounded away from zero whenever $\mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2]$ is bounded, and converges to one if $\mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2] = 1 + o(1)$. Therefore, to rule out detection it suffices to control the second moment of the likelihood ratio under $\mathcal{H}_0$. To conclude

$$\mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2] = 1 + o(1) \implies d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0} \| \mathbb{P}_{\mathcal{H}_1}) = o(1), \tag{31}$$

$$\mathbb{E}_{\mathcal{H}_0}[\mathcal{L}^2] = O(1) \implies d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0} \| \mathbb{P}_{\mathcal{H}_1}) \leq 1 - \Omega(1). \tag{32}$$

Let us introduce some notation. Let $\mathbf{X}$ denote the $n \times d$ matrix formed by stacking n i.i.d. standard Gaussian vectors, and let $\mathbf{W} = \mathbf{X}\mathbf{X}^T$ be the associated $n \times n$ Wishart matrix. Observe that $\mathbf{W}_{ii} = \|\mathbf{x}_i\|^2$, and that

$$\left\langle \mathbf{x}_i / \|\mathbf{x}_i\|,\ \mathbf{x}_j / \|\mathbf{x}_j\| \right\rangle = \mathbf{W}_{ij} / \sqrt{\mathbf{W}_{ii}\mathbf{W}_{jj}}. \tag{33}$$

Define

$$\sigma_{ij} \triangleq \mathbb{1}\{\langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq t_{q,d}\}, \tag{34}$$

or equivalently,

$$\sigma_{ij} = \mathbb{1}\left\{\mathbf{W}_{ij}/\sqrt{\mathbf{W}_{ii}\mathbf{W}_{jj}} \geq t_{q,d}\right\}. \tag{35}$$

For $\mathcal{K} \in \binom{[n]}{k}$, we write $E(\mathcal{K}) = \{\{i,j\} : i,j \in \mathcal{K},\ i \neq j\}$ and $E^c(\mathcal{K}) = E([n]) \setminus E(\mathcal{K})$. With these notations in place, we may write the likelihood ratio explicitly. Recall the detection problem in (1). Under the null hypothesis

$$\mathbb{P}_{\mathcal{H}_0}(\mathbf{A}) = \prod_{i<j} q^{\mathbf{A}_{i,j}}(1-q)^{1-\mathbf{A}_{i,j}}, \tag{36}$$

while under the alternative

$$\mathbb{P}_{\mathcal{H}_1|\mathcal{K}}(\mathbf{A}|\mathcal{K}) = \prod_{\{i,j\}\in E^c(\mathcal{K})} q^{\mathbf{A}_{i,j}}(1-q)^{1-\mathbf{A}_{i,j}} \cdot \mathbb{E}_{\mathbf{W}|\mathcal{K}}\left[\prod_{\{i,j\}\in E(\mathcal{K})} \sigma_{ij}^{\mathbf{A}_{i,j}}(1-\sigma_{ij})^{1-\mathbf{A}_{i,j}}\right]. \tag{37}$$

It follows that the likelihood ratio takes the form

$$\mathcal{L}(\mathbf{A}) = \frac{\mathbb{E}_{\mathcal{K}}\mathbb{P}_{\mathcal{H}_1|\mathcal{K}}(\mathbf{A}|\mathcal{K})}{\mathbb{P}_{\mathcal{H}_0}(\mathbf{A})} = \mathbb{E}_{\mathcal{K}}\mathbb{E}_{\mathbf{W}|\mathcal{K}}\left[\prod_{\{i,j\}\in E(\mathcal{K})}\left(\frac{\sigma_{ij}}{q}\right)^{\mathbf{A}_{i,j}}\left(\frac{1-\sigma_{ij}}{1-q}\right)^{1-\mathbf{A}_{i,j}}\right]. \tag{38}$$

In this expression, the outer expectation over $\mathcal{K}$ is uniform over $\binom{[n]}{k}$, and the inner expectation is with respect to the Wishart distribution.

As discussed above, our objective is to upper bound the second moment of the likelihood ratio under $\mathcal{H}_0$. Nevertheless, in the present setting, certain rare events under $\mathcal{H}_1$ may cause this second moment to diverge, even though $d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0}, \mathbb{P}_{\mathcal{H}_1})$ remains bounded away from one. To circumvent the effect of such atypical events, we instead analyze the second moment conditioned on events that are typical under $\mathcal{H}_1$. This refined approach, originally proposed by Ingster (1997); Butucea and Ingster (2013), is based on controlling the first and second moments of a truncated likelihood ratio Arias-Castro and Verzelen (2014); Wu et al. (2023).

We begin by outlining the truncated second-moment method in general. Let $\mathcal{F}$ be an event such that $\mathbb{P}(\mathcal{F}) = 1 - o(1)$. Define the *truncated/conditional planted model* as

$$\begin{aligned}\tilde{\mathbb{P}}_{\mathcal{H}_1,\mathcal{K},\mathbf{W}}(\mathbf{A},\mathcal{K},\mathbf{W}) &= \frac{\mathbb{P}_{\mathcal{H}_1,\mathcal{K},\mathbf{W}}(\mathbf{A},\mathcal{K},\mathbf{W})\mathbb{1}\{(\mathcal{K},\mathbf{W})\in\mathcal{F}\}}{\mathbb{P}(\mathcal{F})} \\ &= (1+o(1))\cdot\mathbb{P}_{\mathcal{H}_1,\mathcal{K},\mathbf{W}}(\mathbf{A},\mathcal{K},\mathbf{W})\mathbb{1}\{(\mathcal{K},\mathbf{W})\in\mathcal{F}\},\end{aligned} \tag{39}$$

and note that this is a legitimate probability measure. Then, define the *truncated likelihood ratio* as

$$\tilde{\mathcal{L}}(\mathbf{A}) \triangleq \frac{\tilde{\mathbb{P}}_{\mathcal{H}_1}(\mathbf{A})}{\mathbb{P}_{\mathcal{H}_0}(\mathbf{A})} = \frac{1}{\mathbb{P}(\mathcal{F})}\mathbb{E}_{\mathcal{K},\mathbf{W}}\left[\frac{\mathbb{P}_{\mathcal{H}_1|\mathcal{K},\mathbf{W}}(\mathbf{A}|\mathcal{K},\mathbf{W})\mathbb{1}\{(\mathcal{K},\mathbf{W})\in\mathcal{F}\}}{\mathbb{P}_{\mathcal{H}_0}(\mathbf{A})}\right] \tag{40}$$

$$= (1+o(1))\cdot\mathbb{E}_{\mathcal{K},\mathbf{W}}\left[\frac{\mathbb{P}_{\mathcal{H}_1|\mathcal{K},\mathbf{W}}(\mathbf{A}|\mathcal{K},\mathbf{W})\mathbb{1}\{(\mathcal{K},\mathbf{W})\in\mathcal{F}\}}{\mathbb{P}_{\mathcal{H}_0}(\mathbf{A})}\right]. \tag{41}$$

Now, by the data processing inequality for the total variation distance, we know that

$$d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_1}, \tilde{\mathbb{P}}_{\mathcal{H}_1}) \leq d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_1,\mathcal{K},\mathbf{W}}, \tilde{\mathbb{P}}_{\mathcal{H}_1,\mathcal{K},\mathbf{W}}) = \mathbb{P}[\mathcal{F}^c] = o(1). \tag{42}$$

Accordingly, combining (31)–(32) with (42), the triangle inequality implies that

$$\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2] = 1 + o(1) \implies d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0} \| \mathbb{P}_{\mathcal{H}_1}) = o(1), \tag{43}$$

$$\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2] = O(1) \implies d_{\mathsf{TV}}(\mathbb{P}_{\mathcal{H}_0} \| \mathbb{P}_{\mathcal{H}_1}) \leq 1 - \Omega(1). \tag{44}$$

Therefore, with a carefully chosen high probable truncation set $\mathcal{F}$, it suffices to analyze the second moment of the truncated likelihood ratio.

We now proceed to define the truncation set and likelihood ratio in our setting. For any given $\mathcal{K} \subseteq [n]$, define the truncation set as

$$\Gamma_{\mathcal{K}} \triangleq \bigcap_{z=2}^{|\mathcal{K}|} \left\{ \Gamma_{z,\mathcal{K}}^{\mathsf{fro}} \cap \Gamma_{z,\mathcal{K}}^{\mathsf{op}} \right\}, \tag{45}$$

where for $z \geq 2$

$$\Gamma_{z,\mathcal{K}}^{\mathsf{fro}} \triangleq \bigcap_{T \subseteq \mathcal{K}, |T| = z} \bigcap_{i \in \mathcal{T}} \left\{ \sum_{j \in T \setminus \{i\}} A_{ij}^2 \leq L_{k,z,d}^{\mathsf{fro}} \right\}, \tag{46}$$

and

$$\Gamma_{z,\mathcal{K}}^{\mathsf{op}} \triangleq \bigcap_{T \subseteq \mathcal{K}, |T| = z} \left\{ \|A|_T - d\mathbf{I}_z\|_{\mathsf{op}} \leq L_{k,z,d}^{\mathsf{op}} \right\}, \tag{47}$$

where we recall that $A|_T$ is the restriction of A to the indices in T, $L_{k,z,d}^{\mathsf{fro}} \triangleq (1 + C_1)d(z-1)\log k$, and $L_{k,z,d}^{\mathsf{op}} \triangleq C_2(\sqrt{dz} + \sqrt{d \log \binom{k}{z}})$, for some $C_1, C_2 > 0$, specified later on. Then, using (36)–(38) and (39)–(40), we see that the truncated planted model on $\Gamma_{\mathcal{K}}$ is given by

$$\tilde{\mathbb{P}}_{\mathcal{H}_1|\mathcal{K}}(\mathbf{A}|\mathcal{K}) = \frac{1}{\mathbb{P}[\Gamma_{\mathcal{K}}]} \prod_{\{i,j\} \in E^c(\mathcal{K})} q^{\mathbf{A}_{i,j}} (1-q)^{1-\mathbf{A}_{i,j}} \cdot \mathbb{E}_{\mathbf{W}|\mathcal{K}} \left[\prod_{\{i,j\} \in E(\mathcal{K})} \sigma_{ij}^{\mathbf{A}_{i,j}} (1-\sigma_{ij})^{1-\mathbf{A}_{i,j}} \mathbb{1}_{\Gamma_{\mathcal{K}}}(\mathbf{W}) \right], \tag{48}$$

and thus

$$\tilde{\mathcal{L}}(\mathbf{A}) = \frac{1}{\mathbb{P}[\Gamma_{\mathcal{K}}]} \mathbb{E}_{\mathcal{K}} \mathbb{E}_{\mathbf{W}|\mathcal{K}} \left[\prod_{\{i,j\} \in E(\mathcal{K})} \left(\frac{\sigma_{ij}}{q} \right)^{\mathbf{A}_{i,j}} \left(\frac{1-\sigma_{ij}}{1-q} \right)^{1-\mathbf{A}_{i,j}} \mathbb{1}_{\Gamma_{\mathcal{K}}}(\mathbf{W}) \right]. \tag{49}$$

We also denote the conditional likelihood

$$\tilde{\mathcal{L}}_{\mathcal{K}}(\mathbf{A}|\mathcal{K}) \triangleq \frac{\tilde{\mathbb{P}}_{\mathcal{H}_1|\mathcal{K}}(\mathbf{A}|\mathcal{K})}{\mathbb{P}_{\mathcal{H}_0}(\mathbf{A})} \tag{50}$$

$$= \frac{\tilde{\mathbb{P}}_{\mathcal{H}_1|\mathcal{K}}(\mathbf{A}_{\mathcal{K}})}{\mathbb{P}_{\mathcal{H}_0}(\mathbf{A}_{\mathcal{K}})} \tag{51}$$

$$= \frac{1}{\mathbb{P}[\Gamma_{\mathcal{K}}]} \mathbb{E}_{\mathbf{W}|\mathcal{K}} \left[\prod_{\{i,j\} \in E(\mathcal{K})} \left(\frac{\sigma_{ij}}{q} \right)^{\mathbf{A}_{i,j}} \left(\frac{1-\sigma_{ij}}{1-q} \right)^{1-\mathbf{A}_{i,j}} \mathbb{1}_{\Gamma_{\mathcal{K}}}(\mathbf{W}) \right], \tag{52}$$

and we note that $\tilde{\mathcal{L}}(\mathbf{A}) = \mathbb{E}_{\mathcal{K}}\left[\tilde{\mathcal{L}}_{\mathcal{K}}(\mathbf{A}|\mathcal{K})\right]$. Next, we show that the truncation set defined above has high probability, and then derive conditions under which $\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2]$ is bounded or converges to unity.

C.1. High probability truncation set

We show that $\mathbb{P}[\Gamma_{\mathcal{K}}] = 1 - o(1)$. We have

$$\mathbb{P}[\Gamma_{\mathcal{K}}] = \mathbb{E}_{\mathcal{K}}\left[\mathbb{P}_{\mathbf{W}|\mathcal{K}}[\Gamma_{\mathcal{K}}]\right] = \mathbb{P}_{\mathbf{W}}[\Gamma_{\mathcal{K}_0}] \tag{53}$$

where $\mathcal{K}_0$ is any fixed subset of size k in $[n]$, the last equality is because $\mathbb{P}_{\mathbf{W}|\mathcal{K}}[\Gamma_{\mathcal{K}}]$ does not depend on the particular choice of $\mathcal{K}$ by symmetry, and $\mathbb{P}_{\mathbf{W}}$ denotes the Wishart distribution. By the union bound we have

$$\mathbb{P}_{\mathbf{W}}(\Gamma_{\mathcal{K}_0}^c) \leq \mathbb{P}_{\mathbf{W}}\left[\left(\Gamma_{\mathcal{K}_0}^{\mathsf{fro}}\right)^c\right] + \mathbb{P}_{\mathbf{W}}\left[\left(\Gamma_{\mathcal{K}_0}^{\mathsf{op}}\right)^c\right], \tag{54}$$

where $\Gamma_{\mathcal{K}_0}^{\mathsf{fro}} \triangleq \cap_{z=2}^{k}\Gamma_{z,\mathcal{K}}^{\mathsf{fro}}$ and $\Gamma_{\mathcal{K}_0}^{\mathsf{op}} \triangleq \cap_{z=2}^{k}\Gamma_{z,\mathcal{K}}^{\mathsf{op}}$, and we next prove that both terms at the right-hand side of (54) converge to zero as $n \to \infty$.

We start with $\mathbb{P}_{\mathbf{W}}\left[\left(\Gamma_{\mathcal{K}_0}^{\mathsf{fro}}\right)^c\right]$. Fix $z \in [k]$ and $T \subseteq \mathcal{K}_0$ with $|T| = z$. Recall that $\mathbf{W}_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$, where $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are i.i.d. d-dimensional standard Gaussian vectors. Conditioning on $\mathbf{x}_i = x$, we have for any $j \neq i$,

$$\langle x, \mathbf{x}_j \rangle = \sum_{\ell=1}^{d} x_\ell \mathbf{x}_{j\ell} \sim \mathcal{N}(0, \|x\|^2), \tag{55}$$

and these random variables are independent across j. Consequently, given $\mathbf{x}_i = x$,

$$\sum_{j \in T \setminus \{i\}} \mathbf{W}_{ij}^2 \stackrel{d}{=} \|x\|^2 \cdot \mathsf{A}, \tag{56}$$

where $\mathsf{A} \sim \chi^2_{z-1}$. Since $\|\mathbf{x}_i\|^2 \sim \chi^2_d$ and $\mathbf{x}_i$ is independent of $\{\mathbf{x}_j\}_{j \neq i}$, it follows that, unconditionally,

$$\sum_{j \in T \setminus \{i\}} \mathbf{W}_{ij}^2 \stackrel{d}{=} \mathsf{A} \cdot \mathsf{B}, \tag{57}$$

where $\mathsf{B} \sim \chi^2_d$, and $\mathsf{A} \perp\!\!\!\perp \mathsf{B}$. Recall that for a chi-square random variable $Y \sim \chi^2_m$, the following standard Chernoff bound holds (see, e.g., (Boucheron et al., 2013, Ch. 2)):

$$\mathbb{P}\left(Y \geq (1+\varepsilon)m\right) \leq \exp\left(-\frac{m}{2}\left(\varepsilon - \log(1+\varepsilon)\right)\right), \tag{58}$$

for any $\varepsilon > 0$. Furthermore, note that $\varepsilon - \log(1+\varepsilon) \geq \frac{\varepsilon^2}{2(1+\varepsilon)}$, and that for $\varepsilon \geq 1$ we have $\frac{\varepsilon^2}{2(1+\varepsilon)} \geq \frac{\varepsilon}{2}$. Thus, in this regime, (58) simplifies to

$$\mathbb{P}\left(Y \geq (1+\varepsilon)m\right) \leq \exp\left(-\frac{m\varepsilon}{8}\right). \tag{59}$$

Therefore, applying the union bound together with (59), we obtain

$$\mathbb{P}_{\mathbf{W}}\left[\left(\Gamma^{\mathsf{fro}}_{\mathcal{K}_0}\right)^c\right] \leq \sum_{z=2}^{k}\sum_{\substack{T\subseteq\mathcal{K}_0\\|T|=z}}\sum_{i\in T}\mathbb{P}\left(\mathsf{A}\cdot\mathsf{B}\geq L^{\mathsf{fro}}_{k,z,d}\right) \tag{60}$$

$$\leq \sum_{z=2}^{k} z\binom{k}{z}\cdot\left[\mathbb{P}\left(\mathsf{A}\geq(1+C_1')(z-1)\log k\right)+\mathbb{P}\left(\mathsf{B}\geq(1+C_1'')d\right)\right] \tag{61}$$

$$\leq \sum_{z=2}^{k}\exp\left[z\left(2\log\left(\frac{ek}{z}\right)-C_1'\log k\right)\right]+\sum_{z=2}^{k}\exp\left[z\left(\log\left(\frac{ek}{z}\right)-C_1''\frac{d}{8z}\right)\right] \tag{62}$$

$$\leq \sum_{z=2}^{k}\exp\left[z\left(2\log(ek)-C_1'\log k\right)\right]+\sum_{z=2}^{k}\exp\left[z\left(\log(ek)-C_1''\frac{d}{8k}\right)\right]. \tag{63}$$

where the second inequality follows from the trivial inclusion $\{\mathsf{A}\cdot\mathsf{B}\geq(1+C_1)d(z-1)\log k\}\subseteq\{\mathsf{A}\geq(1+C_1')(z-1)\log k\}\cup\{\mathsf{B}\geq(1+C_1'')d\}$ where $C_1',C_1''>0$ are such that $(1+C_1')(1+C_1'')\leq(1+C_1)$. Now, define $\zeta\triangleq\max\left\{2\log(ek)-C_1'\log k,\log(ek)-C_1''\frac{d}{8k}\right\}$. Then, for $C_1'>2$ (achieved by choosing C_1 sufficiently large) the first term in the maximum converges to $-\infty$ as $k\to\infty$. Furthermore, in the impossibility regime of Theorem 2, the condition $d/(k^2\log^2 k)\to\infty$ clearly implies that the second term in the maximum also converges to $-\infty$. Thus, $\zeta\to-\infty$, and

$$\mathbb{P}_{\mathbf{W}}((\Gamma^{1}_{\mathcal{K}_0})^c)\leq 2\sum_{z=2}^{\infty}e^{z\zeta}\leq 2\frac{e^{\zeta}}{1-e^{\zeta}}\to 0, \tag{64}$$

eventually.

