Proceedings of Machine Learning Research vol 247:1–71, 2024 37th Annual Conference on Learning Theory

Detection of L_∞ Geometry in Random Geometric Graphs: Suboptimality of Triangles and Cluster Expansion

Kiril Bangachev KIRILB@MIT.EDU
Guy Bresler GUY@MIT.EDU
MIT EECS

Editors: Shipra Agrawal and Aaron Roth

Abstract

In this paper we study the random geometric graph $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ with L_q distance where each vertex is sampled uniformly from the d-dimensional torus and where the connection radius is chosen so that the marginal edge probability is p. In addition to results addressing other questions, we make progress on determining when it is possible to distinguish $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ from the Erdős-Rényi graph $\mathsf{G}(n, p)$.

Our strongest result is in the setting $q = \infty$, in which case $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$ is the AND of d 1-dimensional random geometric graphs. We derive a formula similar to the *cluster-expansion* from statistical physics, capturing the compatibility of subgraphs from each of the d 1-dimensional copies, and use it to bound the signed expectations of small subgraphs. We show that counting signed 4-cycles is optimal among all low-degree tests, succeeding with high probability if and only if $d = \tilde{o}(np)$. In contrast, the signed triangle test is suboptimal and only succeeds when $d = \tilde{o}((np)^{3/4})$. Our result stands in sharp contrast to the existing literature on random geometric graphs (mostly focused on L_2 geometry) where the signed triangle statistic is optimal.

Keywords: High-Dimensional Random Geometric Graphs; Cluster Expansion.

1. Introduction

Networks arising in the sciences are often modeled as latent space graphs. Each node in a network has a latent feature vector and the probability of connection between two nodes is a function of the two feature vectors. One instance is the case of (random) geometric graphs in which each feature vector is a (random) element of a metric space and the connection probability is determined by the distance between the two vectors. Applications include protein-protein interactions and viral spread in the biological sciences Higham et al. (2008); Preciado and Jadbabaie (2009), wireless networks and motion planning in engineering Haenggi et al. (2009); Solovey et al. (2018), consensus dynamics and citation networks in the social sciences Xie et al. (2016); Estrada and Sheerin (2016).

Formally, a random geometric graph is defined as follows.

Definition 1 (Random Geometric Graph) *Given are a metric space (Ω, μ), a distribution $\mathcal{D}$ over Ω, and connection function $\sigma : \Omega \times \Omega \longrightarrow [0, 1]$ such that $\sigma(\mathbf{x}, \mathbf{y})$ only depends on $\mu(\mathbf{x}, \mathbf{y})$. Let $\mathbf{E}[\sigma(\mathbf{x}, \mathbf{y})] = p$. Then, $\mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p)$ is the following distribution over n-vertex graphs:*

$$\mathbf{P}[\mathbf{G} = A] = \mathbf{E}_{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n \overset{\text{i.i.d.}}{\sim} \mathcal{D}} \left[\prod_{1 \leq i < j \leq n} \sigma(\mathbf{x}^i, \mathbf{x}^j)^{A_{i,j}} (1 - \sigma(\mathbf{x}^i, \mathbf{x}^j))^{1 - A_{i,j}} \right].$$

When σ is monotone in μ, we say that $\mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p)$ is a monotone random geometric graph.

In words, each node i has an associated independent latent vector $\mathbf{x}^i$ in Ω distributed according to $\mathcal{D}$. Conditioned on $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n$, each pair of nodes (i,j) independently forms an edge with probability $\sigma(\mathbf{x}^i, \mathbf{x}^j)$. We focus on the monotone increasing case which has the natural interpretation that closer nodes are more likely to be adjacent. In practice, the feature vectors are oftentimes not fully available. In this work, we assume that the vectors are fully hidden.

Associated to random geometric graphs with latent vectors are a wide range of statistical and computational tasks such as: 1) *Clustering and Embedding* of the nodes in a way that captures the distances between latent vectors Li and Schramm (2023); O'Connor et al. (2020); Ma et al. (2020); 2) *Estimating the dimension* of the underlying space Ω in the case when dimension is naturally defined such as $\Omega \in \{\mathbb{S}^{d-1}, \mathbb{T}^d, \{\pm 1\}^d\}$ Bubeck et al. (2014); Friedrich et al. (2023b); 3) *Testing* whether the network has a geometric structure against a "pure noise" (i.e., Erdős-Rényi)[1] null hypothesis Devroye et al. (2011); Bubeck et al. (2014); Brennan et al. (2020); Liu and Rácz (2023a); Liu et al. (2022); Brennan et al. (2026); Bangachev and Bresler (2023, 2024) and others.

The current work is mostly focused on the hypothesis-testing question which can be formalised as follows (e.g. Bangachev and Bresler (2023)): Given G, decide between

$$H_0 : G \sim \mathsf{G}(n,p) \quad \text{versus} \quad H_1 : G \sim \mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p). \tag{P1}$$

Associated to these hypotheses are (at least) two different questions:

1. *Statistical:* When is there a consistent test? We aim to characterize the parameter regimes in which the total variation between the two distributions tends to zero or instead to one.
2. *Computational:* When is there a computationally efficient test? In particular, when does there exist a polynomial-time test solving (**P1**) with high probability?

Question (**P1**) has received significant attention in recent years in the case when (Ω, μ) captures an L_2 geometry. Concretely, μ is the induced L_2 distance from $\mathbb{R}^d$ and $(\Omega, \mathcal{D})$ is either the unit sphere $\mathbb{S}^{d-1}$ with its uniform (Haar) measure Devroye et al. (2011); Bubeck et al. (2014); Brennan et al. (2020); Liu et al. (2022); Bangachev and Bresler (2024) or Euclidean space $\mathbb{R}^d$ with a Gaussian measure Liu and Rácz (2023a,b); Brennan et al. (2026). In all of the above *monotone* models, the conjectured information-theoretically optimal statistic is the signed triangle statistic, computable in polynomial time. For a summary of results on L_2 models, we refer the reader to Duchemin and de Castro (2022); Bangachev and Bresler (2023). Most relevant to our work is the case when $\Omega = \mathbb{S}^{d-1}, \mathcal{D} = \mathsf{Unif}$, and $\sigma(\mathbf{x}, \mathbf{y}) = \mathbb{1}[\langle \mathbf{x}, \mathbf{y}\rangle \geq \rho_p^d]$, where ρ_p^d is chosen so that the expected density is p. The state of the art results are as follows. When $d = \tilde{O}(n^3p^3)$, by counting signed triangles one can distinguish between the RGG model and $\mathsf{G}(n,p)$ with high probability Bubeck et al. (2014); Liu et al. (2022) and this regime is optimal with respect to low-degree tests Bangachev and Bresler (2024). There is a matching information-theoretic lower bound when $p = \Theta(n^{-1})$ Liu et al. (2022) and when $p = \tilde{\Theta}(1)$ Bubeck et al. (2014). The case $n^{-1} \ll p \ll \tilde{\Theta}(1)$ remains open and the best known lower bound is $d = \tilde{\Omega}(n^3p^2)$ Liu et al. (2022).

In Bubeck et al. (2014), the authors also show that the signed triangle statistic is optimal for *exact* recovery of the dimension in the model $\Omega = \mathbb{S}^{d-1}, \mathcal{D} = \mathsf{Unif}$ and $\sigma(\mathbf{x}^i, \mathbf{x}^j) = \mathbb{1}[\langle \mathbf{x}^i, \mathbf{x}^j\rangle \geq 0]$. The (signed) triangle statistic in monotone models is intuitive as it captures the axiomatic triangle inequality: If x and y are close and y and z are close, then so are x and z Bubeck et al. (2014).

1. In the Erdős-Rényi distribution $\mathsf{G}(n,p)$, each of the $\binom{n}{2}$ edges appears independently with probability p. As there is no underlying dependence structure, this is a natural null model.

These results and intuition have led to the conventional wisdom that (signed) triangles are most informative, at least in monotone random geometric graphs.[2] Subsequent works in very different geometries have also used triangle-based statistics, for example to estimate the hidden dimension Almagro and M. & Serrano (2022); Friedrich et al. (2023b).

In this paper, we go against this conventional wisdom and demonstrate that the (signed) triangle statistic can be *suboptimal*. More concretely, we study the hypothesis testing problem under L_q geometry for $q \in [1, \infty) \cup \{\infty\}$ and show that different values of q yield both quantitatively and qualitatively different behaviours (see Figures 1 and 2). In particular, when $q = \infty$, triangle-based tests are always suboptimal. The suboptimality of triangle-based estimators extends to the task of dimension estimation as well. We use the (unweighted version of the) model of Friedrich et al. (2023a,b) with L_q geometry over $\mathbb{T}^d$. The model can be viewed as a high-dimensional analogue of the planted dense cycle model, which has also been of recent interest to the combinatorial statistics community Mao et al. (2023, 2024).

Definition 2 (L_q-Hard Thresholds Model on $\mathbb{T}^d$) *Consider the torus $\mathbb{T}^d \cong (2\mathbb{S}^1)^{\times d}$, which is a product of d circles of circumference 2.*[3] *Let Unif be the uniform (Haar) measure over $\mathbb{T}^d$. For $x_1, y_1 \in 2\mathbb{S}^1$, denote by $|x_1 - y_1|_C \in [0,1]$ the circular distance, i.e. the length of the shorter arc connecting x_1 and y_1. For $1 \le q < +\infty$, introduce the L_q distance on $\mathbb{T}^d$ given by*

$$\|\mathbf{x} - \mathbf{y}\|_q := \Big(\sum_{i=1}^d |x_i - y_i|_C^q\Big)^{1/q}.$$

Also, $\|\mathbf{x} - \mathbf{y}\|_\infty := \lim_{q \longrightarrow +\infty} \|\mathbf{x} - \mathbf{y}\|_q = \max_i |x_i - y_i|_C$. Let $1 \ge p \ge 0, \tau_p^q \ge 0$ be such that $\mathbf{E}_{\mathbf{x},\mathbf{y} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}(\mathbb{T}^d)}\Big[\mathbb{1}[\|\mathbf{x} - \mathbf{y}\|_q \le \tau_p^q]\Big] = p$ and $\sigma_p^q(\mathbf{x}, \mathbf{y}) := \mathbb{1}[\|\mathbf{x} - \mathbf{y}\|_q \le \tau_p^q]$. Then, $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ is the random geometric graph over $\mathbb{T}^d$ with expected density p in which two vertices are adjacent whenever the L_q distance between their latent vectors is at most τ_p^q.

To the best of our knowledge, the work of Friedrich et al. (2023a) is the first to explore (**P1**) for random geometric graphs in non-L_2 geometries. They showed that in the L_q model of Definition 2 (as well as for an inhomogeneous generalization of it) for fixed p, n,

$$\lim_{d \to \infty} \mathsf{TV}\Big(\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p), \mathsf{G}(n,p)\Big) = 0.$$

Their approach, based on a multidimensional Berry-Esseen theorem and mimicking Devroye et al. (2011), however, only yields TV distance of order $o(1)$ when $d = \exp(\Omega(n^2))$. Improving this bound is posed as an open problem, which is also one of the main motivations of the current work.

Friedrich et al. (2023a) also estimate the probability with which a given set of edges appears in $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$ (and its inhomogeneous generalizations). They show that for edge subsets $\mathcal{A}$ of constant size and $d = \omega(\log^2 n)$, the probability that all edges of $\mathcal{A}$ appear in $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$ is $p^{|\mathcal{A}|}(1 + o(1))$. This also allows the authors to bound the clique number of $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$. In a subsequent paper, the authors use these quantities for estimating the dimension of a random geometric graph Friedrich et al. (2023b).

2. Bangachev and Bresler (2023) do give several geometric examples in which signed triangles are not the optimal statistical test for (**P1**). However, in all of them, either the connection probabilities are not monotone or they do not correspond to true "distances" (but, for example, to a non-PSD inner product as in their Theorem 6.17).

3. We choose the circumference to be equal to 2 simply for convenience. One can equivalently define $\mathbb{T}^d = \mathbb{R}^d / \sim$, where $\mathbf{x} \sim \mathbf{y}$ if and only if $\mathbf{x} - \mathbf{y} \in 2\mathbb{Z}^d$.

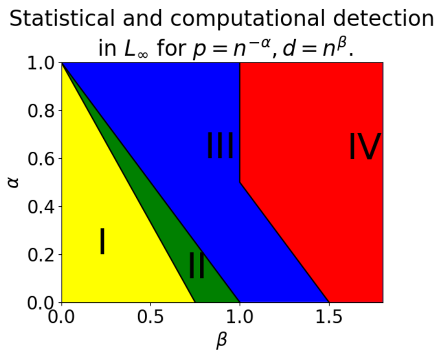

Figure 1: Visualizing Theorems 4, 5 and 6. In region I, the signed triangle test solves (**P1**) for $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$ with high probability. In region $\mathrm{I} + \mathrm{II}$, the signed 4-cycle test succeeds with high probability. In region $\mathrm{III} + \mathrm{IV}$, no low-degree polynomial test succeeds. In IV, it is information theoretically impossible to solve (**P1**) with high probability. The last region is potentially suboptimal.

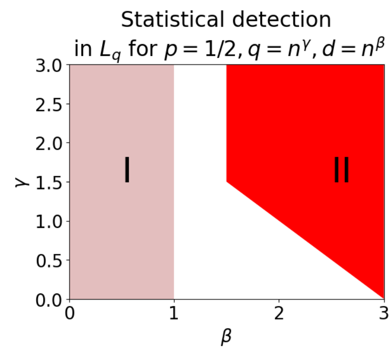

Figure 2: Visualizing Theorems 9 and 10. In region I, the entropy of $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^q_{1/2}, 1/2)$ is much lower than that of $\mathsf{G}(n, 1/2)$. Yet, we do not know any efficient test that distinguishes the two graph models in this region (even though, we believe the signed 4-cycle count does in a strictly larger region). In region II, it is information theoretically impossible to solve (**P1**) with high probability. Both regions are potentially suboptimal.

2. Main Results

Throughout, we frequently refer to signed subgraph counts and low-degree polynomial tests. As these are by now standard in the literature on latent space graphs, we defer the full definitions to Appendix A. Now, we only informally recount the signed triangle count test. For expected density p, it is defined by $\mathsf{SC}_\triangle(G) := \sum_{1 \le i < j < k} (G_{ij} - p)(G_{jk} - p)(G_{ki} - p)$ over all triangles i, j, k on input graph G. If $\mathbf{G} \sim \mathsf{G}(n, p)$, this sum has expectation zero and, by Chebyshev's inequality, with high probability $\mathsf{SC}_\triangle(\mathbf{G}) \in [-v, v]$, where v is any value asymptotically larger than the standard deviation. $\mathsf{SC}_\triangle(\mathbf{H})$ for $\mathbf{H} \sim \mathsf{RGG}$ similarly concentrates (via Chebyshev's inequality) in some interval $[-w + \theta, w + \theta]$. If the two intervals are disjoint, the value $\mathsf{SC}_\triangle$ distinguishes between $\mathsf{G}(n, p)$ and RGG. On the other hand, when the standard-deviation-width intervals overlap, we say that the signed triangle test fails. One can similarly reason with polynomials other than $\mathsf{SC}_\triangle(\mathbf{G})$, in particular the signed four-cycle count $\mathsf{SC}_\square(\mathbf{G})$.

Throughout the rest of the paper, we make the following assumption:

$$\text{There exist some absolute constants } \delta, \epsilon > 0 \text{ such that } n^{-1+\epsilon} \le p \le 1/2, n^\delta \le d. \qquad \textbf{(A)}$$

2.1. Main Results for L_∞ Geometry

The L_∞ case is special because of the following factorization property over coordinates: $\|\mathbf{x} - \mathbf{y}\|_\infty \le \tau$ *holds if and only if* $|x_i - y_i| \le \tau$ *holds for each* $i \in [d]$. This means that each edge (u, v) is the AND of d *independent* edges in the 1-dimensional random geometric graphs over the different coordinates. In comparison, previously studied L_2 models have a (weighted) $\mathsf{MAJORITY}$ combinatorics. For instance, in the spherical case $\langle \mathbf{x}, \mathbf{y} \rangle \ge \rho$ if and only if $\sum_{i=1}^d |x_i| \times |y_i| \times \operatorname{sign}(x_i y_i) \ge \rho$. Each $\operatorname{sign}(x_i y_i)$ is an independent 1-dimensional edge and the values $|x_i| \times |y_i|$ are the corresponding weights. Similarly, over $\{\pm 1\}^d$, weights equal 1 and $\mathsf{MAJORITY}$ is unweighted.

Factorization over the induced independent 1-dimensional random geometric graphs makes the computation of expected signed subgraph counts tractable as computations in one dimension are naturally simpler. Signed subgraph counts are fundamental in studying random graph distributions as they are the Fourier coefficients of the probability mass function. The factorization property, also utilized in Friedrich et al. (2023a), is the first main ingredient in our results in the L_∞ case.

The second ingredient is combining the induced 1-dimensional structures via the AND function. While in certain special cases this step is nearly trivial (e.g., in Theorem 4 we only need to do it for $K_{2,t}$ subgraphs and in Theorem 5 for triangles and 4-cycles), in full generality it requires a careful analysis of the *compatibility* of induced 1-dimensional structures. We carry out such an analysis in Section 3 by viewing each 1-dimensional structure as a *polymer* and expanding the product over the d coordinates. A rearrangement of terms yields a tremendous amount of cancellations that leaves us with an expression for the expected signed subgraph counts similar to the celebrated cluster expansion formula (e.g., Mayer and Mayer (1940); Koteckÿ and Preiss (1986); Friedli and Velenik (2017)) from statistical physics (which has found many other applications in combinatorics and theoretical computer science, e.g. Scott and Sokal (2005); Helmuth et al. (2019); Jenssen and Perkins (2020)). In our case, the compatibility criterion is given by the size of the overlap of different 1-dimensional structures. What makes a cluster-expansion-like formula appealing is a rapid decay of terms which means that terms corresponding to small clusters determine its asymptotics (as in the Koteckỳ-Preiss theorem Koteckÿ and Preiss (1986)). The derivation and analysis of this formula is our technical and conceptual highlight in the L_∞ case.

This gives the following bound on signed subgraph weights of $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$. For a set of edges $H = \{(i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k)\}$, denote

$$\begin{aligned} \mathsf{SW}_H(G) &\coloneqq \prod_{(ij)\in E(H)} (G_{ij} - p) \quad \text{(the signed weight of } H), \\ \mathsf{W}_H(G) &\coloneqq \prod_{(ij)\in E(H)} G_{ij} \quad \text{(the unsigned weight of } H). \end{aligned} \tag{1}$$

Theorem 3 *Suppose that $H \subseteq K_n$ is a graph on $|E(H)| \le (\log d)^{5/4}/(\log\log d)$ edges. Under Assumption (A), there exists a universal constant C such that*

$$\left|\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\Big[\mathsf{SW}_H(\mathbf{G})\Big]\right| = O\left(p^{|E(H)|}\left(\frac{(\log d)^C}{d}\right)^{|V(H)|/2}\right).$$

The quantity $p^{|E(H)|}$ appears naturally as each of the $|E(H)|$ edges has marginal expectation p. An exponentially small quantity in the number of vertices, i.e. $\big((\log d)^C/d\big)^{|V(H)|/2}$, appears frequently in the computation of Fourier coefficients of latent space graphs as it corresponds to events determined by the $|V(H)|$ latent vectors (e.g., Hopkins (2018) for planted clique and Kothari et al. (2023); Rush et al. (2023) for instances of the stochastic block model). While we do not have an intuitive explanation of why $|V(H)|/2$ is the correct dependence, it is crucial to the proof of Theorem 6. An exponent of the form $|V(H)|/(2+\xi)$ for any constant $\xi > 0$ would not suffice.

In a subsequent work by the same authors Bangachev and Bresler (2024), a similar statement to Theorem 3 is derived for spherical random geometric graphs. In that setting, the signed subgraph count of H is bounded by $(8p)^{|E(H)|} \times \big(\frac{(\log d)^C}{d}\big)^{\mathsf{OEI}(H)/2}$, where $\mathsf{OEI}(H)$ is a function of H for which $\mathsf{OEI}(H) \in [\lceil(|V(H)|-1)/2\rceil, |V(H)|-1]$. In particular, in the spherical case the bound might be as large as $(8p)^{|E(H)|} \times \big(\frac{(\log d)^C}{d}\big)^{|V(H)|/4}$. Hence, the Fourier coefficients of d-dimensional spherical random geometric graphs might be polynomially larger than the Fourier coefficients of d-dimensional random geometric graphs over the torus with L_∞ metric. This explains why detection is (information-theoretically and with respect to low-degree polynomials) possible in a larger range of dimensions d in the spherical case.

Our proof of Theorem 3 also yields improved estimates for the unsigned subgraph weights studied in Friedrich et al. (2023a). We discuss this in Appendix B. Now, we present the algorithmic implications of Theorem 3.

2.1.1. DETECTING L_∞ GEOMETRY

The first approach to (**P1**) is information-theoretic. An argument of Liu and Rácz (2023a) (stated in Appendix A) reduces this question to bounding signed weights of $K_{2,t}$ subgraphs. We obtain:

Theorem 4 (Information-Theoretic Lower Bound for L_∞ Model) *If (**A**) holds and $d \geq (\log n)^C \max(n^{3/2}p, n)$ for some universal constant C, then*

$$\mathsf{TV}\Big(\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p), \mathsf{G}(n,p)\Big) = o(1).$$

Theorem 4 already highlights a quantitative difference between L_∞ random geometric graphs over $\mathbb{T}^d$ and L_2 models over $\mathbb{S}^{d-1}$(recall the results of Liu et al. (2022)): The former converge to Erdős-Rényi at a polynomially smaller dimension. Much more interesting, however, is the following qualitative difference in the relative performances of signed triangle and 4-cycle counts.

Theorem 5 *Under Assumption (**A**), consider problem (**P1**) with $H_1 : \mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$. There exists some universal constant $C > 0$ such that:*

1. *The signed 4-cycle test distinguishes the two graph models successfully with high probability if $d \leq (\log n)^{-C} np$ and fails with high probability if $d > (\log n)^C np$.*
2. *The signed triangle test distinguishes the two graph models sucessfully with high probability if $d \leq (\log n)^{-C}(np)^{3/4}$ and fails with high probability if $d \geq (\log n)^C (np)^{3/4}$.*

We provide intuition behind the suboptimality of signed triangles and its consequences in Section 2.1.2. Now, we address the gap between the upper and lower bounds in Theorems 5 and 4.

Theorem 6 (Computational Lower Bound for L_∞ Model) *If (**A**) holds and $d \geq (np)^{1+\eta}$ for any absolute constant $\eta > 0$, then no polynomial test of degree at most $(\log n)^{5/4}/(\log \log n)$ can distinguish $\mathsf{G}(n,p)$ and $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$ with high probability.*

A popular conjecture is that "sufficiently noisy" statistical problems in high-dimension can be solved in polynomial time only if there is an $O(\log n)$-degree polynomial test that solves them Hopkins (2018). In this light, our result suggests that: 1) Either, there is a statistical-computational gap for detecting L_∞ geometry; 2) Or, Theorem 4 is suboptimal. Resolving the presence of a statistical-computational gap is an exciting question for future research. Closely related models provide examples of both positive and negative answers to this question. Spherical random geometric graphs do not exhibit a statistical-computational gap in the dense case $p = 1/2$ Bubeck et al. (2014). A certain quiet planted coloring model (Kothari et al., 2023, Definition 2.18) (which, in particular, can be realized as a random algebraic graph (see Definition 11) over a discrete torus) is shown to exhibit an information-computation gap within the low-degree polynomial tests framework.

2.1.2. TRIANGLES AND 4-CYCLES IN L_∞ GEOMETRY

We end our discussion of the L_∞ model with a further comparison between signed triangle counts and signed four-cycle counts. First, we illustrate with an example.

Example 1 *Consider the case $p = \frac{1}{2}$ in which $\tau^{\infty}_{1/2} = 1 - \lambda$, where $\lambda = \Theta(1/d)$ (see Definition 2).*

First, we interpret $\mathbf{E}[(2\mathbf{G}_{12} - 1)(2\mathbf{G}_{23} - 1)(2\mathbf{G}_{31} - 1)]$, the signed expectation of triangle $\{1, 2, 3\}$. It measures the correlation between the events "2 is a neighbour of 1" *(captured by the term $(2\mathbf{G}_{12} - 1)$) and* "2 is a two-step neighbour of 1 via 3" *(the term $(2\mathbf{G}_{13} - 1)(2\mathbf{G}_{32} - 1)$). Over the unit sphere, these two notions have perfect rank correlation as both are monotone in the distance between vectors $\mathbf{x}^1, \mathbf{x}^2$. The closer $\mathbf{x}^1, \mathbf{x}^2$ are, the larger the probability that $\mathbf{x}^3$ is a common neighbor or a neighbor of neither. This is not the case in the L_∞ model. Consider $\mathbf{x}^1 = (0, 0, \ldots, 0), \mathbf{x}^{2_a} = (1, 0, 0, \ldots, 0)$, and $\mathbf{x}^{2_b} = (\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$. Clearly, $\|\mathbf{x}^1 - \mathbf{x}^{2_a}\|_\infty = 1$, so vertices 1 and 2_a are not adjacent. Still, the set of latent vectors adjacent to $\mathbf{x}^1, \mathbf{x}^{2_a}$ has measure $(1 - 2\lambda) \times (1 - \lambda)^{d-1} = \frac{1}{2}(1 + o(1))$ since a point $\mathbf{x}^3$ is adjacent to $\mathbf{x}^1$ and $\mathbf{x}^{2_a}$ if and only if $(\mathbf{x}^3)_1 \notin (-\lambda, \lambda) \cup (1 - \lambda, 1 + \lambda)$, and $(\mathbf{x}^3)_i \notin (1 - \lambda, 1 + \lambda)$ for $i \in \{2, 3, \ldots, d\}$. In contrast, $\mathbf{x}^1$ and $\mathbf{x}^{2_b}$ are adjacent and only at distance $1/2$, but the set of latent vectors adjacent to $\mathbf{x}^1, \mathbf{x}^{2_b}$ has the much smaller measure $(1 - 2\lambda)^d = \frac{1}{4}(1 + o(1))$ ($(\mathbf{x}^3)_i \notin (\frac{3}{2} - \lambda, \frac{3}{2} + \lambda) \cup (1 - \lambda, 1 + \lambda)\ \forall i$).*

The 4-cycle statistic on cycle $\{1, 3, 2, 4\}$ measures the correlation between two-step paths 1–3–2 and 1–4–2 from 1 to 2. This statistic does not suffer from the same issue as signed triangle counts because the two two-step paths are the same function of $\mathbf{x}^1 - \mathbf{x}^2$.

The advantage of counting signed four cycles over counting signed triangles in the L_∞ model extends to other tasks beyond testing against Erdős-Rényi, for example estimating the dimension in $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^{\infty}_p, p)$. The existing literature on dimension estimation is fully focused on triangle-based estimators Bubeck et al. (2014); Almagro and M. & Serrano (2022); Friedrich et al. (2023b). Not much is known about the optimality of these estimators beyond the case of L_2 geometry. We consider the following problem.

On input n, p and $\mathbf{G}$, where $\mathbf{G} \sim \mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^{\infty}_p, p)$, find the unknown dimension d. **(P2)**

One can also study variants of this problem, such as when the expected density p is unknown or when one allows for a small error in estimating d. We focus on this simplest version as our goal is to demonstrate the advantage of counting signed four-cycles over counting signed three-cycles.

Theorem 7 (Simple Estimators for Dimension Recovery) *Consider (**P2**) under assumption (**A**) with known value of δ such that $d \geq n^\delta$. There exists an absolute constant $C > 0$ such that:*

1. *The signed 4-cycle statistic recovers d exactly with high probability when $d \leq (\log n)^{-C}(np)^{2/3}$ and fails with high probability when $d \geq (\log n)^{C}(np)^{2/3}$.*
2. *The signed triangle statistic recovers d exactly with high probability when $d \leq (\log n)^{-C}(np)^{1/2}$ and fails with high probability when $d \geq (\log n)^{C}(np)^{1/2}$.*

It is important to note that Theorem 7 holds under the assumption (**A**) requiring np and d to be polynomial in n. The setting of Friedrich et al. (2023b) in which the authors use a (weighted) signed triangle count is in a disjoint regime $np = \Theta(1), d = o(\log n)$.