Next, we bound $\mathbb{P}_{\mathbf{W}}\left[\left(\Gamma^{\mathsf{op}}_{\mathcal{K}_0}\right)^c\right]$. By the union bound,

$$\mathbb{P}_{\mathbf{W}}\left[\left(\Gamma^{\mathsf{op}}_{\mathcal{K}_0}\right)^c\right]\leq\sum_{z=2}^{k}\sum_{\substack{T\subseteq\mathcal{K}_0\\|T|=z}}\mathbb{P}\left[\|\mathbf{W}|_T-d\mathbf{I}_z\|_{\mathsf{op}}>L^{\mathsf{op}}_{k,z,d}\right] \tag{65}$$

$$\leq\sum_{z=2}^{k}\binom{k}{z}\mathbb{P}\left[\|\mathbf{W}|_T-d\mathbf{I}_z\|_{\mathsf{op}}>L^{\mathsf{op}}_{k,z,d}\right] \tag{66}$$

$$\leq\sum_{z=2}^{k}e^{(1-C_2')\log\binom{k}{z}}, \tag{67}$$

where in the last inequality we applied standard concentration inequalities for sample covariance matrices (see, e.g., (Vershynin, 2018, Remark 4.7.3). Choosing $C_2'>1$ (which can be ensured by taking C_2 sufficiently large) makes the exponent $(1-C_2')\log\binom{k}{z}$ negative for every z, and therefore the sum in (67) converges to zero. Combining (54), (64), and (67), we conclude that $\mathbb{P}[\Gamma_{\mathcal{K}}]=1-o(1)$, as required.

C.2. Truncated second moment

We begin by deriving a simplified formula for the truncated second moment $\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2]$, and then derive conditions under which $\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2] = O(1)$ and $\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2] = 1 + o(1)$. Let $\mathcal{K}$ and $\bar{\mathcal{K}}$ be two independent copies drawn uniformly at random over $\binom{[n]}{k}$, i.e., $\mathcal{K} \perp\!\!\!\perp \bar{\mathcal{K}} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}\left(\binom{[n]}{k}\right)$. Similarly let $\mathbf{W} \perp\!\!\!\perp \bar{\mathbf{W}} \overset{\text{iid}}{\sim} \mu_{\mathbf{W}_\mathbf{k}}$ be two independent copies of a Wishart matrix of size $n \times n$. Recall (49), and define

$$\theta(\mathbf{A}_{ij}, \mathbf{W}_{ij}) \triangleq \left(\frac{\sigma_{ij}}{q}\right)^{\mathbf{A}_{i,j}} \left(\frac{1-\sigma_{ij}}{1-q}\right)^{1-\mathbf{A}_{i,j}}, \tag{68}$$

for $(i,j) \in \binom{[n]}{2}$. Below we denote by $\mathbb{E}_{(\mathcal{K},\bar{\mathcal{K}},\mathbf{W},\bar{\mathbf{W}})}$ the expectation with respect to the distribution $\mathbb{P}_{\mathcal{K}} \times \mathbb{P}_{\bar{\mathcal{K}}} \times \mathbb{P}_{\mathbf{W}\perp\!\!\!\perp\bar{\mathbf{W}}|(\mathcal{K},\bar{\mathcal{K}})}$, as defined above. Also, let $\bar{\sigma}_{ij}$ denote the analogue of σ_{ij} constructed from the independent copy $\bar{\mathbf{W}}$, for all $(i,j) \in \binom{[n]}{2}$. Then, by Fubini's theorem, we have

$$\begin{aligned}\mathbb{E}_{\mathcal{H}_0}\left[\tilde{\mathcal{L}}^2\right] &= \frac{1}{(\mathbb{P}[\Gamma_{\mathcal{K}}])^2}\mathbb{E}_{(\mathcal{K},\bar{\mathcal{K}},\mathbf{W},\bar{\mathbf{W}})}\mathbb{E}_{\mathcal{H}_0}\Bigg[\prod_{\{i,j\}\in E(\mathcal{K})}\theta(\mathbf{A}_{ij},\mathbf{W}_{ij})\mathbb{1}_{\Gamma_{\mathcal{K}}}(\mathbf{W}) \\ &\qquad \prod_{\{i,j\}\in E(\bar{\mathcal{K}})}\theta(\mathbf{A}_{ij},\bar{\mathbf{W}}_{ij})\mathbb{1}_{\Gamma_{\bar{\mathcal{K}}}}(\bar{\mathbf{W}})\Bigg] \\ &= \frac{1}{(\mathbb{P}[\Gamma_{\mathcal{K}}])^2}\mathbb{E}_{(\mathcal{K},\bar{\mathcal{K}},\mathbf{W},\bar{\mathbf{W}})}\left[\mathbb{1}_{\Gamma_{\mathcal{K}}}(\mathbf{W})\mathbb{1}_{\Gamma_{\bar{\mathcal{K}}}}(\bar{\mathbf{W}})\eta(\mathcal{K},\bar{\mathcal{K}},\mathbf{W},\bar{\mathbf{W}})\right],\end{aligned} \tag{69}$$

where

$$\eta(\mathcal{K},\bar{\mathcal{K}},\mathbf{W},\bar{\mathbf{W}}) \triangleq \mathbb{E}_{\mathcal{H}_0}\left[\prod_{\{i,j\}\in E(\mathcal{K})}\theta(\mathbf{A}_{ij},\mathbf{W}_{ij})\prod_{\{i,j\}\in E(\bar{\mathcal{K}})}\theta(\mathbf{A}_{ij},\bar{\mathbf{W}}_{ij})\right]. \tag{70}$$

If $|\mathcal{K}\cap\bar{\mathcal{K}}| \leq 1$, namely, at most one vertex shared between $\mathcal{K}$ and $\bar{\mathcal{K}}$, then the set of edges formed among nodes in $\mathcal{K}$ are disjoint from the set of edges formed among nodes in $\bar{\mathcal{K}}$. Using the fact that $\mathbb{E}_{\mathcal{H}_0}[\theta(\mathbf{A}_{ij},\mathbf{W}_{ij})] = 1$, for all $(i,j) \in \binom{[n]}{2}$, we get that

$$\eta(\mathcal{K},\bar{\mathcal{K}},\mathbf{W},\bar{\mathbf{W}}) = \mathbb{1}\left\{|\mathcal{K}\cap\bar{\mathcal{K}}| \leq 1\right\}. \tag{71}$$

Otherwise, if $|\mathcal{K}\cap\bar{\mathcal{K}}| \geq 2$, we have

$$\begin{aligned}\eta(\mathcal{K},\bar{\mathcal{K}},\mathbf{W},\bar{\mathbf{W}}) &= \mathbb{E}_{\mathcal{H}_0}\Bigg[\mathbb{1}\left\{|\mathcal{K}\cap\bar{\mathcal{K}}| \geq 2\right\}\prod_{\{i,j\}\in E(\mathcal{K}\cap\bar{\mathcal{K}})}\theta(\mathbf{A}_{ij},\mathbf{W}_{ij})\theta(\mathbf{A}_{ij},\bar{\mathbf{W}}_{ij}) \\ &\qquad \prod_{\{i,j\}\in E(\mathcal{K})\setminus E(\mathcal{K}\cap\bar{\mathcal{K}})}\theta(\mathbf{A}_{ij},\mathbf{W}_{ij})\prod_{\{i,j\}\in E(\bar{\mathcal{K}})\setminus E(\mathcal{K}\cap\bar{\mathcal{K}})}\theta(\mathbf{A}_{ij},\bar{\mathbf{W}}_{ij})\Bigg]\end{aligned} \tag{72}$$

$$= \mathbb{E}_{\mathcal{H}_0}\left[\prod_{\{i,j\}\in E(\mathcal{K}\cap\bar{\mathcal{K}})}\theta(\mathbf{A}_{ij},\mathbf{W}_{ij})\theta(\mathbf{A}_{ij},\bar{\mathbf{W}}_{ij})\mathbb{1}\left\{|\mathcal{K}\cap\bar{\mathcal{K}}| \geq 2\right\}\right], \tag{73}$$

where the the second equality follows from the fact that $\mathbb{E}_{\mathcal{H}_0}[\theta(\mathbf{A}_{ij}, \mathbf{W}_{ij})] = 1$ and $\mathbb{E}_{\mathcal{H}_0}[\theta(\mathbf{A}_{ij}, \bar{\mathbf{W}}_{ij})] = 1$ for all $(i,j) \in \binom{[n]}{2}$. Combining (69), (71), and (73), we get

$$\begin{aligned}\mathbb{E}_{\mathcal{H}_0}\left[\tilde{\mathcal{L}}^2\right] &= \frac{\mathbb{P}\left[\mathbf{W} \in \Gamma_{\mathcal{K}}, \bar{\mathbf{W}} \in \Gamma_{\bar{\mathcal{K}}}, |\mathcal{K} \cap \bar{\mathcal{K}}| \leq 1\right]}{(\mathbb{P}[\Gamma_{\mathcal{K}}])^2} \\ &\quad + \mathbb{E}_{\mathcal{K} \perp\!\!\!\perp \bar{\mathcal{K}}} \mathbb{E}_{\mathcal{H}_0}\left[g(\mathbf{A}, \mathcal{K}, \bar{\mathcal{K}}) \mathbb{1}\left\{|\mathcal{K} \cap \bar{\mathcal{K}}| \geq 2\right\}\right] \qquad (74) \\ &= \mathbb{P}\left[|\mathcal{K} \cap \bar{\mathcal{K}}| \leq 1 \middle| \mathbf{W} \in \Gamma_{\mathcal{K}}, \bar{\mathbf{W}} \in \Gamma_{\bar{\mathcal{K}}}\right] \\ &\quad + \mathbb{E}_{\mathcal{K} \perp\!\!\!\perp \bar{\mathcal{K}}} \mathbb{E}_{\mathcal{H}_0}\left[g(\mathbf{A}, \mathcal{K}, \bar{\mathcal{K}}) \mathbb{1}\left\{|\mathcal{K} \cap \bar{\mathcal{K}}| \geq 2\right\}\right], \qquad (75)\end{aligned}$$

where

$$g(\mathbf{A}, \mathcal{K}, \bar{\mathcal{K}}) \triangleq \frac{1}{(\mathbb{P}[\Gamma_{\mathcal{K}}])^2} \mathbb{E}_{\mathbf{W} \perp\!\!\!\perp \bar{\mathbf{W}} | \mathcal{K}, \bar{\mathcal{K}}} \left(\prod_{\{i,j\} \in E(\mathcal{K} \cap \bar{\mathcal{K}})} \theta(\mathbf{A}_{ij}, \mathbf{W}_{ij}) \mathbb{1}_{\Gamma_{\mathcal{K}}}(\mathbf{W}) \prod_{\{i,j\} \in E(\mathcal{K} \cap \bar{\mathcal{K}})} \theta(\mathbf{A}_{ij}, \bar{\mathbf{W}}_{ij}) \mathbb{1}_{\Gamma_{\bar{\mathcal{K}}}}(\bar{\mathbf{W}}) \right). \qquad (76)$$

Recall the definition of $\Gamma_{\mathcal{K}}$ in (45)–(47), and for brevity, for any $T \subseteq [n]$, define the sets $\mathcal{Q}_T^{\mathsf{fro}} \triangleq \bigcap_{i \in T} \left\{A : \sum_{j \in T \setminus \{i\}} A_{ij}^2 \leq L_{k,|T|,d}^{\mathsf{fro}}\right\}$, $\mathcal{Q}_T^{\mathsf{op}} \triangleq \left\{A : \left\|A|_T - d\mathbf{I}_{|T|}\right\|_{\mathsf{op}} \leq L_{k,||T||,d}^{\mathsf{op}}\right\}$, and $\mathcal{Q}_T \triangleq \mathcal{Q}_T^{\mathsf{fro}} \cap \mathcal{Q}_T^{\mathsf{op}}$. Then, by definition, we note that $\Gamma_{\mathcal{K}} \subseteq \mathcal{Q}_{\mathcal{K} \cap \bar{\mathcal{K}}}$ and $\Gamma_{\bar{\mathcal{K}}} \subseteq \mathcal{Q}_{\mathcal{K} \cap \bar{\mathcal{K}}}$, and as so, $\mathbb{1}_{\Gamma_{\mathcal{K}}}(\mathbf{W}) \leq \mathbb{1}_{\mathcal{Q}_{\mathcal{K} \cap \bar{\mathcal{K}}}}(\mathbf{W})$, and similarly $\mathbb{1}_{\Gamma_{\bar{\mathcal{K}}}}(\bar{\mathbf{W}}) \leq \mathbb{1}_{\mathcal{Q}_{\mathcal{K} \cap \bar{\mathcal{K}}}}(\bar{\mathbf{W}})$. Denoting $\mathcal{U} \triangleq \mathcal{K} \cap \bar{\mathcal{K}}$, we get

$$\begin{aligned}g(\mathbf{A}, \mathcal{K}, \bar{\mathcal{K}}) &\leq \frac{1}{(\mathbb{P}[\Gamma_{\mathcal{K}}])^2} \mathbb{E}_{\mathbf{W} \perp\!\!\!\perp \bar{\mathbf{W}} | \mathcal{K}, \bar{\mathcal{K}}} \left(\prod_{\{i,j\} \in E(\mathcal{U})} \theta(\mathbf{A}_{ij}, \mathbf{W}_{ij}) \mathbb{1}_{\mathcal{Q}_{\mathcal{U}}}(\mathbf{W}) \prod_{\{i,j\} \in E(\mathcal{U})} \theta(\mathbf{A}_{ij}, \bar{\mathbf{W}}_{ij}) \mathbb{1}_{\mathcal{Q}_{\mathcal{U}}}(\bar{\mathbf{W}}) \right) \qquad (77) \\ &= \frac{(\mathbb{P}[\mathcal{Q}_{\mathcal{U}}])^2}{(\mathbb{P}[\Gamma_{\mathcal{K}}])^2} \left[\frac{1}{\mathbb{P}[\mathcal{Q}_{\mathcal{U}}]} \mathbb{E}_{\mathbf{W}|\mathcal{U}} \left(\prod_{\{i,j\} \in E(\mathcal{U})} \theta(\mathbf{A}_{ij}, \mathbf{W}_{ij}) \mathbb{1}_{\mathcal{Q}_{\mathcal{U}}}(\mathbf{W}) \right) \right]^2 \qquad (78) \\ &\leq (1 + o(1)) \cdot \left[\frac{1}{\mathbb{P}[\mathcal{Q}_{\mathcal{U}}]} \mathbb{E}_{\mathbf{W}|\mathcal{U}} \left(\prod_{\{i,j\} \in E(\mathcal{U})} \theta(\mathbf{A}_{ij}, \mathbf{W}_{ij}) \mathbb{1}_{\mathcal{Q}_{\mathcal{U}}}(\mathbf{W}) \right) \right]^2, \qquad (79)\end{aligned}$$

where the last inequality is because we have proved that $1 - o(1) = \mathbb{P}[\Gamma_{\mathcal{K}}] \leq \mathbb{P}[\mathcal{Q}_{\mathcal{U}}] \leq 1$. Now recalling (48) and (52) we readily see that the the right-hand side of (79) can be written as

$$g(\mathbf{A}, \mathcal{K}, \bar{\mathcal{K}}) \leq (1 + o(1)) \cdot \left[\frac{\tilde{\mathbb{P}}_{\mathcal{H}_1|\mathcal{K} \cap \bar{\mathcal{K}}}(\mathbf{A}|_{\mathcal{K} \cap \bar{\mathcal{K}}})}{\mathbb{P}_{\mathcal{H}_0}(\mathbf{A}|_{\mathcal{K} \cap \bar{\mathcal{K}}})} \right]^2 = (1 + o(1)) \tilde{\mathcal{L}}_{\mathcal{K} \cap \bar{\mathcal{K}}}^2(\mathbf{A}|_{\mathcal{K} \cap \bar{\mathcal{K}}}). \qquad (80)$$

Combining the above we get that

$$\mathbb{E}_{\mathcal{H}_0}\left[\tilde{\mathcal{L}}^2\right] \leq \mathbb{P}\left[|\mathcal{K}\cap\bar{\mathcal{K}}| \leq 1 \middle| \mathbf{W}\in\Gamma_{\mathcal{K}}, \bar{\mathbf{W}}\in\Gamma_{\bar{\mathcal{K}}}\right] + \mathbb{E}_{\mathcal{K}\perp\!\!\!\perp\bar{\mathcal{K}}}\mathbb{E}_{\mathcal{H}_0}\left[\left(\frac{\tilde{\mathbb{P}}_{\mathcal{H}_1|\mathcal{K}\cap\bar{\mathcal{K}}}(\mathbf{A}|_{\mathcal{K}\cap\bar{\mathcal{K}}})}{\mathbb{P}_{\mathcal{H}_0}(\mathbf{A}|_{\mathcal{K}\cap\bar{\mathcal{K}}})}\right)^2 \mathbb{1}\left\{|\mathcal{K}\cap\bar{\mathcal{K}}|\geq 2\right\}\right]. \tag{81}$$

At this point, we observe that the conditional likelihood in (80) corresponds exactly to a hypothesis test between the Erdős–Rényi distribution $\mathcal{G}(|\mathcal{K}\cap\bar{\mathcal{K}}|, q)$ and the *truncated geometric random graph distribution* $\tilde{\mathcal{G}}_d(|\mathcal{K}\cap\bar{\mathcal{K}}|, q)$, which is defined as the standard geometric random graph model with the additional constraint that the latent vectors lie in the intersection of $\mathbb{S}^{d-1}$ and $\mathcal{Q}_{\mathcal{U}}$. Accordingly, let $Z \triangleq |\mathcal{K}\cap\bar{\mathcal{K}}|$, and note that $Z \sim \mathsf{Hypergeometric}(n,k,k)$. Then,

$$\begin{aligned}\mathbb{E}_{\mathcal{H}_0}\left[\tilde{\mathcal{L}}^2\right] &\leq \mathbb{P}\left[Z\leq 1 \middle| \mathbf{W}\in\Gamma_{\mathcal{K}}, \bar{\mathbf{W}}\in\Gamma_{\bar{\mathcal{K}}}\right] \\ &\quad + \mathbb{E}_{\mathcal{U}}\left[(1+\chi^2(\tilde{\mathcal{G}}_d(Z,q)\|\mathcal{G}(Z,q)))\mathbb{1}\left\{Z\geq 2\right\}\right] &(82)\\ &\leq 1 + o(1) + \mathbb{E}_{\mathcal{U}}\left[\chi^2(\tilde{\mathcal{G}}_d(Z,q)\|\mathcal{G}(Z,q))\mathbb{1}\left\{Z\geq 2\right\}\right]. &(83)\end{aligned}$$

Next, for a given $Z = z \geq 2$, we uniformly upper bound $\chi^2\big(\tilde{\mathcal{G}}_d(z,q)\|\mathcal{G}(z,q)\big)$ over z. To this end, following the approach of Bubeck et al. (2016), it is convenient to view the graph ensemble distributions $\tilde{\mathcal{G}}_d(z,q)$ and $\mathcal{G}(z,q)$ as being generated through a truncation procedure applied to Wishart and Gaussian orthogonal ensemble (GOE) random matrices, as described below.