2.2. Additional Results

2.2.1. L_q GEOMETRY FOR $q < \infty$

So far, we have shown that random geometric graphs with L_∞ geometry behave qualitatively and quantitatively differently from L_2 models with respect to (**P1**). This motivates the question of studying (**P1**) in other geometries as well, in particular L_q. The analysis of L_q models, however, turns out to be much more challenging when $q < \infty$. The factorization over 1-dimensional random geometric graphs does not hold any longer. This makes the computation of signed subgraph counts much more difficult. We have not succeeded to perform such a computation even for triangles.

One special case in which we manage to bound the signed subgraph count is the case of bipartite graphs $K_{2,t}$, which is enough to prove an analogue of Theorem 4. What makes this calculation simpler is that the signed expectation of $K_{2,t}$ is given by the t-th centered moment of the self-convolution of $\sigma^q_{1/2}$. Using the Bernstein-McDiarmid inequality (in Appendix A), we bound the centered moments of σ by revealing the d coordinates one at a time. The technical highlight of this argument is proving that each coordinate (say x_d) is marginally nearly uniform on $\mathbb{T}^1$ even conditioned on the value of $\sigma^q_{1/2}(\mathbf{x}, \mathbf{y})$ when $q \ll d$. The reason for this phenomenon is that the contribution of the remaining $d-1$ coordinates, i.e. $\sum_{i=1}^{d-1} |x_i - y_i|^q_C$, is sufficiently anticoncentrated and, thus, there are no spikes in its distribution that would bias x_d strongly when conditioning on $\sigma^q_{1/2}(\mathbf{x}, \mathbf{y})$. We derive the following general anticoncentration result by extending the work of Bobkov and Chistyakov (2014) to random variables with potentially unbounded density.

Corollary 8 *Suppose that X is a non-negative real-valued random variable that is absolutely continuous with respect to the Lebesgue density with pdf f. Let $d \in \mathbb{N}$ and $\rho \in (0, 1]$ be such that $d > \rho^{-1}$. Let m be such that $\int_{\{f(x)>m\}} f(y)dy = 1 - \rho$. Then, for any interval $[a, b] \subseteq \mathbb{R}$, if $X_1, X_2, \ldots, X_d$ are independent copies of X,*

$$\mathbf{P}\big[X_1 + X_2 + \cdots + X_d \in [a, b]\big] \le \exp(-d\rho/8) + \sqrt{2e}\frac{m}{\sqrt{\rho^3 d}}(b - a).$$

We fix $p = 1/2$ and vary q so that we obtain a meaningful comparison of different geometries.

Theorem 9 *Suppose that $q \ge 1$. There exists an absolute constant $C > 0$ such that:*

1. *If $q = o(d/\log d)$ and $dq \ge n^3(\log n)^C$, $\mathsf{TV}\Big(\mathsf{RGG}(n, \mathbb{T}^d, \sigma^q_{1/2}, 1/2), \mathsf{G}(n, 1/2)\Big) = o(1)$.*
2. *If $q = \Omega(d/\log d)$ and $d^2 \ge n^3(\log n)^C$, $\mathsf{TV}\Big(\mathsf{RGG}(n, \mathbb{T}^d, \sigma^q_{1/2}, 1/2), \mathsf{G}(n, 1/2)\Big) = o(1)$.*

This statement interpolates between known results for L_2 models where convergence to $\mathsf{G}(n, 1/2)$ occurs when $d = \tilde{\omega}(n^3)$ (for example, in the spherical case Bubeck et al. (2014)) and L_∞ models when convergence occurs for $d^2 = \tilde{\omega}(n^3)$ (see Theorem 4). Our corresponding lower bound is:

Theorem 10 *Take any $q \in [1, +\infty]$ and any p such that $1/2 \ge p \ge 1/n$. If $d = o(np/\log n)$, then*

$$\mathsf{TV}\Big(\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^q_p, p), \mathsf{G}(n, p)\Big) = 1 - o(1).$$

Interestingly, this gives the same bound as the signed 4-cycle test when $q = \infty$ (Theorem 5). The proof proceeds by discretizing $\mathbb{T}^d$ and applying an entropy argument similar to (Bangachev and Bresler, 2023, Theorem 7.5) which shows that the support of $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ is concentrated on a set of size $\exp(O(dn\log d))$. Developing algorithmic upper bounds for general L_q remains open. We present some ideas and conjectures in Appendix G, based on a Fourier-analytic interpretation of signed subgraph counts similar to (Bangachev and Bresler, 2023, Observation 2.1).

2.2.2. RANDOM ALGEBRAIC GRAPHS

What makes the Bernstein-McDiarmid analysis feasible in the case of Theorem 9 is that the coordinates of $\mathbb{T}^d$ are independent. This method can be extended to other cases of a product structure.

Definition 11 (Random Algebraic Graph over Tori Bangachev and Bresler (2023)) *Suppose that $\mathcal{G}$ is a finite Abelian group or a finite-dimensional torus $\mathbb{T}^d$. Let Unif be the uniform (Haar) measure over $\mathcal{G}$ and let $\sigma : \mathcal{G} \longrightarrow [0,1]$ be a measurable function such that $\sigma(\mathbf{g}) = \sigma(-\mathbf{g})$ holds a.s. and $\mathbf{E}_{\mathbf{g}\sim\mathsf{Unif}(\mathcal{G})}[\sigma(\mathbf{g})] = p$. $\mathsf{RAG}(n, \mathcal{G}, \sigma, p)$ is the following distribution over n-vertex graphs:*

$$\mathbf{P}[\mathbf{G} = A] = \mathbf{E}_{\mathbf{x}^1,\mathbf{x}^2,\ldots,\mathbf{x}^n \overset{\text{i.i.d.}}{\sim} \mathcal{D}}\Bigg[\prod_{1\le i<j\le n} \sigma(\mathbf{x}^i - \mathbf{x}^j)^{A_{i,j}}(1-\sigma(\mathbf{x}^i - \mathbf{x}^j))^{1-A_{i,j}}\Bigg].$$

For any $n, d, q, p,$ the random geometric graph $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ is also a random algebraic graph under the choice $\mathcal{G} = \mathbb{T}^d$ and $\sigma(\mathbf{g}) = \mathbb{1}[\|\mathbf{g}\|_q \le \tau_p^q]$. Overloading notation, we will also use σ_p^q as one function-argument, that is $\sigma_p^q(\mathbf{x}, \mathbf{y}) = \sigma_p^q(\mathbf{x} - \mathbf{y})$.

In Bangachev and Bresler (2023), the authors study in detail the case $\mathcal{G} = \{\pm 1\}^d$ and derive a general criterion based on the sizes of Fourier coefficients on each level of σ that guarantee $\mathsf{TV}\Big(\mathsf{RAG}(n, \{\pm1\}^d, \sigma, p), \mathsf{G}(n,p)\Big) = o(1)$. Using a technically much simpler argument, based on the combination of (Liu and Rácz, 2023a, (4)) and Bernstein's inequality, we also recover such a criterion. It relates statistical convergence to Erdős-Rényi with influences of Boolean functions.

Theorem 12 *Suppose that $\sigma : \{\pm1\}^d \longrightarrow [0,1]$ is a connection with expectation p. Then,*

$$\mathsf{TV}\Big(\mathsf{RAG}(n, \{\pm1\}^d, \sigma, p), \mathsf{G}(n,p)\Big)^2 = O\Bigg(\frac{n^3\sum_{i=1}^d \mathbf{Inf}_i[\sigma]^2}{p^2(1-p)^2}\Bigg).$$

We thoroughly compare Theorem 12 and (Bangachev and Bresler, 2023, Theorem 3.1) in Appendix H. For now, we reprove two results from Bangachev and Bresler (2023) using Theorem 12.

Corollary 13 $\mathsf{TV}\Big(\mathsf{RAG}(n, \{\pm1\}^d, \sigma, p), \mathsf{G}(n,p)\Big) = o(1)$ *in the following cases:*

1. *If σ is $\frac{1}{r\sqrt{d}}$-Lipschitz and $d = \omega\big(\frac{n^3}{p^2r^4}\big)$.*
2. *If $\sigma(\mathbf{g}) = \mathbb{1}\Big[\sum_{i=1}^d g_i \ge \tau_p^{\{\pm1\}^d}\Big]$, where $\tau_p^{\{\pm1\}^d}$ is such that $\mathbf{E}[\sigma] = p$, $d \ge (\log n)^C n^3 p^2$.*

Proof For part 1, observe that whenever σ is $(r\sqrt{d})^{-1}$-Lipschitz, by the definition of influence, $\mathbf{Inf}_i[\sigma] = \mathbf{E}_{\mathbf{x}\sim\mathsf{Unif}(\{\pm1\}^d)}\Bigg[\Big(\frac{\sigma(\mathbf{x})-\sigma(\mathbf{x}^{\oplus i})}{2}\Big)^2\Bigg] \le \frac{1}{r^2 d}$, where $\mathbf{x}^{\oplus i}$ denotes the vector $\mathbf{x}$ with the i-th coordinate flipped. We used $|\sigma(\mathbf{x}) - \sigma(\mathbf{x}^{\oplus i})| \le \frac{2}{r\sqrt{d}}$ which follows from the Lipschitzness assumption. The conclusion follows from Theorem 12.

For part 2, again consider $\mathbf{Inf}_i[\sigma]$. The expression $\sigma(\mathbf{x}) - \sigma(\mathbf{x}^{\oplus i})$ is non-zero only if $\mathbf{x}$ has $\frac{d+\tau_p}{2}$ or $\frac{d+\tau_p}{2} - 1$ ones. A simple calculation (carried out, for example, in (Bangachev and Bresler, 2023, Proof of Proposition 4.7)) shows that the probability of this happening is $O(p\sqrt{\log\frac{1}{p}}/\sqrt{d})$. Each influence is of order $\tilde{O}(p^2/d)$ and the conclusion follows. ■

3. Cluster Expansion in L_∞ Random Geometric Graphs

Here, we describe our conceptual and technical highlight: a "cluster-expansion" formula for $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$ which yields Theorem 3.

Preliminaries: Recall Assumption (A). Note that τ_p^∞ satisfies $(\tau_p^\infty)^d = p$. Indeed, this is the case since $p = \mathbf{P}[\|\mathbf{x}\|_\infty \leq \tau_p^\infty] = \mathbf{P}[|x_1|_C \leq \tau_p^\infty]^d$. This immediately implies that $\tau_p^\infty = 1 - \lambda_p^\infty$, where $\lambda_p^\infty = \frac{\log(1/p)}{d}(1+o(1))$. We will write σ, λ, τ instead of $\sigma_p^\infty, \lambda_p^\infty, \tau_p^\infty$ for brevity.

Fix some subgraph $H \subseteq K_n$ defined by edges $e_1, e_2, \ldots, e_k$. We want to bound $\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{SW}_H(\mathbf{G})]$ and $\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{W}_H(\mathbf{G})]$. We utilize the AND structure of L_∞ random geometric graphs, described in the introduction, towards this goal. This is done in several steps, which can be similarly applied in other instances of AND structure (another random graph family exhibiting AND structure is random intersection graphs, see Brennan et al. (2020)).

Step 1: Factorizing Expected Weights over Independent Coordinates. A simple but crucial observation about the L_∞ model is that the different coordinates factorize. Namely, $e_\ell = (i_\ell, j_\ell)$ is an edge if and only if $|x_u^{i_\ell} - x_u^{j_\ell}|_C \leq 1 - \lambda$ for each coordinate $u \in [d]$. Using the independence of coordinates under the distribution $\mathsf{Unif}(\mathbb{T}^d)$,

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{W}_H(\mathbf{G})] = \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma_{1-\lambda}^\infty,1-\lambda)}[\mathsf{W}_H(\mathbf{G})]^d. \tag{2}$$

Step 2: Computations Over a Single Coordinate via Inclusion-Exclusion. Computing the one-dimensional quantities over the graph complement $\overline{\mathbf{G}}$ is simpler than computing them over $\mathbf{G}$. The intuitive reason is that in the complement each edge appears only with very low probability $\lambda = \tilde{\Theta}(1/d)$. In other words, the appearance of an edge is a very restrictive event that largely determines the configuration of latent vectors. Concretely, for a set of edges A, denote by $\chi(A)$ the probability that no edge of A appears in $\mathbf{G} \sim \mathsf{RGG}(n, \mathbb{T}^1, \mathsf{Unif}, \sigma_{1-\lambda}^\infty, 1-\lambda)$, i.e.,

$$\chi(A) := \mathbf{P}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma_{1-\lambda}^\infty,1-\lambda)}[G_{ij} = 0 \text{ for all } ij \in A]\,. \tag{3}$$

Equivalently, $\chi(A)$ is the probability that each edge in set A appears in the random geometric graph over $\mathbb{T}^1$ with connection $\sigma(x,y)_\lambda^{1,>} = \mathbb{1}[|x-y|_C \geq 1-\lambda]$ and expected density λ.

The event $\{\sigma(x,y)_\lambda^{1,>} = 1\}$ significantly constrains the relative locations of x, y on $\mathbb{T}^1$: They are at distance $1 - \tilde{O}(d^{-1})$, so they are nearly diametrically opposite. The principle of inclusion-exclusion converts the computations in the complement to computations over the original graph:

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma_{1-\lambda}^\infty,1-\lambda)}[\mathsf{W}_H(\mathbf{G})] = \sum_{A\subseteq E(H)} (-1)^{|E(A)|}\chi(A). \tag{4}$$

Step 3: Measuring Perturbations From Erdős-Rényi . We take an approach inspired by statistical - physics of measuring perturbations from the "ground state" Erdős-Rényi graph.[4] Measuring perturbations from Erdős-Rényi is natural as that is the null model against which we are testing. We first do this at the level of single subgraphs appearing in the 1-dimensional complements, as in (4):

$$\psi(A) \coloneqq \chi(A) - \lambda^{|E(A)|}. \tag{5}$$

This is the deviation from the probability of all edges in A appearing in $\mathsf{G}(n, \lambda)$. Recalling (4), we immediately get a perturbative expression for $\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^{\infty}_{1-\lambda},1-\lambda)}[\mathsf{W}_H(\mathbf{G})]$:

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^{\infty}_{1-\lambda},1-\lambda)}[\mathsf{W}_H(\mathbf{G})] = \sum_{A\subseteq H}(-1)^{|E(A)|}\chi(A) = \sum_{A\subseteq H}(-1)^{|E(A)|}(\psi(A) + \lambda^{|E(A)|})$$

$$= (1-\lambda)^{|E(H)|} + \mathsf{Err}(H,\lambda), \qquad \text{where} \quad \mathsf{Err}(H,\lambda) \coloneqq \sum_{A\subseteq H}(-1)^{|E(A)|}\psi(A). \tag{6}$$

We interpret each subgraph A of H as a *polymer* and the quantity $(-1)^{|E(A)|}\psi(A)$ as the *weight* of the polymer. In that view, the expression $\mathsf{Err}(H, \lambda)$ is the sum of the weights of polymers which captures "the first order" deviation from the ground state $(1-\lambda)^{|E(H)|}$. The quantity $(1-\lambda)^{|E(H)|}$ is a natural ground state for the expected weight of H in one dimension as it corresponds to the expected weight when edges are independent. Expanding $\Big((1-\lambda)^{|E(H)|} + \mathsf{Err}(H,\lambda)\Big)^d$ in (2),

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma^{\infty}_{p},p)}[\mathsf{W}_H(\mathbf{G})] = \sum_{i=0}^{d}\binom{d}{i}(1-\lambda)^{(d-i)|E(H)|}\mathsf{Err}(H,\lambda)^i. \tag{7}$$

Again, the term $(1-\lambda)^{d|E(H)|} = p^{|E(H)|}$ corresponding to $i = 0$ is the "ground state" weight of H in $\mathsf{G}(n, p)$. Each term of the form $\mathsf{Err}(H, \lambda)^i$ is composed of products of i-tuples of polymer weights, and, thus, can be interpreted as "the i-th order" perturbation from the ground state.

Step 4: Bounds on Polymer Weights. To derive a bound from (7), one needs to bound the polymer weights and, subsequently, the $\mathsf{Err}(H, \lambda)$ term. In Appendix A and B, we show that perturbations $\psi(A)$ are indeed small. Relatively straightforward computations (as they are all over a single dimension, recall (3)) yield:

Lemma 14 *For every set of edges A such that $|V(A)| \leq 1/(8\lambda)$, the following hold:*
1. *If A can be decomposed as $A_1 \cup A_2$, where $|V(A_1) \cap V(A_2)| \leq 1$, then $\chi(A) = \chi(A_1)\chi(A_2)$.*
2. *If A is a forest, then $\chi(A) = \lambda^{|E(A)|}$ and $\psi(A) = 0$.*
3. *$\chi(A) \leq \lambda^{|V(A)|-1}$ whenever A is connected.*
4. *If A is not bipartite, $\chi(A) = 0$. In particular, $\psi(C_{2m+1}) = -\lambda^{2m+1}$.*
5. *$|\psi(A)| \leq 2 \cdot \lambda^{\max\{|V(A)|/2+1, |V(A)|-\mathsf{numc}(A)\}}$, where $\mathsf{numc}(A)$ denotes the number of connected components of A.*
6. *If $m \leq 1/8\lambda$, then $\chi(C_m) = \lambda^{m-1}\phi(m-1)$, where $\phi(m) \coloneqq \mathbf{P}[U_1 + U_2 + \cdots + U_{m-1} \in [-1, 1]]$ for $U_1, U_2, \ldots, U_{m-1} \overset{\text{i.i.d.}}{\sim} [-1, 1]$.*[5]

4. While no familiarity with statistical physics is needed to follow the argument, we will borrow some terminology with the purpose of explaining our approach in familiar language.
5. One can easily check that $\phi(1) = 1, \phi(2) = 3/4, \phi(3) = 2/3, \phi(m-1) = \Theta(m^{-1/2})$.

Step 5: From Unsigned Weights to Signed Weights - Again Inclusion-Exclusion. Signed subgraph weights do not immediately factorize over the independent coordinates. That is, while in the unsigned case we have $\mathbb{1}[\|x^i - x^j\|_\infty \leq 1-\lambda] = \prod_{u=1}^d \mathbb{1}[\|x^i_u - x^j_u\|_\infty \leq 1-\lambda]$, no such expression holds for $(\mathbb{1}[\|x^i - x^j\|_\infty \leq 1-\lambda] - p)$.[6] Instead, we reduce to what we know about unsigned weights:

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{SW}_H(\mathbf{G})] = \mathbf{E}\Big[\prod_{i=1}^k(\mathbf{G}_{e_i}-p)\Big]\\
&= \sum_{A\subseteq E(H)} (-p)^{|E(H)|-|E(A)|}\mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{W}_A(\mathbf{G})] \\
&= \sum_{A\subseteq E(H)} (-1)^{|E(H)|-|E(A)|}(1-\lambda)^{d(|E(H)|-|E(A)|))}\mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{W}_A(\mathbf{G})].
\end{aligned} \tag{8}$$

Using (7) for any $A \subseteq H$, we obtain

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{SW}_H(\mathbf{G})] = \\
&= \sum_{A\subseteq H}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{d(|E(H)|-|E(A)|)}\sum_{i=0}^d\binom{d}{i}(1-\lambda)^{(d-i)|E(A)|}\mathsf{Err}(A,\lambda)^i \\
&= \sum_{i=0}^d\binom{d}{i}(1-\lambda)^{(d-i)|E(H)|}\sum_{A\subseteq H}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\mathsf{Err}(A,\lambda)^i.
\end{aligned} \tag{9}$$

Step 6: The Cluster Expansion Perspective on Signed Subgraph Counts. $\mathsf{Err}(A,\lambda)^i$ is the sum of products of i-tuples of weights of polymers, equivalently "the i-th order" deviation from the ground state. When we sum over $A \subseteq H$, each i-tuple appears with some coefficient which captures the compatibility of this i-tuple. Expanding (9) (full detail in Appendix B.3.2) in the style of the formal derivation of the cluster expansion formula (e.g. (Friedli and Velenik, 2017, Chapter 5)):

$$\begin{aligned}
&\sum_{A\subseteq E(H)}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\mathsf{Err}(A,\lambda)^i\\
&= \sum_{K_1,K_2,\ldots,K_i\subseteq E(H)}(1-(1-\lambda)^i)^{|E(H)|-|E(K_1\cup\cdots\cup K_i)|}\prod_{j=1}^i(-1)^{|E(K_j)|}\psi(K_j).
\end{aligned} \tag{10}$$

Expression (10) is the i-th order of the "cluster expansion" for signed subgraph weights. Since $\sum_{A\subseteq H}(-1)^{|E(H)|-|E(A)|} = \sum_{j=0}^{|E(H)|}\binom{|E(H)|}{j}(-1)^{|E(H)|-j} = 0$, the ground state captured by the terms appearing when $i=0$ vanishes. This is intuitive because in the ground state case of independent edges each expected signed subgraph weight is 0. It remains to interpret the "soft compatibility criterion" captured by the coefficient $(1-(1-\lambda)^i)^{|E(H)|-|E(K_1\cup K_2\cdots\cup K_i)|}$. Whenever $|E(K_1\cup K_2\cdots\cup K_i)|$ is small, this coefficient is very small as $1-(1-\lambda)^i = \tilde{O}(d^{-1})$. Thus,

6. One cannot expect $(\mathbb{1}[\|x^i - x^j\|_\infty \leq 1-\lambda] - p)$ to always be a d-th power, for example because $\mathbb{1}[\|x^i - x^j\|_\infty \leq 1-\lambda] - p$ might be negative while a d'th power is always positive when d is even.

polymers $K_1, K_2, \ldots, K_i$ are more compatible when $|E(K_1 \cup K_2 \cdots \cup K_i)|$ is smaller. Such a compatibility criterion should not be surprising—subgraphs K_j corresponding to different coordinates are more compatible when they are more similar (so that their union does not blow up).

The final step towards Theorem 3 is to bound the i-th order deviations from the ground state:

Lemma 15 *Recall the definition of* $\mathsf{Err}(A, \lambda)$ *in* (7). *For* $1 \le i \le d$, *the following holds:*

$$\left| \sum_{A \subseteq H} (-1)^{|E(H)|-|E(A)|} (1-\lambda)^{i(|E(H)|-|E(A)|)} \mathsf{Err}(A,\lambda)^i \right| \le \frac{1}{(4d)^i} \times \Big(\frac{(\log d)^C}{d}\Big)^{|V(H)|/2}.$$

In proving Lemma 15, there are two conceptually distinct regimes for i, as is common in the asymptotic analysis of sums (in particular, in the cluster-expansion formula).

1. **Small values of** i. We use (10). By Lemma 14, $|\psi(K_j)| \le (2\lambda)^{|V(K_j)|/2+1}$. Thus, whenever $\sum_{j=1}^i |V(K_j)|$ is large, the total weight $|\psi(K_1)\psi(K_2)\cdots\psi(K_i)|$ of the i-tuple is low. An energy-entropy trade-off phenomenon occurs—and there are very few i-tuples for which $\sum_{j=1}^i |V(K_j)|$ is small:

   Lemma 16 *Let* $i \ge 2, 0 \le b \le i$ *be integers and* $a > 0$ *be a real number. Then, the number of* i*-tuples* $K_1, K_2, \ldots K_i$ *of* H *such that* $\sum_{j=1}^i |V(K_j)| \le ab$ *is at most* $\exp\big(b(\log i) + a^2 i \log|E(H)| + |E(H)|b\big)$.

   To handle the few potentially "high-energy" terms – for which $\sum_{j=1}^i |V(K_j)|$ is small – we use a comparison inequality. Namely, $|\psi(K_j)| \le (2\lambda)^{|V(K_j)|-\mathsf{numc}(K_j)}$ from Lemma 14 for all j and the fact that the quantity $|V(K)| - \mathsf{numc}(K)$ is subadditive under edge unions (proved in Appendix B) allows us to bound $|\psi(K_1)\psi(K_2)\cdots\psi(K_i)|$ by $|\psi(K_1 \cup K_2 \cdots \cup K_i)|$. This makes all quantities in (10) functions of $K_1 \cup K_2 \cdots \cup K_i$ (up to signs).
2. **Large values of** i. "High degree" terms are asymptotically irrelevant due to a rapid enough decay of $\mathsf{Err}(A,\lambda)^i$. Specifically, one can prove that for all $A \subset E(H)$, $|\mathsf{Err}(A,\lambda)| \le d^{-3+o_d(1)}$ by applying triangle inequality over all subgraphs K of A (recall the definition of $\mathsf{Err}(A,\lambda)$ in (7)) and using that $|\psi(K)| \le (2\lambda)^{\max(3,|V(K)|/2+1)}$ from Lemma 14.

4. Discussion and Future Directions

Testing for Different Geometries. For different values of q, not only the limits of computational and statistical detection of L_q geometry vary, but also the optimal algorithms are different. In particular, contrary to previous work, the signed triangle count is not always optimal as the signed 4-cycle test succeeds in a polynomially larger range. This naturally leads to several other questions.

What other tests besides counting signed 3- and 4- cycles can be optimal for detecting high-dimensional latent geometry? We note that the recent work Yu et al. (2024) addresses a similar question for models of a planted dense subgraph in a dense Erdős-Rényi graph (instead of models with latent geometry). They show that for this family of models, the optimal constant-degree test is always a star or an edge count.

Are there instances in which a statistical-computational gap for detecting high-dimensional geometry is present? A positive answer to this question might even be hidden in the $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ models considered in the current paper as there are gaps between the statistical lower-bounds and computationally efficient algorithmic upper bounds.

Other Statistical Tasks on Random Geometric Graphs Over The Torus. Especially intriguing seems the task of efficiently embedding a sample from $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ into $(\mathbb{T}^d, \|\cdot\|_q)$ so that marginal distances are non-trivially approximated.

Problem 17 *On input a sample* $\mathbf{G} \sim \mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ *corresponding to latent vectors* $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n \in \mathbb{T}^d$*, find some vectors* $\widehat{\mathbf{g}_1}, \widehat{\mathbf{g}_2}, \ldots, \widehat{\mathbf{g}_n} \in \mathbb{T}^d$ *such that* $\sum_{1 \le i < j \le n} \Big| \|\mathbf{g}_i - \mathbf{g}_j\|_q - \|\widehat{\mathbf{g}_i} - \widehat{\mathbf{g}_j}\|_q \Big| = o\Big(\sum_{1 \le i < j \le n} \|\mathbf{g}_i - \mathbf{g}_j\|_q \Big)$.

This question has been addressed in prior work for different random geometric graph models, but we believe that the setting of $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ will require substantially different ideas. The spectral approach of Li and Schramm (2023) heavily relies on an inner product structure, which is only present in $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ when $q = 2$. The optimization framework of Ma et al. (2020) works in settings of L_q geometry for general q, but only gives strong poly-time guarantees for connection functions bounded away from 0 and 1, i.e. $c \le \sigma(\mathbf{x}, \mathbf{y}) \le 1 - c$ for some $c > 0$. This however, is not the case in $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ as σ_p^q only takes values 0 and 1. We should mention that Mao et al. (2024) consider the embedding problem for the planted dense cycle model, which resembles $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$ albeit the latent geometry has dimension 1. Nevertheless, their algorithm gives an information-theoretic upper bound and is not obviously efficient.

The Cluster Expansion Approach. The first step (2) in our "cluster-expansion" approach for bounding the Fourier coefficients of small subgraphs is to exploit the AND structure over induced 1-dimensional random geometric graphs in $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$. The same approach can be applied to other random graphs generated by an AND (respectively OR in the complement) structure such as random intersection graphs (e.g. Brennan et al. (2020)).

Nevertheless, other models exhibit different combinatorial structure. As discussed, L_2 geometry gives rise to a (weighted) $\mathsf{MAJORITY}$ structure (as would any L_C when $C = O(1)$). It could be interesting to consider an extension of these constructions for general $f : \{0,1\}^d \longrightarrow \{0,1\}$ beyond AND and $\mathsf{MAJORITY}$. One way to formalize is the following.