Specifically, recall that if $\mathbf{x}$ is a d-dimensional standard Gaussian vector, then $\mathbf{x}/||\mathbf{x}||_2$ is uniformly distributed on the sphere $\mathbb{S}^{d-1}$. Let $\mathbf{X}$ be a $z\times d$ matrix obtained from stacking z i.i.d. standard Gaussian vectors in $\mathbb{R}^d$, and let $\mathbf{W} = \mathbf{X}\mathbf{X}^T$ be the corresponding Wishart matrix. Define the matrix $\mathbf{A}$ as

$$\mathbf{A}_{ij} = \begin{cases} 1, & \mathbf{W}_{ij}/\sqrt{\mathbf{W}_{ii}\mathbf{W}_{jj}} \geq t_{q,d}, \text{ and } i\neq j, \\ 0, & i=j. \end{cases} \tag{84}$$

Then $\mathbf{A}$ has the same law as the adjacency matrix of $\mathcal{G}_d(z,q)$. We denote by $H_{q,d}$ the map defined by $H_{q,d}(\mathbf{W}) = \mathbf{A}$. In the truncated case, we sample $\tilde{\mathbf{W}}$ from the Wishart distribution conditioned on the event $\mathcal{Q}_{\mathcal{U}}$ (that is, we repeatedly sample from the Wishart distribution until $\tilde{\mathbf{W}}\in\mathcal{Q}_{\mathcal{U}}$), and write $H_{q,d}(\tilde{\mathbf{W}}) = \tilde{\mathbf{A}} \sim \tilde{\mathcal{G}}_d(z,q)$.

We next describe an analogous construction for Erdős–Rényi graphs using Gaussian orthogonal ensemble matrices. Let $\mathbf{M}(z)$ be a GOE matrix: a symmetric $z\times z$ matrix with diagonal entries i.i.d. $\mathcal{N}(0,2)$ and off-diagonal entries i.i.d. $\mathcal{N}(0,1)$. Define

$$\mathbf{B}_{ij} = \begin{cases} 1, & [\mathbf{M}(z)]_{ij} \geq \overline{\Phi}^{-1}(q), \text{ and } i\neq j, \\ 0, & i=j. \end{cases} \tag{85}$$

Then $\mathbf{B}$ has the same law as the adjacency matrix of $\mathcal{G}(z,q)$. Since $\mathbf{B}$ depends only on off-diagonal entries, it is convenient to rescale and shift $\mathbf{M}(z)$ so that it matches the normalization of the Wishart ensemble. Define

$$\mathbf{M}(z,d) \triangleq \sqrt{d}\mathbf{M}(z) + d\mathbf{I}_z, \tag{86}$$

and accordingly replace $\overline{\Phi}^{-1}(q)$ with $\sqrt{d}\overline{\Phi}^{-1}(q)$. We denote by $K_{q,d}$ the map taking $\mathbf{M}(z,d)$ to $\mathbf{B}$, i.e., $\mathbf{B} = K_{q,d}(\mathbf{M}(z,d))$. In the truncated case, $\tilde{\mathbf{M}}$ is sampled from the GOE distribution conditioned on the event $\mathcal{Q}_{\mathcal{U}}$, and we write $\tilde{\mathbf{B}} = K_{q,d}(\tilde{\mathbf{M}}(z,d)) \sim \tilde{\mathcal{G}}(z,q)$.

Remark 7 *We are also interested in the probability of $\mathcal{Q}_{\mathcal{U}}$ under the GOE distribution. The proof is essentially identical to the one provided in the previous section. Specifically, the high probability of $\mathcal{Q}_{\mathcal{U}}^{\text{fro}}$ under $\mathbb{P}_{\mathcal{H}_0}$ follows from the concentration of the χ^2_{z-1} distribution around its mean. Similarly, the high probability of $\mathcal{Q}_{\mathcal{U}}^{\text{op}}$ is a consequence of the eigenvalues of $\mathbf{M}(z)$ concentrating in the interval $[-3\sqrt{z}, 3\sqrt{z}]$ (see, e.g., (Vershynin, 2018, Theorem 4.4.3)). This leads us to conclude that $\mathbb{P}_{\mathcal{H}_0}(\Gamma_{\mathcal{K}}) = 1 - o(1)$.*

We will frequently bound moments of likelihood ratios, so it is convenient to introduce the following notation.

Definition 8 (D_m–divergence) *Let $\mathbb{P}, \mathbb{Q}$ be two probability measures such that $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$, i.e., $\mathbb{P} \ll \mathbb{Q}$. For $m \in \mathbb{N}$ define*

$$D_m(\mathbb{P}\|\mathbb{Q}) \triangleq \mathbb{E}_{\mathbb{Q}}\left[\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\right)^m\right] - 1. \tag{87}$$

It is rather a straightforward task to prove that $D_m(P\|Q)$ is non negative and convex, implying that it is an f divergence (Polyanskiy and Wu, 2024, Ch. 7). We will repeatedly use the following two lemmata in our analysis; the first is standard (e.g., Polyanskiy and Wu (2024)), and the second is proved in Appendix F.

Lemma 9 (Data processing inequality) *Let $P_X, Q_X \in \mathcal{P}(\mathcal{X})$ and let $P_{Y|X}$ be a transition kernel from $\mathcal{X}$ to $\mathcal{Y}$. Let P_Y and Q_Y be the resulting marginal measures on $\mathcal{Y}$ such that $P_Y = P_X P_{Y|X}$ and $Q_Y = Q_X P_{Y|X}$. Then, for any $m \in \mathbb{N}$*

$$D_m(P_Y\|Q_Y) \leq D_m(P_X\|Q_X). \tag{88}$$

Lemma 10 (Cauchy–Schwarz inequality) *Let $\mathbb{P}, \mathbb{R}, \mathbb{Q}$ be three measures over the same probability space, such that, $\mathbb{P} \ll \mathbb{R} \ll \mathbb{Q}$. Then, for any $m \in \mathbb{N}$*

$$1 + D_m(\mathbb{P}||\mathbb{Q}) \leq \sqrt{1 + D_{2m}(\mathbb{P}||\mathbb{R})}\sqrt{1 + D_{2m-1}(\mathbb{R}||\mathbb{Q})}. \tag{89}$$

For simplicity of notations, we denote by $\tilde{\mathbb{P}}_1^{(z)}$ and $\mathbb{P}_0^{(z)}$ the probability measures induced by $\tilde{\mathcal{G}}_d(z,q) \stackrel{d}{=} H_{q,d}(\tilde{\mathbf{W}}(z,d))$ and $\mathcal{G}(z,q) \stackrel{d}{=} K_{q,d}(\mathbf{M}(z,d))$, respectively. Furthermore, we denote by $\tilde{\mathbb{P}}_0^{(z)}$ the probability measures induced by $H_{q,d}(\tilde{\mathbf{M}}(z,d))$. Then, using the above notations, our goal is to upper bound $\chi^2(\tilde{\mathbb{P}}_1^{(z)}\|\mathbb{P}_0^{(z)}) = D_2(\tilde{\mathbb{P}}_1^{(z)}\|\mathbb{P}_0^{(z)})$. By Lemma 10 we have

$$1 + \chi^2(\tilde{\mathbb{P}}_1^{(z)}\|\mathbb{P}_0^{(z)}) \leq \sqrt{1 + D_4(\tilde{\mathbb{P}}_1^{(z)}\|\tilde{\mathbb{P}}_0^{(z)})}\sqrt{1 + D_3(\tilde{\mathbb{P}}_0^{(z)}\|\mathbb{P}_0^{(z)})}. \tag{90}$$

We now bound each of the two terms on the right-hand side separately. Throughout, $z \in [k]$ is fixed, and to lighten the notation, we suppress the superscript (z) from the measures.

Bounding $D_4(\tilde{\mathbb{P}}_1^{(z)}\|\tilde{\mathbb{P}}_0^{(z)})$. Observe that under both $\tilde{\mathbb{P}}_1$ and $\tilde{\mathbb{P}}_0$ we map the random matrices $\tilde{\mathbf{W}}(z,d)$ and $\tilde{\mathbf{M}}(z,d)$, respectively, using the *same* map $H_{q,d}$. Accordingly, by applying the DPI in Lemma 9, we obtain

$$1 + D_4(\tilde{\mathbb{P}}_1^{(z)}\|\tilde{\mathbb{P}}_0^{(z)}) = 1 + D_4(H_{q,d}(\tilde{\mathbf{W}}(z,d))\|H_{q,d}(\tilde{\mathbf{M}}(z,d))) \tag{91}$$

$$\leq 1 + D_4(\tilde{\mathbf{W}}(z,d)||\tilde{\mathbf{M}}(z,d)). \tag{92}$$

Let $\mathcal{P} \subset \mathbb{R}^{z^2}$ denote the cone of positive semidefinite matrices, and let $f_{z,d}$ be the density of $\mathbf{W}(z,d)$ with respect to the Lebesgue measure on $\mathcal{P}$ when $d \geq z$. This condition holds since $z \leq k$ and, in the impossible regime, $d/(k^2 \log^2 k) \to \infty$. Furthermore, let $g_{z,d}$ denote the density of $\mathbf{M}(z,d)$ with respect to the Lebesgue measure on $\mathbb{R}^{z^2}$. Accordingly, let $\tilde{f}_{z,d}$ and $\tilde{g}_{z,d}$ be the corresponding densities of $\tilde{\mathbf{W}}(z,d)$ and $\tilde{\mathbf{M}}(z,d)$, respectively, that is, the conditional densities given $\mathcal{Q}_{\mathcal{U}}$, i.e., $\tilde{f}_{z,d}(\cdot) = f_{z,d}(\cdot|\mathcal{Q}_{\mathcal{U}})$ and $\tilde{g}_{z,d}(\cdot) = g_{z,d}(\cdot|\mathcal{Q}_{\mathcal{U}})$. We observe that

$$\tilde{f}_{z,d}(\cdot) = \frac{1}{c_1} f_{z,d}(\cdot)\mathbb{1}\{\cdot \in \mathcal{Q}_{\mathcal{U}}\} \quad \text{and} \quad \tilde{g}_{z,d}(\cdot) = \frac{1}{c_2} g_{z,d}(\cdot)\mathbb{1}\{\cdot \in \mathcal{Q}_{\mathcal{U}}\}, \tag{93}$$

where $c_1 = \int_{\mathcal{Q}_{\mathcal{U}}} f_{z,d}(\mathbf{A})\mathrm{d}\mathbf{A}$ and $c_2 = \int_{\mathcal{Q}_{\mathcal{U}}} g_{z,d}(\mathbf{A})\mathrm{d}\mathbf{A}$ are the normalization constants, and we note that because $\mathbb{P}(\mathcal{Q}_{\mathcal{U}}) = 1 - o(1)$, we have $c_1 = 1 - o(1)$, and by (7), that $c_2 = 1 - o(1)$. Define $\alpha_{z,d}(\mathbf{A}) \triangleq \log(f_{z,d}(\mathbf{A})/g_{z,d}(\mathbf{A}))$ and $\tilde{\alpha}_{z,d}(\mathbf{A}) \triangleq \log(\tilde{f}_{z,d}(\mathbf{A})/\tilde{g}_{z,d}(\mathbf{A}))$. Using the above, we clearly have

$$1 + D_4(\tilde{\mathbf{W}}(z,d)||\tilde{\mathbf{M}}(z,d)) = \mathbb{E}_{\mathbf{A}\sim\tilde{\mathbf{M}}(z,d)}\left[\exp\left(4\tilde{\alpha}_{z,d}(\mathbf{A})\right)\mathbb{1}_{\mathcal{P}}\left(\mathbf{A}\right)\right] \tag{94}$$

$$\leq (1 + o(1)) \cdot \mathbb{E}_{\mathbf{A}\sim\tilde{\mathbf{M}}(z,d)}\left[\exp\left(4\alpha_{z,d}(\mathbf{A})\right)\mathbb{1}_{\mathcal{P}}\left(\mathbf{A}\right)\right]. \tag{95}$$

The following is a core estimate on $\alpha_{z,d}(\mathbf{A})$ used in the proof of (Bubeck et al., 2016, Thm. 7).

Lemma 11 ((Bubeck et al., 2016, eqns. (36)–(37))) *For any $\mathbf{A} \in \mathcal{P}$, let $\{\lambda_i\}_{i=1}^z$ denote the eigenvalues of $\mathbf{A}$. With probability one,*

$$\alpha_{z,d}(\mathbf{A}) = \sum_{i=1}^{z} h(\lambda_i) + O\left(\frac{z^3}{d}\right), \tag{96}$$

as $d - z \to \infty$, where

$$h(x) \triangleq -\frac{z+1}{2d}(x-d) + \frac{z+1}{4d^2}(x-d)^2 + \frac{d-z-1}{6d^3}(x-d)^3 + \frac{d-z-1}{6\xi^4}(x-d)^4, \tag{97}$$

and ξ is some real number between x and d.

Applying Lemma 11 on (95), we obtain

$$1 + D_4(\tilde{\mathbf{W}}(z,d)||\tilde{\mathbf{M}}(z,d)) = e^{O\left(\frac{z^3}{d}\right)} \cdot \mathbb{E}_{\mathbf{A}\sim\tilde{\mathbf{M}}(z,d)}\left[\exp\left(4\sum_{i=1}^{z} h(\lambda_i)\right)\mathbb{1}_{\{\mathbf{A}\in\mathcal{P}\}}\right]. \tag{98}$$

Next, recall that under $\mathcal{Q}_{\mathcal{U}}$ the operator norm constraint in (47) implies that $|\lambda_i - d| \leq L_{k,z,d}^{\mathsf{op}} = C_2(\sqrt{dz} + \sqrt{d\log\binom{k}{z}}) \leq 2C_2\sqrt{dz\log k}$, for all $i \in [z]$. Furthermore, because ξ_i is between λ_i and d for all $i \in [z]$, we get

$$\sum_{i=1}^{z}\left|\frac{z+1}{4d^2}(\lambda_i - d)^2 - \frac{d-z-1}{8\xi^4}(\lambda_i - d)^4\right| \leq O\left(\frac{z^3\log^2 k}{d}\right). \tag{99}$$

It remains to bound the linear and cubic terms in (97). Define $h_1(x) \triangleq -\frac{z+1}{2d}(x-d)$ and $h_3(x) \triangleq \frac{d-z-1}{6d^3}(x-d)^3$. Then, combining (98) and (99), we have

$$1 + D_4(\tilde{\mathbf{W}}(z,d)||\tilde{\mathbf{M}}(z,d)) \leq e^{O\left(\frac{z^3\log^2 k}{d}\right)} \cdot \mathbb{E}_{\mathbf{A}\sim\tilde{\mathbf{M}}(z,d)}\left[\exp\left(4\sum_{i=1}^{z} h_1(\lambda_i) + h_3(\lambda_i)\right)\mathbb{1}_{\{\mathbf{A}\in\mathcal{P}\}}\right] \tag{100}$$

$$\leq e^{O\left(\frac{z^3\log^2 k}{d}\right)} \cdot \mathsf{A}^{1/2}\cdot\mathsf{B}^{1/2}, \tag{101}$$

where the last passage follows from Cauchy-Schwarz inequality and we define

$$\mathsf{A} \triangleq \mathbb{E}_{\mathbf{A}\sim\tilde{\mathbf{M}}(z,d)}\left[\exp\left(8\sum_{i=1}^{z} h_1(\lambda_i)\right)\mathbb{1}_{\{\mathbf{A}\in\mathcal{P}\}}\right], \tag{102}$$

and

$$\mathsf{B} \triangleq \mathbb{E}_{\mathbf{A}\sim\tilde{\mathbf{M}}(z,d)}\left[\exp\left(8\sum_{i=1}^{z} h_3(\lambda_i)\right)\mathbb{1}_{\{\mathbf{A}\in\mathcal{P}\}}\right]. \tag{103}$$

Let us upper bound A and B, starting with the former. First, note that

$$\sum_{i=1}^{z} h_1(\lambda_i) = -\frac{z+1}{2d}\mathrm{T}r\big(\mathbf{A}-d\mathbf{I}_z\big), \tag{104}$$

and so,

$$\mathsf{A} \leq (1+o(1))\cdot\mathbb{E}_{\mathbf{A}\sim\mathbf{M}(z)}\left[\exp\left(-8\frac{z+1}{2\sqrt{d}}\mathrm{T}r(\mathbf{A})\right)\right] \tag{105}$$

$$= (1+o(1))\cdot\left(\mathbb{E}_{X\sim\mathcal{N}(0,2)}\left[\exp\left(-8\frac{z+1}{2\sqrt{d}}X\right)\right]\right)^z \tag{106}$$

$$= \exp\left(O\left(\frac{z^3}{d}\right)\right), \tag{107}$$

where in the first inequality we have removed the truncation and use the fact that $c_2 = 1-o(1)$, the first equality is because $[\mathbf{M}(z)]_{ii} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,2)$, and in the second equality we used the fact that $\mathbb{E}[\exp(tX)] = \exp(\sigma^2 t^2/2)$ for $X\sim\mathcal{N}(0,\sigma^2)$. Next, we bound B. Denote $G(\mathbf{A}) \triangleq \mathrm{T}r(\mathbf{A}^3)$, and note that

$$\sum_{i=1}^{z} h_3(\lambda_i) = \frac{d-z-1}{6d^3}G(\mathbf{A}-d\mathbf{I}_z). \tag{108}$$

Before continuing, we collect the statements that will be used in the proof. We first recall the following definition.

Definition 12 (Log–Sobolev inequality) *Let μ be a probability measure on $\mathbb{R}^n$ that is absolutely continuous with respect to Lebesgue measure. We say that μ satisfies a* log–Sobolev inequality *with constant $C_{\mathsf{LSI}} > 0$ if for all smooth functions $f : \mathbb{R}^n \to \mathbb{R}$ with f^2 integrable,*

$$\mathsf{Ent}_\mu(f^2) \leq 2C_{\mathsf{LSI}} \int_{\mathbb{R}^n} \|\nabla f(x)\|^2 \mathrm{d}\mu(x), \tag{109}$$

where the entropy with respect to μ is defined by

$$\mathsf{Ent}_\mu(\mathsf{G}) \triangleq \int g \log g \mathrm{d}\mu - \left(\int g \mathrm{d}\mu\right) \log\left(\int g \mathrm{d}\mu\right). \tag{110}$$

The smallest constant C_{LSI} for which this inequality holds is called the log–Sobolev (LSI) constant *of μ.*

Lemma 13 ((Anderson et al., 2009, Lemma 2.3.3)) *Let P be a probability measure satisfying the LSI on $\mathbb{R}^M$ with constant C_{LSI}. Let G be a Lipschitz function on $\mathbb{R}^M$, with Lipschitz constant $|G|_{\mathcal{L}}$. Then for all $\beta \in \mathbb{R}$,*

$$\mathbb{E}_P\left[\exp\left(\beta\left(G - \mathbb{E}_P[G]\right)\right)\right] \leq \exp\left(C_{\mathsf{LSI}}\beta^2 \left|G\right|_{\mathcal{L}}^2/2\right). \tag{111}$$

Lemma 14 ((Anderson et al., 2009, Lemma 2.3.2)) *Let P be a Gaussian law with mean zero and variance σ^2. Then, the $C_{\mathsf{LSI}} = \sigma^2$, additionally, if $P = \bigotimes_{i=1}^n P_i$ such that each probability measure P_i has LSI constant C_i then $C_{\mathsf{LSI}} = \max_i C_i$.*

Lemma 15 (McShane-Whitney extension theorem (e.g., (Gutev, 2025, Thm. 1.1))) *Let (X, d) be a metric space and $A \subset X$. Then every L-Lipschitz function $f : A \to \mathbb{R}$ can be extended to an L-Lipschitz function $f^\star : X \to \mathbb{R}$.*

We are now in a position to bound B. First, note that

$$\mathsf{B} = \mathbb{E}_{\mathbf{A}\sim\tilde{\mathbf{M}}(z)}\left[\exp\left(\frac{d-z-1}{6d^{3/2}}G(\mathbf{A})\right)\mathbb{1}_{\{\mathbf{A}\in\mathcal{P}\}}\right] \tag{112}$$

$$= \mathbb{E}_{\mathbf{A}\sim\tilde{\mathbf{M}}(z)}\left[\exp\left(\frac{C}{\sqrt{d}}G(\mathbf{A})\right)\mathbb{1}_{\{\mathbf{A}\in\mathcal{P}\}}\right] \tag{113}$$

$$\leq (1+o(1))\cdot\mathbb{E}_{\mathbf{A}\sim\mathbf{M}(z)}\left[\exp\left(\frac{C}{\sqrt{d}}G(\mathbf{A})\right)\mathbb{1}_{\bar{\mathcal{Q}}_{\mathcal{U}}^{\mathsf{op}}}(\mathbf{A})\right], \tag{114}$$

where the second equality follows from the fact that $d \gg z$ and $C > 0$, and the inequality holds since $c_2 = 1 - o(1)$. Moreover, we retain only the (rescaled, as in (86)) operator-norm truncation set in (47), namely, $\bar{\mathcal{Q}}_{\mathcal{U}}^{\mathsf{op}} \triangleq \left\{\|\mathbf{A}\|_{\mathsf{op}} \leq \bar{L}_{k,z,d}^{\mathsf{op}}\right\}$ and $\bar{L}_{k,z,d}^{\mathsf{op}} \triangleq L_{k,z,d}^{\mathsf{op}}/\sqrt{d} = C_2(\sqrt{z} + \sqrt{\log\binom{k}{z}}) = O(\sqrt{z\log k})$. Now, by the mean value theorem and the Cauchy–Schwarz inequality, for any $\mathbf{A}, \mathbf{B}$ there exists a matrix $\mathbf{C}$ on the line segment connecting $\mathbf{A}$ and $\mathbf{B}$ such that

$$|G(\mathbf{A}) - G(\mathbf{B})| \leq \langle\nabla G(\mathbf{C}), \mathbf{A} - \mathbf{B}\rangle \leq \|\nabla G(\mathbf{C})\|_F \cdot \|\mathbf{A} - \mathbf{B}\|_F. \tag{115}$$

Moreover, since $\bar{\mathcal{Q}}_{\mathcal{U}}^{\mathsf{op}}$ is convex, the matrix $\mathbf{C}$ also belongs to this set, and hence

$$\|\nabla G(\mathbf{C})\|_F \leq \sqrt{z}\|\nabla G(\mathbf{C})\|_{\mathsf{op}} = 3\sqrt{z}\|\mathbf{C}\|_{\mathsf{op}}^2 = O(z^{3/2}\log k), \tag{116}$$

which implies that $|G|_{\mathcal{L}} \leq O(z^{3/2}\log k)$. We now apply Lemma 15 to extend G from $\bar{\mathcal{Q}}_{\mathcal{U}}^{\mathsf{op}}$ to a function $G^{\star}$ defined on the space of GOE matrices, such that $G^{\star} = G$ on $\bar{\mathcal{Q}}_{\mathcal{U}}^{\mathsf{op}}$ and $|G^{\star}|_{\mathcal{L}} = |G|_{\mathcal{L}}$. By Lemma 14, the LSI constant for the GOE is $C_{\mathsf{LSI}} = 2$. Using the fact that $\mathbb{E}_{\mathbf{A}\sim\mathbf{M}(z)}[\mathsf{Tr}(\mathbf{A}^3)] = 0$ and the positivity of the exponential function, Lemma 13 gives

$$\mathsf{B} \leq (1+o(1))\cdot\mathbb{E}_{\mathbf{A}\sim\mathbf{M}(z)}\left[\exp\left(\frac{C}{\sqrt{d}}G^{\star}(\mathbf{A})\right)\right] \leq \exp\left(O\left(\frac{z^3\log^2 k}{d}\right)\right). \tag{117}$$

Combining (107) and (117) with (101) and (91), we conclude that

$$1 + D_4(\tilde{\mathbb{P}}_1^{(z)}\|\tilde{\mathbb{P}}_0^{(z)}) \leq \exp\left(O\left(\frac{z^3\log^2 k}{d}\right)\right). \tag{118}$$

Bounding $D_3(\tilde{\mathbb{P}}_0^{(z)}\|\mathbb{P}_0^{(z)})$. We again introduce an intermediate probability measure. Recall that $\tilde{\mathbb{P}}_0^{(z)}$ and $\mathbb{P}_0^{(z)}$ denote the probability measures induced by the measurable mappings $H_{q,d}(\tilde{\mathbf{M}}(z,d))$ and $K_{q,d}(\mathbf{M}(z,d))$, respectively. For notational convenience, we denote the corresponding random matrices by $\tilde{\mathbf{X}}$ and $\mathbf{Y}$, respectively. Furthermore, define

$$\mathbf{D}_{ij} \triangleq \sqrt{\left(1+[\mathbf{M}(z)]_{ii}/\sqrt{d}\right)\left(1+[\mathbf{M}(z)]_{jj}/\sqrt{d}\right)}, \tag{119}$$

and

$$\tilde{\mathbf{D}}_{ij} \triangleq \sqrt{\left(1+[\tilde{\mathbf{M}}(z)]_{ii}/\sqrt{d}\right)\left(1+[\tilde{\mathbf{M}}(z)]_{jj}/\sqrt{d}\right)}. \tag{120}$$

With this notation, the entries of $\tilde{\mathbf{X}}$ and $\mathbf{Y}$ can be written as

$$\tilde{\mathbf{X}}_{ij} = \mathbb{1}\{\tilde{\mathbf{D}}_{ij}^{-1}\tilde{\mathbf{M}}_{ij} \geq \sqrt{d}t_{q,d}\}, \qquad \mathbf{Y}_{ij} = \mathbb{1}\{\mathbf{M}_{ij} \geq \overline{\Phi}^{-1}(q)\}. \tag{121}$$

Additionally, let $\tilde{\mathbf{Z}}$ denote the $z\times z$ matrix defined by

$$\tilde{\mathbf{Z}}_{ij} = \mathbb{1}\{\tilde{\mathbf{M}}_{ij} \geq \sqrt{d}\,t_{q,d}\}, \tag{122}$$

for all $1 \leq i < j \leq z$, and let $\tilde{\mathbb{Q}}^{(z)}$ denote its law. We denote by $\mathbf{Z}$ and $\mathbb{Q}^{(z)}$ the corresponding non-truncated matrix and its law, respectively. By Lemma 10, we have

$$1 + D_3(\tilde{\mathbb{P}}_0^{(z)}\|\mathbb{P}_0^{(z)}) \leq \sqrt{1 + D_6(\tilde{\mathbb{P}}_0^{(z)}\|\tilde{\mathbb{Q}}^{(z)})}\sqrt{1 + D_5(\tilde{\mathbb{Q}}^{(z)}\|\mathbb{P}_0^{(z)})}. \tag{123}$$

We now bound each of the two terms appearing on the right-hand side separately.

Bounding $D_5(\tilde{\mathbb{Q}}^{(z)}\|\mathbb{P}_0^{(z)})$. By definition, it is clear that $\tilde{\mathbf{Z}}$ are stochastically dominated by $\mathbf{Z}$. Therefore

$$D_5(\tilde{\mathbb{Q}}^{(z)}\|\mathbb{P}_0^{(z)}) \leq D_5(\mathbb{Q}^{(z)}\|\mathbb{P}_0^{(z)}). \tag{124}$$

Since both $\mathbb{Q}^{(z)}$ and $\mathbb{P}_0^{(z)}$ are product measures, we have

$$1 + D_5(\mathbb{Q}^{(z)}\|\bar{\mathbb{P}}^{(z)}) = \prod_{1\leq i<j\leq z}\left(1 + D_5(\mathbb{Q}_{ij}^{(z)}\|\bar{\mathbb{P}}_{ij}^{(z)})\right), \tag{125}$$

where $\mathbb{Q}_{ij}^{(z)} = \mathsf{Bern}(q+\delta_{q,d})$ and $\bar{\mathbb{P}}_{ij}^{(z)} = \mathsf{Bern}(q)$, where $\delta_{q,d} \triangleq \bar{\Phi}(\sqrt{d}t_{q,d}) - \Psi_d(\sqrt{d}t_{q,d})$. We utilize the following result.

Lemma 16 ((Bubeck et al., 2016, Lemma 7)) *For any fixed $q \in (0,1)$, there exists a constant C_q such that:*

$$|\delta_{q,d}| = |\bar{\Phi}(\sqrt{d}t_{q,d}) - \Psi_d(\sqrt{d}t_{q,d})| \leq C_q d^{-1}. \tag{126}$$

Expanding $D_5(\mathbb{Q}_{ij}^{(z)} \| \bar{\mathbb{P}}_{ij}^{(z)})$, we obtain

$$1 + D_5(\mathbb{Q}_{ij}^{(z)} \| \bar{\mathbb{P}}_{ij}^{(z)}) = q\left(1 + \frac{\delta_{q,d}}{q}\right)^5 + (1-q)\left(1 - \frac{\delta_{q,d}}{1-q}\right)^5 \tag{127}$$

$$= 1 + \frac{55\delta_{q,d}^2}{q(1-q)} + O_q(\delta_{q,d}^3). \tag{128}$$

Applying Lemma 16, we have $\delta_{q,d}^2 \leq C_q^2 d^{-2}$. Substituting this bound into the product over edges and using the inequality $1 + x \leq e^x$, we obtain

$$1 + D_5(\mathbb{Q}^{(z)} \| \bar{\mathbb{P}}^{(z)}) \leq \left(1 + \frac{55C_q^2}{q(1-q)} d^{-2} + O(d^{-3})\right)^{\binom{z}{2}} \tag{129}$$

$$\leq \exp\left(\frac{z(z-1)}{2}\left[\frac{55C_q^2}{q(1-q)d^2}\right]\right) \tag{130}$$

$$\leq \exp\left(\bar{C}_q \frac{z^2}{d^2}\right), \tag{131}$$

for a constant $\bar{C}_q$ depending only on q.

Bounding $D_6(\tilde{\mathbb{P}}_0^{(z)} \| \tilde{\mathbb{Q}}^{(z)})$. We begin by introducing some notation. Recall that $\tilde{\mathbf{X}} = H_{q,d}(\tilde{\mathbf{M}}(z,d)) \sim \tilde{\mathbb{P}}_0^{(z)}$ and $\tilde{\mathbf{Z}} \sim \tilde{\mathbb{Q}}^{(z)}$, where the entries of $\tilde{\mathbf{Z}}$ are defined in (122). Let $\mathbf{X}$ and $\mathbf{Z}$ denote the untruncated versions of $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Z}}$, respectively, that is, $\mathbf{X} = H_{q,d}(\mathbf{M}(z,d))$, and the entries of $\mathbf{Z}$ are given by (122) with $\tilde{\mathbf{M}}$ replaced by $\mathbf{M}$.

For any matrix $\mathbf{A}$, let $D'(\mathbf{A})$ denote the diagonal matrix formed by retaining only the diagonal entries of $\mathbf{A}$, and define the off-diagonal component by $D(\mathbf{A}) = \mathbf{A} - D'(\mathbf{A})$. Introduce the transformation

$$f(\mathbf{A}) \triangleq \left(\mathbf{I}_z + \frac{1}{\sqrt{d}} D'(\mathbf{A})\right)^{-1/2} \mathbf{A} \left(\mathbf{I}_z + \frac{1}{\sqrt{d}} D'(\mathbf{A})\right)^{-1/2}, \tag{132}$$

Observe that the random matrix $\mathbf{Z}$ is obtained by applying an entrywise thresholding procedure to $D(\mathbf{M})$, whereas $\mathbf{X}$ is obtained analogously from $D(f(\mathbf{M}))$. Similarly, $\tilde{\mathbf{Z}}$ and $\tilde{\mathbf{X}}$ are obtained from $D(\tilde{\mathbf{M}})$ and $D(f(\tilde{\mathbf{M}}))$, respectively, via the same entrywise thresholding procedure. Hence, Lemma 9 implies that

$$1 + D_6(\mathbf{X} \| \mathbf{Z}) \leq 1 + D_6(D(f(\mathbf{M})) \| D(\mathbf{M})), \tag{133}$$

where, with a slight abuse of notation, we use the matrix-valued random variables to denote the corresponding probability measures. Similarly,

$$1 + D_6(\tilde{\mathbb{P}}_0^{(z)} \| \tilde{\mathbb{Q}}^{(z)}) = 1 + D_6(\tilde{\mathbf{X}} \| \tilde{\mathbf{Z}}) \leq 1 + D_6(D(f(\tilde{\mathbf{M}})) \| D(\tilde{\mathbf{M}})). \tag{134}$$

Thus, to control the right-hand side of the above inequalities, it suffices to study the densities of $D(\mathbf{M})$ and $D(f(\mathbf{M}))$, denoted by $q(x)$ and $w(x)$, respectively, together with their truncated counterparts, whose densities are denoted by $\tilde{q}(x)$ and $\tilde{w}(x)$.

Let $\Omega = \mathbb{R}^{(n^2-n)/2}$, identified with the space of symmetric $n \times n$ matrices having zero diagonal, and let $\Omega' = \mathbb{R}^n$, corresponding to the space of diagonal entries. Slightly abusing notation, we treat D and D' as maps from symmetric matrices into Ω and Ω', respectively. Under this identification, the mapping $D \circ f$ acts from $\Omega \oplus \Omega'$ into Ω.

Accordingly, $w(x)$ arises as the push-forward of the Gaussian measure with density $\gamma(x, y)$ under the transformation $D \circ f$, where $\gamma(x, y)$ is the GOE density on $\Omega \oplus \Omega'$, i.e.,

$$\gamma(x, y) \triangleq \frac{1}{2^{n/2}(2\pi)^{(n^2+n)/2}} \exp\left(-\frac{1}{2}\sum_{1\le i<j\le n} x_{ij}^2 - \frac{1}{4}\sum_{i=1}^{n} y_i^2\right), \tag{135}$$

for $(x, y) \in \Omega \oplus \Omega'$. Similarly, since $D(\mathbf{M})$ is the image of the same Gaussian measure under the map D, the density $q(x)$ is given by

$$q(x) = \int_{\Omega'} \gamma(x, y)\mathrm{d}y = \frac{1}{(2\pi)^{(n^2-n)/2}} \exp\left(-\frac{1}{2}\sum_{1\le i<j\le n} x_{ij}^2\right), \tag{136}$$

which is the standard Gaussian density on Ω. A closed-form expression for $w(x)$ was derived in (Bubeck et al., 2016, Lemma 6).

Lemma 17 ((Bubeck et al., 2016, Lemma 6)) *The law of $D(f(\mathbf{M}))$ is absolutely continuous with respect to the law of $D(\mathbf{M})$. Moreover, the Radon–Nikodym derivative is given by*

$$\frac{\mathrm{d}\mathcal{L}(D(f(\mathbf{M})))}{\mathrm{d}\mathcal{L}(D(\mathbf{M}))}(x) = \frac{w(x)}{q(x)}, \qquad x \in \Omega, \tag{137}$$

where w and q denote the densities of $D(f(\mathbf{M}))$ and $D(\mathbf{M})$, respectively. Moreover,

$$\frac{w(x)}{q(x)} = \mathbb{E}_\Lambda\left[\prod_{i=1}^{n}\left|1 + \frac{\Lambda_i}{\sqrt{d/2}}\right|^{\frac{z-1}{2}} \exp\left(-\frac{1}{2}\sum_{1\le i<j\le z} x_{ij}^2\left(\frac{\Lambda_i + \Lambda_j}{\sqrt{d/2}} + \frac{2\Lambda_i\Lambda_j}{d}\right)\right)\right], \tag{138}$$

where $\Lambda = (\Lambda_1, \ldots, \Lambda_z) \sim N(0, \mathbf{I}_z)$.

Let us now characterize the truncated densities $\tilde{q}(x)$ and $\tilde{w}(x)$, starting with the former. Define the conditional (truncated) law

$$\tilde{\gamma}(x, y) \triangleq \frac{\gamma(x, y)\mathbb{1}_{\mathcal{Q}_\mathcal{U}}(x, y)}{\mathbb{P}_\gamma(\mathcal{Q}_\mathcal{U})}, \qquad \mathbb{P}_\gamma(\mathcal{Q}_\mathcal{U}) \triangleq \int_{\Omega\times\Omega'} \gamma(x, y)\mathbb{1}_{\mathcal{Q}_\mathcal{U}}(x, y)\,\mathrm{d}x\,\mathrm{d}y. \tag{139}$$

Define the random variables $X \triangleq D(\mathbf{M}) \in \Omega$ and $Z \triangleq D(f(\mathbf{M})) \in \Omega$. Then

$$\tilde{q}(x) = \int_{\Omega'} \tilde{\gamma}(x, y)\mathrm{d}y \tag{140}$$

$$= \frac{1}{\mathbb{P}_\gamma(\mathcal{Q}_\mathcal{U})}\int_{\Omega'} \gamma(x, y)\mathbb{1}_{\mathcal{Q}_\mathcal{U}}(x, y)\mathrm{d}y. \tag{141}$$

Define the conditional density of $D'(\mathbf{M})$ given $X = x$ by $\gamma(y|x) \triangleq \frac{\gamma(x,y)}{q(x)}$. Then the conditional probability of $\mathcal{Q}_{\mathcal{U}}$ given $X = x$ is

$$\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}}|X = x) = \int_{\Omega'} \mathbb{1}_{\mathcal{Q}_{\mathcal{U}}}(x,y)\gamma(y|x)\mathrm{d}y \tag{142}$$

$$= \frac{1}{q(x)} \int_{\Omega'} \gamma(x,y)\mathbb{1}_{\mathcal{Q}_{\mathcal{U}}}(x,y)\mathrm{d}y. \tag{143}$$

Rearranging,

$$\int_{\Omega'} \gamma(x,y)\mathbb{1}_{\mathcal{Q}_{\mathcal{U}}}(x,y)\mathrm{d}y = q(x)\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}}|X = x). \tag{144}$$

Substituting into the expression for $\tilde{q}(x)$ yields

$$\tilde{q}(x) = q(x)\frac{\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}}|X = x)}{\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}})}. \tag{145}$$

In a similar fashion, we obtain

$$\tilde{w}(x) = w(x)\frac{\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}}|Z = x)}{\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}})}. \tag{146}$$

Therefore, for all $x \in \Omega$,

$$\frac{\tilde{w}(x)}{\tilde{q}(x)} = \frac{w(x)}{q(x)} \cdot \frac{\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}}|Z = x)}{\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}}|X = x)} \leq \frac{w(x)}{q(x)} \cdot \frac{1}{\mathbb{P}_\gamma(\mathcal{Q}_{\mathcal{U}}|X = x)}. \tag{147}$$