Definition 18 (Coordinate-Factorizabe Graph Distributions) *Given are a "1-dimensional" distribution* $\mathcal{G}$ *over* n*-vertex graphs, an integer* $d \ge 1$*, and a function* $f : \{0,1\}^d \longrightarrow \{0,1\}$*. To generate a sample* $\mathbf{G}$ *from the coordinate-factorizable graph* $\mathsf{CFG}(\mathcal{G}, d, f)$*, one first samples* $\mathbf{G}^1, \mathbf{G}^2, \ldots, \mathbf{G}^d \overset{\text{i.i.d.}}{\sim} \mathcal{G}$ *and then forms the* d *dimensional graph* $\mathbf{G}$ *in which* $\mathbf{G}_{ij} = f(\{\mathbf{G}^u_{ij}\}_{u=1}^d)$.

For $\mathcal{G} = \mathsf{RGG}(n, \mathbb{T}^1, \mathsf{Unif}, \sigma_{1-\lambda}^\infty, 1-\lambda)$ and $f = \mathsf{AND}$, this gives $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$.

When do coordinate-factorizabe graph distributions converge to Erdős-Rényi information- theoretically? When are they distinguishable from Erdős-Rényi via low-degree polynomial tests? One approach towards the low-degree question is to imitate our "cluster-expansion" using the Fourier expansion of f in an analogue of (2). Yet, for choices of f more complicated than AND (AND being simply a product of the coordinates), this approach will likely require new technical insights.

Acknowledgments

We want to thank Will Perkins for insightful conversations on the cluster expansion formula and for suggesting that it might be useful in understanding random geometric graphs. We also want to thank Tselil Schramm for stimulating conversations on random geometric graphs. KB was supported by a Siebel Scholarship. GB is supported by NSF Career Award CCF-1940205.

References

Boguñá Almagro, P. and M.Á. M. & Serrano. Detecting the ultra low dimensionality of real networks. *Nature Communications*, 2022.

Kiril Bangachev and Guy Bresler. Random Algebraic Graphs and Their Convergence to Erdos-Renyi, 2023.

Kiril Bangachev and Guy Bresler. On The Fourier Coefficients of High-Dimensional Random Geometric Graphs. In *To Appear In 56th ACM Symposium on Theory of Computing (STOC 2024)*, 2024.

Thomas F. Bloom and Olof Sisask. The Kelley–Meka bounds for sets free of three-term arithmetic progressions. *Essential Number Theory*, 2023.

Sergey Bobkov and G. Chistyakov. Bounds on the Maximum of the Density for Sums of Independent Random Variables. *Journal of Mathematical Sciences*, 199:100–106, 05 2014.

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

Matthew Brennan, Guy Bresler, and Dheeraj Nagaraj. Phase transitions for detecting latent geometry in random graphs. *Probability Theory and Related Fields*, 178:1215–1289, 12 2020.

Matthew Brennan, Guy Bresler, Samuel B. Hopkins, Jerry Li, and Tselil Schramm. Statistical query algorithms and low-degree tests are almost equivalent. *Conference on Learning Theory*, 2021.

Matthew Brennan, Guy Bresler, and Brice Huang. Threshold for Detecting High Dimensional Geometry in Anisotropic Random Geometric Graphs. *Random Structures and Algorithms*, 64: 125–137, 2026.

Sébastien Bubeck, Jian Ding, Ronen Eldan, and Miklós Rácz. Testing for high-dimensional geometry in random graphs. *Random Structures & Algorithms*, 49:503–532, 11 2014.

F. Chung, R. Graham, and R. Wilson. Quasi-random graphs. *Combinatorica*, pages 345–362, 1988.

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.

Quentin Duchemin and Yohann de Castro. Random Geometric Graph: Some recent developments and perspectives, 2022.

Ernesto Estrada and Matthew Sheerin. Consensus dynamics on random rectangular graphs. *Physica D: Nonlinear Phenomena*, 323-324:20–26, 2016. ISSN 0167-2789. Nonlinear Dynamics on Interconnected Networks.

Sacha Friedli and Yvan Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, 2017. ISBN 978-1-107-18482-4.

Tobias Friedrich, Andreas Göbel, Maximilian Katzmann, and Leon Schiller. Cliques in high-dimensional geometric inhomogeneous random graphs, 2023a.

Tobias Friedrich, Andreas Göbel, Maximilian Katzmann, and Leon Schiller. A simple statistic for determining the dimensionality of complex networks, 2023b.

Martin Haenggi, Jeffrey G. Andrews, Francois Baccelli, Olivier Dousse, and Massimo Franceschetti. Stochastic geometry and random graphs for the analysis and design of wireless networks. *IEEE Journal on Selected Areas in Communications*, 27(7):1029–1046, 2009. doi: 10.1109/JSAC.2009.090902.

Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic pirogov–sinai theory. *Probability Theory and Related Fields*, 176(3–4):851–895, June 2019. ISSN 1432-2064.

Desmond J. Higham, Marija Rašajski, and Nataša Pržulj. Fitting a geometric graph to a protein–protein interaction network. *Bioinformatics*, 24(8):1093–1099, 03 2008. ISSN 1367-4803.

Samuel Hopkins. STATISTICAL INFERENCE AND THE SUM OF SQUARES METHOD, 2018.

Matthew Jenssen and Will Perkins. Independent sets in the hypercube revisited. *Journal of the London Mathematical Society*, 102(2):645–669, April 2020. ISSN 1469-7750.

Z. Kelley and R. Meka. Strong bounds for 3-progressions. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 933–973, Los Alamitos, CA, USA, nov 2023. IEEE Computer Society.

R. Kotecký and D. Preiss. Cluster expansion for abstract polymer models. *Communications in Mathematical Physics*, 103(3):491 – 498, 1986.

Pravesh Kothari, Santosh S. Vempala, Alexander S. Wein, and Jeff Xu. Is planted coloring easier than planted clique? In *Annual Conference Computational Learning Theory*, 2023.

Dmitriy Kunisky, Alexander S. Wein, and Afonso S. Bandeira. Notes on Computational Hardness of Hypothesis Testing: Predictions Using the Low-Degree Likelihood Ratio. In Paula Cerejeiras and Michael Reissig, editors, *Mathematical Analysis, its Applications and Computation*, pages 1–50, Cham, 2022. Springer International Publishing. ISBN 978-3-030-97127-4.

Shuangping Li and Tselil Schramm. Spectral clustering in the Gaussian mixture block model, 2023.

Siqi Liu, Sidhanth Mohanty, Tselil Schramm, and Elizabeth Yang. Testing thresholds for high-dimensional sparse random geometric graphs. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450392648.

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

Suqi Liu and Miklós Z. Rácz. Phase transition in noisy high-dimensional random geometric graphs. *Electronic Journal of Statistics*, 17(2):3512 – 3574, 2023b.

Zhuang Ma, Zongming Ma, and Hongsong Yuan. Universal latent space model fitting for large networks with edge covariates. *J. Mach. Learn. Res.*, 21:4:1–4:67, 2020.

Cheng Mao, Alexander S. Wein, and Shenduo Zhang. Detection-recovery gap for planted dense cycles. In Gergely Neu and Lorenzo Rosasco, editors, *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pages 2440–2481. PMLR, 12–15 Jul 2023.

Cheng Mao, Alexander S. Wein, and Shenduo Zhang. Information-theoretic thresholds for planted dense cycles, 2024.

Maria Mayer and Joseph Mayer. *Statistical Mechanics*. John Wiley, 1940.

Colin McDiarmid. *Probabilistic Methods for Algorithmic Discrete Mathematics*, chapter Concentration. Springer, 1998.

Andrea Montanari and Alexander S. Wein. Equivalence of approximate message passing and low-degree polynomials in rank-one matrix estimation, 2022.

Luke O'Connor, Muriel Médard, and Soheil Feizi. Maximum likelihood embedding of logistic random dot product graphs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34: 5289–5297, 04 2020.

Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press;, 2014.

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

Victor M. Preciado and Ali Jadbabaie. Spectral analysis of virus spreading in random geometric networks. *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 4802–4807, 2009.

Cynthia Rush, Fiona Skerman, Alexander S. Wein, and Dana Yang. Is it easier to count communities than find them? In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA*, volume 251 of *LIPIcs*, pages 94:1–94:23, 2023.

Tselil Schramm and Alexander S. Wein. Computational barriers to estimation from low-degree polynomials. *The Annals of Statistics*, 50(3), 6 2022.

Alexander Scott and Alan Sokal. The repulsive lattice gas, the independent-set polynomial, and the lovasz local lemma. *Journal of Statistical Physics*, 118:1151–1261, 01 2005.

Kiril Solovey, Oren Salzman, and Dan Halperin. New perspective on sampling-based motion planning via random geometric graphs. *The International Journal of Robotics Research*, 37(10): 1117–1133, 2018.

Zheng Xie, Zhenzheng Ouyang, Qi Liu, and Jianping Li. A geometric graph model for citation networks of exponentially growing scientific papers. *Physica A: Statistical Mechanics and its Applications*, 456:167–175, 2016. ISSN 0378-4371.

Xifan Yu, Ilias Zadik, and Peiyuan Zhang. Counting Stars is Constant-Degree Optimal For Detecting Any Planted Subgraph, 2024.

Appendix A. Preliminaries and Notation

Graph Notation. Denote by K_n the clique on n vertices, by $K_{a,b}$ the complete bipartite graph with parts of sizes a and b, and by C_m the cycle on m vertices. For a set of edges $H = \{(i_1, j_1), \ldots, (i_k, j_k)\} \in [n] \times [n]$, denote by H the subgraph of K_n with vertex set $\{i_1, j_1, i_2, j_2, \ldots, i_k, j_k\}$ and edge set $\{(i_1, j_1), \ldots, (i_k, j_k)\}$.

A graph is 2-connected if it is connected and for any $v \in V(H)$, the induced subgraph of H on vertex set $V(H)\backslash\{v\}$ is connected.

A.1. Statistical Detection of Latent Space Structure

Information Theory. We use the standard notions for Total Variation and KL-distance (for example, Polyanskiy and Wu (Forthcoming)). Specifically, for two distributions $\mathbf{P}, \mathbf{Q}$ over the same measurable spaces $(\Omega, \mathcal{F})$, such that $\mathbf{P}$ is absolutely continuous with respect to $\mathbf{Q}$,

$$\begin{aligned} \mathsf{TV}(\mathbf{P}, \mathbf{Q}) &= \sup_{A \in \mathcal{F}} |\mathbf{P}(A) - \mathbf{Q}(A)| = \frac{1}{2} \int_\Omega \Big| \frac{d\mathbf{P}(\omega)}{d\mathbf{Q}(\omega)} - 1 \Big| d\mathbf{Q}(\omega), \\ \mathsf{KL}(\mathbf{P} \| \mathbf{Q}) &= \int_\Omega \frac{d\mathbf{P}(\omega)}{d\mathbf{Q}(\omega)} \log \frac{d\mathbf{P}(\omega)}{d\mathbf{Q}(\omega)} d\mathbf{Q}(\omega). \end{aligned} \tag{11}$$

Total variation appears naturally in hypothesis testing settings as $1 - \mathsf{TV}(\mathbf{P}, \mathbf{Q})$ is the minimal sum of Type I and Type II errors when testing between $\mathbf{P}$ and $\mathbf{Q}$ with a single sample (e.g. Polyanskiy and Wu (Forthcoming)). In practice, it is usually more convenient to work and compute with KL. Importantly, this is enough for proving convergence in total variation due to the celebrated inequality of Pinsker stating that $\mathsf{TV}(\mathbf{P}, \mathbf{Q})^2 \le \frac{1}{2}\mathsf{KL}(\mathbf{P}, \mathbf{Q})$.

A Bound on the KL divergence due to Liu and Racz. In Liu and Rácz (2023a), the authors give the following convenient bound on the KL divergence between $\mathsf{G}(n, p)$ and a probabilistic latent space graph. Specialized to random algebraic graphs (which encompass $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^q, p)$), their bound[7] reads as follows:

$$\begin{aligned} &\mathsf{KL}\Big(\mathsf{RAG}(n, \mathcal{G}, \sigma, p) \| \mathsf{G}(n, p)\Big) \le \sum_{k=0}^{n-1} \log \Big(\mathbf{E}_{\mathbf{x} \sim \mathsf{Unif}(\mathcal{G})} \Big[\Big(1 + \frac{\gamma(\mathbf{x})}{p(1-p)}\Big)^k \Big] \Big), \\ &\text{where } \gamma(\mathbf{x}) := \mathbf{E}_{\mathbf{z} \sim \mathsf{Unif}\mathcal{G}} \Big[(\sigma(\mathbf{x} - \mathbf{z}) - p)(\sigma(\mathbf{z}) - p) \Big] = \mathbf{E}_{\mathbf{z} \sim \mathsf{Unif}\mathcal{G}} \Big[\sigma(\mathbf{x} - \mathbf{z}) \sigma(\mathbf{z}) \Big] - p^2. \end{aligned} \tag{12}$$

Over random algebraic graphs, $\gamma(\mathbf{x}) = \sigma * \sigma(\mathbf{x}) - p^2$, where $\sigma * \sigma$ is the self-convolution $\sigma * \sigma(\mathbf{x}) := \mathbf{E}_{\mathbf{z} \sim \mathsf{Unif}\mathcal{G}} \Big[\sigma(\mathbf{x} - \mathbf{z}) \sigma(\mathbf{z}) \Big]$. Thus, one can expand the left hand-side of (12) either in terms of the moments of $\sigma * \sigma$ or in terms of the moments of $\sigma * \sigma - p^2$. It turns out that in the case of $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$, one can easily compute (up to lower-order terms) the moments of $\sigma * \sigma$ and this is enough to prove Theorem 4.

7. In (Liu and Rácz, 2023a, p.2427-2428) the authors prove this bound for a specific Gaussian random geometric graph, but the proof reads verbatim for random algebraic graphs (and in fact any probabilistic latent space graph as defined in Bangachev and Bresler (2023)).

Remark 19 We briefly discuss two combinatorial interpretations of (12) which connect the bound of Liu and Racz to different notions of pseudorandomness appearing in the literature. *3-Term Arithmetic Progressions:* Expanding the left-hand side of (12), we conclude that small (centered) moments of the self-convolution imply a certain randomness of σ, respectively of $A \subseteq \mathcal{G}$ when $\sigma(\mathbf{g}) := \mathbb{1}[\mathbf{g} \in A]$. We note that the same notion of pseudorandomness was recently used by Kelley and Meka in their breakthrough paper Kelley and Meka (2023) on 3-term arithmetic progressions, in the case $\mathcal{G} = \mathbf{F}_q^n$ (see also the exposition Bloom and Sisask (2023)). One simplification in our setup is that $\sigma(\mathbf{g}) = \sigma(-\mathbf{g})$ in the context of random algebraic graphs, so $\sigma * \sigma(\mathbf{g}) := \mathbf{E_h}\sigma(\mathbf{g}-\mathbf{h})\sigma(\mathbf{h}) = \mathbf{E_h}\sigma(\mathbf{g}+\mathbf{h})\sigma(\mathbf{h}) =: \sigma \star \sigma(\mathbf{g})$. *Quasi-Randomness:* The left-hand side of (12) can be expanded either in terms of the moments of $\sigma * \sigma$ or in terms of the moments of $(\sigma - p) * (\sigma - p)$. However, one can easily observe that $\mathbf{E}[(\sigma * \sigma)^k]$ is exactly the probability that each edge of a fixed copy of $K_{2,t}$ appears in $\mathsf{RAG}(n, \mathcal{G}, \sigma, p)$. In other words, one interpretation of (12) is that if all subgraphs of the form $K_{2,t}$ appear with probability sufficiently close to p^{2t} in $\mathsf{RAG}(n, \mathcal{G}, \sigma, p)$, then $\mathsf{RAG}(n, \mathcal{G}, \sigma, p)$ is (up to $o(1)$ total variation) the same as $\mathsf{G}(n,p)$. This can be viewed as a certain analogue of the celebrated theorem due to Chung-Graham-Wilson Chung et al. (1988). It (among other things) states that if a graph simultaneously has a number of edges and 4-cycles close to that of $\mathsf{G}(n,p)$, it is quasirandom and in particular every other subgraph count is close to that of $\mathsf{G}(n,p)$. Similarly, the t-th moment of $(\sigma - p) * (\sigma - p)/(p(1-p))$ is the Fourier coefficient corresponding to $K_{2,t}$ and one can equivalently interpret for signed copies of $K_{2,t}$.

The Bernstein-McDiarmid Approach. In the case of L_q geometry for $q < \infty$, calculating the moments of $\sigma * \sigma$ is technically challenging. Our proof of Theorem 9 instead exploits the product structure of $\mathbb{T}^d$ to bound the moments of γ via the Bernstein-McDiarmid inequality.

Lemma 20 ((McDiarmid, 1998, Theorem 3.8)) *Let $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_d$ be independent random variables and γ a function of $(\mathbf{g}_1, \ldots, \mathbf{g}_d)$. Denote $\mathbf{g}_{-i} := (\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{i-1}, \mathbf{g}_{i+1}, \ldots, \mathbf{g}_d)$ and*

$$D_i\gamma(\mathbf{g}_{-i}) := \sup_{\mathbf{g}_i^+} \gamma((\mathbf{g}_1, \ldots, \mathbf{g}_{i-1}, \mathbf{g}_i^+, \mathbf{g}_{i+1}, \ldots, \mathbf{g}_d)) - \inf_{\mathbf{g}_i^-} \gamma((\mathbf{g}_1, \ldots, \mathbf{g}_{i-1}, \mathbf{g}_i^-, \mathbf{g}_{i+1}, \ldots, \mathbf{g}_d)),$$

$$\mathbf{Var}_i[\gamma(\mathbf{g}_{-i})] := \mathbf{Var}_{\mathbf{g}_i}[\gamma((\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{i-1}, \mathbf{g}_i, \mathbf{g}_{i+1}, \ldots, \mathbf{g}_d))|\mathbf{g}_{-i}].$$

Then, for any positive t,

$$\mathbf{P}\Big[\gamma(\mathbf{g}) \geq t + \mathbf{E}[\gamma(\mathbf{g})]\Big] \leq \exp\Bigg(- \min\bigg(\frac{t^2}{4\sum_{j=1}^d \|\mathbf{Var}_i[\gamma]\|_\infty}, \frac{t}{2\max_i \|D_i\gamma\|_\infty}\bigg)\Bigg).$$

An immediate corollary is the following.

Lemma 21 ((Boucheron et al., 2013, Theorem 2.3)) *In the setup of Lemma 20, there exists some absolute constant C such that*

$$\|\gamma - \mathbf{E}[\gamma]\|_k \leq C\left(\sqrt{k}\sqrt{\sum_{i=1}^d \|\mathbf{Var}_i[\gamma]\|_\infty} + k \max_i \|D_i\gamma\|_\infty\right).$$

We bound $\mathbf{Var}_i[\sigma], D_i[\gamma]$ for γ defined as in (12) via a careful combination of Fourier-theoretic and anticoncentration arguments to obtain Theorem 9. We also derive Theorem 12 as a combination of (12) and Lemma 21.

A.2. Computational Detection of Latent Space Structure

To solve (**P1**), one observes a certain n-vertex graph G and needs to compute a function $f(G)$ based on which to decide between H_0 and H_1. The graph G is simply a sequence of $\binom{n}{2}$ bits. It is well-known that any function of $0/1$ vectors is simply a polynomial O'Donnell (2014). For computationally efficient tests, one needs to be able to compute f in time polynomial in n.

Signed Subgraph Counts. Most important to the current paper are polynomials corresponding to signed-subgraph counts. Namely, suppose that we want to test between two graph distributions over n vertices in which each edge appears with a marginal probability p. Let $H = \{(i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k)\}$ be any subgraph of K_n. Then, we define the signed weight of H as the polynomial

$$\mathsf{SW}_H(G) := \sum_{(ij)\in E(H)} (G_{ij} - p). \tag{13}$$

For brevity and uniformity with the SW notation, for a set of edges $H = \{(i_1, j_1), \ldots, (i_k, j_k)\}$, denote the unsigned weight $\mathsf{W}_H(G) = \prod_{(ij)\in H} G_{ij} = \mathbb{1}[G_{ij} = 1\ \forall (ij) \in H]$. The signed count of H in G is

$$\mathsf{SC}_H(G) = \sum_{H_1 \subseteq E(K_n)\,:H_1 \sim H} \mathsf{SW}_{H_1}(G), \tag{14}$$

where the sum is over all subgraphs of K_n isomorphic to H. Note that whenever H has a constant number of edges, the polynomial $\mathsf{SC}_H(G)$ is certainly efficiently computable.

Clearly $\mathbf{E}_{\mathbf{G}\sim\mathsf{G}(n,p)}\mathsf{SC}_H(\mathbf{G}) = 0$, which leads to the following approach to (**P1**) appearing in Bubeck et al. (2014). Upon observing G, compute $\mathsf{SC}_H(G)$ and, if sufficiently close to 0, report H_0. Else report H_1. Using Chebyshev's inequality, this can be formalized as follows.

Definition 22 (Success of the Signed Subgraph Count) *We say that signed H-count statistical test $\mathsf{SC}_H(G)$ succeeds in distinguishing between $\mathsf{G}(n,p)$ and RGG if*

$$\big|\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\big[\mathsf{SC}_H(\mathbf{G})\big]\big| = \omega\Big(\sqrt{\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}\big[\mathsf{SC}_H(\mathbf{K})\big] + \mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}}\big[\mathsf{SC}_H(\mathbf{G})\big]}\Big). \tag{15}$$

*Indeed, if this is the case, one can solve (**P1**) with Type I and Type II errors both of order $o(1)$ by comparing $\mathsf{SC}_H(G)$ to $\frac{1}{2}\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\Big[\mathsf{SC}_H(\mathbf{G})\Big]$.*

If, on the other hand,

$$\big|\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\big[\mathsf{SC}_H(\mathbf{G})\big]\big| = o\Big(\sqrt{\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}\big[\mathsf{SC}_H(\mathbf{K})\big] + \mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}}\big[\mathsf{SC}_H(\mathbf{G})\big]}\Big), \tag{16}$$

we say that the signed H-count statistical test fails with high probability.

In this work, we are mostly interested in the case of triangles, $H = C_3$, and 4-cycles, $H = C_4$.

Low-Degree Tests. In Definition 22, one can replace $\mathsf{SW}_H(\cdot)$ with any polynomial $f(\cdot)$ and compare

$$\big|\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\big[f(\mathbf{G})\big] - \mathbf{E}_{\mathbf{G}\sim\mathsf{G}(n,p)}\big[f(\mathbf{G})\big]\big| \text{ and } \sqrt{\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}\big[f(\mathbf{K})\big] + \mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}}\big[f(\mathbf{G})\big]}.$$

High-probability success and failure are similarly defined.

A popular conjecture Hopkins (2018) states that all polynomial time algorithms for solving (sufficiently noisy) hypothesis testing questions in high-dimension are captured by polynomials of degree $O(\log n)$. Indeed, there is growing evidence in support of this conjecture. Clearly, low degree polynomial tests capture (signed) counts of small subgraphs (note that one can even capture the first $O((\log n)/k)$ moments of the (signed) counts of a graph H with k edges and hence a lot more about the distribution of signed counts), which have proven powerful in detecting random geometric graphs Bubeck et al. (2014), planted cliques and colorings Kothari et al. (2023), the number of communities in a stochastic block model Rush et al. (2023) and others. Low-degree polynomials further capture spectral methods Kunisky et al. (2022), constant round approximate message passing algorithms Montanari and Wein (2022), and statistical query algorithms Brennan et al. (2021). Thus, a lot of recent work on the complexity of problems in high-dimensional statistics has focused on ruling out low-degree polynomial algorithms for statistical problems. This constitutes strong evidence that the respective statistical problems cannot be solved in polynomial time.

Formally, in the case of (**P1**) one needs to show that there exists some function $D(n) = \omega(\log n)$ such that for all degree $D = D(n)$ polynomials f, it is the case that

$$\left|\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\big[f(\mathbf{G})\big] - \mathbf{E}_{\mathbf{G}\sim\mathsf{G}(n,p)}\big[f(\mathbf{G})\big]\right| = o\Big(\sqrt{\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}\big[f(\mathbf{K})\big] + \mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}}\big[f(\mathbf{G})\big]}\Big).$$

One way to prove such an inequality is by bounding the following quantity Hopkins (2018):

$$\mathsf{ADV}_{\leq D} := \max_{f\,:\,deg(f)\leq D} \frac{\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\big[f(\mathbf{G})\big]}{\sqrt{\mathbf{E}_{\mathbf{K}\sim\mathsf{G}(n,p)}\big[f(\mathbf{K})^2\big]}}. \tag{17}$$

In particular, if $\mathsf{ADV}_{\leq D} = 1 + o(1)$, then statistical test $f(\cdot)$ fails with large probability (e.g. Rush et al. (2023)).

The product structure of $\mathsf{G}(n,p)$ yields a convenient formula for $\mathsf{ADV}_{\leq D}$. The set of polynomials $\{\mathsf{SW}_H \times (p(1-p))^{-|E(H)|/2}\}_{H\subseteq E(K_n)\,:\,0\leq|E(H)|\leq D}$ forms an orthonormal basis of the polynomials of degree up to D with respect to $\mathsf{G}(n,p)$. A standard application of the Cauchy-Schwartz inequality (e.g. Hopkins (2018)) shows that

$$\mathsf{ADV}^2_{\leq D} - 1 = \sum_{H\subseteq E(K_n)\,:\,1\leq|E(H)|\leq D} \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\big[\mathsf{SW}_H \times (p(1-p))^{-|E(H)|/2}\big]^2.$$

We summarize in the following proposition.

Lemma 23 *If there exists some $D = \omega(\log n)$ such that*

$$\sum_{H\subseteq E(K_n)\,:\,1\leq|E(H)|\leq D} \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\big[\mathsf{SW}_H \times (p(1-p))^{-|E(H)|/2}\big]^2 = o(1),$$

*then the Type I plus Type II error of any degree D polynomial in solving (**P1**) is of order $1 - o(1)$.*

We use the bounds from Theorem 3 and this proposition to prove Theorem 6. We note that low-degree polynomials are similarly used in the literature for estimation and refutation tasks (e.g. Schramm and Wein (2022); Rush et al. (2023)). We discuss this in more detail in Section C.2 in the context of estimating the dimension of a graph sampled from $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$.