We are now in a position to bound $D_6(D(f(\tilde{\mathbf{M}}))\|D(\tilde{\mathbf{M}}))$. By definition, we have

$$1 + D_6(D(f(\tilde{\mathbf{M}}))\|D(\tilde{\mathbf{M}})) = \mathbb{E}_{X\sim\tilde{q}}\left[\frac{\tilde{w}(X)}{\tilde{q}(X)}\right]^6 \tag{148}$$

$$\leq \mathbb{E}\left[\left[\frac{w(X)}{q(X)}\right]^6 \frac{1}{\mathbb{P}_\gamma^6(\mathcal{Q}_{\mathcal{U}}|X)}\right] \tag{149}$$

$$= \int_{\mathcal{Q}_{\mathcal{U}}} \left[\frac{w(x)}{q(x)}\right]^6 \frac{1}{\mathbb{P}_\gamma^6(\mathcal{Q}_{\mathcal{U}}|X = x)}\tilde{q}(x)\mathrm{d}x \tag{150}$$

$$= \int_{\mathcal{Q}_{\mathcal{U}}} \left[\frac{w(x)}{q(x)}\right]^6 \tilde{q}(x)\mathrm{d}x. \tag{151}$$

Next, we derive a uniform upper bound on $\frac{w(x)}{q(x)}$, for all $x \in \mathcal{Q}_{\mathcal{U}}$. Specifically, using Lemma 17 and the fact that $|1+u| \leq e^{u+u^2}$ (applied with $u = \Gamma_i/\sqrt{d/2}$), we obtain

$$\frac{w(x)}{q(x)} \leq \mathbb{E}_\Lambda\Bigg[\exp\Bigg(\frac{z-1}{\sqrt{2d}}\sum_{i=1}^{z}\Lambda_i + \frac{z-1}{d}\sum_{i=1}^{z}\Lambda_i^2$$

$$- \frac{1}{\sqrt{2d}}\sum_{i<j} x_{ij}^2(\Lambda_i + \Lambda_j) - \frac{2}{d}\sum_{i<j} x_{ij}^2\Lambda_i\Lambda_j\Bigg)\Bigg] \tag{152}$$

$$= \mathbb{E}_\Lambda\left[\exp\left(\mathbf{L}^T\Lambda + \Lambda^T\mathbf{Q}\Lambda\right)\right], \tag{153}$$

where the entries of the vector $\mathbf{L}$ are defined by

$$\mathbf{L}_i \triangleq \frac{1}{\sqrt{2d}}\left((z-1) - \sum_{j=1,j\neq i}^{z} x_{ij}^2\right), \tag{154}$$

for $i \in [z]$, and the entries of the matrix $\mathbf{Q}$ are given by

$$\mathbf{Q}_{ij} = \begin{cases} \frac{z-1}{d}, & \text{if} i = j \\ -\frac{1}{d}x_{ij}^2, & \text{if} i \neq j, \end{cases} \tag{155}$$

for $1 \leq i, j \leq z$. Crucially, the expectation in (153) is finite only if the matrix $\mathbf{I}_z - 2\mathbf{Q}$ is positive definite. By the Gershgorin circle theorem, the eigenvalues of $\mathbf{Q}$ lie in the union of the discs

$$\mathcal{D}_i = \left\{\lambda \in \mathbb{R} : \left|\lambda - \frac{z-1}{d}\right| \leq \frac{1}{d}\sum_{j=1,j\neq i}^{z} x_{ij}^2\right\}, \tag{156}$$

for $i \in [z]$. Consequently, the operator norm of $\mathbf{Q}$ satisfies

$$\|\mathbf{Q}\|_{\mathsf{op}} \leq \frac{z-1}{d} + \frac{1}{d}\max_{1\leq i\leq z}\sum_{j=1,j\neq i}^{z} x_{ij}^2. \tag{157}$$

Furthermore,

$$\|\mathbf{L}\|_2^2 = \frac{1}{2d}\sum_{i=1}^{z}\left(z-1-\sum_{j\neq i} x_{ij}^2\right)^2 \leq \frac{z(z-1)^2}{2d} + \frac{z}{2d}\left(\max_{1\leq i\leq z}\sum_{j\neq i} x_{ij}^2\right)^2. \tag{158}$$

Now, for any $x \in \mathcal{Q}_{\mathcal{U}}$, by the definition of $\mathcal{Q}_{\mathcal{U}}$, we have

$$\sum_{j=1,j\neq i} x_{ij}^2 \leq (1+C_1)d(z-1)\log k, \tag{159}$$

for all $i \in [z]$. Consequently,

$$\|\mathbf{L}\|_2^2 \leq \frac{z^3}{2d} + \frac{z^3\log^2 k}{2d} \leq C'\frac{z^3\log^2 k}{d}, \tag{160}$$

and

$$\|\mathbf{Q}\|_{\mathsf{op}} \leq \frac{z-1}{d} + (1+C_1)\frac{(z-1)\log k}{d} \leq C''\frac{z\log k}{d}, \tag{161}$$

for some positive constants C' and C''. In the impossible regime, where $d \gg k^2\log^2 k$, it follows that for d sufficiently large we have $\|\mathbf{Q}\|_{\mathsf{op}} < \frac{1}{2}$, and we may therefore apply the Gaussian integral formula to obtain

$$\mathbb{E}_\Lambda\left[\exp\left(\mathbf{L}^T\Lambda + \Lambda^T\mathbf{Q}\Lambda\right)\right] = \det(\mathbf{I}_z - 2\mathbf{Q})^{-1/2}\exp\left(\frac{1}{2}\mathbf{L}^T(\mathbf{I}_z - 2\mathbf{Q})^{-1}\mathbf{L}\right). \tag{162}$$

Moreover, since all eigenvalues of $\mathbf{Q}$ are of order $\Theta(z\log k/d)$, the eigenvalues of $\mathbf{I}_z - 2\mathbf{Q}$ lie in the interval $[1 - O(z\log k/d),\ 1 + O(z\log k/d)]$. In the regime where $d \gg k^2\log^2 k$, this interval is contained in $[1/2, 3/2]$. Hence, letting $\{\lambda_i\}_{i=1}^z$ denote the eigenvalues of $\mathbf{I}_z - 2\mathbf{Q}$, we obtain

$$\det(\mathbf{I}_z - 2\mathbf{Q}) = \prod_{i=1}^{z}\lambda_i \geq (1 - 2\|\mathbf{Q}\|_{\mathsf{op}})^z. \tag{163}$$

This implies that

$$\begin{aligned} -\frac{1}{2}\log\det(\mathbf{I}_z - 2\mathbf{Q}) &\leq -\frac{z}{2}\log(1 - 2\|\mathbf{Q}\|_{\mathsf{op}}) && (164)\\ &\leq \frac{z}{2}\cdot\frac{2\|\mathbf{Q}\|_{\mathsf{op}}}{1 - 2\|\mathbf{Q}\|_{\mathsf{op}}} && (165)\\ &\leq \bar{C}\frac{z^2\log^2 k}{d}, && (166) \end{aligned}$$

for some constant $\bar{C} > 0$. Here, the first inequality follows from (163), while the second uses the bound $\log(1-x) \leq x/(1-x)$ for $0 < x < 1$, together with (161). Turning to the second factor in (162), let $\{v_i\}_{i=1}^z$ denote the eigenvectors of $\mathbf{I}_z - 2\mathbf{Q}$. Then

$$\begin{aligned} \frac{1}{2}\mathbf{L}^T(\mathbf{I}_z - 2\mathbf{Q})^{-1}\mathbf{L} &= \frac{1}{2}\sum_{i=1}^{z}\lambda_i^{-1}\left|\langle v_i, \mathbf{L}\rangle\right|^2 && (167)\\ &\leq \frac{1}{2}\frac{\|\mathbf{L}\|_2^2}{1 - 2\|\mathbf{Q}\|_{\mathsf{op}}} && (168)\\ &\leq \bar{C}'\|\mathbf{L}\|_2^2 && (169)\\ &\leq \bar{C}''\frac{z^3\log^2 k}{d}, && (170) \end{aligned}$$

for some positive constants $\bar{C}', \bar{C}'' > 0$. Finally, combining (151), (153), (162), (166), and (170), we get

$$1 + D_6(\tilde{\mathbb{P}}_0^{(z)}\|\tilde{\mathbb{Q}}^{(z)}) \leq \exp\left(C\left(\frac{z^2}{d} + \frac{z^3}{d}\right)\log^2 k\right), \tag{171}$$

for some constant $C > 0$.

Completing the second-moment calculation. After bounding each divergence separately, we are ready to assemble the final bound on $\mathbb{E}_{\mathcal{H}_0}\left[\tilde{\mathcal{L}}^2\right]$. To this end, we substitute (118), (123), (131), and (171) into (90), and then into (83). We obtain

$$\begin{aligned} \mathbb{E}_{\mathcal{H}_0}\left[\tilde{\mathcal{L}}^2\right] &\leq 1 + o(1) + \mathbb{E}_{\mathcal{U}}\left[\chi^2(\tilde{\mathcal{G}}_d(Z,q)\|\mathcal{G}(Z,q))\mathbb{1}\{Z \geq 2\}\right] && (172)\\ &\leq 1 + o(1) + \mathbb{E}_Z\left[\left(\exp\left(C\frac{Z^3\log^2 k}{d}\right) - 1\right)\mathbb{1}\{Z \geq 2\}\right] && (173)\\ &\leq \mathbb{E}_Z\left[\exp\left(C\frac{Z^3\log^2 k}{d}\right)\right] + o(1), && (174) \end{aligned}$$

for some constant $C > 0$, and the third inequality follows from the fact that $\mathbb{P}[Z \geq 0] = 1$. Let us find the conditions under which the right-hand side of (174) is converging to unity or bounded.

Next, we analyze the expectation on the right-hand side of (174). We decompose the expectation into two regimes: the first corresponding to $Z \leq z_1$, and its complement, where we define $z_1 \triangleq C' \lceil (d/\log^2 k)^{1/3} \rceil$, for some $C' > 0$. We have

$$\mathbb{E}_Z \left[\exp\left(C\frac{Z^3 \log^2 k}{d} \right) \right] = \mathbb{E}_Z \left[\exp\left(C\frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{Z \leq z_1\}} \right] + \mathbb{E}_Z \left[\exp\left(C\frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{Z > z_1\}} \right]. \tag{175}$$

As for the first expectation on the right-hand side of (175), we consider the impossibility of strong and weak detection separately. For the former, we simply have

$$\mathbb{E}_Z \left[\exp\left(C\frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{z \leq z_1\}} \right] \leq \mathbb{E}_Z \left[\exp\left(C\frac{z_1^3 \log^2 k}{d} \right) \right] = O(1). \tag{176}$$

Next, we turn to the impossibility of weak detection. We assume that $\frac{dn^3}{k^6 \log^2 k\, f(k)} \to \infty$ for any function $f(k) = \omega(1)$. We again split the analysis into two regimes: $Z \leq z_0 \triangleq (k^2 f(k))/n$ and $z_0 < Z \leq z_1$. For $Z \leq z_0$, we bound the expectation as follows:

$$\mathbb{E}_Z \left[\exp\left(C\frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{Z \leq z_0\}} \right] \leq \exp\left(C\frac{k^6 \log^2 k f(k)}{n^3 d} \right) = 1 + o(1). \tag{177}$$

For the regime $z_0 < Z \leq z_1$, we use the fact that $Z \sim \mathsf{Hypergeometric}(n,k,k)$ is stochastically dominated by $B \sim \mathsf{Binomial}(k,\rho)$ with $\rho = k/(n-k)$. Together with the Chernoff bound $\mathbb{P}(B \geq z) \leq \exp\big(-k d_{\mathsf{KL}}(z/k \| \rho)\big)$, this yields

$$\mathbb{E}_Z \left[\exp\left(C\frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{z_0 < z \leq z_1\}} \right] \leq \sum_{z=z_0+1}^{z_1} \exp\left[C\frac{z^3 \log^2 k}{d} - k d_{\mathsf{KL}}\left(\frac{z}{k} \| \rho \right) \right] \tag{178}$$

$$\leq \sum_{z=z_0+1}^{z_1} \exp\left[z\left(C\frac{z^2 \log^2 k}{d} - \log\left(\frac{z}{k\rho} \right) + 1 \right) \right], \tag{179}$$

where the last inequality follows from the identity $(1-x)\log(1-x) \geq -x$. For any fixed $a > 0$, the function $f(x) = ax^2 - \log x$ is decreasing on $(0, 1/\sqrt{2a})$ and increasing on $(1/\sqrt{2a}, \infty)$. Consequently, for any $z \in \{z_0+1, \ldots, k\}$, we have

$$C\frac{z^2 \log^2 k}{d} - \log\left(\frac{z}{k\rho} \right) \leq -\omega, \tag{180}$$

where

$$\omega \triangleq \min\left\{ C\frac{z_0^2 \log^2 k}{d} - \log\left(\frac{z_0}{k\rho} \right), C\frac{z_1^2 \log^2 k}{d} - \log\left(\frac{z_1}{k\rho} \right) \right\}. \tag{181}$$

Focusing first on the first quantity in the minimum, we note that in our regime $(k^2 \log^2 k)/d \to 0$, and hence $(z_0^2 \log^2 k)/d \to 0$ as well. Moreover, $\log(z_0/(k\rho)) = \log(f(k) + o(1)) \to \infty$, and therefore the first term diverges. Turning to the second quantity, its first term converges to zero by the same reasoning. In addition, since $z_1/(k\rho) = (dn^3/(k^6 \log^2 k))^{1/3}$ and $dn^3/(k^6 \log^2 k) \to \infty$, it follows that $\log(z_1/(k\rho)) \to \infty$. This implies that

$$\mathbb{E}_Z \left[\exp\left(C \frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{z_0 < Z \le z_1\}} \right] \le \sum_{z=z_0+1}^{z_1} e^{-\omega z} \le \frac{e^{-\omega/2}}{1 - e^{-\omega/2}} = o(1), \tag{182}$$

which consequently leads to

$$\mathbb{E}_Z \left[\exp\left(C \frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{Z \le z_1\}} \right] \le 1 + o(1). \tag{183}$$

We now turn to the second expectation on the right-hand side of (175). We again use the stochastic dominance of the binomial distribution over the hypergeometric distribution, together with a Chernoff bound, to obtain

$$\mathbb{E}_Z \left[\exp\left(C \frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{Z \ge z_1\}} \right] \le \sum_{z=z_1+1}^{k} e^{-\omega z}, \tag{184}$$

where

$$\omega \triangleq \min\left\{ C \frac{z_1^2 \log^2 k}{d} - \log\left(\frac{z_1}{k\rho}\right), C \frac{k^2 \log^2 k}{d} - \log\left(\frac{1}{\rho}\right) \right\}. \tag{185}$$

We have already shown that the first term in the minimum diverges under the condition $dn^3/(k^6 \log^2 k) \to \infty$. For the second term, we distinguish between the regimes $k = o(n)$ and $k = \Theta(n)$.

For $k = o(n)$, the second term clearly diverges due to the growth of $\log(1/\rho) = \log(n/k + o(1))$ together with the assumption $d \gg k^2 \log^2 k$. Hence $\omega \to \infty$, which implies

$$\mathbb{E}_Z \left[\exp\left(C \frac{Z^3 \log^2 k}{d} \right) \mathbb{1}_{\{Z > z_1\}} \right] = o(1). \tag{186}$$

Combining (183) and (186), we conclude that under the conditions of Theorem 2 but with $\frac{dn^3}{k^6 \log^2 k} \to \infty$ replaced by $\frac{dn^3}{k^6 f(k)} \to \infty$ for some function $f(k) = \omega(\log^2 k)$, $\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2] \le 1 + o(1)$, and hence weak detection is impossible. Similarly, under the conditions of Theorem 2, equations (176) and (186) yield $\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2] = O(1)$, and therefore strong detection is impossible.

Finally, when $k = \Theta(n)$, we use the fact that $Z \le k$ to obtain

$$\mathbb{E}_Z \left[\exp\left(C \frac{Z^3 \log^2 k}{d} \right) \right] \le \exp\left(C \frac{k^3 \log^2 k}{d} \right). \tag{187}$$

Then, under the conditions of Theorem 2, namely $(dn^3)/(k^6 \log^2 k) \to \infty$, when $k = \Theta(n)$ we have $k^3 \log^2 k/d \to 0$. It follows that $\mathbb{E}_{\mathcal{H}_0}[\tilde{\mathcal{L}}^2] \le 1 + o(1)$, and hence both weak and strong detection are impossible. This concludes the proof of Theorem 2.

Appendix D. Proofs of Upper Bounds

In this section, we prove Theorem 3. We begin by analyzing the Type-I and II error probabilities of the signed triangle test, and then turn to the triangle scan test.

D.1. Signed triangle test

Recall the signed triangle test in (9). To analyze its performance we use the second moment technique. To that end, we start by finding the first and second moments of $\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n)$ in (7), under $\mathcal{H}_0$ and $\mathcal{H}_1$. Specifically, it is easy to show that (Bubeck et al., 2016, Section 3.1)

$$\mathbb{E}_{\mathcal{H}_0}\left[\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n)\right] = 0, \tag{188}$$

$$\mathsf{Var}_{\mathcal{H}_0}\left(\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n)\right) = \binom{n}{3} q^3(1-q)^3 \le n^3. \tag{189}$$

On the other hand, define $\mathcal{T}_{\mathcal{K}} \triangleq \left\{\{i,j,\ell\} \subset \binom{[n]}{3} : i,j,\ell \in \mathcal{K}\right\}$, that is, the collection of all triples whose indices lie entirely within the planted set $\mathcal{K}$. Then, observe that

$$\mathbb{E}_{\mathcal{H}_1}\left[\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n)\right] = \mathbb{E}_{\mathcal{H}_1}\left[\sum_{\{i,j,\ell\}\subset\binom{[n]}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell)\right] \tag{190}$$

$$= \sum_{\{i,j,\ell\}\in\mathcal{T}_{\mathcal{K}}} \mathbb{E}_{\mathcal{H}_1}\left[\mathsf{T}_{\mathsf{G}_n}(i,j,\ell)\right] \tag{191}$$

$$\ge \binom{k}{3}\frac{\mathsf{C}_q}{\sqrt{d}}, \tag{192}$$

where the second equality follows from the fact that the expectation of a signed triangle $\mathsf{T}_{\mathsf{G}_n}(i,j,\ell)$ is zero unless all three indices (i,j,ℓ) belong to the planted set $\mathcal{K}$. The inequality then follows from (Bubeck et al., 2016, Lemma 3), where C_q denotes a constant depending only on q. Finally,

$$\mathsf{Var}_{\mathcal{H}_1}\left(\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n)\right) = \mathsf{Var}_{\mathcal{H}_1}\left(\sum_{\{i,j,\ell\}\subset\binom{[n]}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell)\right) \tag{193}$$

$$\le 2\cdot\mathsf{Var}_{\mathcal{H}_1}\left(\sum_{\{i,j,\ell\}\in\mathcal{T}_{\mathcal{K}}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell)\right) + 2\cdot\mathsf{Var}_{\mathcal{H}_1}\left(\sum_{\{i,j,\ell\}\in\mathcal{T}_{\mathcal{K}}^c} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell)\right) \tag{194}$$

$$\le 2k^3 + \frac{6k^4}{d} + 2n^3, \tag{195}$$

where first inequality follows from the fact that $\mathsf{Var}(X+Y) \le 2\cdot\mathsf{Var}(X)+2\cdot\mathsf{Var}(Y)$, for any pair of random variables (X,Y). The second inequality is obtained by bounding the first variance term in (194) using (Bubeck et al., 2016, eq. (29)), and the second variance term using (189). Consequently, by Chebyshev's inequality, we obtain

$$\mathbb{P}_{\mathcal{H}_0}\left(\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n) \ge \frac{\mathbb{E}_{\mathcal{H}_1}\left[\mathsf{T}(\mathsf{G}_n)\right]}{2}\right) \le \mathsf{C}_1\cdot\frac{dn^3}{k^6}, \tag{196}$$

for some constant $\mathsf{C}_1 > 0$, and

$$\mathbb{P}_{\mathcal{H}_1}\left(\mathsf{T}_{\mathsf{triangle}}(\mathsf{G}_n) < \frac{\mathbb{E}_{\mathcal{H}_1}\left[\mathsf{T}(\mathsf{G}_n)\right]}{2}\right) \leq \mathsf{C}_2 \frac{dk^3 + 3k^4 + dn^3}{k^6}, \tag{197}$$

for some constant $\mathsf{C}_2 > 0$. Therefore, we conclude that $\mathsf{R}(\mathcal{A}_{\mathsf{triangle}}) \to 0$ as $n \to \infty$ provided that $dn^3/k^6 \to 0$ and $d/k^3 \to 0$. Since the latter condition is implied by the former, this reduces to the requirement $dn^3/k^6 \to 0$, as stated in item 1 of Theorem 3.

D.2. Triangle scan test

We analyze the Type-I and Type-II error probabilities associated with the triangle scan test defined in (8) and (10). We define $\tau_{\mathsf{scan}} = \frac{1}{2}\binom{k}{3}\frac{\mathsf{C}_q}{\sqrt{d}}$, and denote the underlying planted set by $\mathcal{K}$. Note that

$$\zeta \triangleq \mathbb{E}_{\mathcal{H}_1}\left[\sum_{\{i,j,\ell\}\subset\binom{\mathcal{K}}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell)\right] \geq \binom{k}{3}\frac{\mathsf{C}_q}{\sqrt{d}}, \tag{198}$$

where the inequality follows from (Bubeck et al., 2016, Lemma 3), and C_q, as before, denotes a constant depending only on q. We now turn to the Type-II error probability; applying Chebyshev's inequality yields

$$\begin{aligned}
\mathbb{P}_{\mathcal{H}_1}\left(\mathsf{T}_{\mathsf{scan}}(\mathsf{G}_n) < \tau_{\mathsf{scan}}\right) &= \mathbb{P}_{\mathcal{H}_1}\left(\max_{\mathcal{S}\subset[n]:|\mathcal{S}|=k} \sum_{\{i,j,\ell\}\subset\binom{\mathcal{S}}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell) < \tau_{\mathsf{scan}}\right) && (199)\\
&\leq \mathbb{P}_{\mathcal{H}_1}\left(\sum_{\{i,j,\ell\}\subset\binom{\mathcal{K}}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell) < \tau_{\mathsf{scan}}\right) && (200)\\
&\leq \frac{\mathsf{Var}_{\mathcal{H}_1}\left(\sum_{\{i,j,\ell\}\subset\binom{\mathcal{K}}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell)\right)}{(\zeta - \tau_{\mathsf{scan}})^2} && (201)\\
&\leq \mathsf{C}_3 \frac{dk^3 + 3k^4}{k^6}, && (202)
\end{aligned}$$

for some constant $\mathsf{C}_3 > 0$, and the last inequality follows from (Bubeck et al., 2016, eq. (29)). Thus, we see that the Type-II converges to zero if $d/k^3 \to 0$ and $k \to \infty$.

Next, we analyze the Type-I error probability. By the union bound, we have

$$\mathbb{P}_{\mathcal{H}_0}\left(\mathsf{T}_{\mathsf{scan}}(\mathsf{G}_n) \geq \tau_{\mathsf{scan}}\right) \leq \binom{n}{k} \cdot \mathbb{P}_{\mathcal{H}_0}\left(\sum_{\{i,j,\ell\}\subset\binom{\mathcal{K}}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell) \geq \tau_{\mathsf{scan}}\right), \tag{203}$$

and thus, it remains to upper bound the probability term on the right-hand side of (203). The following result provides the required bound.

Lemma 18 *Fix $\mathcal{K} \subset [n]$ with $|\mathcal{K}| = k \geq 3$. Let τ_{scan} be any threshold of the form $\tau_{\mathsf{scan}} \geq c_0 k^3/\sqrt{d}$ for a constant $c_0 = c_0(q) > 0$. Then, for some $C = C(q) > 0$ and $C' = C'(q) > 0$,*

$$\mathbb{P}_{\mathcal{H}_0}\left(\sum_{\{i,j,\ell\}\subset\binom{\mathcal{K}}{3}} \mathsf{T}_{\mathsf{G}_n}(i,j,\ell) \geq \tau_{\mathsf{scan}}\right) \leq \exp\left(-C\frac{k^3}{d\alpha\log n}\right) + \exp\left(-\alpha C' k\log n\right), \tag{204}$$

for any $\alpha > 0$.

Assuming the validity of Lemma 18, we can bound (203) as follows:

$$\mathbb{P}_{\mathcal{H}_0}\left(\mathsf{T}_{\mathsf{scan}}(\mathsf{G}_n) \geq \tau_{\mathsf{scan}}\right) \leq \binom{n}{k}\cdot\left[\exp\left(-C\frac{k^3}{\alpha d\log n}\right) + \exp\left(-C' k\log n\right)\right] \tag{205}$$

$$\leq \exp\left[k\cdot\left(\log n - C\frac{k^2}{\alpha d\log n}\right)\right] + \exp\left(-(\alpha C' - 1)\log n\right). \tag{206}$$

Thus, we see that the first term at the right-hand side of (206) converges to zero provided that $k^2 > \frac{\alpha d\log^2 n}{C}$ and $k\to\infty$, while the second term converges to zero by taking any fixed constant $\alpha > 1/C'$, and $k\to\infty$. It therefore remains to establish Lemma 18. To this end, we prove a sequence of auxiliary results, beginning with the following lemma.

Lemma 19 *Fix $\mathcal{K} \subset [n]$ with $|\mathcal{K}| = k \geq 3$, and let $\mathsf{G} \sim \mathcal{G}(n,q)$ under $\mathcal{H}_0$. For an edge $e = (i,j) \subset \binom{\mathcal{K}}{2}$, set*

$$\bar{\mathbf{A}}_e \triangleq \mathbf{A}_e - q, \tag{207}$$

$$W_e \triangleq \sum_{\ell\in\mathcal{K}\setminus\{i,j\}} \bar{\mathbf{A}}_{i\ell}\bar{\mathbf{A}}_{j\ell}, \tag{208}$$

$$\mathsf{T}_{\mathcal{K}}(\mathsf{G}) \triangleq \sum_{\{i,j,\ell\}\subset\binom{\mathcal{K}}{3}} \bar{\mathbf{A}}_{ij}\bar{\mathbf{A}}_{i\ell}\bar{\mathbf{A}}_{j\ell} = \frac{1}{3}\sum_{e\subset\binom{\mathcal{K}}{2}} \bar{\mathbf{A}}_e W_e. \tag{209}$$

There exists an absolute constant $C_{\mathrm{dec}} \in (0,\infty)$ (one may take $C_{\mathrm{dec}} = 8$) such that for every $\theta \in \mathbb{R}$,

$$\mathbb{E}\left[\exp\left(\theta\mathsf{T}_{\mathcal{K}}(\mathsf{G})\right)\right] \leq \mathbb{E}\left[\exp\left(\frac{C_{\mathrm{dec}}^2\theta^2}{72}\sum_{e\subset\binom{\mathcal{K}}{2}} W_e^2\right)\right]. \tag{210}$$

Proof [Proof of Lemma 19] Introduce three *independent* copies of the edges $\{\mathbf{A}_{ij}\}$ under $\mathcal{H}_0$, denoted by $\{\mathbf{A}_{ij}^{(1)}\}, \{\mathbf{A}_{ij}^{(2)}\}, \{\mathbf{A}_{ij}^{(3)}\}$, each distributed as $G(n,q)$, and write $\bar{\mathbf{A}}_{ij}^{(r)} \triangleq \mathbf{A}_{ij}^{(r)} - q$. Define the fully decoupled statistic

$$\mathsf{T}_{\mathcal{K}}^{\mathrm{dec}}(\mathsf{G}) \triangleq \sum_{\{i,j,\ell\}\subset\binom{\mathcal{K}}{3}} \bar{\mathbf{A}}_{ij}^{(1)}\bar{\mathbf{A}}_{i\ell}^{(2)}\bar{\mathbf{A}}_{j\ell}^{(3)}. \tag{211}$$

We use the following classical decoupling inequality for *canonical* U-statistics of order 3. We start with the following definition.

Definition 20 *Let $X_1, X_2, \ldots$ be i.i.d. random variables with values in a measurable space $(\mathcal{X}, \mathcal{A})$. For an integer $m \geq 1$, let $h : \mathcal{X}^m \to \mathbb{R}$ be a symmetric measurable function (the kernel). The kernel h is said to be* canonical *if*

$$\mathbb{E}\left[h(X_1, \ldots, X_m) | X_r\right] = 0 \qquad \textit{almost surely for every } r = 1, \ldots, m. \tag{212}$$

A U-statistic of order m built from a canonical kernel is called a canonical U-statistic.

Lemma 21 (Decoupling for U-statistics (de la Pena and Giné, 1999, Thm. 3.1.1)) *Let U be a canonical U-statistic of order m built from i.i.d. inputs. Let U^{dec} denote the corresponding fully decoupled version, constructed by evaluating the kernel on m independent copies of the input sequence. Then there exists a constant C_m depending only on the order m such that, for every convex nondecreasing function $\Phi : \mathbb{R} \to \mathbb{R}$,*

$$\mathbb{E}\left[\Phi(U)\right] \leq \mathbb{E}\left[\Phi\left(C_m U^{\mathrm{dec}}\right)\right]. \tag{213}$$

One may take $C_m = 2^{m-1}$.

Applying this lemma with $\Phi(x) = e^{\theta x}$ and $U = \mathsf{T}_{\mathcal{K}}(\mathsf{G})$ we get

$$\mathbb{E}\left[\exp\left(\theta \mathsf{T}_{\mathcal{K}}(\mathsf{G})\right)\right] \leq \mathbb{E}\left[\exp\left(C_{\mathrm{dec}} \theta \mathsf{T}_{\mathcal{K}}^{\mathrm{dec}}(\mathsf{G})\right)\right], \tag{214}$$

where C_{dec} is an absolute constant. Now, for each edge $e = (i,j)$, define the *decoupled coefficient*

$$S_e \triangleq \sum_{\ell \in K \setminus \{i,j\}} \bar{\mathbf{A}}_{i\ell}^{(2)} \bar{\mathbf{A}}_{j\ell}^{(3)}. \tag{215}$$

Then we can rewrite

$$\mathsf{T}_{\mathcal{K}}^{\mathrm{dec}}(\mathsf{G}) = \frac{1}{3} \sum_{e \subset \binom{\mathcal{K}}{2}} \bar{\mathbf{A}}_e^{(1)} S_e. \tag{216}$$

Crucially, conditionally on the arrays $\{\mathbf{A}^{(2)}\}, \{\mathbf{A}^{(3)}\}$, the family $\{\bar{\mathbf{A}}_e^{(1)} : e \subset \binom{\mathcal{K}}{2}\}$ is independent, mean-zero, and each $\bar{\mathbf{A}}_e^{(1)} \in [-q, 1-q]$, while the coefficients $\{S_e\}$ are deterministic constants. Therefore, by Hoeffding's lemma[2] applied edgewise and then multiplied,

$$\mathbb{E}\left[\exp\left(\frac{C_{\mathrm{dec}}\theta}{3} \sum_{e \subset \binom{\mathcal{K}}{2}} \bar{\mathbf{A}}_e^{(1)} S_e\right) \middle| \mathbf{A}^{(2)}, \mathbf{A}^{(3)}\right] \leq \exp\left(\frac{(C_{\mathrm{dec}}\theta)^2}{72} \sum_{e \subset \binom{\mathcal{K}}{2}} S_e^2\right). \tag{217}$$

Taking expectation over $(\mathbf{A}^{(2)}, \mathbf{A}^{(3)})$ and using (214),

$$\mathbb{E}\left[\exp\left(\theta \mathsf{T}_{\mathcal{K}}(\mathsf{G})\right)\right] \leq \mathbb{E}\left[\exp\left(\frac{C_{\mathrm{dec}}^2 \theta^2}{72} \sum_{e \subset \binom{\mathcal{K}}{2}} S_e^2\right)\right]. \tag{218}$$

2. For mean-zero $X \in [a,b]$: $\mathbb{E}e^{\lambda X} \leq \exp\{\lambda^2 (b-a)^2/8\}$.

For fixed $e = (i,j)$, the random sum $S_e = \sum_{\ell \neq i,j} \bar{\mathbf{A}}^{(2)}_{i\ell} \bar{\mathbf{A}}^{(3)}_{j\ell}$ is a sum of independent products of independent, centered, bounded Bernoulli variables; hence S_e has the *same distribution* as $W_e = \sum_{\ell \neq i,j} \bar{\mathbf{A}}_{i\ell} \bar{\mathbf{A}}_{j\ell}$ (the only difference is that the two factors in each product come from independent copies, which does not change the one-dimensional law, since $\bar{\mathbf{A}}_{i\ell}$ and $\bar{\mathbf{A}}_{j\ell}$ are independent already). Consequently, the random vectors $(S_e)_{e \subset \binom{\mathcal{K}}{2}}$ and $(W_e)_{e \subset \binom{\mathcal{K}}{2}}$ have the same joint law under the product Erdős–Rényi measure on arrays $\{\mathbf{A}^{(2)}\} \times \{\mathbf{A}^{(3)}\}$ and on $\{\mathbf{A}\}$, respectively, *up to relabeling of underlying independent coordinates*. In particular,

$$\sum_{e \subset \binom{\mathcal{K}}{2}} S_e^2 \overset{d}{=} \sum_{e \subset \binom{\mathcal{K}}{2}} W_e^2, \tag{219}$$

and the right-hand sides of (218) and (210) coincide in distribution. This proves (210). ■

Next, we have the following high probability upper bound on $\sum_{e \subset \binom{\mathcal{K}}{2}} W_e^2$.

Lemma 22 *Let $\mathcal{K} \subset [n]$ with $|\mathcal{K}| = k \geq 3$, and for each $e = (i,j) \subset \binom{\mathcal{K}}{2}$ define*

$$\mathsf{W}_e = \sum_{\ell \in \mathcal{K} \setminus \{i,j\}} (\mathbf{A}_{i\ell} - q)(\mathbf{A}_{j\ell} - q). \tag{220}$$

There exist constants $C = C(q) > 0$ $C = C'(q) > 0$ such that for sufficiently large n and $k = \Omega(\log n)$

$$\mathbb{P}\left(\sum_{e \in \binom{\mathcal{K}}{2}} W_e^2 > \alpha C k^3 \log n \right) \leq 3 \exp\left(-\alpha C' k \log n\right), \tag{221}$$

for any $\alpha > 0$.

Proof [Proof of Lemma 22] For simplicity of notations, let $\mathbf{B} \in \mathbb{R}^{k \times k}$ be the symmetric matrix indexed by $\mathcal{K}$ with

$$\mathbf{B}_{ij} \triangleq \mathbf{A}_{ij} - q, \tag{222}$$

for $i \neq j$ and $\mathbf{B}_{ij} = 0$, for $i = j$. Now for $i \neq j$,

$$[\mathbf{B}^2]_{ij} = \sum_{m \in \mathcal{K}} \mathbf{B}_{im} \mathbf{B}_{mj}. \tag{223}$$

Since $\mathbf{B}_{ii} = \mathbf{B}_{jj} = 0$ and $\mathbf{B}_{mj} = \mathbf{B}_{jm}$, this becomes

$$[\mathbf{B}^2]_{ij} = \sum_{m \in \mathcal{K} \setminus \{i,j\}} (\mathbf{A}_{im} - q)(\mathbf{A}_{jm} - q) = W_{ij}. \tag{224}$$

So $W_{ij} = [\mathbf{B}^2]_{ij}$ for all $i \neq j$. Therefore the "energy"

$$\mathcal{E} \triangleq \sum_{e \subset \binom{\mathcal{K}}{2}} W_e^2 \tag{225}$$

satisfies

$$\mathcal{E} = \sum_{i<j}([\mathbf{B}^2]_{ij})^2 = \frac{1}{2}\sum_{i\neq j}([\mathbf{B}^2]_{ij})^2 \leq \frac{1}{2}\sum_{i,j}([\mathbf{B}^2]_{ij})^2 = \frac{1}{2}\|\mathbf{B}^2\|_F^2. \tag{226}$$

But $\|\mathbf{B}^2\|_F^2 = \mathsf{trace}([\mathbf{B}^2]^\top[\mathbf{B}^2]) = \mathsf{trace}(\mathbf{B}^4)$ since $\mathbf{B}$ is symmetric. Hence

$$\mathcal{E} \leq \frac{1}{2}\mathsf{trace}(\mathbf{B}^4) = \frac{1}{2}\|\mathbf{B}^2\|_F^2 \leq \frac{1}{2}\|\mathbf{B}\|_{\text{op}}^2\,\|\mathbf{B}\|_F^2, \tag{227}$$

where we use the Frobenius-operator norm inequality.

Next, we derive high probability upper bounds for the above Frobenius and norm operators, starting with the former. Note that for $i \neq j$,

$$\mathbf{B}_{ij} \in [-q, 1-q] \subset [-1,1], \quad \mathbb{E}[\mathbf{B}_{ij}] = 0, \quad \mathbb{E}[\mathbf{B}_{ij}^2] = \mu, \tag{228}$$

where $\mu \triangleq q(1-q)$. We have

$$\|\mathbf{B}\|_F^2 = \sum_{i,j}\mathbf{B}_{ij}^2 = 2\sum_{1\leq i<j\leq k}\mathbf{B}_{ij}^2 \triangleq 2\sum_{i<j} Z_{ij}, \tag{229}$$

where $Z_{ij} \triangleq \mathbf{B}_{ij}^2 \in [0,1]$, i.i.d. over $i<j$, and $\mathbb{E}[Z_{ij}] = \mu$. Let $N \triangleq \binom{k}{2}$. Then $\sum_{i<j} Z_{ij}$ is a sum of N i.i.d. $[0,1]$-bounded variables with mean μ. Let us apply Bernstein's inequality to $X_r \triangleq Z_r - \mu$. We note that $|X_r| \leq 1$ and $\mathsf{Var}(X_r) \leq \mathbb{E}[Z_r^2] \leq \mathbb{E}[Z_r] = \mu$ since $0 \leq Z_r \leq 1$. Thus, Bernstein's inequality yields

$$\mathbb{P}\left(\sum_{i<j} Z_{ij} \geq 2N\mu\right) = \mathbb{P}\left(\sum_{i<j}(Z_{ij}-\mu) \geq N\mu\right) \tag{230}$$

$$\leq \exp\left(-\frac{(N\mu)^2}{2N\mu + \frac{2}{3}N\mu}\right) \tag{231}$$

$$= \exp\left(-\frac{3}{8}N\mu\right) \tag{232}$$

$$\leq \exp\left(-C\mu k^2\right) \tag{233}$$

for an absolute $C > 0$. Therefore

$$\mathbb{P}\left(\|\mathbf{B}\|_F^2 > 2\mu k^2\right) \leq \exp(-C\mu k^2). \tag{234}$$

As for $\|\mathbf{B}\|_{\text{op}}$ we can use classical results, e.g., (Vershynin, 2018, Thm. 4.4.3), and get for any $\alpha > 0$

$$\mathbb{P}\left(\|\mathbf{B}\|_{\text{op}} \geq C\sqrt{\alpha k \log n}\right) \leq 2\exp(-C'\alpha k \log n), \tag{235}$$

for universal constants $C, C' > 0$. Combining (227), (234), (235), and the union bound, we get

$$\mathbb{P}\left(\sum_{e\subset\binom{\mathcal{K}}{2}} W_e^2 > \alpha C''(q)k^3\log n\right) \leq e^{-C\mu k^2} + 2e^{-C'\alpha k\log n} \leq 3e^{-C'\alpha k\log n}, \tag{236}$$

for an constant $C'' = C''(q) > 0$, and in the last inequality we used that fact that $k = \Omega(\log n)$. ■

Using the above lemmata we prove the following general tail bound on $\mathsf{T}_{\mathcal{K}}(\mathsf{G})$.