Generic Bounds of Subgraph Weights in Random Algebraic Graphs. We end this section by proving several simple generic facts about signed weights in random algebraic graphs which will be useful throughout. In particular, they immediately yield parts 1, 2, and 3 of Lemma 14 as χ is the unsigned weight of the random algebraic graph $\mathsf{RGG}(n, \mathbb{T}^1, \mathsf{Unif}, \sigma_\lambda^{1,>}, \lambda)$

Lemma 24 *Consider any random algebraic graph* $\mathsf{RAG}(n, \mathcal{G}, \sigma, p)$*. Let A be any subgraph. Then,*

1. *If A can be decomposed as $A_1 \cup A_2$ such that $E(A_1) \cup E(A_2) = E(A)$ and $|V(A_1) \cap V(A_2)| \leq 1$, the edge sets $\{\mathbf{G}_e\}_{e \in E(A_1)}$ and $\{\mathbf{G}_e\}_{e \in E(A_2)}$ are independent over $\mathbf{G} \sim \mathsf{RAG}(n, \mathcal{G}, \sigma, p)$. In particular, for any two functions f, g on those edge sets,*
$$\mathbf{E}[f(\{\mathbf{G}_e\}_{e\in E(A_1)})g(\{\mathbf{G}_e\}_{e\in E(A_2)})] = \mathbf{E}[f(\{\mathbf{G}_e\}_{e\in E(A_1)})] \times \mathbf{E}[g(\{\mathbf{G}_e\}_{e\in E(A_2)})].$$
2. *If A is a forest, $\mathbf{E}_{\mathbf{G}\sim \mathsf{RAG}(n,\mathcal{G},\sigma,p)}[\mathsf{W}_A(\mathbf{G})] = p^{|E(A)|}$ and $\mathbf{E}_{\mathbf{G}\sim \mathsf{RAG}(n,\mathcal{G},\sigma,p)}[\mathsf{SW}_A(\mathbf{G})] = 0.$*
3. *If A is connected, then $\Big|\mathbf{E}_{\mathbf{G}\sim \mathsf{RAG}(n,\mathcal{G},\sigma,p)}[\mathsf{W}_A(\mathbf{G})]\Big| \leq p^{|V(A)|-1}.$*

Proof

Item 1. If A_1 and A_2 do not share a vertex, $\{\mathbf{G}_e\}_{e\in E(A_1)}$ and $\{\mathbf{G}_e\}_{e\in E(A_2)}$ are independent as they are fully determined by disjoint sets of latent vectors. If $|V(A_1) \cap V(A_2)| = 1$, we use the measure-preserving transitive group of translations in $\mathcal{G}$ as follows. Let $V(A_1) = \{u_0, u_1, \ldots, u_k\}$, $V(A_2) = \{v_0, v_1, \ldots, v_r\}$, where $u_0 = v_0$. Note that A_1, A_2 have no common edges. Then

$$\begin{aligned}
&\mathbf{P}[\{\mathbf{G}_e\}_{e\in E(A_1)} = \{g_e\}_{e\in E(A_1)}, \{\mathbf{G}_e\}_{e\in E(A_2)} = \{g_e\}_{e\in E(A_2)}]\\
&= \mathbf{E}_{\mathbf{x}^{u_0},\mathbf{x}^{u_1},\ldots,\mathbf{x}^{u_k},\mathbf{x}^{v_1},\ldots,\mathbf{x}^{v_r} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}(\mathcal{G})}\Big[\prod_{(u_s,u_t)\in E(A_1)} \sigma(\mathbf{x}^{u_s} - \mathbf{x}^{u_t})^{g_{(u_s,u_t)}} \times\\
&\quad \times (1-\sigma(\mathbf{x}^{u_s} - \mathbf{x}^{u_t}))^{1-g_{(u_s,u_t)}} \prod_{(v_k,v_\ell)\in E(A_2)} \sigma(\mathbf{x}^{v_k} - \mathbf{x}^{v_\ell})^{g_{(v_k,v_\ell)}}(1-\sigma(\mathbf{x}^{v_k} - \mathbf{x}^{v_\ell}))^{1-g_{(v_k,v_\ell)}}\Big]\\
&= \mathbf{E}_{\mathbf{z},\mathbf{x}^{u_0},\mathbf{x}^{u_1},\ldots,\mathbf{x}^{u_k},\mathbf{x}^{v_1},\ldots,\mathbf{x}^{v_r} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}(\mathcal{G})}\Big[\prod_{(u_s,u_t)\in E(A_1)} \sigma(\mathbf{x}^{u_s} - \mathbf{x}^{u_t})^{g_{(u_s,u_t)}} \times\\
&\quad \times (1-\sigma(\mathbf{x}^{u_s} - \mathbf{x}^{u_t}))^{1-g_{(u_s,u_t)}} \prod_{(v_k,v_\ell)\in E(A_2)} \sigma((\mathbf{x}^{v_k}+\mathbf{z}) - (\mathbf{x}^{v_\ell}+\mathbf{z}))^{g_{(v_k,v_\ell)}} \times\\
&\quad \times (1-\sigma((\mathbf{x}^{v_k}+\mathbf{z}) - (\mathbf{x}^{v_\ell}+\mathbf{z})))^{1-g_{(v_k,v_\ell)}}\Big]\\
&= \mathbf{E}_{\mathbf{x}^{u_0},\mathbf{x}^{u_1},\ldots,\mathbf{x}^{u_k} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}(\mathcal{G})}\Big[\prod_{(u_s,u_t)\in E(A_1)} \sigma(\mathbf{x}^{u_s} - \mathbf{x}^{u_t})^{g_{(u_s,u_t)}}(1-\sigma(\mathbf{x}^{u_s} - \mathbf{x}^{u_t}))^{1-g_{(u_s,u_t)}}\Big] \times\\
&\times \mathbf{E}_{\mathbf{z},\mathbf{x}^{v_1},\ldots,\mathbf{x}^{v_r} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}(\mathcal{G})}\Big[\prod_{(v_k,v_\ell)\in E(A_2)} \sigma((\mathbf{x}^{v_k}+\mathbf{z}) - (\mathbf{x}^{v_\ell}+\mathbf{z}))^{g_{(v_k,v_\ell)}} \times\\
&\qquad\qquad \times (1-\sigma((\mathbf{x}^{v_k}+\mathbf{z}) - (\mathbf{x}^{v_\ell}+\mathbf{z})))^{1-g_{(v_k,v_\ell)}}\Big]\\
&= \mathbf{P}[\{\mathbf{G}_e\}_{e\in E(A_1)} = \{g_e\}_{e\in E(A_1)}] \times \mathbf{P}[\{\mathbf{G}_e\}_{e\in E(A_2)} = \{g_e\}_{e\in E(A_2)}].
\end{aligned} \tag{18}$$

We used the fact that the vectors $\mathbf{x}^{u_0}, \mathbf{x}^{u_1}, \ldots, \mathbf{x}^{u_k}, \mathbf{x}^{u_0}+\mathbf{z}, \mathbf{x}^{v_1}+\mathbf{z}, \ldots, \mathbf{x}^{v_r}+\mathbf{z}$ are independent.

Item 2. Follows from an inductive application of item 1 and the fact that each edge appears marginally with probability p in $\mathbf{G}$ for the functions $\mathsf{SW}_A, \mathsf{W}_A$.

Item 3. Let T be a spanning tree of A with $V(A)-1$ edges. The simple fact $\mathsf{W}_T(\mathbf{G}) \geq \mathsf{W}_A(\mathbf{G})$ (as edge-indicators are in $[0,1]$) and item 2 give the desired inequality. ■

Appendix B. Signed and Unsigned Weights in L_∞: Theorem 3 and Extensions

The main part of this appendix is Section B.3 where we complete the argument in Section 3. Before that, we set things up by finishing the proof of Lemma 14 in Section B.1 and bounding the unsigned weights of subgraphs in Section B.2. These arguments are relatively straightforward and an impatient reader is welcome to read the statement of Corollary 27 and continue to the more involved Section B.3. Throughout, we use the notation and bounds introduced in Section 3.

B.1. Preliminaries: The Proof of Lemma 14

We now prove the remaining parts 4 - 6 of Lemma 14. Let $\mathbf{H} \sim \mathsf{RGG}(n, \mathbb{T}^1, \mathsf{Unif}, \sigma(x,y)^{1,>}_\lambda, \lambda)$.

Item 4. Suppose that A is not bipartite. Then it has an odd cycle formed by vertices $i_1, i_2, \cdots i_{2k+1}, i_{2k+2} = i_1$ of length $2k+1 \leq \frac{1}{8\lambda}$. We will show that for any latent vectors $\mathbf{x}^{i_1}, \mathbf{x}^{i_2}, \ldots, \mathbf{x}^{i_{2k+1}} \in \mathbb{T}^1$ it is the case that there exists some $t \in [2k+1]$ for which $\sigma(\mathbf{x}^{i_t}, \mathbf{x}^{i_{t+1}})^{1,>}_\lambda = 0$. Indeed, otherwise $|\mathbf{x}^{i_t} - \mathbf{x}^{i_{t+1}}|_C \geq 1-\lambda$ and $|\mathbf{x}^{i_{t+1}} - \mathbf{x}^{i_{t+2}}|_C \geq 1-\lambda$ imply that $|\mathbf{x}^{i_t} - \mathbf{x}^{i_{t+2}}|_C \leq 2\lambda$ holds for each t. However, this means that $|\mathbf{x}^1 - \mathbf{x}^{2k+1}| \leq k\cdot 2\lambda < 1-\lambda$, which means that $\sigma(\mathbf{x}^{i_1}, \mathbf{x}^{i_{2k+1}})^{1,>}_\lambda = 0$.

Item 5. Observe that A has a spanning forest T on $V(A) - \mathsf{numc}(A)$ edges. This gives the bound

$$|\psi(A)| \leq |\chi(A)| + \lambda^{|E(A)|} \leq |\chi(T)| + \lambda^{|E(T)|} = 2\lambda^{|E(T)|} = 2\lambda^{|V(A)| - \mathsf{numc}(A)}.$$

The only remaining case is when $|V(A)| - \mathsf{numc}(A) < |V(A)|/2 + 1$ or, equivalently, $\mathsf{numc}(A) > |V(A)|/2 - 1$. Note, however, that since A is defined by a set of edges, there are no isolated vertices and, so, $\mathsf{numc}(A) \leq |V(A)|/2$. Thus, we have two cases. First, $\mathsf{numc}(A) = |V(A)|/2$, in which case A must be the union of $|V(A)|/2$ disjoint edges, but then $\psi(A) = 0$ by item 2. Or, $\mathsf{numc}(A) = |V(A)|/2 - 1/2$, so A must be the union of a triangle and $(|V(A)|-3)/2$ disjoint edges. In that case, using items 1, 2, and 4, $\chi(A) = 0$, so $\psi(A) = -\lambda^{|E(A)|} = -\lambda^{|V(A)|/2 + 3/2}$.

Item 6. Let C_m be the cycle on m vertices $1, 2, \ldots, m$. Note that whenever (ij) is an edge in $\mathbf{H}$, $x_i = x_j + 1 + \lambda_{ij}$, where $\lambda_{ij} \in [-\lambda, \lambda]$. Using that the path $1, 2, \ldots, m$ is a tree and item 2,

$$\begin{aligned}
&\chi(C_m) = \mathbf{E}[\mathsf{W}_{C_m}(\mathbf{H})] = \mathbf{P}[H_{1,m} = 1, H_{1,2} = 1, \ldots, H_{m-1,m}] \\
&= \mathbf{P}\big[H_{1,m} = 1 \big| H_{1,2} = 1, \ldots, H_{m-1,m}\big] \mathbf{P}\big[H_{1,2} = 1, \ldots, H_{m-1,m}\big] \\
&= \mathbf{P}\big[|x_1 - x_m|_C \geq 1-\lambda \big| x_{i+1} = x_i + 1 + \lambda_{i,i+1}, |\lambda_{i,i+1}| \leq \lambda \forall i\big] \cdot \lambda^{m-1} \\
&= \mathbf{P}\Big[|x_1 - x_m|_C \geq 1-\lambda \Big| x_m = 1 + x_1 + \sum_{i=1}^{m-1} \lambda_{i,i+1}, |\lambda_{i,i+1}| \leq \lambda \forall i\Big] \cdot \lambda^{m-1} \\
&= \mathbf{P}\Big[\sum_{i=1}^{m-1} \lambda_{i,i+1} \in [-\lambda, \lambda] \Big| |\lambda_{i,i+1}| \leq \lambda \forall i\Big] \cdot \lambda^{m-1} = \lambda^{m-1}\phi(m-1).
\end{aligned}$$

This completes the proof.

B.2. Warm-Up: Unsigned Weights of Small Subgraphs

First, we compute the unsigned weight of a cycle as this is the simplest and most important case to our work, in particular used in the proof of Theorem 3.

Lemma 25 *Suppose that* $1 \leq m \leq 1/(8\lambda)$. *Let* $\mathbf{G} \sim \mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$. *Then,*

$$\mathbf{E}\Big[\mathsf{W}_{C_m}(\mathbf{G})\Big] = \begin{cases} p^m\Big(1 + \frac{d\lambda^m}{(1-\lambda)^m} + O\big(d^2\lambda^{2m}\big)\Big) & \text{when } m \text{ is odd,} \\ p^m\Big(1 + \frac{d(\lambda^{m-1}\phi(m-1)-\lambda^m)}{(1-\lambda)^m} + O\big(d^2\lambda^{2(m-1)}\big)\Big) & \text{when } m \text{ is even.} \end{cases} \tag{19}$$

Proof We use (6). Note that any $A \subsetneq C_m$ is acyclic, so $\chi(A) = \lambda^m, \psi(A) = 0$ and

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma_{1-\lambda}^\infty,1-\lambda)}[\mathsf{W}_{C_m}(\mathbf{G})] = (1-\lambda)^m + (-1)^m\psi(C_m).$$

In the odd case C_m is not bipartite and item 4 of Lemma 14 applies, so $\psi(C_m) = -\lambda^m$:

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\Big[\mathsf{W}_{C_m}(\mathbf{G})\Big] = \big((1-\lambda)^m + \lambda^m\big)^d = (1-\lambda)^{md}(1+\lambda^m/(1-\lambda)^m)^d$$
$$= p^m\Big(1 + d\frac{\lambda^m}{(1-\lambda)^m} + \sum_{k=2}^{d}\binom{d}{k}\frac{\lambda^{mk}}{(1-\lambda)^{mk}}\Big).$$

The statement follows as the sum can be bounded by $\sum_{j=2}^\infty (d\lambda^m)^k/(1-\lambda)^{mk}$. Now, clearly, there is exponential decay in the sum as $d\lambda^m/(1-\lambda)^m \leq d\lambda^3/(1-\lambda)^3 = o(1)$. Finally, note that $(1-\lambda)^m \geq (1-\lambda)^{8/\lambda} = \Omega(1)$. The even case is the same, except that we use item 6 of Lemma 14, which gives $\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{W}_{C_m}(\mathbf{G})\big] = \big((1-\lambda)^m + \lambda^{m-1}\phi(m-1) - \lambda^m\big)^d$. ■

Remark 26 *We get arbitrarily better precision in Lemma 25 by keeping* $k \geq 2$ *terms in the expansion of* $(1+\lambda^m/(1-\lambda)^m)^d$.

Corollary 27 *Suppose that* $1 \leq m \leq 1/8\lambda$. *Let* $\mathbf{G} \sim \mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$. *Then,*

$$\mathbf{E}\Big[\mathsf{SW}_{C_m}(\mathbf{G})\Big] = \begin{cases} p^m\Big(\frac{d\lambda^m}{(1-\lambda)^m} + O\big(d^2\lambda^{2m}\big)\Big) & \text{when } m \text{ is odd,} \\ p^m\Big(\frac{d(\lambda^{m-1}\phi(m-1)-\lambda^m)}{(1-\lambda)^m} + O\big(d^2\lambda^{2(m-1)}\big)\Big) & \text{when } m \text{ is even.} \end{cases} \tag{20}$$

Proof Using the definition of SW_{C_m} in (13) and Lemma 24,

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\Big[\mathsf{SW}_{C_m}(\mathbf{G})\Big] = \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\Big[\mathsf{W}_{C_m}(\mathbf{G})\Big] + \sum_{F\subsetneq E(C_m)} (-p)^{m-|E(F)|}\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\Big[\mathsf{W}_F(\mathbf{G})\Big]$$
$$= \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\Big[\mathsf{W}_{C_m}(\mathbf{G})\Big] + \sum_{j=1}^m\binom{m}{j}(-p)^{m-j}p^j = \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\Big[\mathsf{W}_{C_m}(\mathbf{G})\Big] - p^m,$$

which is enough by Lemma 25. ■

Finally, to derive the bounds on the weights of arbitrary subgraphs, we use the truncated inclusion exclusion inequality in place of (4). Namely, for any odd number t,

$$\begin{aligned}&\sum_{A\subseteq E(H)\,:\,|A|\le t}(-1)^{|E(A)|}\chi(A)\le \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}[\mathsf{W}_H(\mathbf{G})]\\ &\le\sum_{A\subseteq E(H)\,:\,|A|\le t+1}(-1)^{|E(A)|}\chi(A).\end{aligned}\tag{21}$$

This yields the following statement.

Lemma 28 *Let $H=\{(i_1,j_1),(i_2,j_2),\dots(i_k,j_k)\}$ be a set of edges and let m be the girth of H (i.e. the length of the shortest cycle). Let $N(u)$ be the number of cycles of length u in H. Suppose further that Assumption (A) holds and $k^{m+2}=o(1/\lambda)=o(d/\log(1/p))$ and let $\phi(u)=\Theta(u^{-1/2})$ be defined as in Lemma 14. Then, for $\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)$,*

$$\Big|\mathbf{E}\Big[\mathsf{W}_H(\mathbf{G})\Big]\Big|=\begin{cases}p^k\Big(1+d\big(N(m)+\phi(m+1)N(m+1)\big)\lambda^m(1+o(1))\Big) & \text{when } m \text{ is odd,}\\ p^k\Big(1+d\phi(m)N(m)\lambda^{m-1}(1+o(1))\Big) & \text{when } m \text{ is even.}\end{cases}\tag{22}$$

The only truly restrictive condition in this theorem is $k^{m+2}=o(1/\lambda)$. However, it still covers a wide range of cases. Indeed, suppose that $d=\mathsf{poly}(n)$. As $k\le m$, it can be applied whenever $k=|E(H)|=o(\log d/\log\log d)$. If, furthermore, m is a constant (say $m\in\{3,4\}$), it can be applied to very large graphs with polynomial number of edges, i.e. $|E(H)|=d^{1/(m+2)-o(1)}$.

Proof

Case 1) The smallest cycle of H is of even size. From (21), we have

$$\sum_{A\subseteq E(H)\,:\,|A|\le m+1}(-1)^{|E(A)|}\chi(A)\le\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{W}_H(\mathbf{G})]\le\sum_{A\subseteq E(H)\,:\,|A|\le m}(-1)^{|E(A)|}\chi(A).\tag{23}$$

First, consider the upper bound. Since each subgraph of H on at most $m-1$ edges is acyclic and there are exactly $N(m)$ cycles on m edges, from Theorem 14,

$$\sum_{A\subseteq H\,:\,|A|\le m}(-1)^{|E(A)|}\chi(A)=\sum_{j=0}^m(-1)^j\binom{|E(H)|}{j}\lambda^j+N(m)\times\Big(\lambda^{m-1}\phi(m-1)-\lambda^m\Big).$$

Similarly, we can carry out the calculation for the lower bound in (23). Note that all $(m+1)$-edge subgraphs of H have one of three structures: 1) Acyclic, in which case $\chi(A)=\lambda^{m+1}$, 2) An m cycle with an extra edge not creating a cycle, in which case $\chi(A)=\lambda^m\phi(m-1)$, 3) An $m+1$ cycle in which case $\chi(A)=0$. In all three cases, importantly, $|\chi(A)|\le\lambda^m$. Altogether,

$$\begin{aligned}&\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}[\mathsf{W}_H(\mathbf{G})]\\ &=\sum_{j=0}^m(-1)^j\binom{|E(H)|}{j}\lambda^j+N(m)\times\Big(\lambda^{m-1}\phi(m-1)-\lambda^m\Big)+O\bigg(\binom{|E(H)|}{m+1}\lambda^m\bigg).\end{aligned}\tag{24}$$

Again, using the truncated principle of inclusion-exclusion,

$$\sum_{j=0}^{m}(-1)^j\binom{|E(H)|}{j}\lambda^j \geq (1-\lambda)^{|E(H)|} \geq \sum_{j=0}^{m+1}(-1)^j\binom{|E(H)|}{j}\lambda^j,$$

so $\sum_{j=0}^{m}(-1)^j\binom{|E(H)|}{j}\lambda^j = (1-\lambda)^{|E(H)|} + O\Big(\binom{|E(H)|}{m+1}\lambda^m\Big)$. Using (24),

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^{\infty}_{1-\lambda},1-\lambda)}[\mathsf{W}_H(\mathbf{G})]\\
&= (1-\lambda)^{|E(H)|} + N(m)\times\Big(\lambda^{m-1}\phi(m-1)-\lambda^m\Big) + O\Bigg(\binom{|E(H)|}{m+1}\lambda^m\Bigg)\\
&= (1-\lambda)^{|E(H)|} + N(m)\times\lambda^{m-1}\phi(m-1) + O\Bigg(\binom{|E(H)|}{m+1}\lambda^m + \binom{|E(H)|}{m}\lambda^m\Bigg),
\end{aligned}$$

where we used the trivial observation that $N(m)\leq\binom{|E(H)|}{m+1}$. Now, using the simple fact that that $\phi(m-1)=\Theta(m^{-1/2})$ (see Lemma 14) and the assumption that $|E(H)|^{m+1}=o(\lambda^{-1})$, one can easily see that the last expression is of order

$$(1-\lambda)^{|E(H)|} + N(m)\times\lambda^{m-1}\phi(m-1)(1+o(1)).$$

Finally,

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma^{\infty}_{p},p)}[\mathsf{W}_H(\mathbf{G})]\\
&= \Bigg((1-\lambda)^{|E(H)|} + N(m)\times\lambda^{m-1}\phi(m-1)(1+o(1))\Bigg)^d\\
&= (1-\lambda)^{d\times|E(H)|}\times\Bigg(1+\frac{N(m)\times\lambda^{m-1}\phi(m-1)(1+o(1))}{(1-\lambda)^{|E(H)|}}\Bigg)^d.
\end{aligned}$$

Since $|E(H)|\leq\lambda^{1/(m+1)}$, clearly $(1-\lambda)^{|E(H)|}=1+o(1)$. Furthermore, $|N(m)\lambda^{m-1}| = O(\binom{|E(H)|}{m}\lambda^{m-1}) = O(\lambda^{m-2}) = O(1/d)$ as $\lambda = O(\log n/d) = o(d^{-1/2})$. Using also the fact that $(1-\lambda)^d=p$, we conclude that

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma^{\infty}_{p},p)}[\mathsf{W}_H(\mathbf{G})] = p^{|E(H)|}\times\Bigg(1+dN(m)\phi(m-1)\lambda^{m-1}\big(1+o(1)\big)\Bigg).$$

Case 2) The smallest cycle of H is of even size. We repeat the same steps as in the even case. The only difference is that when considering cycles of length $m+1$, one needs to take extra care of cycles of length $m+1$ as $\chi(C_{m+1})=\lambda^m\phi(m)$, which is of the same order as $\psi(C_m)=-\lambda^m$.

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^{\infty}_{1-\lambda},1-\lambda)}[\mathsf{W}_H(\mathbf{G})] \leq \sum_{A\subseteq E(H)\,:\,|A|\leq m+1}(-1)^{|E(A)|}\chi(A)\\
&= \sum_{j=0}^{m+1}(-\lambda)^j\binom{|E(H)|}{m+1} + N(m)\lambda^m + N(m+1)(\phi(m)\lambda^m-\lambda^{m+1}) + O\Bigg(\lambda^{m+1}\binom{|E(H)|}{m+1}\Bigg).
\end{aligned} \tag{25}$$

Again, we used that all subgraphs of H on at most m edges are acyclic, except for $N(m)$ isomorphic to C_m. The subgraphs on $m+1$ vertices have one of three structures: 1) Acyclic, in which case $\chi(A) = \lambda^{m+1}$, 2) An m cycle with an extra edge not creating a cycle, in which case $\chi(A) = 0$, 3) An $m+1$ cycle in which case $\chi(A) = \phi(m)\lambda^m$.

Similarly, the subgraphs on $m+2$ vertices have one of four structures: 1) Acyclic, in which case $\chi(A) = \lambda^{m+2}$, 2) An m cycle with two extra edges, in which case $\chi(A) = 0$, 3) An $m+2$ cycle, in which case $\chi(A) = 0$, 4) An $m+1$ cycle with an extra edge, in which case $\chi(A) = \phi(m)\lambda^{m+1}$. In all cases, $|\chi(A)|$ is at most λ^{m+1}. Thus,

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^{\infty}_{1-\lambda},1-\lambda)}[\mathsf{W}_H(\mathbf{G})]\\
&= \sum_{j=0}^{m+1}(-\lambda)^j\binom{|E(H)|}{m+1} + N(m)\lambda^m + N(m+1)\phi(m)\lambda^m +\\
&+ O\Bigg(\binom{|E(H)|}{m+1}\lambda^{m+1} + \binom{|E(H)|}{m+2}\lambda^{m+1}\Bigg)\\
&= (1-\lambda)^{|E(H)|} + \Bigg(N(m)\lambda^m + N(m+1)\phi(m)\lambda^m\Bigg)\big(1+o(1)\big).
\end{aligned}$$

As in the even case, we used $\sum_{j=0}^{m+1}(-1)^j\binom{|E(H)|}{j}\lambda^j = (1-\lambda)^{|E(H)|} + O\Bigg(\binom{|E(H)|}{m+2}\lambda^{m+1}\Bigg)$. The desired conclusion follows as in the even case. ■

B.3. Signed Weights of Small Subgraphs: Theorem 3

We now fill in the details in Section 3 and prove Theorem 3.

Fix H with at most $(\log d)^{5/4}/(\log\log d)$ edges. We also assume H is 2-connected. Otherwise H can be decomposed into two graphs H_1, H_2 which share at most one vertex and no edges. By Lemma 24,

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma^{\infty}_p,p)}\Big[\mathsf{SW}_H(\mathbf{G})\Big] = \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\Big[\mathsf{SW}_{H_1}(\mathbf{G})\Big] \times \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\Big[\mathsf{SW}_{H_2}(\mathbf{G})\Big]$$

and we can induct as $|V(H_1)| + |V(H_2)| \geq |V(H)|, |E(H_1)| + |E(H_2)| \geq |E(H)|$. In particular, the 2-connectivity assumption implies that $|V(H)| \leq |E(H)|$.

We also assume that H has at least 4 edges as the other cases are covered in Lemma 14 (for acyclic graphs, the signed expectation is 0) and Corollary 27 (for triangles, we get $\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\big[\mathsf{SW}_{C_3}(\mathbf{G})\big] = p^3(\log(1/p)/d)^2$).

B.3.1. Proof of Theorem 3 Assuming Lemma 15

We first show how Lemma 15 implies Theorem 3.

Proof Using (9) and the fact that the 0-order deviation vanishes, we compute:

$$
\begin{aligned}
&\Big|\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_H(\mathbf{G})]\Big|\\
&\leq \Big|\sum_{i=0}^{d}\binom{d}{i}(1-\lambda)^{(d-i)|E(H)|}\sum_{A\subseteq H}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\mathsf{Err}(A,\lambda)^i\Big|\\
&\leq \sum_{i=1}^{d}\binom{d}{i}(1-\lambda)^{(d-i)|E(H)|}\Big|\sum_{A\subseteq H}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\mathsf{Err}(A,\lambda)^i\Big|\\
&\leq \sum_{i=1}^{d} d^i(1-\lambda)^{d|E(H)|}(1-\lambda)^{-i|E(H)|}\frac{1}{(4d)^i}\times\Big(\frac{(\log d)^C}{d}\Big)^{|V(H)|/2}\\
&= (1-\lambda)^{d|E(H)|}\Big(\frac{(\log d)^C}{d}\Big)^{|V(H)|/2}\sum_{i=1}^{d}\Big(\frac{1}{4(1-\lambda)^{|E(H)|}}\Big)^i\\
&= p^{|E(H)|}\Big(\frac{(\log d)^C}{d}\Big)^{|V(H)|/2}\sum_{i=1}^{d}\Big(\frac{1}{4(1-\lambda)^{|E(H)|}}\Big)^i.
\end{aligned}
\tag{26}
$$

Now, observe that $(1-\lambda)^{|E(H)|}\geq 1-\lambda|E(H)|\geq 1-O((\log d)^{9/4}/d)\geq 1/2$ for all large enough d. Thus, $4(1-\lambda)^{|E(H)|}>2$ and so $\sum_{i=1}^{d}\Big(\frac{1}{4(1-\lambda)^{|E(H)|}}\Big)^i\leq 1$, which completes the proof. ■

What remains is to prove Lemma 15. As described in Step 6 of Section 3, there are two conceptually different regimes.

B.3.2. PROOF OF THEOREM 15 FOR SMALL VALUES OF i.

Suppose that $i<11|V(H)|/41$.[8] The first step towards proving Lemma 15 is expanding (9).