Proposition 1 (Upper tail for $\mathsf{T}_{\mathcal{K}}(\mathsf{G})$) *Let $\mathsf{T}_{\mathcal{K}}(\mathsf{G})$ be as above. Then for all $t > 0$ and $u > 0$,*

$$\mathbb{P}\left(\mathsf{T}_{\mathcal{K}}(\mathsf{G}) \geq t, \sum_{e\subset K} W_e^2 \leq u\right) \leq \exp\left(-\theta t + \frac{C_{\text{dec}}^2\theta^2}{72}u\right) \qquad \textit{for all } \theta > 0. \tag{237}$$

Optimizing in θ gives

$$\mathbb{P}\left(\mathsf{T}_{\mathcal{K}}(\mathsf{G}) \geq t, \sum_{e\subset K} W_e^2 \leq u\right) \leq \exp\left(-\frac{18t^2}{C_{\text{dec}}^2 u}\right). \tag{238}$$

Consequently,

$$\mathbb{P}\left(\mathsf{T}_{\mathcal{K}}(\mathsf{G}) \geq t\right) \leq \exp\left(-\frac{18t^2}{C_{\text{dec}}^2 u}\right) + \mathbb{P}\left(\sum_{e\subset K} W_e^2 > u\right). \tag{239}$$

Proof [Proof of Proposition 1] By Markov's inequality and Lemma 19,

$$\begin{aligned}
\mathbb{P}\left(\mathsf{T}_{\mathcal{K}}(\mathsf{G}) \geq t \sum_{e\subset K} W_e^2 \leq u\right) &= \mathbb{E}\left[\mathbb{1}\left\{\mathsf{T}_{\mathcal{K}}(\mathsf{G}) \geq t, \sum_{e\subset K} W_e^2 \leq u\right\}\right] &(240)\\
&\leq e^{-\theta t}\mathbb{E}\left[e^{\theta \mathsf{T}_{\mathcal{K}}(\mathsf{G})}\mathbb{1}\left\{\sum_{e\subset K} W_e^2 \leq u\right\}\right] &(241)\\
&\leq \exp\left(-\theta t + \tfrac{C_{\text{dec}}^2\theta^2}{72}u\right). &(242)
\end{aligned}$$

The choice $\theta^\star = 36t/(C_{\text{dec}}^2 u)$ minimizes the exponent and yields the result. ■

Finally, we are in a position to prove Lemma 18. Specifically, apply Proposition 1 with $u = \alpha C k^3 \log n$ from Lemma 22. Then

$$\mathbb{P}\left(\mathsf{T}_{\mathcal{K}}(\mathsf{G}) \geq \tau_{\mathsf{scan}}\right) \leq \exp\left(-\frac{18\tau_{\mathsf{scan}}^2}{C_{\text{dec}}^2 \alpha C k^3 \log n}\right) + 3\exp\left(-\alpha C' k \log n\right). \tag{243}$$

For $\tau_{\mathsf{scan}} \geq c_0 k^3/\sqrt{d}$ and $k \geq 3$, the first term is $\leq \exp\{-ck^3/(\alpha d \log n)\}$ for some $c = c(q) > 0$. This concludes the proof.

Appendix E. Computational Lower bounds

In this subsection we prove Theorem 6, starting with the lower bound component, showing when $\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0}^2$ remains bounded.

Low-degree computational lower bound. Recall that $L^2(\mathcal{H}_0)$ is the Hilbert space of random variables over the probability space on which $\mathbb{P}_{\mathcal{H}_0}$ is defined, with a finite second moment, and equipped with the inner product

$$\langle \varphi(\mathsf{G}), \psi(\mathsf{G})\rangle_{\mathcal{H}_0} \triangleq \mathbb{E}_{\mathcal{H}_0}\left[\varphi(\mathsf{G})\cdot\psi(\mathsf{G})\right]. \tag{244}$$

For an edge $e = \{i, j\}$ define the centered and normalized edge variable

$$\psi_{ij}(\mathsf{G}) \triangleq \frac{\mathsf{G}_{ij} - q}{\sqrt{q(1-q)}}. \tag{245}$$

For any non-empty subset of edges in the complete graph $\alpha \subseteq \binom{[n]}{2}$, define the Fourier character χ_α as

$$\chi_\alpha(\mathsf{G}) \triangleq \prod_{e\in\alpha} \psi_e(\mathsf{G}) = \prod_{\{i,j\}\in\alpha} \frac{\mathsf{G}_{ij} - q}{\sqrt{q(1-q)}}. \tag{246}$$

and $\chi_\emptyset(\mathsf{G}) \equiv 1$, for each $\mathsf{G} \in \{0,1\}^{\binom{n}{2}}$. Note that any subset of edges $\alpha \subseteq \binom{[n]}{2}$ induces a subgraph $H_\alpha = (V(\alpha), \alpha)$ containing no isolated vertices. In fact, there is a one-to-one correspondence between subgraphs without isolated vertices and subsets of edges. We therefore identify each character χ_α with a subgraph H_α. Observe that χ_α is polynomial in the entries of G, with degree $|e(\alpha)|$, which, with slight abuse of notation, we also denote by $|\alpha|$. It is well known (and easy to verify) that the set $\{\chi_\alpha\}_{\alpha\subseteq\binom{[n]}{2}}$ forms an orthonormal basis for $L^2(\mathcal{H}_0)$.

Define the likelihood ratio $\mathcal{L}_n \triangleq \frac{\mathrm{d}\mathbb{P}_{\mathcal{H}_1}}{\mathrm{d}\mathbb{P}_{\mathcal{H}_0}}$. For $D \geq 0$, let $\mathcal{V}_{n,\leq D}$ be the subspace of polynomials of total degree at most D with respect to the edge coordinates, and let $\mathcal{P}_{\leq D}$ be the orthogonal projection onto $\mathcal{V}_{n,\leq D}$ in $L^2(\mathbb{P}_{\mathcal{H}_0})$. The degree-$D$ truncated likelihood ratio is

$$\mathcal{L}_{n,\leq D} \triangleq \mathcal{P}_{\leq D}\mathcal{L}_n. \tag{247}$$

By Parseval's identity we have

$$\|\mathcal{L}_{n,\leq D}\|^2_{\mathcal{H}_0} = \sum_{|\alpha|\leq D} \langle \mathcal{L}_n, \chi_\alpha\rangle^2_{\mathcal{H}_0} = \sum_{|\alpha|\leq D} \left(\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})]\right)^2 = 1 + \sum_{1\leq|\alpha|\leq D} \left(\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})]\right)^2. \tag{248}$$

Thus it suffices to control

$$\mathcal{E}_D \triangleq \sum_{1\leq|\alpha|\leq D} \left(\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})]\right)^2. \tag{249}$$

Fix $\alpha \subseteq \binom{[n]}{2}$, and let $V(\alpha) \subseteq [n]$ denote the set of vertices incident to edges in α, with $v(\alpha) \triangleq |V(\alpha)|$. For $i, j \in \mathcal{K}$ define the planted, centered and normalized variable

$$Z_{ij} \triangleq \frac{\sigma_{ij} - q}{\sqrt{q(1-q)}}. \tag{250}$$

We have the following result.

Lemma 23 (Coefficient formula) *For every $\alpha \subseteq \binom{[n]}{2}$,*

$$\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})] = \frac{(k)_{v(\alpha)}}{(n)_{v(\alpha)}} \mathbb{E}\left[\prod_{\{i,j\}\in\alpha} Z_{ij}\right], \tag{251}$$

where the expectation on the right is taken over i.i.d. latent vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_{v(\alpha)}\} \sim \mathsf{Unif}(\mathbb{S}^{d-1})$ and $\sigma_{ij} = \mathbb{1}\{\langle \mathbf{x}_i, \mathbf{x}_j\rangle \geq t_{q,d}\}$.

Proof Condition on $(\mathcal{K}, \{\mathbf{x}_i\}_{i\in\mathcal{K}})$. Under this conditioning, if an edge $e = \{i,j\}$ is not fully contained in $\mathcal{K}$, then $\mathsf{G}_{ij} \sim \mathsf{Bern}(q)$ and hence $\mathbb{E}[\psi_{ij}(\mathsf{G})|\mathcal{K},\mathbf{x}] = 0$. If $i,j \in \mathcal{K}$ then $\mathsf{G}_{ij} = \sigma_{ij}$ deterministically and $\psi_{ij}(\mathsf{G}) = Z_{ij}$. Therefore,

$$\mathbb{E}\left[\chi_\alpha(\mathsf{G})|\mathcal{K},\mathbf{x}\right] = \begin{cases} \prod_{e\in\alpha} Z_e, & \text{if} V(\alpha) \subseteq \mathcal{K}, \\ 0, & \text{otherwise.} \end{cases} \tag{252}$$

Taking expectation gives

$$\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})] = \mathbb{E}\left[\mathbb{1}\{V(\alpha)\subseteq\mathcal{K}\}\prod_{e\in\alpha} Z_e\right] = \mathbb{P}(V(\alpha)\subseteq\mathcal{K})\cdot\mathbb{E}\prod_{e\in\alpha} Z_e, \tag{253}$$

using independence of $\mathcal{K}$ and the latent vectors. Finally, for any fixed $S \subseteq [n]$ with $|S| = v(\alpha)$,

$$\mathbb{P}(S\subseteq\mathcal{K}) = \frac{\binom{n-v(\alpha)}{k-v(\alpha)}}{\binom{n}{k}} = \frac{(k)_{v(\alpha)}}{(n)_{v(\alpha)}}, \tag{254}$$

and by exchangeability the latent expectation depends only on the unlabeled structure of α. This yields the stated formula. ■

Combining (248) and Lemma 23 yields

$$\mathcal{E}_D = \sum_{1\le|\alpha|\le D} \left(\frac{(k)_{v(\alpha)}}{(n)_{v(\alpha)}}\right)^2 \left(\mathbb{E}\prod_{e\in\alpha} Z_e\right)^2. \tag{255}$$

We proceed by considering two regimes: $k \le n^{1/2-\epsilon}$, for arbitrary $\epsilon > 0$, and the complementary regime. We start with the former and show that $\mathcal{E}_D = O(1)$ uniformly over all d. To this end, since $\sigma_e \le 1$, we have

$$|Z_e| \le \frac{1-q}{\sqrt{q(1-q)}} = \sqrt{\frac{1-q}{q}}. \tag{256}$$

Consequently,

$$\mathcal{E}_D \le \sum_{1\le|\alpha|\le D} \left(\frac{(k)_{v(\alpha)}}{(n)_{v(\alpha)}}\right)^2 \left(\frac{1-q}{q}\right)^{|\alpha|}. \tag{257}$$

The right-hand side of (257) is exactly the same expression that arises in the low-degree analysis of the planted clique model. Indeed, this is not surprising: by bounding $|Z_e|$ by $\sqrt{\frac{1-q}{q}}$, we effectively dominate the geometric model by planted clique model. It is well-known (see, e.g., Hopkins (2018)) that for the planted clique model, the low-degree second moment is bounded, i.e., $\mathcal{E}_D = O(1)$, if and only if $k \le n^{1/2-\epsilon}$, uniformly over all d. This establishes the first part of Theorem 6.

Next, we consider the more challenging regime where $k = \Omega(\sqrt{n})$. a key property of the centered variables Z_{ij} is that products over graphs with *leaves* have zero expectation, as proved next.

Lemma 24 *Let $H = (V, E)$ be a finite simple graph on vertex set V and edge set E. Let $\{\mathbf{x}_v\}_{v \in V}$ be i.i.d.* $\mathsf{Unif}(\mathbb{S}^{d-1})$ *and define $\sigma_{uv} = \mathbb{1}\{\langle \mathbf{x}_u, \mathbf{x}_v \rangle \geq t_{q,d}\}$ and $Z_{uv} = (\sigma_{uv} - q)/\sqrt{q(1-q)}$. Then if H has a leaf vertex, i.e., there exists $u \in V$ with $\deg_H(u) = 1$, we have*

$$\mathbb{E}\left[\prod_{\{i,j\} \in E} Z_{ij}\right] = 0. \tag{258}$$

Proof [Proof of Lemma 24] Let u be a leaf and let v be its unique neighbor, so that $e = \{u, v\}$ is the unique edge incident to u. Condition on all latent vectors except $\mathbf{x}_u$. Then the product factors as

$$\prod_{f \in E} Z_f = Z_{uv} \cdot \prod_{f \in E \setminus \{e\}} Z_f, \tag{259}$$

where $\prod_{f \in E \setminus \{e\}} Z_f$ is measurable with respect to $\{\mathbf{x}_w\}_{w \in V \setminus \{u\}}$. Therefore,

$$\mathbb{E}\left[\prod_{f \in E} Z_f\right] = \mathbb{E}\left[\left(\prod_{f \in E \setminus \{e\}} Z_f\right) \cdot \mathbb{E}\left[Z_{uv} | \{\mathbf{x}_w\}_{w \in V \setminus \{u\}}\right]\right]. \tag{260}$$

By rotational symmetry, conditional on $\mathbf{x}_v$, the random vector $\mathbf{x}_u$ is uniform on $\mathbb{S}^{d-1}$, and the event $\{\langle \mathbf{x}_u, \mathbf{x}_v \rangle \geq t_{q,d}\}$ has probability q. Hence

$$\mathbb{E}[\sigma_{uv} | \mathbf{x}_v] = q \quad \Longrightarrow \quad \mathbb{E}[Z_{uv} | \mathbf{x}_v] = \frac{q - q}{\sqrt{q(1-q)}} = 0. \tag{261}$$

Thus $\mathbb{E}[Z_{uv} | \{\mathbf{x}_w\}_{w \in V \setminus \{u\}}] = 0$, and the whole expectation is zero. ■

Accordingly, only edge sets α whose associated graph has minimum degree at least two contribute to $\mathcal{E}_D$. The challenge is to bound the remaining contributions. Fortunately enough, in a recent paper Bangachev and Bresler (2024), the following result was proved.

Lemma 25 ((Bangachev and Bresler, 2024, Thm. 1.1 & Prop. 1.2)) *There exist constants $A, B > 0$ depending only on q such that for every connected graph $H = (V, E)$ with $v \triangleq |V| \geq 2$ and $m \triangleq |E| \geq 1$, we have*

$$\left|\mathbb{E}\left[\prod_{e \in E(H)} Z_e\right]\right| \leq A^m \left(B \frac{v \cdot m \cdot (\log d)^{3/2}}{\sqrt{d}}\right)^{\lceil \frac{v-1}{2} \rceil}. \tag{262}$$

Now, denote the connected component of α by $\alpha_1, \ldots, \alpha_c$. Since these components are supported on disjoint vertex sets, then the latent vectors on different components are independent, and so

$$\mathbb{E} \prod_{e \in \alpha} Z_e = \prod_{r=1}^{c} \mathbb{E} \prod_{e \in \alpha_r} Z_e. \tag{263}$$

Applying Lemma 25 to each component gives

$$\left|\mathbb{E}\prod_{e\in E(H)} Z_e\right| \le A^{|\alpha|}\left(B\frac{|\alpha|v(\alpha)(\log d)^{3/2}}{\sqrt{d}}\right)^{\sum_{r=1}^{c}\lceil (v(\alpha_r)-1)/2\rceil}. \tag{264}$$

Note that $\sum_{r=1}^{c}\lceil (v(\alpha_r)-1)/2\rceil \ge \frac{v(\alpha)-c}{2}$. Because we assume that α has minimum degree at least two, then each connected component has at least three vertices, so $c \le v(\alpha)/3$ and therefore

$$\left|\mathbb{E}\prod_{e\in E(H)} Z_e\right| \le A^{|\alpha|}\left(B\frac{|\alpha|v(\alpha)(\log d)^{3/2}}{\sqrt{d}}\right)^{v(\alpha)/3}. \tag{265}$$

We are now in a position to bound $\mathcal{E}_D$. Define

$$\beta_n \triangleq \frac{k^6}{n^3 d}(\log d)^3. \tag{266}$$

Suppose that there exists $\varepsilon > 0$ such that for all sufficiently large n,

$$\beta_n = \frac{k^6}{n^3 d}(\log d)^3 \le n^{-\varepsilon}. \tag{267}$$

Let $D = D(n)$ satisfy $D \le c\log n$ for a sufficiently small constant $c > 0$ (depending only on p and ε). Fix integers $v \ge 3$ and $m \ge 1$. Consider any α with $v(\alpha) = v$ and $|\alpha| = m$ and such that $H_\alpha = (V(\alpha), \alpha)$ has minimum degree at least 2. First, Stirling's approximation gives

$$\frac{(k)_v}{(n)_v} \le \left(\frac{k}{n}\right)^v. \tag{268}$$

Then, (265) gives

$$\left(\frac{(k)_v}{(n)_v}\right)^2\left(\mathbb{E}\prod_{e\in\alpha} Z_e\right)^2 \le \left(\frac{k}{n}\right)^{2v} A^{2m}\left(B^2\frac{v^2m^2(\log d)^3}{d}\right)^{v/3}. \tag{269}$$

Next, we bound the number of edge-sets α with $v(\alpha) = v$ and $|\alpha| = m$. Choose the v vertices in $\binom{n}{v}$ ways, and then choose the m edges among the $\binom{v}{2}$ possible edges on those vertices in $\binom{\binom{v}{2}}{m}$ ways. Hence

$$\left|\left\{\alpha\subseteq\binom{[n]}{2} : v(\alpha) = v, |\alpha| = m\right\}\right| \le \binom{n}{v}\binom{\binom{v}{2}}{m}. \tag{270}$$

Note that

$$\binom{n}{v} \le \left(\frac{en}{v}\right)^v, \quad \binom{\binom{v}{2}}{m} \le \left(\frac{e\binom{v}{2}}{m}\right)^m \le \left(\frac{ev^2}{2m}\right)^m. \tag{271}$$

Because we assume that H_α has minimum degree at least two, then $2m \ge 2v$ and hence $m \ge v$. Therefore $v^2/m \le v$ and thus there exists an absolute constant $C_0 > 0$ such that

$$\binom{\binom{v}{2}}{m} \le (C_0 v)^m. \tag{272}$$

Combining these gives in our case

$$\left|\{\alpha \subseteq \binom{[n]}{2} : v(\alpha) = v, |\alpha| = m\}\right| \leq \left(\frac{en}{v}\right)^v (C_0 v)^m. \tag{273}$$

Let us sum (269) over all α with fixed (v, m):

$$\sum_{\substack{\alpha:\, v(\alpha)=v \\ |\alpha|=m}} \left(\frac{(k)_v}{(n)_v}\right)^2 \left(\mathbb{E}\prod_{e\in\alpha} Z_e\right)^2 \leq \left(\frac{en}{v}\right)^v (C_0 v)^m \left(\frac{k}{n}\right)^{2v} A^{2m} \left(B^2 \frac{v^2 m^2 (\log d)^3}{d}\right)^{v/3} \tag{274}$$

$$\leq \left(\frac{k^2}{n}\right)^v \left(B^2 \frac{v^2 m^2 (\log d)^3}{d}\right)^{v/3} (C_1 v)^m, \tag{275}$$

where $C_1 = C_0 A^2$ and we used the fact that $v \geq 3$. Note that

$$\left(\frac{k^2}{n}\right)^v \left(\frac{1}{d}\right)^{v/3} (\log d)^v = \left(\frac{k^6}{n^3 d}(\log d)^3\right)^{v/3} = \beta_n^{v/3}. \tag{276}$$

Therefore,

$$\sum_{\substack{\alpha:\, v(\alpha)=v \\ |\alpha|=m}} \left(\frac{(k)_v}{(n)_v}\right)^2 \left(\mathbb{E}\prod_{e\in\alpha} Z_e\right)^2 \leq \left(\frac{k^6}{n^3 d}(\log d)^3\right)^{v/3} \left(B^2 v^2 m^2\right)^{v/3} (C_1 v)^m \tag{277}$$

$$= \beta_n^{v/3} (B^2 v^2 m^2)^{v/3} (C_1 v)^m. \tag{278}$$

Finally, we sum over m and v. Since the minimum degree is two, we must have $m \geq v$, and we also have $m \leq D$ because we only consider $|\alpha| \leq D$. Thus

$$\mathcal{E}_D \leq \sum_{v=3}^{D} \sum_{m=v}^{D} \beta_n^{v/3} (B^2 v^2 m^2)^{v/3} (C_1 v)^m \leq \sum_{v=3}^{D} \beta_n^{v/3} (B^2 v^2 D^2)^{v/3} \sum_{m=v}^{D} (C_1 v)^m. \tag{279}$$

Since $v \leq D$, we have $(C_1 v)^m \leq (C_1 D)^m$, and hence

$$\sum_{m=v}^{D} (C_1 v)^m \leq \sum_{m=0}^{D} (C_1 D)^m \leq (D+1)(C_1 D)^D \leq (C_2 D)^D, \tag{280}$$

for a constant $C_2 > 0$ depending only on C_1. Therefore,

$$\mathcal{E}_D \leq (C_2 D)^D \sum_{v=3}^{D} \beta_n^{v/3} (B^2 v^2 D^2)^{v/3} \leq (C_3 D^{2.5})^D \cdot \frac{\beta_n}{1 - \beta_n^{1/3}} \leq 2 (C_3 D^{2.5})^D \beta_n, \tag{281}$$

for a constant $C_3 > 0$, and for all sufficiently large n since $\beta_n \to 0$. Assume now (267), i.e., $\beta_n \leq n^{-\varepsilon}$ for all large n. Let $D \leq c \log n / \log\log n$ where $c > 0$ is chosen sufficiently small (as specified below). Then

$$(C_3 D^{2.5})^D \beta_n \leq \exp\left(3D \log D\right) \cdot n^{-\varepsilon} \leq \exp\left(3c \frac{\log n}{\log\log n} \cdot \log\log n\right) \cdot n^{-\varepsilon} = n^{3c-\varepsilon}. \tag{282}$$

Thus for a sufficiently small absolute constant c such that $c < \varepsilon/3$, we have $(C_3 D^{2.5})^D \beta_n = o(1)$, and hence $\mathcal{E}_D = o(1)$. Therefore

$$\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0}^2 = 1 + \mathcal{E}_D \leq 1 + o(1), \tag{283}$$

which implies $\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0} = O(1)$.

Low-degree computational upper bound. Finally we prove that when the dimension is below the signed-triangle threshold, the low-degree norm diverges. The proof uses only the first nonzero coefficients (degree-1 vanishes, and graphs with leaves vanish), and in particular it suffices to lower bound the triangle contribution. Specifically, let Tri denote the set of edge sets $\alpha \subseteq \binom{[n]}{2}$ that form a triangle on three vertices, i.e. $\alpha = \{\{i,j\},\{i,\ell\},\{j,\ell\}\}$, for some distinct i, j, ℓ.

For any $D \geq 3$, by (248),

$$\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0}^2 \geq 1 + \sum_{\alpha \in \mathsf{Tri}} \left(\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})]\right)^2. \tag{284}$$

Thus it suffices to show that the triangle sum diverges. Fix $\alpha = \{\{1,2\},\{1,3\},\{2,3\}\}$. Then $v(\alpha) = 3$, so by Lemma 23,

$$\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})] = \frac{(k)_3}{(n)_3}\mathbb{E}[Z_{12}Z_{13}Z_{23}]. \tag{285}$$

By exchangeability, the same value holds for every labeled triangle $\alpha \in \mathsf{Tri}$. Since $|\mathsf{Tri}| = \binom{n}{3}$, we obtain

$$\sum_{\alpha \in \mathsf{Tri}} \left(\mathbb{E}_{\mathcal{H}_1}[\chi_\alpha(\mathsf{G})]\right)^2 = \binom{n}{3}\left(\frac{(k)_3}{(n)_3}\right)^2 \left(\mathbb{E}[Z_{12}Z_{13}Z_{23}]\right)^2. \tag{286}$$

Therefore, the proof reduces to a quantitative lower bound on the *signed triangle moment* $\mathbb{E}[Z_{12}Z_{13}Z_{23}]$ for the spherical random geometric graph. It is rather classical that for every $0 < q < 1$, there exists a constant $c_q > 0$ such that for all n and d we have (see, e.g., (Bubeck et al., 2016, Lemma 3))

$$\mathbb{E}[Z_{12}Z_{13}Z_{23}] \geq \frac{c_q}{\sqrt{d}}. \tag{287}$$

Fix any $D \geq 3$. By (284), (286), and (287),

$$\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0}^2 \geq 1 + \binom{n}{3}\left(\frac{(k)_3}{(n)_3}\right)^2 \frac{c_q^2}{d}. \tag{288}$$

Note that $\binom{n}{3} \geq n^3/6$ for all $n \geq 3$, and

$$\frac{(k)_3}{(n)_3} = \frac{k(k-1)(k-2)}{n(n-1)(n-2)} = (1+o(1))\left(\frac{k}{n}\right)^3, \tag{289}$$

whenever $k, n \to \infty$. Hence

$$\binom{n}{3}\left(\frac{(k)_3}{(n)_3}\right)^2 = \frac{(1+o(1))}{6}\frac{k^6}{n^3}. \tag{290}$$

Combining,

$$\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0}^2 \geq 1 + \frac{(1+o(1))c_q^2}{6}\frac{k^6}{n^3 d}. \tag{291}$$

Thus, if $d \ll k^6/n^3$, then the right-hand side diverges to ∞, implying $\|\mathcal{L}_{n,\leq D}\|_{\mathcal{H}_0} = \omega(1)$.

Appendix F. Proof of Lemma 10

Proof [Proof of Lemma 10] Let μ be a measure such that $\mathbb{P},\mathbb{R},\mathbb{Q}\ll\mu$, and write $p\triangleq\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mu}$, $r\triangleq\frac{\mathrm{d}\mathbb{R}}{\mathrm{d}\mu}$, and $q\triangleq\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mu}$. Then, $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}=\frac{p}{q}$, $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{R}}=\frac{p}{r}$, and $\frac{\mathrm{d}\mathbb{R}}{\mathrm{d}\mathbb{Q}}=\frac{r}{q}$. Hence, by definition

$$1+D_m(\mathbb{P}\|\mathbb{Q})=\int\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\right)^m\mathrm{d}\mathbb{Q} \tag{292}$$

$$=\int\left(\frac{p}{q}\right)^m q\mathrm{d}\mu \tag{293}$$

$$=\int p^m q^{1-m}\mathrm{d}\mu \tag{294}$$

$$=\int p^m r^{1-m}\left(\frac{r}{q}\right)^{m-1}\mathrm{d}\mu \tag{295}$$

$$=\int\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{R}}\right)^m\left(\frac{\mathrm{d}\mathbb{R}}{\mathrm{d}\mathbb{Q}}\right)^{m-1}\mathrm{d}\mathbb{R}. \tag{296}$$

Applying Cauchy–Schwarz we get

$$1+D_m(P\|Q)\leq\sqrt{\int\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{R}}\right)^{2m}\mathrm{d}\mathbb{R}}\sqrt{\int\left(\frac{\mathrm{d}\mathbb{R}}{\mathrm{d}\mathbb{Q}}\right)^{2m-2}\mathrm{d}\mathbb{R}}. \tag{297}$$

Now, we note that

$$\int\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{R}}\right)^{2m}\mathrm{d}\mathbb{R}=1+D_{2m}(\mathbb{P}\|\mathbb{R})\leq 1+D_{2m}(\mathbb{P}\|\mathbb{Q}), \tag{298}$$

and

$$\int\left(\frac{\mathrm{d}\mathbb{R}}{\mathrm{d}\mathbb{Q}}\right)^{2m-2}\mathrm{d}\mathbb{R}=\int\left(\frac{\mathrm{d}\mathbb{R}}{\mathrm{d}\mathbb{Q}}\right)^{2m-1}\mathrm{d}\mathbb{Q}=1+D_{2m-1}(\mathbb{R}\|\mathbb{Q}). \tag{299}$$

Thus

$$1+D_m(\mathbb{P}\|\mathbb{Q})\leq\sqrt{1+D_{2m}(\mathbb{P}\|\mathbb{Q})}\sqrt{1+D_{2m-1}(\mathbb{R}\|\mathbb{Q})}, \tag{300}$$

as claimed. ■

Appendix G. Densest subgraph test

In this appendix, we consider the case where the edge probability inside the planted set is taken to be $p\neq q$, and propose an optimal detection algorithm. Specifically, consider the following statistic:

$$\mathsf{T}_{\mathsf{dense}}(\mathsf{G}_n)\triangleq\max_{\mathcal{S}\subset[n]:|\mathcal{S}|=k}\sum_{i,j\in\mathcal{S}:i<j}\mathbf{A}_{i,j}. \tag{301}$$

This statistic scans over all k-vertex subgraphs and selects the densest one in terms of edge count. Define the following detection algorithm:

$$\mathcal{A}_{\mathsf{dense}}(\mathsf{G}_n)\triangleq\mathbb{1}\left\{\mathsf{T}_{\mathsf{dense}}(\mathsf{G}_n)\geq\tau_{\mathsf{dense}}\right\}, \tag{302}$$

where $\tau_{\mathsf{dense}}\in\mathbb{R}_+$ is specified below. We have the following result.

Theorem 26 (Upper bounds) *Consider the detection problem in (1), and the detection algorithm in (302), with* $\tau_{\mathsf{dense}} = \binom{k}{2}\frac{p+q}{2}$. *If* $\frac{(p-q)^2}{q(1-q)} = \omega\left(\frac{\log n}{k}\right)$, *then* $\mathsf{R}(\mathcal{A}_{\mathsf{dense}}) \to 0$, *as* $k, n \to \infty$.

When $p \neq q$ is fixed, Theorem 26 shows that strong detection is achievable whenever $k = \omega(\log n)$.
Proof [Proof of Theorem 26] Recall the densest subgraph test defined in (301) and (302), and let us analyze its Type-I and Type-II error probabilities. Set $\tau_{\mathsf{dense}} = \binom{k}{2}\frac{p+q}{2}$. We begin with the analysis of the Type-I error probability. For any subset $\mathcal{S} \subset [n]$ with $|\mathcal{S}| = k$, define $e(\mathcal{S}) \triangleq \sum_{i,j\in\mathcal{S}:i<j} \mathbf{A}_{i,j}$. Under the null hypothesis $\mathcal{H}_0$, we have $e(\mathcal{S}) \sim \mathsf{Binomial}\left(\binom{k}{2}, q\right)$. Consequently, by the union bound together with Bernstein's inequality, we obtain

$$\mathbb{P}_{\mathcal{H}_0}\left[\mathcal{A}_{\mathsf{dense}}(\mathsf{G}_n) = 1\right] \leq \binom{n}{k} \cdot \mathbb{P}\left[\mathsf{Binomial}\left(\binom{k}{2}, q\right) \geq \tau_{\mathsf{dense}}\right] \tag{303}$$

$$\leq \binom{n}{k} \exp\left(-\frac{\binom{k}{2}^2 (p-q)^2/4}{2\binom{k}{2}q + \binom{k}{2}(p-q)/3}\right) \tag{304}$$

$$\leq \exp\left[k \log n - \Omega\left(k^2 \frac{(p-q)^2}{q(1-q)}\right)\right], \tag{305}$$

where the last inequality follows from Stirling's approximation. We now turn to the analysis of the Type-II error probability. Denoting by $\mathcal{K}$ the underlying planted vertex set, an application of Chebyshev's inequality yields

$$\mathbb{P}_{\mathcal{H}_1}\left[\mathcal{A}_{\mathsf{dense}}(\mathsf{G}_n) = 0\right] \leq \mathbb{P}_{\mathcal{H}_1}\left[e(\mathcal{K}) < \tau_{\mathsf{dense}}\right] \tag{306}$$

$$\leq \frac{\mathsf{Var}_{\mathcal{H}_1}(e(\mathcal{K}))}{\binom{k}{2}^2 \frac{(p-q)^2}{4}}. \tag{307}$$

Now,

$$\mathsf{Var}_{\mathcal{H}_1}(e(\mathcal{K})) = \mathsf{Var}_{\mathcal{H}_1}\left(\sum_{i,j\in\mathcal{K}:i<j} \mathbf{A}_{i,j}\right) \tag{308}$$

$$= \sum_{i,j\in\mathcal{K}:i<j} \mathsf{Var}_{\mathcal{H}_1}(\mathbf{A}_{i,j}) + \sum_{(i,j)\neq(i',j')} \mathsf{cov}_{\mathcal{H}_1}(\mathbf{A}_{i,j}, \mathbf{A}_{i',j'}). \tag{309}$$

By definition, for any $i, j \in \mathcal{K}$ we have $\mathsf{Var}_{\mathcal{H}_1}(\mathbf{A}_{i,j}) = p(1-p)$. We claim that for any $i, i', j, j' \in \mathcal{K}$ with $i < j$ and $i' < j'$, such that $(i, j) \neq (i', j')$, it holds that

$$\mathsf{Cov}_{\mathcal{H}_1}(\mathbf{A}_{i,j}, \mathbf{A}_{i',j'}) = 0. \tag{310}$$

There are two cases to consider. First, if the indices i, i', j, j' are all distinct, then $\mathbf{A}_{i,j}$ and $\mathbf{A}_{i',j'}$ are independent, and hence their covariance is zero. The second, less immediate case is when the two edges share a vertex, for instance when $i = i'$ and $j \neq j'$. In this case, we have

$$\mathsf{cov}_{\mathcal{H}_1}(\mathbf{A}_{i,j}, \mathbf{A}_{i,j'}) = \mathbb{E}_{\mathcal{H}_1}\left[\mathbf{A}_{i,j} \cdot \mathbf{A}_{i,j'}\right] - p^2 \tag{311}$$

$$= \mathbb{P}_{\mathcal{H}_1}\left[\langle \mathbf{x}_i, \mathbf{x}_j \rangle \geq t_{q,d}, \langle \mathbf{x}_i, \mathbf{x}_{j'} \rangle \geq t_{q,d}\right] - p^2. \tag{312}$$

By rotation invariance on the sphere, we can fix $\mathbf{x}_i = \mathbf{e}_1$, where $\mathbf{e}_1$ is the d-dimensional unit vector $\mathbf{e}_1 = (1, \ldots, 0)$. Then

$$\mathbb{P}_{\mathcal{H}_1}\left[\langle \mathbf{e}_1, \mathbf{x}_j\rangle \geq t_{q,d}, \langle \mathbf{e}_1, \mathbf{x}_{j'}\rangle \geq t_{q,d}\right] = \mathbb{P}_{\mathcal{H}_1}\left[\langle \mathbf{e}_1, \mathbf{x}_j\rangle \geq t_{q,d}\right] \cdot \mathbb{P}_{\mathcal{H}_1}\left[\langle \mathbf{e}_1, \mathbf{x}_{j'}\rangle \geq t_{q,d}\right] \tag{313}$$

$$= p^2. \tag{314}$$

Combined with (312) we obtain that $\mathsf{cov}_{\mathcal{H}_1}(\mathbf{A}_{i,j}, \mathbf{A}_{i,j'}) = 0$. Therefore

$$\mathsf{Var}_{\mathcal{H}_1}(e(\mathcal{K})) = \binom{k}{2} p(1-p), \tag{315}$$

and by (307), we have

$$\mathbb{P}_{\mathcal{H}_1}\left[\mathcal{A}_{\mathsf{dense}}(\mathsf{G}_n) = 0\right] \leq \frac{\binom{k}{2} p(1-p)}{\binom{k}{2}^2 \frac{(p-q)^2}{4}} \tag{316}$$

$$= \mathsf{C}\frac{p(1-p)}{k^2(p-q)^2}, \tag{317}$$

for some constant $\mathsf{C} > 0$. Combining (305) and (317), we conclude that $\mathsf{R}(\mathcal{A}_{\mathsf{dense}}) \to 0$ provided that

$$\frac{k^2(p-q)^2}{p(1-p)} = \omega(1) \quad \text{and} \quad \frac{(p-q)^2}{q(1-q)} = \omega\left(\frac{\log n}{k}\right). \tag{318}$$

Since the latter condition implies the former, it follows that $\frac{(p-q)^2}{q(1-q)} = \omega\left(\frac{\log n}{k}\right)$ is the governing requirement. ■

Appendix H. Conclusion and Outlook

This work provides a sharp characterization of the statistical and computational limits of detecting a planted high-dimensional geometric subgraph in a random graph. While our results focus on a specific geometric model, they open several directions for future research.

- It would be natural to extend our analysis beyond the Euclidean sphere to other probabilistic metric spaces. For example, one could consider latent positions drawn from more general manifolds, anisotropic distributions, or non-Euclidean geometries, and investigate how curvature, intrinsic dimension, or metric structure affect detectability and computational hardness.

- Our model, as is standard in the literature, assumes that the latent feature vectors are completely unobserved. In many practical settings, however, one has partial access to side information about the latent geometry. For instance, one might observe a quantized or noisy version of the feature vectors, or only a subset of their coordinates. Understanding how such partial observations alter the statistical thresholds and computational barriers is an interesting open problem. In particular, it is unclear whether even coarse or noisy geometric side information can significantly reduce the sample complexity or break the computational barriers identified here.

- While we focus on detection, the recovery problem remain largely unexplored in the geometric setting. This includes identifying the planted vertex set, estimating latent positions, or designing adaptive querying procedures that exploit sequential access to the graph. These problems are likely to exhibit their own statistical–computational tradeoffs, distinct from those governing detection.

- Although we primarily study the geometry-only regime in which the null and alternative hypotheses have matching edge marginals, it is also of interest to understand settings where the edge densities differ under the null and alternative. In the dense regime with fixed probabilities, this reduces to classical planted dense subgraph detection, where geometry plays a limited role. However, when edge probabilities depend on the graph size, such as in sparse or vanishing density regimes, the interaction between density contrast and geometry becomes substantially more intricate. Analyzing these regimes, where both sparsity and latent geometry are present, poses significant challenges and may reveal new phase transitions beyond those identified in this work.