Detailed Derivation of the Cluster Expansion

$$
\begin{aligned}
&\sum_{A\subseteq H}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\mathsf{Err}(A,\lambda)^i\\
&=\sum_{A\subseteq H}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\Big(\sum_{K\subseteq A}(-1)^{|E(K)|}\psi(K)\Big)^i\\
&=\Bigg(\sum_{K_1,K_2,\dots,K_i\subseteq H}(-1)^{|E(K_1)|+|E(K_2)|+\dots+|E(K_i)|}\psi(K_1)\psi(K_2)\cdots\psi(K_i)\times\\
&\qquad\times\sum_{A\subseteq H\,:\,K_j\subseteq A\;\forall j}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\Bigg).
\end{aligned}
\tag{27}
$$

Let $\mathcal{K}=K_1\cup K_2\cdots K_i$. Then, in the last sum, we perform a summation over all A such that $\mathcal{K}\subseteq A\subseteq H$. In particular, we obtain

8. In principle, any constant in the interval $(1/4,1/2)$ would work for the proof, but constants less than $3/10$ reduce the amount of case work, hence the peculiar choice of $11/41$.

$$\begin{aligned}
&\sum_{A\subseteq H\,:\,K_j\subseteq A\ \forall j} (-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\\
&= \sum_{\mathcal{K}^c\subseteq H\setminus \mathcal{K}} (-1)^{|E(\mathcal{K}^c)|}(1-\lambda)^{i|E(\mathcal{K}^c)|}\\
&= \sum_{t=0}^{|E(H)|-|E(\mathcal{K})|} \binom{|E(H)|-|E(\mathcal{K})|}{t}(-1)^t(1-\lambda)^{it}\\
&= \big(1-(1-\lambda)^i\big)^{|E(H)|-|E(\mathcal{K})|} \leq (\lambda i)^{|E(H)|-|E(\mathcal{K})|},
\end{aligned} \tag{28}$$

where in the last line we used Bernoulli's inequality $(1-\lambda)^i \geq 1-\lambda i$. Now, using (28), we can rewrite the right-hand side of (27) as

$$\sum_{K_1,K_2,\ldots,K_i\subseteq H} (1-(1-\lambda)^i)^{|E(H)|-|E(\mathcal{K})|}(-1)^{|E(K_1)|+|E(K_2)|+\cdots+|E(K_i)|}\psi(K_1)\psi(K_2)\cdots\psi(K_i).$$

Using the triangle-inequality, we bound this quantity by

$$\sum_{K_1,K_2,\ldots,K_i\subseteq H} (1-(1-\lambda)^i)^{|E(H)|-|E(\mathcal{K})|}\Big|\psi(K_1)\psi(K_2)\cdots\psi(K_i)\Big|. \tag{29}$$

Bounds Based on Combinatorial Inequalities. Now, we will bound the quantity $|\psi(K_1)\psi(K_2)\ldots\psi(K_i)|$ in two different ways.

Lemma 29 *The value of* $\big|\psi(K_1)\psi(K_2)\cdots\psi(K_i)\big|$ *is less than each of*

1. $\prod_{j=1}^{i}(2\lambda)^{|V(K_j)|/2+1}$, *and*
2. $(2\lambda)^{|V(\mathcal{K})|-\mathsf{numc}(\mathcal{K})}$.

To prove Lemma 29, we will need the following subadditivity property.

Claim 30 *Suppose that G is a graph and G_1 and G_2 are two (not necessarily induced) subgraphs such that $E(G_1)\cup E(G_2) = E(G)$. Then,* $|V(G)|-\mathsf{numc}(G) \leq |V(G_1)|-\mathsf{numc}(G_1)+|V(G_2)|-\mathsf{numc}(G_2)$.

Proof Let G_1 have a connected components with vertex sets $D_1, D_2, \ldots, D_a$ and let G_2 have b connected components with vertex sets $F_1, F_2, \ldots, F_b$. Consider the bipartite graph $\mathcal{G}$ on parts $\mathcal{D}, \mathcal{F}$ with vertex sets respectively $D_1, D_2, \ldots, D_a$ and $F_1, F_2, \ldots, F_b$. Draw an edge between D_i and F_j if and only if they have a common vertex and, if so, label this edge with one of their common vertices. Clearly, each edge is labeled by a different vertex.

Note that $\mathsf{numc}(G) = \mathsf{numc}(\mathcal{G})$. On the other hand, as each edge is labelled by a different repeated vertex, $|V(G)| \leq |V(G_1)| + |V(G_2)| - |E(\mathcal{G})|$. Trivially,

$$\mathsf{numc}(\mathcal{G}) \geq |V(\mathcal{G})| - |E(\mathcal{G})| = \mathsf{numc}(G_1) + \mathsf{numc}(G_2) - |E(\mathcal{G})|.$$

Combining these, $|V(G_1)| + |V(G_2)| - |V(G)| \geq |E(\mathcal{G})| \geq \mathsf{numc}(G_1) + \mathsf{numc}(G_2) - \mathsf{numc}(G)$. ■

Proof [Proof of Lemma 29] Using part 5 of Lemma 14 on each $\psi(K_i)$ yields the first bound. For the second bound, we again apply part 5 of Lemma 14 on each $\psi(K_i)$ and then we repeatedly apply subadditivity: $\Big|\psi(K_1)\psi(K_2)\cdots\psi(K_i)\Big| \leq \prod_{j=1}^{i}(2\lambda)^{|V(K_j)|-\mathsf{numc}(K_j)} \leq (2\lambda)^{|V(\mathcal{K})|-\mathsf{numc}(\mathcal{K})}$. ■

We now proceed to bound (29) for a fixed fixed i-tuple $K_1, K_2, \ldots, K_i$.

Lemma 31 *The value of* $(1-(1-\lambda)^i)^{|E(H)|-|E(\mathcal{K})|}\Big|\psi(K_1)\psi(K_2)\cdots\psi(K_i)\Big|$ *is less than each of*

1. $(2\lambda)^{i+\sum_{j=1}^{i}|V(K_i)|/2}$, *and*
2. $(2i\lambda)^{|V(H)|-1}$.

We will need yet another combinatorial inequality.

Claim 32 *For any 2-connected graph H and any (not necessarily induced) subgraph $\mathcal{K}$ of H,*

$$|E(H)| - |E(\mathcal{K})| \geq \mathsf{numc}(\mathcal{K}) + |V(H)| - |V(\mathcal{K})| - 1.$$

Proof First, suppose that $\mathcal{K}$ is connected and $V(H) = V(\mathcal{K})$. Then, the right-hand side of the desired inequality equals 0 and the left-hand side is non-negative.

Otherwise, let the connected components of $\mathcal{K}$ be $F_1, F_2, \ldots, F_a$, where $a = \mathsf{numc}(\mathcal{K})$. Consider the multigraph (with multiedges, but no self-loops) H' on $\mathsf{numc}(K) + |V(H)| - V(\mathcal{K})$ vertices $[a] \cup \big(V(H)\backslash V(\mathcal{K})\big)$. In H', two vertices in $V(H)\backslash V(\mathcal{K})$ are adjacent with multiplicity 1 if and only if they are adjacent in H. The multiplicity of an edge between a connected component F_j and a vertex $u \in \big(V(H)\backslash V(\mathcal{K})\big)$ equals the number of neighbours of u in F_j with respect to H. Finally, the multiplicity between F_j and $F_{j'}$ equals the number of edges between them in H.

Clearly, the number of edges (with multiplicities) in H is at most $|E(H)| - |E(\mathcal{K})|$. On the other hand, it must be at least $a + |V(H)\backslash V(\mathcal{K})| = \mathsf{numc}(\mathcal{K}) + |V(H)| - |V(\mathcal{K})|$. Indeed, otherwise there is a vertex of degree (counted with multiplicities) 1 or 0 in H'. If it is of degree 0, clearly H cannot be connected. If it is of degree 1, suppose that the corresponding edge in H is (u, v) and the vertex of degree 1 is either the vertex u or a connected component F_j containing u. Since H is 2-connected, v has at least one more neighbour u_1 other than u in H. In $H|_{V(H)\backslash\{v\}}$, there is no path between u and u_1. This is a contradiction with the 2-connectivity of H. Thus, it must be the case that $|E(H)| - |E(\mathcal{K})| \geq \mathsf{numc}(\mathcal{K}) + |V(H)| - |V(\mathcal{K})|$. ■

Proof [Proof of Lemma 31] The first bound follows directly from part 1 in Lemma 29 and the fact that $|1-(1-\lambda)^i| \leq 1$. For the second bound, we use the above combinatorial inequality to obtain

$$\begin{aligned}
&(1-(1-\lambda)^i)^{|E(H)|-|E(\mathcal{K})|}\Big|\psi(K_1)\psi(K_2)\cdots\psi(K_i)\Big| \\
&\leq (2\lambda)^{|V(\mathcal{K})|-\mathsf{numc}(\mathcal{K})}(i\lambda)^{|V(H)|-|V(\mathcal{K})|+\mathsf{numc}(\mathcal{K})-1} \\
&\leq (2i\lambda)^{|V(H)|-1}.
\end{aligned}$$

■

Bounding the expression in (29) First, note that H has $2^{|E(H)|}$ subgraphs. Thus,

$$\begin{aligned}&\sum_{K_1,K_2,\dots,K_i\subseteq H}(1-(1-\lambda)^i)^{|E(H)|-|E(\mathcal{K})|}\Big|\psi(K_1)\psi(K_2)\cdots\psi(K_i)\Big|\\&=2^{|E(H)|i}\mathbf{E}\Big[(1-(1-\lambda)^i)^{|E(H)|-|E(\mathbf{K})|}|\psi(\mathbf{K}_1)\psi(\mathbf{K}_2)\cdots\psi(\mathbf{K}_i)|\Big],\end{aligned}\tag{30}$$

where each $\mathbf{K}_j$ is sampled independently of the others by independently including each edge of H with probability $1/2$ and $\mathbf{K}=\mathbf{K}_1\cup\mathbf{K}_2\ldots\cup\mathbf{K}_i$.

Case 1.1) First, suppose that $i=1$. We will use the first bound in Lemma 31. We have to show that

$$2^{|E(H)|}(2\lambda)^{|V(H)|-1}\le\frac{1}{4d}\Big(\frac{(\log d)^C}{d}\Big)^{|V(H)|/2}$$

for some absolute constant C. Taking a logarithm on both sides, it is enough to show that

$$\begin{aligned}&|E(H)|+(\log d-C\log\log d)(|V(H)|/2+1)\\&\le(\log d-C'''\log\log d)(|V(H)|-1)\end{aligned}$$

holds, where C''' is the hidden constant in $\log(1/\lambda)=\log d-O(\log\log d)$. If $|E(H)|=o(\log\log d)$, choosing a large enough C, we need to show that

$$(\log d-(C+2)\log\log d)(|V(H)|/2+1)\le(\log d-C'''\log\log d)(|V(H)|-1).$$

This clearly holds for large enough C as $|V(H)|\ge 4$, so $|V(H)|-1\ge|V(H)|/2+1$.

On the other hand, if $|E(H)|=\Omega(\log\log d)$, this means that $|V(H)|=\Omega((\log\log d)^{1/2})$. Thus, for large enough d, the inequality becomes equivalent to

$$|E(H)|\le\frac{1}{2}(\log d)|V(H)|(1-o_d(1)).$$

This clearly holds since $\sqrt{|E(H)|}\le 2|V(H)|$ for any graph H and $\sqrt{|E(H)|}<(\log d)^{5/8}$.

Case 1.2) Now, suppose that $2<i\le\dfrac{3(\log d)|V(H)|}{10(|E(H)|+\log d)}$. In particular, this case is non-trivial if and only if $2\le\dfrac{3(\log d)|V(H)|}{10(|E(H)|+\log d)}$, which implies that $|V(H)|\ge 6$ for large enough values of d. Thus, we assume that $|V(H)|\ge 6$. This, combined with $i\le\dfrac{3(\log d)|V(H)|}{10(|E(H)|+\log d)}$ implies $(2i\lambda)^{|V(H)|-1}2^{i|E(H)|}\le d^{-i-|V(H)|/2}$ for large enough values of d. One concludes from the second bound in Lemma 31 that

$$\begin{aligned}&2^{|E(H)|i}\mathbf{E}\Big[(1-(1-\lambda)^i)^{|E(H)|-|E(\mathbf{K})|}|\psi(\mathbf{K}_1)\psi(\mathbf{K}_2)\cdots\psi(\mathbf{K}_i)|\Big]\\&\le 2^{|E(H)|i}(2i\lambda)^{|V(H)|-1}\\&\le d^{-i-|V(H)|/2}.\end{aligned}\tag{31}$$

Case 1.3) $\frac{3(\log d)|V(H)|}{10(|E(H)|+\log d)} \leq i \leq \frac{11|V(H)|}{41}$. In particular, such i exist if and only if $|E(H)| \geq 13\log d/110 = \Omega(\log d)$. We assume that $|E(H)| = \Omega(\log d)$ in the rest of this case. We will use Lemma 16 which we restate in a slightly different form for convenience.

Lemma 33 *Let $i \geq 2, 0 \leq b \leq i$ be integers and $a > 0$ be a real number. Then,*

$$\mathbf{P}\Big[\sum_{j=1}^{i} |V(\mathbf{K}_j)| \leq ab\Big] \leq \frac{1}{2^{i\times|E(H)|}} \exp\Big(b(\log i) + a^2 i \log|E(H)| + |E(H)|b\Big).$$

Proof Note that

$$\mathbf{P}[|V(\mathbf{K}_1)| \leq a] \leq \mathbf{P}\Big[|E(\mathbf{K}_1)| \leq a^2\Big] \leq \frac{1}{2^{|E(H)|}} \sum_{j=0}^{a^2} \binom{|E(H)|}{j} \leq \frac{|E(H)|^{a^2}}{2^{|E(H)|}}. \tag{32}$$

It follows that

$$\begin{aligned}
&\mathbf{P}\Big[\sum_{j=1}^{i} |V(\mathbf{K}_j)| \leq ab\Big] \\
&\leq \mathbf{P}[\exists j_1 < j_2 < \ldots < j_{i-b} \text{ s.t. } |V(\mathbf{K}_{j_u})| \leq a \ \forall\ u \in [i-b]] \\
&\leq \sum_{1\leq j_1, j_2, \ldots, j_{i-b} \leq i} \prod_{u=1}^{i-b} \mathbf{P}[|V(\mathbf{K}_{j_u})| \leq a] \\
&\leq \binom{i}{i-b} \left(\frac{|E(H)|^{a^2}}{2^{|E(H)|}}\right)^{i-b} \\
&= \frac{1}{2^{i|E(H)|}} \binom{i}{b} |E(H)|^{a^2(i-b)} 2^{|E(H)|b} \\
&\leq \frac{1}{2^{i|E(H)|}} i^b |E(H)|^{ia^2} e^{|E(H)|b} \\
&\leq \frac{1}{2^{i|E(H)|}} \exp\Big(b(\log i) + ia^2 \log|E(H)| + b|E(H)|\Big),
\end{aligned} \tag{33}$$

which finishes the proof. ■

We will apply the claim with the choices

$$a = \frac{|V(H)|^{1/2}(\log d)^{1/2}}{i^{1/2}} (\log\log d)^{-1} \quad \text{and} \quad b = \left\lfloor \frac{|V(H)|\log d}{|E(H)|} (\log\log d)^{-1} \right\rfloor.$$

The condition $b \leq i$ holds for large enough d since $|E(H)| = \Omega(\log d)$ and $\frac{3(\log d)|V(H)|}{10(|E(H)|+\log d)} \leq i$.

Now, we can write

$$
\begin{aligned}
&2^{|E(H)|i}\mathbf{E}\Big[(1-(1-\lambda)^i)^{|E(H)|-|E(\mathbf{K})|}|\psi(\mathbf{K}_1)\psi(\mathbf{K}_2)\cdots\psi(\mathbf{K}_i)|\Big]\\
&\le 2^{|E(H)|i}\mathbf{E}\Big[(1-(1-\lambda)^i)^{|E(H)|-|E(\mathbf{K})|}|\psi(\mathbf{K}_1)\psi(\mathbf{K}_2)\cdots\psi(\mathbf{K}_i)|\;\Big|\;\sum_{j=1}^{i}|V(\mathbf{K}_j)|\le ab\Big]\times\\
&\qquad\qquad \times\mathbf{P}\Big[\sum_{j=1}^{i}|V(\mathbf{K}_j)|\le ab\Big]+\\
&+2^{|E(H)|i}\mathbf{E}\Big[(1-(1-\lambda)^i)^{|E(H)|-|E(\mathbf{K})|}|\psi(\mathbf{K}_1)\psi(\mathbf{K}_2)\cdots\psi(\mathbf{K}_i)|\;\Big|\;\sum_{j=1}^{i}|V(\mathbf{K}_j)|> ab\Big]\times\\
&\qquad\qquad \times\mathbf{P}\Big[\sum_{j=1}^{i}|V(\mathbf{K}_j)|> ab\Big]\\
&\le 2^{|E(H)|i}\frac{1}{2^{i|E(H)|}}\exp\Big(b(\log i)+ia^2\log|E(H)|+b|E(H)|\Big)(2i\lambda)^{|V(H)|-1}\\
&+2^{|E(H)|i}(2\lambda)^{i+ab/2},
\end{aligned}
\tag{34}
$$

where we used the second bound from Lemma 31 in the case $\sum_{j=1}^{i}|V(\mathbf{K}_j)|\le ab$ and the first bound in the case $\sum_{j=1}^{i}|V(\mathbf{K}_j)|> ab$. We now analyze the two terms separately.

Case 1.3.1) We show that

$$\exp\Big(b(\log i)+ia^2\log|E(H)|+b|E(H)|\Big)(2i\lambda)^{|V(H)|-1}\le d^{-i-|V(H)|/2}.$$

This is equivalent to

$$b(\log i)+ia^2\log|E(H)|+b|E(H)|+(i+|V(H)|/2)\log d\le(|V(H)|-1)(\log d-O(\log\log d)).$$

We compare as follows:

1. $b(\log i)+b|E(H)|\le 2b|E(H)|=2|V(H)|\log d/(\log\log d)\le(|V(H)|-1)(\log d-O(\log\log d))/41$ for all large enough d. We used the fact that $i\le|V(H)|\le|E(H)|$ and $|V(H)|\ge|E(H)|^{1/2}=\omega_d(1)$.
2. $ia^2\log|E(H)|=\frac{|V(H)|(\log d)(\log|E(H)|)}{(\log\log d)^2}\le\frac{(|V(H)|-1)(\log d-O(\log\log d))}{41}$ for all large enough d, where we used the fact that $|E(H)|\le(\log d)^{5/4}$, so $\log|E(H)|=O(\log\log d)$.
3. $(i+|V(H)|/2)\log d\le 33(|V(H)|-1)(\log d-O(\log\log d))/41$ for all large enough d as $i\le 11V(H)/41$ and $|V(H)|\ge\sqrt{|E(H)|}=\Omega(\sqrt{\log d})=\omega_d(1)$.

Altogether, this implies that

$$b(\log i)+ia^2\log|E(H)|+b|E(H)|+(i+|V(H)|/2)\log d\le\frac{35}{41}(|V(H)|-1)(\log d-O(\log\log d)),$$

which is enough.

Case 1.3.2) We show that

$$2^{|E(H)|i}(2\lambda)^{i+ab/2} \leq d^{-i-|V(H)|/2}.$$

Bounding $2^{|E(H)|i} \leq e^{|E(H)|i}$ and using $ab = \frac{|V(H)|^{3/2}(\log d)^{3/2}}{i^{1/2}|E(H)|}(\log\log d)^{-2}$, the inequality becomes

$$\begin{aligned}
&\log d(i + |V(H)|/2) + |E(H)|i \\
&\leq (\log d - O(\log\log d))\frac{|V(H)|^{3/2}(\log d)^{3/2}}{i^{1/2}|E(H)|}(\log\log d)^{-2} \\
&+ i(\log d - O(\log\log d)) \iff \\
&O(i\log\log d) + (\log d)|V(H)|/2 + |E(H)|i \\
&\leq (\log d - O(\log\log d))\frac{|V(H)|^{3/2}(\log d)^{3/2}|}{2i^{1/2}|E(H)|}(\log\log d)^{-2}
\end{aligned}$$

We now handle the terms separately.

1. $O(i\log\log d) \leq \frac{1}{3}(\log d - O(\log\log d))\frac{|V(H)|^{3/2}(\log d)^{3/2}}{2i^{1/2}|E(H)|}(\log\log d)^{-2}$. For large enough d, it is enough to show that

$$8i^{3/2}|E(H)|(\log\log d)^3 \leq (\log d)^{5/2}|V(H)|^{3/2}.$$

This clearly holds as $|E(H)| \leq (\log d)^{5/4}, i \leq 11|V(H)|/41$.

2. $(\log d)|V(H)|/2 \leq \frac{1}{3}(\log d - O(\log\log d))\frac{|V(H)|^{3/2}(\log d)^{3/2}}{2i^{1/2}|E(H)|}(\log\log d)^{-2}$. Again, for large enough d, it is enough to show that

$$4(\log d)|V(H)|i^{1/2}|E(H)|(\log\log d)^2 \leq |V(H)|^{3/2}(\log d)^{5/2}.$$

Again, this holds as $i^{1/2} \leq |V(H)|^{1/2}$ and $|E(H)| \leq (\log d)^{5/4}$.

3. $|E(H)|i \leq \frac{1}{3}(\log d - O(\log\log d))\frac{|V(H)|^{3/2}(\log d)^{3/2}|}{2i^{1/2}|E(H)|}(\log\log d)^{-2}$. For large enough d, it is sufficient to show that

$$8|E(H)|^2 i^{3/2}(\log\log d)^2 \leq (\log d)^{5/2}|V(H)|^{3/2}.$$

Again, this holds as $|V(H)| \geq 3i, |E(H)| \leq (\log d)^{5/4}/(\log\log d)$.[9]

B.3.3. PROOF OF LEMMA 15 FOR LARGE VALUES OF i.

Suppose that $d \geq i \geq 11|V(H)|/41$. The main idea behind proving Lemma 15 in that case is to bound each term $\mathsf{Err}(A,\lambda)$ and then sum over the $2^{|E(H)|}$ subgraphs of H.

Claim 34 *If* $|E(A)| \leq (\log d)^{5/4}/(\log\log d)$, *then* $|\mathsf{Err}(A,\lambda)| \leq d^{-3+o_d(1)}$.

9. This is the only place in the proof where we need $|E(H)| \ll (\log d)^{5/4}$ rather than $|E(H)| \ll (\log d)^2$. Improving Lemma 33 would, potentially, improve the result for polynomials of degree up to $(\log d)^{2-\epsilon}$ for any constant $\epsilon > 0$.

Proof We first prove the statement in the case when $|V(A)| \leq |E(A)|$. Note that $|\psi(K)| \leq \lambda^3$ when $|V(K)| \leq 3$ and $|\psi(K)| \leq 2 \times \lambda^{|V(K)|/2+1}$ otherwise by Lemma 14.

$$
\begin{aligned}
|\mathsf{Err}(A,\lambda)| &= \sum_{E(K)\subseteq E(A)} (-1)^{|E(K)|}\psi(K) \leq \sum_{E(K)\subseteq E(A)} |\psi(K)| \\
&= \sum_{K\colon |V(K)|\leq 3} \lambda^3 + 2\lambda \sum_{(\log d)^{5/7}\geq |V(K)|>3} \lambda^{|V(K)|/2} + 2\lambda \sum_{|V(A)|\geq |V(K)|>(\log d)^{5/7}} \lambda^{|V(K)|/2}
\end{aligned}
\tag{35}
$$

We analyse the three sums separately. The constant $5/7$ is chosen arbitrarily in $(5/8, 1)$.

Case 1) $|V(K)| \leq 3$. There are $O(|V(A)|^3) = O((\log d)^{15/4})$ subgraphs of A on at most three vertices. Thus,

$$\sum_{K\colon |V(K)|\leq 3} \lambda^3 = O\Big((\log d)^{15/4}(\log d/d)^3\Big) = d^{-3+o_d(1)}.$$

Case 2) $(\log d)^{5/7} \geq |V(K)| > 3$. When $|V(K)| = t$, one can choose $V(K)$ in $\binom{|V(A)|}{t}$ ways and, once $V(K)$ is chosen, choose $E(A)$ in $2^{\binom{t}{2}}$ ways at most. This leads to

$$
\begin{aligned}
&2\lambda \sum_{(\log d)^{5/7}\geq |V(K)|>3} \lambda^{|V(A)|/2} \\
&\leq 2\lambda \sum_{(\log d)^{5/7}\geq t>3} \binom{|V(A)|}{t} 2^{\binom{t}{2}} \lambda^{t/2} \\
&\leq 2\lambda \sum_{(\log d)^{5/7}\geq t>3} \Big(\frac{e^2|V(A)|^2 e^t \lambda}{t^2}\Big)^{t/2}
\end{aligned}
\tag{36}
$$

Each value $e^2|V(A)|^2 2^t\lambda/t^2$ is bounded by

$$e^2((\log d)^{5/4})^2 e^{(\log d)^{5/7}}\lambda = O\Big((\log d)^{5/2} \times d^{o(1)} \times (\log d) \times d^{-1}\Big) = d^{-1+o_d(1)},$$

where we used the fact that $5/7 < 1$. As each exponent $t/2$ is at least 2, the sum is bounded by

$$2\lambda(\log d)^{5/7} \times (d^{-1+o_d(1)})^2 = \lambda d^{-2+o_d(1)} = d^{-3+o_d(1)}.$$

Case 3) $|V(A)| \geq |V(K)| \geq (\log d)^{5/7}$. Note that when $|V(K)| = t$, one can choose $V(K)$ in $\binom{|V(A)|}{t} \leq \binom{(\log d)^{5/4}}{t}$ ways and, once $V(K)$ is chosen, choose $E(A)$ in $\sum_{j=0}^{(\log d)^{5/4}} \binom{\binom{t}{2}}{j} \leq$

$(\log d)^{5/4}\binom{\binom{t}{2}}{(\log d)^{5/4}}$ ways at most (as A and, thus, K has at most $(\log d)^{5/4}$ edges). This leads to

$$\begin{aligned}
&\sum_{(\log d)^{5/4}\geq |V(K)|>(\log d)^{5/7}} \lambda^{|V(K)|/2}\\
&\leq (\log d)^{5/4} \sum_{(\log d)^{5/4}\geq t>(\log d)^{5/7}} \binom{(\log d)^{5/4}}{t}\binom{\binom{t}{2}}{(\log d)^{5/4}}\lambda^{t/2}\\
&\leq (\log d)^{5/4} \sum_{(\log d)^{5/4}\geq t>(\log d)^{5/7}} \Big(\frac{e(\log d)^{5/4}}{t}\Big)^{t}\Big(\frac{et^2}{(\log d)^{5/4}}\Big)^{(\log d)^{5/4}}\lambda^{t/2} \qquad (37)\\
&= (\log d)^{5/4}e^{(\log d)^{5/4}} \sum_{(\log d)^{5/4}\geq t>(\log d)^{5/7}} \Big(\frac{t^2}{(\log d)^{5/4}}\Big)^{(\log d)^{5/4}-t}(\lambda e^2t^2)^{t/2}\\
&\leq (\log d)^{5/4}e^{(\log d)^{5/4}} \sum_{(\log d)^{5/4}\geq t>(\log d)^{5/7}} ((\log d)^{5/4})^{(\log d)^{5/4}}(\lambda e^2t^2)^{t/2}.
\end{aligned}$$

Now, consider the expression $(\log d)^{5/4}e^{(\log d)^{5/4}}((\log d)^{5/4})^{(\log d)^{5/4}}(\lambda e^2t^2)^{t/2}$. It can be rewritten as

$$\begin{aligned}
&\exp\Big(O(\log\log d) + ((\log d)^{5/4}+1)\log((\log d)^{5/4}) + t(\log t+1) + (t/2)\log\lambda\Big)\\
&= \exp\Big(O((\log d)^{5/4}\log\log d) - \Omega((\log d\log d^{5/7})\Big)\\
&= \exp\Big(-\Omega(\log d^{1+5/7})\Big),
\end{aligned}$$

since $5/4 < 1+5/7$. Since each of the $O((\log d)^{5/4})$ summands is of order $\exp(-\Omega(\log d^{12/7}))$, the sum is clearly of order $\exp(-\Omega(\log d^{12/7})) = O(d^{-3})$.

Combining the two cases, we obtain that for graphs A satisfying $|V(A)| \leq |E(A)|$, $|\mathsf{Err}(A,\lambda)| \leq d^{-3+o_d(1)}$ as desired.

Now, suppose that $|V(A)| > |E(A)|$. If A is acyclic, by Lemma 14 $|\mathsf{Err}(A,\lambda)| = 0$ as $\psi(K) = 0$ for all $K\subseteq H$ as subgraphs are also acyclic. If A is not acyclic, then, it can be partitioned into two vertex-disjoint graphs $A = A_1\cup A_2$, where A_1 satisfies $|V(A_1)|\leq |E(A_1)|$ and A_2 is acyclic. As in the proof of Lemma 24, this implies that

$$\begin{aligned}
&|\mathsf{Err}(A,\lambda)| = \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}\Big[\mathsf{W}_A(\mathbf{G})\Big] - (1-\lambda)^{|E(A)|}\\
&= \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}\Big[\mathsf{W}_{A_1}(\mathbf{G})\Big]\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}\Big[\mathsf{W}_{A_2}(\mathbf{G})\Big]\\
&\qquad - (1-\lambda)^{|E(A_1)|}(1-\lambda)^{|E(A_2)|} \qquad (38)\\
&= \Big(\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}\Big[\mathsf{W}_{A_2}(\mathbf{G})\Big] - (1-\lambda)^{|E(A_2)|}\Big)\times\\
&\times \qquad \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}\Big[\mathsf{W}_{A_1}(\mathbf{G})\Big] \qquad (39)\\
&\qquad + \Big(\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}\Big[\mathsf{W}_{A_1}(\mathbf{G})\Big] - (1-\lambda)^{|E(A_1)|}\Big)(1-\lambda)^{|E(A_2)|}\\
&= \mathsf{Err}(A_2,\lambda)\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^1,\mathsf{Unif},\sigma^\infty_{1-\lambda},1-\lambda)}\Big[\mathsf{W}_{A_2}(\mathbf{G})\Big] + \mathsf{Err}(A_1,\lambda)(1-\lambda)^{|E(A_2)|}.
\end{aligned}$$

Since A_2 is acyclic, $\mathsf{Err}(A_2, \lambda) = 0$, so thee first term vanishes. Thus,

$$|\mathsf{Err}(A, \lambda)| = |\mathsf{Err}(A_1, \lambda)(1-\lambda)^{|E(A_2)|}| \le |\mathsf{Err}(A_1, \lambda)| = d^{-3+o_d(1)}.$$

■

Since the graph H has at most $2^{|E(H)|}$ subgraphs, the left-hand side in Lemma 15 can be bounded as

$$\begin{aligned}
&\Big|\sum_{A\subseteq H}(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\mathsf{Err}(A,\lambda)^i\Big| \\
&\le \sum_{A\subseteq H}\Big|(-1)^{|E(H)|-|E(A)|}(1-\lambda)^{i(|E(H)|-|E(A)|)}\mathsf{Err}(A,\lambda)^i\Big| \\
&\le 2^{|E(H)|}d^{-i(3+o_d(1))}.
\end{aligned} \tag{40}$$

To prove Lemma 15, it is enough to show that $e^{|E(H)|}d^{-i(3+o_d(1))} \le \dfrac{1}{(4d)^i} \times \Big(\dfrac{(\log d)^C}{d}\Big)^{|V(H)|/2}$. This would follow from

$$|E(H)| - 3(1-o_d(1))i\log d \le -i(\log d + 2) - (\log d)|V(H)|/2$$

or, equivalently,

$$|E(H)| + 2i + (\log d)|V(H)|/2 \le 2i \log d.$$

We analyse each of the terms separately:

1. $|E(H)| \le 2i(\log d)/(\log d)^{3/8}$ for large enough i. Indeed, this follows since

$$\begin{aligned}
|E(H)| &\le \sqrt{|V(H)|^2} \times \sqrt{(\log d)^{5/4}/(\log\log d)} \\
&= |V(H)|(\log d)^{5/8} \times (\log\log d)^{-1/2} \le 2i(\log d)/(\log d)^{3/8}.
\end{aligned}$$

   The last inequality holds for all large enough d since $i \ge 11|V(H)|/41$.

2. $i \le 2i(\log d)/(2\log d)$.

3. $(\log d)|V(H)|/2 \le 2i(\log d) \times \frac{41}{44}$ since $i \ge 11|V(H)|/41$.

Altogether, this gives

$$|E(H)| + 2i + (\log d)|V(H)|/2 \le 2i\log d\Big(\frac{1}{(\log d)^{3/8}} + \frac{1}{2\log d} + \frac{41}{44}\Big) \le 2i\log d$$

for large enough d. With this, the case for large i is also completed and so is the proof of Theorem 3.

B.4. Comparison of Signed and Unsigned Weights

We end with a brief comparison of our bounds on signed and unsigned weights. In particular, it demonstrates why the signed weight bounds are much more subtle and require a more sophisticated argument.

Example 2 *Consider any connected graph H with girth m much smaller than its number of vertices (note that an overwhelming fraction of the connected graphs have girth of constant size). Then, by Theorem 3 and Lemma 28 for $\mathbf{G} \sim \mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$,*

$$\Big|\mathbf{E}[\mathsf{W}_H(\mathbf{G})] - p^{|E(H)|}\Big| = \Omega\bigg(p^{|E(H)|}\Big(\frac{\mathsf{polylog}(d)}{d}\Big)^{m-1}\bigg)$$
$$\gg \Theta\bigg(p^{|E(H)|}\Big(\frac{\mathsf{polylog}(d)}{d}\Big)^{|V(H)|/2}\bigg) \geq \Big|\mathbf{E}[\mathsf{SW}_H(\mathbf{G})]\Big|.$$

In particular, this shows that the much more elementary bound obtained by plugging Lemma 28 into

$$\Big|\mathbf{E}[\mathsf{SW}_H(\mathbf{G})]\Big| = \Big|\sum_{A \subseteq E(H)} (-p)^{|E(H)|-|E(A)|}(\mathbf{E}[\mathsf{W}_A(\mathbf{G})] - p^{|E(A)|})\Big|$$
$$\leq \sum_{A \subseteq E(H)} \Big|(-p)^{|E(H)|-|E(A)|}(\mathbf{E}[\mathsf{W}_A(\mathbf{G})] - p^{|E(A)|})\Big|$$

is wildly suboptimal. Thus, a more refined perturbative analysis as in Section 3 is needed.

Appendix C. Inference and Estimation with Low-Degree Polynomials in L_∞

We now present the proofs of Theorems 5, 6 and 7. The arguments are standard applications of Theorem 3 and Corollary 27.

C.1. Detection Upper Bounds: Theorem 5

Observe that K_n has $\Theta(n^3)$ subgraphs isomorphic to C_3 and $\Theta(n^4)$ subgraphs isomorphic to C_4. From Corollary 27, we conclude that

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\Big[\mathsf{SC}_{C_3}(\mathbf{G})\Big] = \tilde{\Theta}(n^3p^3/d^2),\\
&\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\Big[\mathsf{SC}_{C_4}(\mathbf{G})\Big] = \tilde{\Theta}(n^4p^4/d^2).
\end{aligned} \tag{41}$$

Clearly, $\mathbf{E}_{\mathbf{K}\sim\mathsf{G}(n,p)}\Big[\mathsf{SC}_{C_3}(\mathbf{K})\Big] = \mathbf{E}_{\mathbf{K}\sim\mathsf{G}(n,p)}\Big[\mathsf{SC}_{C_4}(\mathbf{K})\Big] = 0$. We now need to compute the respective variances as in Definition 22.

Triangles. With respect to both the $\mathsf{G}(n,p)$ and $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$ distributions, one can expand the variance as follows (e.g. Liu and Rácz (2023a)). Denote by $\triangle(i,j,k)$ the labelled triangle on vertices i, j, k. Then, taking into account the different possible overlap patterns of two triangles,[10]

$$\begin{aligned}&\mathbf{Var}[\mathsf{SC}_{C_3}(\mathbf{H})]\\&= \Theta(n^3)\times \mathbf{Var}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{H})]\\&+\Theta(n^4)\times \mathbf{Cov}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{H}),\mathsf{SW}_{\triangle(1,2,4)}(\mathbf{H})]\\&+\Theta(n^5)\times \mathbf{Cov}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{H}),\mathsf{SW}_{\triangle(1,4,5)}(\mathbf{H})]\\&+\Theta(n^6)\times \mathbf{Cov}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{H}),\mathsf{SW}_{\triangle(4,5,6)}(\mathbf{H})]].\end{aligned}\tag{42}$$

The product of any two signed weights of subgraphs can be naturally decomposed as a (weighted) sum of signed weights of subgraphs. Thus, we can bound the above expression via Theorem 3 and Corollary 27. We take this approach in Section C.4.1 to show the following.

$$\begin{aligned}&\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SC}_{C_3}(\mathbf{K})]=\Theta(n^3p^3),\text{ and}\\&\mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{SC}_{C_3}(\mathbf{G})]=\Theta(n^3p^3)+\tilde{O}(n^4p^5/d^2).\end{aligned}\tag{43}$$

This is enough to complete part 2 of Theorem 5. According to Definition 22, one can distinguish between $\mathsf{G}(n,p)$ and $\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)$ with high probability using the signed triangle test if and only if

$$\Big|\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\mathsf{SC}_{C_3}(\mathbf{G})\Big|=\omega\Big(\sqrt{\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SC}_{C_3}(\mathbf{K})]+\mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SC}_{C_3}(\mathbf{G})]}\Big).$$

Using (41) and (43), this holds if and only if $d=\tilde{o}((np)^{3/4})$.

4-Cycles. Similarly, in the case of 4-cycles, one obtains

$$\begin{aligned}&\mathbf{Var}[\mathsf{SC}_{C_4}(\mathbf{H})]=\Theta(n^4)\times\mathbf{Var}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H})]\\&+\Theta(n^5)\times\Big(\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(1,2,3,5)}(\mathbf{H})])+\\&+\qquad +\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(1,2,4,5)}(\mathbf{H})]\Big)\\&+\Theta(n^6)\times\Big(\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(1,2,5,6)}(\mathbf{H})]+\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(1,5,2,6)}(\mathbf{H})]\\&\qquad+\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(1,5,3,6)}(\mathbf{H})]\Big)\\&+\Theta(n^7)\times\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(1,5,6,7)}(\mathbf{H})]\\&+\Theta(n^8)\times\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(6,6,7,8)}(\mathbf{H})].\end{aligned}\tag{44}$$

Similarly, we show in Section C.4.2, that

$$\begin{aligned}&\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SC}_{C_4}(\mathbf{K})]=\Theta(n^4p^4)\text{ and,}\\&\mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{SC}_{C_4}(\mathbf{G})]=\Theta(n^4p^4)+\tilde{O}(n^5p^6/d^2+n^6p^7/d^3).\end{aligned}\tag{45}$$

Again,

$$\Big|\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}\mathsf{SC}_{C_4}(\mathbf{G})\Big|=\omega\Big(\sqrt{\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SC}_{C_4}(\mathbf{K})]+\mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SC}_{C_4}(\mathbf{G})]}\Big)$$

holds if and only if $d=\tilde{o}(np)$.

10. Abusing notation, we write $\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{H})$ for the signed weight of the triangle on labelled vertices $1,2,3$. Similarly, $\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H})$ stands for the signed weight of a 4-cycle on labelled vertices $1,2,3,4$.

C.2. Dimesnion Estimation Upper Bounds: Theorem 7

Low degree polynomial statistics are used in the literature not only for testing, but also for estimation (see, for example, Schramm and Wein (2022)). We illustrate with the concrete example of using signed cycles for estimating the dimension of $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$ as in (**P2**).

Suppose that m is a small odd number. The expected signed count of m-cycles is[11] is

$$\frac{(m-1)!}{2}\binom{n}{m} \times \Big(p^m(d\lambda^m/(1-\lambda)^m) + O(p^m d^2\lambda^{2m})\Big)$$

by Corollary 27. Therefore, one can estimate λ from the number of signed m-cycles. Under a sufficiently strong concentration of the number of signed m-cycles, this could allow one to estimate d as $\lambda \approx (\log 1/p)/d$. Similarly, one can perform this for small even numbers. We define the success of a low-degree polynomial test for estimating a parameter (in our case, the dimension) in analogy to Definition 22.

Definition 35 (Success of Polynomial Statistics for Exact Estimation) *Given is a family of random graph distributions $(\mathcal{D}_\theta)_{\theta\in\mathcal{A}}$ over n vertices indexed by a parameter θ taking values in $\mathcal{A}$. Let $f(\cdot)$ be a polynomial in the edges of an n-vertex graph. For each $\theta \in \mathcal{A}$, let $\mathcal{M}_\theta \coloneqq \mathbf{E}_{\mathbf{G}\sim\mathcal{D}_\theta}[f(\mathbf{G})]$. We say that polynomial $f(\cdot)$ succeeds with high probability on exactly recovering θ if the following property holds. There exists some collection of values $\big\{\mathcal{V}_\theta\big\}_{\theta\in\mathcal{A}}$ such that the intervals $\big\{\big[\mathcal{M}_\theta - \mathcal{V}_\theta, \mathcal{M}_\theta + \mathcal{V}_\theta\big], \theta \in \mathcal{A}\big\}$ are disjoint and $\mathcal{V}_\theta = \omega(\mathbf{Var}_{\mathbf{G}\sim\mathcal{D}_\theta}[f(\mathbf{G})]^{1/2})$ for each θ. If, on the other hand, no such intervals exist, we say that the polynomial f fails in the task of exact estimation.*

The interpretation of this definition is simple. Suppose that the true parameter is θ'. Then, by Chebyshev's, inequality with high probability over $\mathbf{G}\sim\mathcal{D}_{\theta'}$, it is the case that $f(\mathbf{G}) \in \big[\mathcal{M}_{\theta'} - \mathcal{V}_{\theta'}, \mathcal{M}_{\theta'} + \mathcal{V}_{\theta'}\big]$. If the intervals are disjoint, this is the unique interval of the form $\big[\mathcal{M}_\theta - \mathcal{V}_\theta, \mathcal{M}_\theta + \mathcal{V}_\theta\big]$ with this property and, thus, one can find θ'. It must be noted that this is the implicit definition used in Bubeck et al. (2014); Friedrich et al. (2023b) for estimating the dimension of random geometric graph models.

To apply this definition to (**P2**), we use the variance bounds (43) and (45) and the following simple estimate of λ, deferred to Section C.4.3

Lemma 36 *Suppose that $d = \omega(\log 1/p)$. Then,*

$$\lambda_p^\infty = 1 - p^{1/d} = \frac{\log 1/p}{d} - \frac{1}{2}\bigg(\frac{\log 1/p}{d}\bigg)^2 + O\Bigg(\bigg(\frac{\log 1/p}{d}\bigg)^3\Bigg),$$

$$\lambda_p^\infty/(1-\lambda_p^\infty) = p^{-1/d} - 1 = \frac{\log 1/p}{d} + \frac{1}{2}\bigg(\frac{\log 1/p}{d}\bigg)^2 + O\Bigg(\bigg(\frac{\log 1/p}{d}\bigg)^3\Bigg).$$

We are now ready to evaluate the intervals in which the signed triangle and 4-cycle statistics succeed with high probability in the exact dimension recovery tasks.

11. The factor $\frac{(m-1)!}{2}\binom{n}{m}$ is the number of undirected m-cycle subgraphs of K_n.

Triangles. Using Corollary 27 and Lemma 36, the expected signed count of three cycles in dimension d is

$$\mathcal{M}_d^{C_3} := \mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{SC}_{C_3}(\mathbf{G})\big] = \binom{n}{3}p^3 \times \left(d\left(\frac{\lambda}{1-\lambda}\right)^3 + O(d^2\lambda^6)\right)$$
$$= \binom{n}{3}p^3 \times \left(\frac{(\log 1/p)^3}{d^2} + \frac{3}{2}\times\frac{(\log 1/p)^4}{d^3} + O\left(\frac{\log(1/p)^5}{d^4}\right)\right).$$

In particular, this means that

$$\begin{aligned}
&\mathcal{M}_d^{C_3} - \mathcal{M}_{d+1}^{C_3} = \mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{SC}_{C_3}(\mathbf{G})\big] - \mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^{d+1},\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{SC}_{C_3}(\mathbf{G})\big]\\
&= \binom{n^3}{p^3}\left(\Theta\left(\frac{(\log(1/p)^3)}{d^3}\right) + O\left(\frac{(\log(1/p)^4)}{d^4}\right)\right)\\
&= \Theta(n^3p^3(\log 1/p)^3/d^3).
\end{aligned} \tag{46}$$

In particular, $\mathcal{M}_{d+1}^{C_3} \leq \mathcal{M}_d^{C_3}$ when $d = \omega(\log 1/p)$. Therefore, numbers $\mathcal{V}_d$ with the desired property from Definition 35 exist if and only if for all $d \in [\omega(\log 1/p), M]$,

$$\begin{aligned}
&\mathcal{M}_d^{C_3} - \mathcal{M}_{d+1}^{C_3}\\
&= \omega\left(\sqrt{\mathbf{Var}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{SC}_{C_3}(\mathbf{G})\big] + \mathbf{Var}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbf{T}^{d+1},\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{SC}_{C_3}(\mathbf{G})\big]}\right).
\end{aligned}$$

Using (43) and (46), this is equivalent to

$$n^3p^3(\log 1/p)^3/d^3 = \tilde{\omega}\left(\sqrt{n^3p^3 + n^4p^5/d^2}\right).$$

One can easily check that this is satisfied if and only if $d = \tilde{o}((np)^{1/2})$.

4-Cycles. In the exact same way we conclude from Corollary 27 and Lemma 36[12]

$$\begin{aligned}
&\mathcal{M}_d^{C_4} = \mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{SC}_{C_4}(\mathbf{G})\big]\\
&= 3\binom{n}{4}p^4\left(\phi(3)\frac{(\log 1/p)^3}{d^2} + \frac{3}{2}\phi(3)\frac{(\log 1/p)^4}{d^3} - \frac{(\log 1/p)^4}{d^3} + O\left(\frac{(\log 1/p)^5}{d^4}\right)\right)\\
&= 3\binom{n}{4}p^4\left(\frac{2}{3}\frac{(\log 1/p)^3}{d^2} + O\left(\frac{(\log 1/p)^5}{d^4}\right)\right).
\end{aligned}$$

Thus, $0 \leq \mathcal{M}_d^{C_4} - \mathcal{M}_{d+1}^{C_4} = \Theta(n^4p^4(\log 1/p)^3/d^3)$. Finally, by (45), the condition

$$\begin{aligned}
&\mathcal{M}_d^{C_4} - \mathcal{M}_{d+1}^{C_4}\\
&= \omega\left(\sqrt{\mathbf{Var}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{SC}_{C_4}(\mathbf{G})\big] + \mathbf{Var}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbf{T}^{d+1},\mathsf{Unif},\sigma_p^\infty,p)}\big[\mathsf{SC}_{C_4}(\mathbf{G})\big]}\right)
\end{aligned}$$

12. Also, from Lemma 14 we recall $\phi(3) = 2/3$, even though the exact value of $\phi(3)$ is irrelevant as long as it is non-zero.

is equivalent to

$$n^4p^4(\log 1/p)^3/d^3 = \tilde{\omega}\Big(\sqrt{n^4p^4 + n^5p^6/d^2 + n^6p^7/d^3}\Big).$$

One can easily check that this is satisfied if and only if $d = \tilde{o}((np)^{2/3})$.

C.3. Detection Lower Bounds: Theorem 6

The proof follows a standard procedure for bounding $\mathsf{ADV}^2_{\leq D}$, e.g in Hopkins (2018).

Suppose that $n^{-1+\epsilon} \leq p \leq 1/2$ and $d \geq np$ for some absolute constant ϵ. In particular, this means that $d \geq n^\epsilon$ and $d \geq p^{-\delta}$ for some absolute constant $\delta > 0$.

Let $D = (\log d)^{5/4}/(\log\log d) = \Theta(\log n/\log\log n)$. Consider the orthonormal basis of $\mathsf{G}(n,p)$ given by the polynomials $\mathsf{p}_H(\cdot) := \mathsf{SW}_H(\cdot)/(p(1-p))^{|E(H)|/2}$ for all subgraphs H of K_n. From Section A.2, we know that to show statistical indistinguishability with respect to degree D polynomials, we simply need to prove the inequality

$$\mathsf{ADV}^2_{\leq D} - 1 := \sum_{H\,:\,1\leq|E(H)|\leq D} \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma^\infty_p,p)}[\mathsf{p}_H(\mathbf{G})]^2 = o(1).$$

We prove this as follows. First, note that if H has a vertex of degree 1, then $\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{p}_H(\mathbf{G})] = 0$ as in Lemma 24. Thus, we can assume that there is no such vertex and, so, $3 \leq |V(H)| \leq |E(H)|$. Using Theorem 3, we obtain the following inequality. In it, we use the standard trick of considering two cases depending on the number of vertices in Fourier coefficients (e.g. Hopkins (2018)).

$$\begin{aligned}
&\sum_{H\,:\,3\leq|E(H)|\leq D} \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{p}_H(\mathbf{G})]^2\\
&= \sum_{H\,:\,3\leq|E(H)|\leq D} \frac{1}{(p(1-p))^{|E(H)|}}\mathbf{E}[\mathsf{SW}_H(\mathbf{G})]^2\\
&\leq \sum_{H\,:\,3\leq|E(H)|\leq D} \frac{1}{(p(1-p))^{|E(H)|}}p^{2|E(H)|}((\log d)^C/d)^{|V(H)|}\\
&\leq \sum_{H\,:\,3\leq|E(H)|\leq D} (2p)^{|E(H)|}((\log d)^C/d)^{|V(H)|}\\
&= \sum_{H\,:\,3\leq|E(H)|,\,|V(H)|\leq D^{2/3}} (2p)^{|V(H)|}((\log d)^C/d)^{|V(H)|}\\
&\qquad + \sum_{H\,:\,3\leq|E(H)|<D,\,|V(H)|>D^{2/3}} (2p)^{|V(H)|}((\log d)^C/d)^{|V(H)|}.
\end{aligned} \tag{47}$$

We used the fact that $1-p \geq 1/2$ and $|V(H)| \leq |E(H)|$. Now, we consider the two sums separately.

Case 1) $|V(H)| \leq D^{2/3}$. When $|V(H)| = t$, there are $\binom{n}{t}$ ways to choose $V(H)$ and then, once $V(H)$ is chosen, at most $2^{\binom{t}{2}}$ ways to choose $E(H)$. This gives

$$\begin{aligned}
&\sum_{H\,:\,0<|E(H)|,\,|V(H)|\leq D^{2/3}} (2p)^{|V(H)|}((\log d)^C/d)^{|V(H)|}\\
&\leq \sum_{t=3}^{D^{2/3}} \binom{n}{t}(2p)^t 2^{\binom{t}{2}}((\log d)^C/d)^t\\
&\leq \sum_{t=3}^{D^{2/3}} \left(\frac{np2^t(\log d)^C}{d}\right)^t\\
&\leq \sum_{t=3}^{D^{2/3}} \left(\frac{np2^{(\log d)^{5/6}}(\log d)^C}{d}\right)^t.
\end{aligned}$$

Clearly, if $d = \max\Big((np)^{1+o_n(1)}), \omega_n(1)\Big)$, one has

$$\left(\frac{np2^{(\log d)^{5/6}}(\log d)^C}{d}\right) = o(1).$$

Thus, there is exponential decay in the sum and it is of order $o(1)$.

Case 2) $|V(H)| \geq D^{2/3}$. When $|V(H)| = t$, there are $\binom{n}{t}$ ways to choose $V(H)$ and then, once $V(H)$ is chosen, at most

$$\sum_{j=0}^{D} \binom{\binom{t}{2}}{j} \leq 2\binom{\binom{t}{2}}{|D|}$$

ways to choose $E(H)$. This gives

$$\begin{aligned}
&\sum_{H\,:\,|E(H)|\le D,\,|V(H)|>D^{2/3}} (2p)^{|V(H)|}((\log d)^C/d)^{|V(H)|}\\
&\le \sum_{t=D^{2/3}}^{D}(2p)^t\binom{n}{t}2\binom{\binom{t}{2}}{D}((\log d)^C/d)^t\\
&\le 2\sum_{t=D^{2/3}}^{D}(2p)^t\Big(\frac{ne}{t}\Big)^t\Big(\frac{et^2}{D}\Big)^D\Big(\frac{(\log d)^C}{d}\Big)^t\\
&\le 2\sum_{t=D^{2/3}}^{D}(2p)^t n^t((t^2)^{D/t})^t\Big(\frac{(\log d)^C}{d}\Big)^t\\
&\le 2\sum_{t=D^{2/3}}^{D}\Big(\frac{2pnt^{2D/t}(\log d)^C}{d}\Big)^t\\
&\le 2\sum_{t=D^{2/3}}^{D}\Big(\frac{2pnD^{2D^{1/3}}(\log d)^C}{d}\Big)^t\\
&\le 2\sum_{t=D^{2/3}}^{D}\Big(\frac{2pne^{O\big((\log\log d)(\log d)^{5/12}\big)}(\log d)^C}{d}\Big)^t.
\end{aligned}$$

Again, under the same conditions $d = \max\big\{(np)^{1+o_n(1)}), \omega_n(1)\big\}$, the expression is of order $o(1)$.

C.4. Omitted Technical Details

C.4.1. PROOF OF (43)

Recall (42). We begin by calculating the variance for Erdős-Rényi . Note that whenever $(i,j,k)\neq(i',j',k')$, there exists some edge in only one of the two triangles $\triangle(i,j,k), \triangle(i',j',k')$. Without loss of generality this is (ij). Then, $\mathbf{E}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SW}_{\triangle(i,j,k)}(\mathbf{K})\mathsf{SW}_{\triangle(i',j',k')}(\mathbf{K})] = 0$ as the product $\mathsf{SW}_{\triangle(i,j,k)}(\mathbf{K})\mathsf{SW}_{\triangle(i',j',k')}(\mathbf{K})$ contains the factor $\mathbf{K}_{(ij)}-p$ which is independent of everything else. Similarly, $\mathbf{E}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SW}_{\triangle(i,j,k)}(\mathbf{K})]\mathbf{E}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SW}_{\triangle(i',j',k')}(\mathbf{K})] = 0$. Thus,

$$\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SW}_{C_3}(\mathbf{K})] = \Theta(n^3)\times\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SC}_{C_3}(\mathbf{K})].$$

We now use the following fact:

$$\text{For any indicator } I, \text{ the equality } (I-p)^2 = (I-p)(1-2p)+(p-p^2) \text{ holds.} \tag{48}$$

We deduce $\mathbf{Var}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SC}_{C_3}(\mathbf{K})] = \mathbf{E}_{\mathbf{K}\sim\mathsf{G}(n,p)}[\mathsf{SC}_{C_3}(\mathbf{K})^2] = \Theta(n^3)\times(p-p^2)^3 = \Theta(n^3p^3)$.

Now, we proceed to bounding $\mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)}[\mathsf{SW}_{C_3}(\mathbf{G})]$. The idea is to split each term $\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\triangle(i,j,k)}(\mathbf{G})\mathsf{SW}_{\triangle(i',j',k')}(\mathbf{G})]$ into a (weighted) sum of signed counts with respect to $\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^\infty,p)$ via (48) and then apply Theorem 3.

1) $\mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G})] = \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G})^2] - \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G})]^2$. By Corollary 27, $\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G})]^2 = \tilde{\Theta}(p^6/d^4) = o(p^3)$. On the other hand,

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G})^2] = \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)-p})^2(\mathbf{G}_{(23)-p})^2(\mathbf{G}_{(13)-p})^2].$$

Using (48), this is equal to $(p-p^2)^3 + (1-2p)^3\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G})]$ and some terms with only one or two factors of the form $\mathbf{G}_{ij} - p$. By Lemma 24, those terms vanish as they form a graph with a leaf. Thus, the result is of order $\Theta(p^3) + \tilde{\Theta}((1-2p)^2p^3/d^2) = \Theta(p^3)$.

2) $\mathbf{Cov}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G}), \mathsf{SW}_{\triangle(1,2,4)}(\mathbf{G})] = \mathbf{E}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G})\mathsf{SW}_{\triangle(1,2,4)}(\mathbf{G})] - \tilde{\Theta}(p^6/d^4)$ by Corollary 27. However, using (48),

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{G})\mathsf{SW}_{\triangle(1,2,4)}(\mathbf{G})]\\
&= (p-p^2)\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(13)}-p)(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(14)}-p)(\mathbf{G}_{(24)}-p)]\\
&\quad + (1-2p)\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)(\mathbf{G}_{(13)}-p)(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(14)}-p)(\mathbf{G}_{(24)}-p)].
\end{aligned}$$

Both summands correspond to the signed weights of graphs on at most 4 vertices. By Theorem 3, the last expression is of order $\tilde{O}(p \times p^4/d^2) = \tilde{O}(p^5/d^2)$.

3) In the cases of $\mathbf{Cov}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{H}), \mathsf{SW}_{\triangle(1,4,5)}(\mathbf{H})], \mathbf{Cov}[\mathsf{SW}_{\triangle(1,2,3)}(\mathbf{H}), \mathsf{SW}_{\triangle(4,5,6)}(\mathbf{H})]]$, the two graphs $\triangle(i,j,k)$ and $\triangle(i',j',k')$ share at most one vertex, so their (signed) weights are independent by Lemma 24 and the covariance is zero.

Combining those estimates via (42), the variance is of order $\Theta(n^3p^3) + \tilde{O}(n^4p^5/d^2)$.

C.4.2. PROOF OF (45)

The estimate for Erdős-Rényi holds in the same way as in the proof of (43). We now estimate each of the terms in (44) for $\mathbf{G} \sim \mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma_p^\infty, p)$.

1) As in the case fro triangles, we estimate

$$\begin{aligned}
&\mathbf{Var}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{G})]\\
&= \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)^2(\mathbf{G}_{(23)}-p)^2(\mathbf{G}_{(34)}-p)^2(\mathbf{G}_{(24)}-p)^2] - \tilde{\Theta}(p^8/d^4)\\
&= (p-p^2)^4 + (1-2p)^4\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(24)}-p)] + \tilde{\Theta}(p^8/d^4)\\
&= (p-p^2)^4 + (1-2p)^4\tilde{\Theta}(p^4/d^2) + \tilde{\Theta}(p^8/d^4) = \Theta(p^4).
\end{aligned}$$

2) Similarly, using (48)

$$\begin{aligned}
&\mathbf{Cov}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{G}), \mathsf{SW}_{\square(1,2,3,5)}(\mathbf{G})]\\
&= \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)^2(\mathbf{G}_{(23)}-p)^2(\mathbf{G}_{(34)}-p)\times\\
&\times \qquad (\mathbf{G}_{(14)}-p)(\mathbf{G}_{(35)}-p)(\mathbf{G}_{(15)}-p)] - \tilde{\Theta}(p^8/d^4)\\
&= (p-p^2)^2\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p)(\mathbf{G}_{(35)}-p)(\mathbf{G}_{(15)}-p)]\\
&\quad + 2(p-p^2)(1-2p)\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p)(\mathbf{G}_{(35)}-p)(\mathbf{G}_{(15)}-p)]\\
&\quad + (1-2p)^2\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p)(\mathbf{G}_{(35)}-p)\times\\
&\times \qquad (\mathbf{G}_{(15)}-p)] - \tilde{\Theta}(p^8/d^4)\\
&= \tilde{O}(p^6/d^2) + \tilde{O}(p^6/d^{5/2}) + \tilde{O}(p^6/d^3) - \tilde{\Theta}(p^8/d^4) = \tilde{O}(p^6/d^2).
\end{aligned}$$

In the second to last line, we used Theorem 3 for each for the corresponding graphs.

3)

$$\begin{aligned}
&\mathbf{Cov}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{G}),\mathsf{SW}_{\square(1,2,4,5)}(\mathbf{G})]\\
&=\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)^2(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p))\times\\
&\times \qquad\qquad (\mathbf{G}_{(14)}-p)(\mathbf{G}_{(35)}-p)(\mathbf{G}_{(15)}-p)(\mathbf{G}_{(25)}-p)]-\tilde{\Theta}(p^8/d^4)\\
&=(p-p^2)\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p)(\mathbf{G}_{(35)}-p)(\mathbf{G}_{(15)}-p)(\mathbf{G}_{(25)}-p)]\\
&\quad+(1-2p)\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p))\times\\
&\times \qquad\qquad (\mathbf{G}_{(35)}-p)(\mathbf{G}_{(15)}-p)(\mathbf{G}_{(25)}-p)]\\
&\quad-\tilde{\Theta}(p^8/d^4)\\
&=\tilde{O}(p^7/d^{5/2})+\tilde{O}(p^7/d^{5/2})-\tilde{\Theta}(p^8/d^4)=\tilde{O}(p^7/d^{5/2})=\tilde{O}(p^6/d^2).
\end{aligned}$$

4)

$$\begin{aligned}
&\mathbf{Cov}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{G}),\mathsf{SW}_{\square(1,2,5,6)}(\mathbf{G})]\\
&=\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)^2(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p))\times\\
&\times \qquad\qquad (\mathbf{G}_{(25)}-p)(\mathbf{G}_{(56)}-p)(\mathbf{G}_{(61)}-p)]-\tilde{\Theta}(p^8/d^4)\\
&=(p-p^2)\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p)(\mathbf{G}_{(25)}-p)(\mathbf{G}_{(56)}-p)(\mathbf{G}_{(61)}-p)]\\
&\quad+(1-2p)\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p)(\mathbf{G}_{(25)}-p))\times\\
&\times \qquad\qquad (\mathbf{G}_{(56)}-p)(\mathbf{G}_{(61)}-p)]-\tilde{\Theta}(p^8/d^4)\\
&=\tilde{O}(p^7/d^3)+\tilde{O}(p^7/d^3)-\tilde{\Theta}(p^8/d^4)=\tilde{O}(p^7/d^3)=\tilde{O}(p^7/d^3)
\end{aligned}$$

5)

$$\begin{aligned}
&\mathbf{Cov}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{G}),\mathsf{SW}_{\square(1,5,2,6)}(\mathbf{G})]\\
&=\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p))\times\\
&\times \qquad\qquad (\mathbf{G}_{(15)}-p)(\mathbf{G}_{(52)}-p)(\mathbf{G}_{(26)}-p)(\mathbf{G}_{(61)}-p)]\\
&\quad-\tilde{\Theta}(p^8/d^4)\\
&=\tilde{O}(p^8/d^3)+\tilde{\Theta}(p^8/d^4)=\tilde{O}(p^8/d^3)=\tilde{O}(p^7/d^3)
\end{aligned}$$

6)

$$\begin{aligned}
&\mathbf{Cov}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{G}),\mathsf{SW}_{\square(1,5,3,6)}(\mathbf{G})]\\
&=\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{(12)}-p)(\mathbf{G}_{(23)}-p)(\mathbf{G}_{(34)}-p)(\mathbf{G}_{(14)}-p))\times\\
&\times \qquad\qquad (\mathbf{G}_{(15)}-p)(\mathbf{G}_{(53)}-p)(\mathbf{G}_{(36)}-p)(\mathbf{G}_{(61)}-p)]\\
&\quad-\tilde{\Theta}(p^8/d^4)\\
&=\tilde{O}(p^8/d^3)+\tilde{\Theta}(p^8/d^4)=\tilde{O}(p^8/d^3)=\tilde{O}(p^7/d^3)
\end{aligned}$$

7) For $\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(1,5,6,7)}(\mathbf{H})],\mathbf{Cov}[\mathsf{SW}_{\square(1,2,3,4)}(\mathbf{H}),\mathsf{SW}_{\square(5,6,7,8)}(\mathbf{H})]]$, the two graphs $\square(i,j,k,\ell)$ and $\square(i',j',k',\ell')$ share at most one vertex, so their (signed) weights are independent by Lemma 24 and the covariance is zero.

Combining those estimates via (44), the variance is of order $\Theta(n^4p^4)+\tilde{O}(n^5p^6/d^2+n^6p^7/d^3)$.

C.4.3. PROOF OF LEMMA 36

We know that λ_p^∞ satisfies $(1-\lambda_p^\infty)^d = p$. Thus,

$$\lambda_p^\infty = 1 - p^{1/d} = 1 - \exp(\log p/d) = 1 - \exp(-(\log 1/p)/d).$$

By the usual Taylor series expansion, $\exp(-x) = 1 - x + x^2/2 + O(x^3)$ when $x = o(1)$. Similarly, $\lambda_p^\infty/(1-\lambda_p^\infty) = p^{-1/d} - 1$ and we argue in the same way.

Appendix D. Statistical Indistinguishability in the L_∞ Model: Theorem 4

In this section, we prove Theorem 4. Recall condition (A). Suppose further that $d = \omega(n(\log n)^2)$. As in Section 3, $\tau_p^\infty = 1 - \lambda_p^\infty$, where $\lambda_p^\infty = \frac{\log(1/p)}{d}(1+o(1))$. We will write σ, λ, τ instead of $\sigma_p^\infty, \lambda_p^\infty, \tau_p^\infty$ for brevity. We can view $\sigma(\mathbf{x}, \mathbf{y})$ as a single argument function of $\mathbf{x} - \mathbf{y}$.

Expanding one of the terms in (12), we obtain

$$\mathbf{E}\Big[\Big(1 + \frac{\gamma(\mathbf{g})}{p(1-p)}\Big)^k\Big] = \mathbf{E}\Big[\Big(\frac{1-2p}{1-p} + \frac{\sigma * \sigma(\mathbf{g})}{p(1-p)}\Big)^k\Big] = \sum_{t=0}^k \binom{k}{t}\left(\frac{1-2p}{1-p}\right)^{k-t} \frac{\mathbf{E}[(\sigma*\sigma)^t]}{p^t(1-p)^t}.$$

We will prove the following bound on the moments of $\sigma * \sigma$.

Lemma 37 *For all $t \geq 1, t = o(1/\lambda) = o(d/(\log d)^2)$, it holds that $\mathbf{E}[(\sigma*\sigma)^t] = p^{2t}(1+\Theta(d\lambda^3 t^2))$. Also, $\mathbf{E}[\sigma*\sigma] = p^2$.*

D.1. Proof of Theorem 4 given Lemma 37

$$\begin{aligned}
&\mathbf{E}\Big[\Big(\frac{1-2p}{1-p} + \frac{\sigma*\sigma(\mathbf{g})}{p(1-p)}\Big)^k\Big]\\
&= \sum_{t=0}^k \binom{k}{t}\Big(\frac{1-2p}{1-p}\Big)^{k-t}\frac{p^{2t}(1+\Theta(d\lambda^3t^2))}{p^t(1-p)^t}\\
&= \sum_{t=0}^k \binom{k}{t}\Big(\frac{1-2p}{1-p}\Big)^{k-t}\Big(\frac{p}{1-p}\Big)^t + d\lambda^3\Theta\left(\sum_{t=0}^k \binom{k}{t}\Big(\frac{1-2p}{1-p}\Big)^{k-t}\Big(\frac{p}{1-p}\Big)^t t^2\right)\\
&= 1 + d\lambda^3\Theta\left(\frac{k(1-2p)^{k-1}p}{(1-p)^k} + k(k-1)\frac{p^2}{(1-p)^2}\sum_{t=2}^k\binom{k-2}{t-2}\Big(\frac{1-2p}{1-p}\Big)^{k-t}\Big(\frac{p}{1-p}\Big)^{t-2}\right)\\
&= 1 + d\lambda^3\Theta(kp + k^2p^2) = 1 + \tilde{\Theta}(d^{-2}kp + d^{-2}k^2p^2).
\end{aligned}$$

Going back to (12),

$$\begin{aligned}
&\sum_{k=0}^{n-1}\log\mathbf{E}\Big[\Big(1+\frac{\gamma(\mathbf{x})}{p(1-p)}\Big)^k\Big] = \sum_{k=0}^{n-1}\log\mathbf{E}\Big[1 + \tilde{\Theta}(d^{-2}kp + d^{-2}k^2p^2)\Big]\\
&\leq \tilde{\Theta}\Big(d^{-2}p\sum_{k=0}^{n-1}k + d^{-2}p^2\sum_{k=0}^{n-1}k\Big) = \tilde{\Theta}\Big(d^{-2}pn^2 + d^{-2}p^2n^3\Big),
\end{aligned}$$

where we used the fact that $t \leq k \leq n = o(d/\log d)$. The last expression is of order $o(1)$ whenever $p \geq 1/n, d \geq n^{3/2}p$ with which the poof follows.

D.2. Proof of Lemma 37

First, note that

$$\sigma * \sigma(\mathbf{x}) = \mathbf{E}_{\mathbf{g}}[\sigma(\mathbf{g})\sigma(\mathbf{x}-\mathbf{g})] = \mathbf{E}\Bigg[\prod_{i=1}^d \mathbb{1}[|g_i|_C \leq 1-\lambda]\mathbb{1}[|x_i - g_i|_C \leq 1-\lambda]\Bigg]$$
$$= \prod_{i=1}^d \mathbf{E}\Big[\mathbb{1}[|g_i|_C \leq 1-\lambda]\mathbb{1}[|x_i - g_i|_C \leq 1-\lambda]\Big].$$

Now, as is easy to see from Figures 3 and 4,

$$f(x) := \mathbf{P}_{g\sim \mathsf{Unif}(\mathbb{T}^1)}[|g|_C \leq 1-\lambda, |x-g|_C \leq 1-\lambda] = \begin{cases} 1-2\lambda & \text{when} |x|_C \geq 2\lambda, \\ 1-\lambda-|x|_C/2 & \text{when} |x|_C \leq 2\lambda. \end{cases} \tag{49}$$

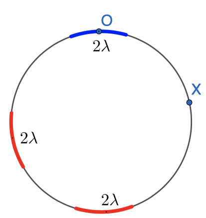

Figure 3: In the case when x is far from the origin (i.e., not within distance 2λ), the antipodal arcs of length 2λ of 0 and x (colored in red) do not intersect. Thus, g should be anywhere outside of the two segments of total length 4λ to satisfy $|g|_C \leq 1-\lambda, |x-g|_C \leq 1-\lambda$.

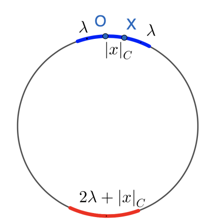

Figure 4: In the case when x is close to the origin (i.e., within distance 2λ), the antipodal arcs of length 2λ of 0 and x (colored in red) intersect. Thus, g should be anywhere outside of the intersection of the two segments of total length $2\lambda + |x|_C$ to satisfy $|g|_C \leq 1-\lambda, |x - g|_C \leq 1-\lambda$.

Thus, $\sigma * \sigma(\mathbf{x}) = \prod_{i=1}^d f(x_i)$. It follows that

$$\mathbf{E}[(\sigma * \sigma(\mathbf{x}))^t] = \mathbf{E}\Big[(\prod_{i=1}^d f(x_i))^t\Big] = \mathbf{E}[f(x_i)^t]^d$$
$$= \left(\int_0^1 f(x)^t\right)^d = \left((1-2\lambda)\times(1-2\lambda)^t + \int_0^{2\lambda}(1-\lambda-s/2)^t ds\right)^d$$
$$= \left((1-2\lambda)^{t+1} + 2\frac{(1-\lambda)^{t+1}-(1-2\lambda)^{t+1}}{(t+1)}\right)^d$$
$$= \left(\frac{2}{t+1}(1-\lambda)^{t+1} + \frac{t-1}{t+1}(1-2\lambda)^{t+1}\right)^d.$$

A simple calculation shows that the last expression is $p^{2t}(1+\Theta(d\lambda^3t^2))$. We expand the brackets on the left-hand side as follows.

$$\frac{2}{t+1}(1-\lambda)^{t+1}+\frac{t-1}{t+1}(1-2\lambda)^{t+1}=$$
$$\sum_{i=0}^{t+1}\binom{t+1}{i}\frac{1}{t+1}(2(-\lambda)^i+(t-1)(-2\lambda)^i)=$$
$$1-2t\lambda+\binom{2t}{2}\lambda^2-\binom{2t}{3}\lambda^3+\frac{2t(t-1)}{6}\lambda^3+\sum_{i=4}^{t}\binom{t+1}{i}\Big(\frac{2(-\lambda)^i+(t-1)(-2\lambda)^i}{t+1}\Big).$$

We claim that the last expression equals $(1-\lambda)^{2t}+\frac{2t(t-1)}{6}\lambda^3(1+o(1))$. This is equivalent to proving that

$$\sum_{i=4}^{k}\Bigg(\binom{t+1}{i}\Big(\frac{2(-\lambda)^i+(t-1)(-2\lambda)^i}{t+1}\Big)-\binom{2t}{i}(-\lambda)^i\Bigg)=o(t^2\lambda^3). \tag{50}$$

We split the sum into two parts, $i\geq 4\log\frac{1}{\lambda}$ and $i<4\log\frac{1}{\lambda}$.

Case 1) Large values of i. We have

$$\begin{aligned}&\Bigg|\sum_{i\geq 4\log\frac{1}{\lambda}}^{t}\Bigg(\binom{t+1}{i}\Big(\frac{2(-\lambda)^i+(t-1)(-2\lambda)^i}{t+1}\Big)-\binom{2t}{i}(-\lambda)^i\Bigg)\Bigg|\\&\leq\sum_{i\geq 4\log\frac{1}{\lambda}}^{t}\Bigg|\binom{t+1}{i}\Big(\frac{2(-\lambda)^i+(t-1)(-2\lambda)^i}{t+1}\Big)\Bigg|+\Bigg|\binom{2t}{i}(-\lambda)^i\Bigg|\\&\leq 2\sum_{i\geq 4\log\frac{1}{\lambda}}(3\lambda t)^i=O((3\lambda t)^{4\log\frac{1}{\lambda}})=O(\lambda^4)=o(\lambda^3t^2),\end{aligned}\tag{51}$$

where we used the fact that $\lambda t=o(1)$.

Case 2) Small values of i. We bound the coefficient in front of $(-\lambda)^i$ as follows.

$$\begin{aligned}&\binom{t+1}{i}\frac{2+2^i(t-1)}{t+1}-\binom{2t}{i}\\&=\frac{t(t-1)(t-2)\cdots(t-i+2)(2^i(t-1)+2)-2t(2t-1)(2t-2)\cdots(2t-i+1)}{i!}\\&=\frac{2t(t-1)(t-2)\cdots(t-i+2)}{i!}+\\&\quad\frac{2t(2t-2)(2t-4)\cdots(2t-2i+4)(2t-2)-2t(2t-1)(2t-2)\cdots(2t-i+1)}{i!}\\&=O(t^{i-1})\\&-\frac{2t(2t-1)(2t-2)\cdots(2t-i+1)}{i!}\Big(1-\frac{2t(2t-2)(2t-4)\cdots(2t-2i+4)(2t-2)}{2t(2t-1)(2t-2)\cdots(2t-i+1)}\Big).\end{aligned}\tag{52}$$

Now, observe that

$$\frac{2t(2t-2)(2t-4)\cdots(2t-2i+4)(2t-2)}{2t(2t-1)(2t-2)\cdots(2t-i+1)} \geq \Big(\frac{2t-2i}{2t}\Big)^i = (1-i/t)^i \geq 1 - i^2/t,$$

where we used that $4 \leq i \leq t$ and Bernoulli's inequality. Furthermore,

$$\frac{2t(2t-2)(2t-4)\cdots(2t-2i+4)(2t-2)}{2t(2t-1)(2t-2)\cdots(2t-i+1)} = \frac{2t-2}{2t-1}\prod_{j=2}^{i-2}\frac{2t-2j}{2t-j-1} < 1.$$

Hence, the desired sum is of order $O(t^{i-1}) - O\Big(\frac{(2t)^i}{i!}\frac{i^2}{t}\Big) = O(t^{i-1})$. It follows that the sum in the small i case is bounded by

$$\sum_{i=4}^{4\log\frac{1}{\lambda}} O(t^{i-1}\lambda^i) = O(t^3\lambda^4) = o(t^2\lambda^3),$$

where again we used $t\lambda = o(1)$.

Altogether, using that $(1-2\lambda)^t \geq 1 - 2\lambda t = \Theta(1)$,

$$\begin{aligned}
&\mathbf{E}[(\sigma * \sigma(\mathbf{g}))^t] = \Big((1-\lambda)^{2t} + \Theta(t^2\lambda^3)\Big)^d \\
&= \Big((1-\lambda)^{2t}(1+\Theta(t^2\lambda^3)\Big)^d \\
&= (1-\lambda)^{2dt}(1+\Theta(t^2\lambda^3))^d \\
&= p^{2t}(1+\Theta(dt^2\lambda^3)),
\end{aligned}$$

where again we used that $t = o(\lambda^{-1})$, $dt^2\lambda^3 = o(1)$ and $(1-\lambda)^d = p$ by definition.

Appendix E. Statistical Indistinguishability in the L_q Model: Theorem 9

Here, we prove Theorem 9. We first give the proof in the case $q = o(d/\log d)$ and then explain the necessary changes in the case $q = \Omega(d/\log d)$. The latter is technically much simpler and does not use any ideas which do not appear in the case $q = o(d/\log d)$.

E.1. The Proof for Small q

Further Notation. Throughout, we fix $q \geq 1, q = o(d\log^{-1} d)$ and consider $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^q_{1/2}, 1/2)$. For simplicity of notation, we denote $\tau^q_{1/2}$ simply by τ and $\sigma^q_{1/2}$ by σ. Note that σ, when viewed as a single argument function, can be equivalently defined as the indicator of $B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})$, where $B_{L_q(\mathbb{T}^d),\tau}(\mathbf{x})$ is the L_q ball of radius τ on $\mathbb{T}^d$ centered at $\mathbf{x}$. Under this notation,

$$\begin{aligned}
&\gamma(\mathbf{g}) = \mathbf{E}_\mathbf{z}\bigg[\Big(\sigma(\mathbf{g}-\mathbf{z}) - \frac{1}{2}\Big)\Big(\sigma(\mathbf{z}) - \frac{1}{2}\Big)\bigg] = \mathbf{E}_\mathbf{z}[\sigma(\mathbf{g}-\mathbf{z})\sigma(\mathbf{z})] - \frac{1}{4} \\
&= \big|B_{L_q(\mathbb{T}^d),\tau}(\mathbf{g}) \cap B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})\big| - \frac{1}{4}.
\end{aligned}$$

Proof Strategy. Our main goal will be to prove that

$$\mathbf{E}_{\mathbf{g}\sim\mathsf{Unif}(\mathbb{T}^d)}\left[|\gamma(\mathbf{g})|^k\right]^{1/k} \le C\frac{k}{\sqrt{dq}}. \tag{53}$$

for an absolute constant C. This is sufficient to conclude Theorem 9 for the following reason. Using (12) and the fact $\mathbf{E}_{\mathbf{g}\sim\mathsf{Unif}(\mathbb{T}^d)}[\gamma(\mathbf{g})] = 0$,

$$\begin{aligned}
\mathsf{KL}\Big(\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma^q_{1/2},1/2)\|\mathsf{G}(n,1/2)\Big) &\le \sum_{k=0}^{n-1}\log\Big(\mathbf{E}_{\mathbf{g}\sim\mathsf{Unif}(\mathbb{T}^d)}[(1+4\gamma(\mathbf{g}))^k]\Big)\\
&= \sum_{k=0}^{n-1}\log\Big(1+\sum_{t\ge 2}\binom{k}{t}4^t\mathbf{E}[\gamma(\mathbf{g})^t]\Big)\\
&\le \sum_{k=0}^{n-1}\sum_{t=2}^{k}\binom{k}{t}4^t\mathbf{E}[\gamma(\mathbf{g})^t]\\
&\le \sum_{k=0}^{n-1}\sum_{t=2}^{k}\binom{k}{t}4^t\mathbf{E}[|\gamma(\mathbf{g})|^t]\\
&\le n\sum_{k=0}^{n}\binom{n}{k}4^k\mathbf{E}[|\gamma(\mathbf{g})|^k]\\
&\le n\sum_{k=2}^{n}\Big(\frac{ne}{k}\Big)^k 4^k C^k\frac{k^k}{(dq)^{k/2}}\\
&\le n\sum_{k=2}^{n}\Big(\frac{4eCn}{\sqrt{dq}}\Big)^k\\
&= n\times O\Big(\frac{n^2}{dq}\Big) = O\Big(\frac{n^3}{dq}\Big) = o(1).
\end{aligned} \tag{54}$$

We used the fact that $dq = \omega(n^3)$ to conclude that there is exponential decay in $\sum_{k=2}^{n}\Big(\frac{4eCn}{\sqrt{dq}}\Big)^k$.

In light of Lemma 21, to prove (53), it is enough to show the following two statements:

1. *Small Marginal Increments:* $\|D_i\gamma\|_\infty = O(\frac{1}{\sqrt{dq}})$ for all i.
2. *Small Marginal Variances:* $\|\mathbf{Var}_i[\gamma]\|_\infty = O(\frac{1}{d^2q})$ for all i.

Due to symmetry, it is enough to prove the statements for $d = i$. In deriving those two quantities, we will need the following anticoncentration result.

Anticoncentration of random L_q-distances. In Section I, we prove Corollary 8 and derive the following bound by specializing to $X_i = U_i^q$, where $U_i \sim \mathsf{Unif}([0,1])$.

Lemma 38 *For any interval* $[a,b]$, $\mathbf{P}\Big[\sum_{i=1}^{d-1}U_i^q \in [a,b]\Big] \le \exp(-\Omega(d/q)) + (b-a)\times\sqrt{q/d}$.

A simple integration, in turn, gives the following statement.

Lemma 39 *Suppose that* $U_1, U_2, \dots, U_{d-1} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}([0,1])$ *and* $q = o(d/\log d)$. *Let* F *be the CDF of* $\sum_{i=1}^{d-1} U_i^q$. *Then, for* $\psi(\ell) := F(\tau^q) - F(\tau^q - \ell^q) = F([\tau^q - \ell^q, \tau^q])$, *we have*

$$\int_0^1 \psi(\ell)d\ell = O\Big(\frac{1}{\sqrt{dq}}\Big) \quad \textit{and} \quad \int_0^1 \psi(\ell)^2 d\ell = O\Big(\frac{1}{d}\Big). \tag{55}$$

E.1.1. MARGINAL INCREMENTS

For any fixed $\mathbf{g}_{-d}$,

$$\begin{aligned} D_d\gamma(\mathbf{g}_{-d}) &= \sup_{\mathbf{g}_d^M} \gamma(\mathbf{g}_{-d}, \mathbf{g}_d^M) - \inf_{\mathbf{g}_d^m} \gamma(\mathbf{g}_{-d}, \mathbf{g}_d^m) \\ &= \sup_{\mathbf{g}_d^M} \big|B_{L_q(\mathbb{T}^d),\tau}(\mathbf{g}_{-d}, \mathbf{g}_d^M) \cap B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})\big| \\ &- \inf_{\mathbf{g}_d^m} \big|B_{L_q(\mathbb{T}^d),\tau}(\mathbf{g}_{-d}, \mathbf{g}_d^m) \cap B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})\big| \\ &= |B_{L_q(\mathbb{T}^d),\tau}(\mathbf{g}_{-d}, 0) \cap B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})\big| - \big|B_{L_q(\mathbb{T}^d),\tau}(\mathbf{g}_{-d}, 1) \cap B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})\big| \\ &\leq |B_{L_q(\mathbb{T}^d),\tau}(\mathbf{g}_{-d}, 0) \backslash B_{L_q(\mathbb{T}^d),\tau}(\mathbf{g}_{-d}, 1)| \\ &= |B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0}) \backslash B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0}_{-d}, 1)|. \end{aligned} \tag{56}$$

Now, observe that a point $(h_1, h_2, \dots, h_d)$ is in $B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0}) \backslash B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0}_{-d}, 1)$ if and only if

$$\sum_{i=1}^{d-1} |h_i|_C^q \leq \tau^q - |h_d|_C^q \quad \text{and} \quad \sum_{i=1}^{d-1} |h_i|_C^q \geq \tau^q - |1 - h_d|_C^q = \tau^q - (1 - |h_d|_C)^q.$$

Clearly, one needs to have $|h_d|_C \in [0, 1/2]$ for this event to occur. Since each $|h_i|_C$ is uniformly distributed on $[0,1]$, we conclude that the probability of this event is

$$\begin{aligned} &\int_0^{1/2} \mathbf{P}\Bigg[\tau^q - (1-\ell)^q \leq \sum_{i=1}^{d-1} |h_i|_C^q \leq \tau^q - \ell^q\Bigg] d\ell \\ &\leq \int_0^1 \mathbf{P}\Bigg[\tau^q - \ell^q \leq \sum_{i=1}^{d-1} |h_i|_C^q \leq \tau^q\Bigg] d\ell \\ &= \int_0^1 F([\tau^q - \ell^q, \tau^q]) d\ell = \int_0^1 \psi(\ell) d\ell = O\left(\frac{1}{\sqrt{qd}}\right). \end{aligned} \tag{57}$$

E.1.2. MARGINAL VARIANCES

For the second moment, we will first rewrite γ. By definition,

$$\begin{aligned} \gamma(\mathbf{g}) + \frac{1}{4} &= \int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\Big[\mathbf{z} \in B_{L_q(\mathbb{T}^d),\tau}(\mathbf{g})\Big] d\mathbf{z} \\ &= \int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\Bigg[\sum_{i=1}^d |g_i - z_i|_C^q \leq \tau^q\Bigg] d\mathbf{z}. \end{aligned} \tag{58}$$

Now, fix $\mathbf{g}_{-d}$ and denote

$$\kappa(u) = \int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i|_C^q + |u - z_d|_C^q \le \tau^q\right] d\mathbf{z}.$$

$\mathbf{Var}_{U\sim\mathsf{Unif}(\mathbb{S}^1)}[\kappa(U)]$ is exactly $\mathbf{Var}_d[\gamma(\mathbf{g}_{-d})]$. By definition,

$$\begin{aligned}\mathbf{E}_{U\sim\mathsf{Unif}(\mathbb{S}^1)}[\kappa(U)] &= \mathbf{E}_U\left[\int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}[\sum_{i=1}^{d-1} |g_i - z_i|_C^q + |U - z_i|_C^q \le \tau^q] d\mathbf{z}\right]\\ &= \mathbf{E}_V\left[\int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}[\sum_{i=1}^{d-1} |g_i - z_i|_C^q + |V|_C^q \le \tau^q] d\mathbf{z}\right] := \mathbf{E}[\rho(V)],\end{aligned}$$

where ρ is defined by the last equation, i.e.,

$$\rho(v) := \int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i|_C^q + |v|_C^q \le \tau^q\right] d\mathbf{z}.$$

On the other hand,

$$\begin{aligned}&\mathbf{E}_{U\sim\mathsf{Unif}(\mathbb{S}^1)}[\kappa^2(U)]\\ &= \mathbf{E}_U\Bigg[\int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i^1|_C^q + |U - z_i^1|_C^q \le \tau^q\right] d\mathbf{z}^1\\ &\qquad \times \int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i^2|_C^q + |U - z_i^2|_C^q \le \tau^q\right] d\mathbf{z}^2\Bigg]\\ &= \mathbf{E}_U\Bigg[\int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i^1|_C^q + |V|_C^q \le \tau^q\right] d\mathbf{z}^1\\ &\qquad \times \int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i^2|_C^q + |V + z_i^1 - z_i^2|_C^q \le \tau^q\right] d\mathbf{z}^2\Bigg]\\ &= \mathbf{E}_{V,R}[\rho(V)\rho(V+R)],\end{aligned}$$

where $V = U - z_d^1 \sim \mathsf{Unif}(\mathbb{S}^1)$ and $R \sim z_d^1 - z_d^2$ and V, R are independent. It follows that

$$\mathbf{Var}[\kappa] = \mathbf{E}_{V,R}[\rho(V)\rho(V+R)] - \mathbf{E}[\rho(V)]^2. \tag{59}$$

Since $\rho : \mathbb{S}^1 \longrightarrow [0,1]$ is clearly L_2-integrable, we can write its Fourier series. Furthermore, as $\rho(v) = \rho(-v)$ and ρ is real, we can write

$$\rho(v) = \widehat{\rho}(0) + \sum_{k\ge 1} 2\widehat{\rho}(k)\cos(\pi k v).$$

In particular, we have $\mathbf{E}_{V\sim\mathsf{Unif}(\mathbb{S}^1)}[\rho(V)] = \widehat{\rho}(0)$ and, using the convolution formula,

$$\mathbf{E}_{V,R}[\rho(V)\rho(V+R)] = \widehat{\rho}(0)^2 + \sum_{k\geq 1} 2\widehat{\rho}(k)^2\mathbf{E}_R[\cos(\pi k R)]. \tag{60}$$

Putting all of this together, we have

$$\begin{aligned}\mathbf{Var}_d[\gamma(\mathbf{g}_{-d})] &= \sum_{k\geq 1} 2\widehat{\rho}(k)^2\mathbf{E}_R[\cos(\pi k R)] \leq \left(\sum_{k\geq 1} 2\widehat{\rho}(k)^2\right) \times \sup_{k\geq 1}\mathbf{E}_R[\cos(\pi k R)] \\ &= \mathbf{Var}_{V\sim\mathsf{Unif}(\mathbb{S}^1)}[\rho(V)] \times \sup_{k\geq 1}\mathbf{E}_R[\cos(\pi k R)].\end{aligned} \tag{61}$$

We will now bound the variance of ρ and the cosine expectations separately. That is, we will show

$$\mathbf{Var}_{V\sim\mathsf{Unif}(\mathbb{S}^1)}[\rho(V)] = O\left(\frac{1}{d}\right) \quad \text{and} \quad \sup_{k\geq 1}\mathbf{E}_R[\cos(\pi k R)] = O\left(\frac{1}{qd}\right),$$

which is enough.

1) Variance of ρ.

$$\begin{aligned}\mathbf{Var}[\rho] &\leq \mathbf{E}_V\left[\left(\rho(0) - \rho(V)\right)^2\right] \\ &= \mathbf{E}_V\left[\left(\int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i|_C^q \leq \tau^q\right] d\mathbf{z} - \right.\right.\\ &\qquad\qquad \left.\left. - \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i|_C^q + |V|_C^q \leq \tau^q\right] d\mathbf{z}\right)^2\right] \\ &= \mathbf{E}_V\left[\left(\int_{B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i|_C^q \in [\tau^q, \tau^q - |V|_C^q]\right] d\mathbf{z}\right)^2\right] \\ &\leq \mathbf{E}_V\left[\left(\int_{\mathbb{T}^d} \mathbb{1}\left[\sum_{i=1}^{d-1} |g_i - z_i|_C^q \in [\tau^q, \tau^q - |V|_C^q]\right] d\mathbf{z}\right)^2\right].\end{aligned}$$

Since the integral is over the entire torus, the variables $\{|g_i - z_i|_C\}_{i=1}^{d-1}$ are iid uniformly distributed over $[0,1]$, just like V. Therefore, the last expression equals

$$\int_0^1 F([\tau^q - \ell^q, \tau^q])^2 d\ell = O\left(\frac{1}{d}\right), \tag{62}$$

where we used Lemma 39.

2) Cosine Expectation. We need to find

$$\sup_{k\geq 1}\mathbf{E}_R[\cos(\pi k R)], \quad R \sim z_d^1 - z_d^2, \tag{63}$$

where z_d^1, z_d^2 are independent copies of the last coordinate of a uniformly random point in $B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})$. First, we will make a few simple observations abound the density of R. Let the density of z_d^1 be $\nu(x)$. Note that

$$\nu(x) \propto F(\tau^q - |x|_C^q)$$

since $(z_1, z_2, \ldots, z_{d-1}, x) \in B_{L_q(\mathbb{T}^d),\tau}(\mathbf{0})$ if and only if

$$\sum_{i=1}^{d-1} |z_i|_C^q \leq \tau^q - x^q,$$

but $|z_1|_C, |z_2|_C, \ldots, |z_{d-1}|_C$ are iid $\mathsf{Unif}([0,1])$ variables. In particular, this has the following implications:

1. $|x|_C \longrightarrow \nu(x)$ is positive and decreasing.
2. ν is even, i.e. $\nu(x) = \nu(-x)$.

Now, $R \sim z_d^1 - z_d^2 \sim z_d^1 + z_d^2$. Thus, if μ is the distribution of R, clearly $\mu = \nu * \nu$. In particular:

$$\begin{aligned} &1.\ \mu \text{ is even, i.e. } \mu(y) = \mu(-y),\\ &2.\ |y|_C \longrightarrow \mu(y) \text{ is decreasing.} \end{aligned} \tag{64}$$

The first fact is trivial. The second fact for $y \in [0,1]$ can be shown as follows. First, note that $\nu'(x) \leq 0$ for $x \in [0,1]$ as $|x|_C \longrightarrow \nu(x)$ is decreasing and $\nu'(x) = -\nu'(-x)$ since x is even. Now,

$$\begin{aligned} \mu'(y) &= (\nu * \nu)'(y) = (\nu' * \nu)(y) = \int_{-1}^{1} \nu'(x)\nu(y-x)dx \\ &= \int_0^1 \nu'(x)\nu(y-x)dx + \int_{-1}^{0} \nu'(x)\nu(y-x)dx \\ &= \int_0^1 \nu'(x)\nu(y-x)dx + \int_0^1 \nu'(-x)\nu(y+x)dx \\ &= \int_0^1 \nu'(x)\big(\nu(y-x) - \nu(y+x)\big)dx. \end{aligned}$$

We know that $\nu'(x) \leq 0$ for $x \in [0,1]$. On the other hand $\nu(y-x) \geq \nu(y+x)$ holds because $|z|_C \longrightarrow \nu(z)$ is decreasing and $|y-x|_C \leq |y+x|_C$ whenever $x, y \in [0,1]$. To show the last part, note that $|y-x|_C \in \{y-x, x-y\}$, and $|y+x|_C \in \{y+x, 2-y-x\}$. However,

1. $y-x, x-y \leq y+x$ whenever $x, y \geq 0$.
2. $y-x, x-y \leq 2-y-x$ whenever $x, y \leq 1$.

We split the rest of the proof into two claims.

Claim 40 $\sup_{k\geq 1} \mathbf{E}_R[\cos(\pi k R)] \leq 2\mathsf{TV}\Big(R, \mathsf{Unif}([-1,1])\Big)$.

Proof Let (R, U) be an optimal coupling of a $\mathsf{Unif}([-1,1])$ random variable U with R. Then,

$$\begin{aligned} \mathbf{E}[\cos(\pi k R)] =& \mathbf{E}[\cos(\pi k R) - \cos(\pi k U)] = \mathbf{E}\Big[\mathbb{1}[U \neq R]\big(\cos(\pi k R) - \cos(\pi k U)\big)\Big] \\ \leq& \|\cos(\pi k R) - \cos(\pi k U)\|_\infty \times \mathbf{E}\Big[\mathbb{1}[U \neq R]\Big] \leq 2\mathsf{TV}\Big(R, \mathsf{Unif}([-1,1])\Big). \end{aligned} \tag{65}$$

■

Claim 41 $\mathsf{TV}\Big(R, \mathsf{Unif}([-1,1])\Big) = O(1/dq)$.

Proof We use properties of the aforementioned density μ. Let $U \sim \mathsf{Unif}([-1,1])$.

$$\begin{aligned}\mathsf{TV}(R,U) &= \sup_A \Big(\mathbf{P}[U\in A] - \mathbf{P}[R\in A]\Big)\\ &= \sup_A \int_A \Big(\frac{1}{2} - \mu(u)\Big)dy \le |A| \times \Big(\frac{1}{2} - \inf_y \mu(y)\Big) \le \Big(\frac{1}{2} - \inf_y \mu(y)\Big).\end{aligned} \tag{66}$$

Our last step will be to show that $\inf_y \mu(y) = \frac{1}{2} - O\Big(\frac{1}{dq}\Big)$. As we know, $\nu(x) \propto F(\tau^q - |x|_C^q) = F(\tau^q) - \psi(|x|_C) = C - \psi(|x|_C)$, where

$$C = F(\tau^q) = \mathbf{P}\Big[\sum_{i=1}^{d-1} U_i^q \le \tau^q\Big] \ge \mathbf{P}\Big[\sum_{i=1}^{d} U_i^q \le \tau^q\Big] \ge \frac{1}{2}.$$

Thus,

$$\mu(y) = (\nu * \nu)(y) \propto \int_{-1}^{1} F(\tau^q - |x|_C^q) F(\tau^q - |y-x|_C^q)dx = \int_{-1}^{1} (C - \psi(|x|_C))(C - \psi(|y-x|_C))dx.$$

Now, we will find the normalizing constant in $\propto$.

$$\begin{aligned}K &= 2\int_{-1}^{1}\int_{-1}^{1} (C - \psi(|x|_C))(C - \psi(|y-x|_C))dxdy = 2\Big(\int_{-1}^{1} (C - \psi(|x|_C))dx\Big)^2\\ &= 2\Big(2C - \int_{-1}^{1}\psi(|x|_C)dx\Big)^2 = 2\Big(4C^2 - 4C\int_{-1}^{1}\psi(|x|_C)dx + \Big(\int_{-1}^{1}\psi(|x|_C)dx\Big)^2\Big)\\ &= 8C^2 - 8C\int_{-1}^{1}\psi(|x|_C)dx + O\Big(\frac{1}{dq}\Big),\end{aligned} \tag{67}$$

where we used Theorem 39. Note that $K = \Theta(1)$ since $C \geq 1/2$ and $\int_{-1}^{1} \psi(|x|_C)dx = O(\sqrt{\frac{1}{qd}}) = o(1)$ by Theorem 39. However, we know that $\inf_y \mu(y) = \mu(1)$ by (64), so

$$\begin{aligned}\mu(1) &= \frac{1}{K}\int_{-1}^{1}(C - \psi(|x|_C))(C - \psi(|1-x|_C))dx\\ &= \frac{1}{K}\Bigg(4C^2 - 2C\Big(\int_{-1}^{1}\psi(|x|_C)dx + \int_{-1}^{1}\psi(|1-x|_C)dx\Big)\\ &+ \int_{-1}^{1}\psi(|x|_C)dx \times \int_{-1}^{1}\psi(|1-x|_C)dx\Bigg)\\ &\geq \frac{1}{K}\Bigg(4C^2 - 2C\Big(\int_{-1}^{1}\psi(|x|_C)dx + \int_{-1}^{1}\psi(|1-x|_C)dx\Big)\Bigg)\\ &= \frac{1}{K}\Bigg(4C^2 - 4C\int_{-1}^{1}\psi(|x|_C)dx\Bigg)\\ &= \frac{K/2 - O\Big(\frac{1}{dq}\Big)}{K} = \frac{1}{2} - O\Big(\frac{1}{dq}\Big),\end{aligned}$$

with which the desired bound on the cosine expectation follows. ■

E.2. The Proof for Large q

When $q = \Omega(d/\log d)$, we will follow a similar strategy as in the proof for the case of $q = o(d/\log d)$. Namely, our goal will be to prove that for any integer k,

$$\mathbf{E}_{\mathbf{g}\sim\mathsf{Unif}(\mathbb{T}^d)}[|\gamma(\mathbf{g})|^k]^{1/k} \leq C(\log d)^C\frac{k}{d}. \tag{68}$$

for some absolute constant C. Following the same steps as in (54), this will be enough to conclude the second part of Theorem 9. Again, we will use the Bernstein-McDiarmid approach to bounding the moments of of γ. Our goal, this time, is to show the following.

1. *Small Marginal Increments:* $\|D_i\gamma\|_\infty = \tilde{O}(\frac{1}{d})$ for all i.
2. *Small Marginal Variances:* $\|\mathbf{Var}_i[\gamma]\|_\infty = \tilde{O}(\frac{1}{d^3})$ for all i.

We use the following anticoncentration results instead of Lemma 39. The rest of the proof is exactly the same.

Lemma 42 *Suppose that $U_1, U_2, \ldots, U_{d-1} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}([0,1])$ and $q = \Omega(d/\log d)$. Let F be the CDF of $\sum_{i=1}^{d-1} U_i^q$. Then, for $\psi(\ell) := F(\tau^q) - F(\tau^q - \ell^q) = F([\tau^q - \ell^q, \tau^q])$, we have*

$$\int_0^1 \psi(\ell)d\ell = \tilde{O}\left(\frac{1}{d}\right) \quad \textit{and} \quad \int_0^1 \psi(\ell)^2 d\ell = \tilde{O}\left(\frac{1}{d}\right). \tag{69}$$

The proof of Lemma 42 is substantially different (and much simpler) than the proof of Lemma 39. As we will need one of the ingredients in the next section as well, we present it in full detail here.

Lemma 43 *Suppose that* $U_1, U_2, \ldots, U_{d-1}$ *are iid* $\mathsf{Unif}([0,1])$ *random variables and* $q \geq 1$. *Then, for any interval* $[a,b]$,

$$\mathbf{P}\Big[\sum_{i=1}^{d-1} U_i^q \in [a,b]\Big] \leq b^{(d-1)/q} - a^{(d-1)/q}.$$

Proof The main idea is to reduce the computation to a computation for $q = \infty$. Let $W = \max(U_1, \ldots, U_{d-1})$ and $V_1, V_2, \ldots, V_{d-2} \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}([0,1])$ be independent of W. Then, $\sum_{i=1}^{d-1} U_i^q$ has the same distribution as $W^q \times \big(1 + \sum_{j=1}^{d-2} V_j^q\big)$. This follows simply by conditioning on the maximal value w of $U_1, U_2, \ldots, U_{d-1}$. Denote $T = 1 + \sum_{j=1}^{d-2} V_j^q$ and observe that $T \geq 1$ a.s. This implies that

$$\mathbf{P}\Big[\sum_{i=1}^{d-1} U_i^q \in [a,b]\Big] = \mathbf{P}\Big[W^q \times T \in [a,b]\Big] \leq \sup_{t \geq 1} \mathbf{P}\Big[W^q \times t \in [a,b]\Big]$$
$$= \sup_{t\geq 1} \mathbf{P}\Big[\big(\frac{a}{t}\big)^{1/q} \leq W \leq \big(\frac{b}{t}\big)^{1/q}\Big].$$

Now, since W is the maximum of $d-1$ iid $\mathsf{Unif}([0,1])$ random variables, $\mathbf{P}[W \leq x] = x^{d-1}$ for any $x \in [0,1]$. Thus, $\mathbf{P}\Big[\big(\frac{a}{t}\big)^{1/q} \leq W \leq \big(\frac{b}{t}\big)^{1/q}\Big] = \big(\frac{b}{t}\big)^{\frac{d-1}{q}} - \big(\frac{a}{t}\big)^{\frac{d-1}{q}} \leq b^{(d-1)/q} - a^{(d-1)/q}$, where we used $t \geq 1$. ■

Now, we are ready to prove Lemma 42.

Proof [Proof of Lemma 42] Suppose that $q \geq d/(C' \log d)$ for some absolute constant C'. We begin by proving the following two simple statements:

1. $\tau \geq 1 - 1/d$. Recall that τ is defined as the radius of a $1/2$ volume ball in $(\mathbb{T}^d, L_q)$. Let $U_1, U_2, \ldots, U_d$ be iid $\mathsf{Unif}([0,1])$ random variables. So, $1/2 = \mathbf{P}\big[\tau \geq \|(U_1, U_2, \ldots, U_d)\|_q\big]$. However,

$$\mathbf{P}\big[1 - 1/d \geq \|(U_1, U_2, \ldots, U_d)\|_q\big]$$
$$\leq \mathbf{P}\big[1 - 1/d \geq \|(U_1, U_2, \ldots, U_d)\|_\infty\big] = (1-1/d)^d \leq 1/e < 1/2,$$

   which means that $\tau \geq 1 - 1/d$.

2. $\tau^q \leq C''(\log d)$ for some constant C'' depending solely on C'. Observe that each variable U_i^q has expectation $1/(q+1)$, variance lass than $\mathbf{E}[U_i^{2q}] = 1/(2q+1)$ and is bounded between 0 and 1. Thus, by Lemma 20,

$$\mathbf{P}\Big[\sum_{j=1}^{d} U_j^q \geq t + d/(q+1)\Big] \leq \exp\Big(-\min\big\{\Theta(t^2/(d/q)), \Theta(t)\big\}\Big).$$

In particular, this means that setting $t = C'' \times \max(1, d/q) \leq C'' \times C' \times (\log d)$ for large enough C'', we obtain a tail bound less than $1/2$. Thus, $\tau^q \leq C'' \log d$ for some C''.

Now, we go back to proving Lemma 42. We begin with the first inequality.

$$\int_0^1 \psi(\ell)d\ell = \int_0^1 \mathbf{P}\Bigg[\tau^q - \ell^q \le \sum_{i=1}^{d-1} U_i^{d-1} \le \tau^q\Bigg]d\ell$$
$$= \int_0^{1-\frac{(\log d)^3}{d}} \mathbf{P}\Bigg[\tau^q - \ell^q \le \sum_{i=1}^{d-1} U_i^{d-1} \le \tau^q\Bigg]d\ell + \int_{1-\frac{(\log d)^3}{d}}^1 \mathbf{P}\Bigg[\tau^q - \ell^q \le \sum_{i=1}^{d-1} U_i^{d-1} \le \tau^q\Bigg]d\ell$$
$$\le \mathbf{P}\Bigg[\tau^q - \Bigg(1 - \frac{(\log d)^3}{d}\Bigg)^q \le \sum_{i=1}^{d-1} U_i^{d-1} \le \tau^q\Bigg] + \frac{(\log d)^3}{d}.$$

All that is left to do is bound $\mathbf{P}\Bigg[\tau^q - \Bigg(1 - \frac{(\log d)^3}{d}\Bigg)^q \le \sum_{i=1}^{d-1} U_i^{d-1} \le \tau^q\Bigg]$. Using Lemma 43,

$$\begin{aligned}
&\mathbf{P}\Bigg[\tau^q - \Bigg(1 - \frac{(\log d)^3}{d}\Bigg)^q \le \sum_{i=1}^{d-1} U_i^{d-1} \le \tau^q\Bigg] \\
&\le (\tau^q)^{(d-1)/q} - \Bigg(\tau^q - \Bigg(1 - \frac{(\log d)^3}{d}\Bigg)^q\Bigg)^{(d-1)/q} \\
&= (\tau^q)^{(d-1)/q} \times \Bigg[1 - \Bigg(1 - \Bigg(\frac{1 - (\log d)^3/d}{\tau}\Bigg)^q\Bigg)^{(d-1)/q}\Bigg].
\end{aligned} \tag{70}$$

Since $\tau \ge 1 - 1/d$, it is the case that $\frac{1-(\log d)^3/d}{\tau} \le 1 - (\log d)^3/(2d)$. Thus,

$$\Bigg(\frac{1 - (\log d)^3/d}{\tau}\Bigg)^q \le \Bigg(1-(\log d)^3/(2d)\Bigg)^q \le \exp\Bigg(-(\log d)^3 q/(2d)\Bigg) \le \exp\Bigg(-\Theta((\log d)^2)\Bigg).$$

It follows that

$$\begin{aligned}
\Bigg(1 - \Bigg(\frac{1 - (\log d)^3/d}{\tau}\Bigg)^q\Bigg)^{(d-1)/q} &\ge \Bigg(1 - \exp\Bigg(-\Theta((\log d)^2)\Bigg)\Bigg)^{(d-1)/q} \\
\ge \Bigg(1 - \exp\Bigg(-\Theta((\log d)^2)\Bigg)\Bigg)^{C'(\log d)} &= 1 - \exp\Bigg(-\Theta((\log d)^2)\Bigg).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&(\tau^q)^{(d-1)/q} \times \Bigg[1 - \Bigg(1 - \Bigg(\frac{1 - (\log d)^3/d}{\tau}\Bigg)^q\Bigg)^{(d-1)/q}\Bigg] \\
&\le (C''(\log d))^{C' \log d} \times \exp\Bigg(-\Theta((\log d)^2)\Bigg) = \exp\Bigg(-\Theta((\log d)^2)\Bigg) = o(1/d).
\end{aligned}$$

With this, the proof of the first inequality is completed. The second inequality follows directly as

$$\int_0^1 \psi(\ell)^2 d\ell = \int_0^1 \mathbf{P}\Big[\tau^q - \ell^q \le \sum_{i=1}^{d-1} U_i^{d-1} \le \tau^q\Big]^2 d\ell$$
$$\le \int_0^1 \mathbf{P}\Big[\tau^q - \ell^q \le \sum_{i=1}^{d-1} U_i^{d-1} \le \tau^q\Big] d\ell = \int_0^1 \psi(\ell) d\ell.$$

■

Appendix F. Entropic Upper Bound in the L_q Model: Theorem 10

Theorem 44 (ε-Net Argument) *There exists some constant C with the following property. Consider a random geometric graph $\mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p)$ over the metric space (Ω, μ), where $1/n \le p \le 1/2$ and $\sigma(\mathbf{x}, \mathbf{y}) = \mathbb{1}[\mu(\mathbf{x}, \mathbf{y}) \le \tau]$ for some τ. Suppose, further, that (Ω, μ) has a finite ε-net $\mathcal{N}(\varepsilon)$ which satisfies the following property. $\mathbf{P}_{\mathbf{x},\mathbf{y} \overset{\text{i.i.d.}}{\sim} \mathcal{D}}[\mu(\mathbf{x}, \mathbf{y}) \in [\tau - 2\varepsilon, \tau + 2\varepsilon]] = o(n^{-2})$. If $|\mathcal{N}(\varepsilon)| \le \exp\Big(Cnp\log 1/p\Big)$, then*

$$\mathsf{TV}\Big(\mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p), \mathsf{G}(n, p)\Big) = 1 - o(1).$$

Proof First, we will show that there exists a graph distribution $\mathcal{Q}$ on support of size at most $|\mathcal{N}(\varepsilon)|^n$ such that

$$\mathsf{TV}\Big(\mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p), \mathcal{Q}\Big) = o(1).$$

Let π be the projection map form Ω to $\mathcal{N}(\varepsilon)$. Let $\mathcal{D}'$ be the distribution over $\mathcal{N}(\varepsilon)$ defined by $\pi \circ \mathcal{D}$. Let $p' = \mathbf{P}_{\mathbf{x},\mathbf{y} \overset{\text{i.i.d.}}{\sim} \mathcal{D}'}\big[\mu(\mathbf{x}, \mathbf{y}) \le \tau\big]$. We will show that $\mathcal{Q} = \mathsf{RGG}(n, \mathcal{N}(\varepsilon), \mathcal{D}', \sigma, p')$ satisfies the desired property. Here, we think of $\mathcal{N}(\varepsilon)$ as a metric space with the induced metric μ.

First, $\mathcal{Q}$ has support of size at most $|\mathcal{N}(\varepsilon)|^n$ as the n latent vectors in $\mathcal{N}(\varepsilon)$ uniquely determine the corresponding geometric graph.

Second, we will form a coupling between $\mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p)$ and $\mathsf{RGG}(n, \mathcal{N}(\varepsilon), \mathcal{D}', \sigma, p')$ as follows. For latent vectors $\mathbf{x}^1, \mathbf{x}^2, \ldots \mathbf{x}^n \in \Omega$, let $\mathbf{gg}_\Omega(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n)$ be the corresponding graph according to $\mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p)$ and $\mathbf{gg}_{\mathcal{N}(\varepsilon)}(\pi(\mathbf{x}^1), \pi(\mathbf{x}^2), \ldots, \pi(\mathbf{x}^n))$ be the corresponding graph according $\mathsf{RGG}(n, \mathcal{N}(\varepsilon), \mathcal{D}', \sigma, p')$. By definition, when we take $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n \overset{\text{i.i.d.}}{\sim} \mathcal{D}$, it is the case that

$$\mathbf{gg}_\Omega(\pi(\mathbf{x}^1), \pi(\mathbf{x}^2), \ldots, \pi(\mathbf{x}^n)) \sim \mathsf{RGG}(n, \Omega, \mathcal{D}, \sigma, p) \text{ and,}$$
$$\mathbf{gg}_{\mathcal{N}(\varepsilon)}(\pi(\mathbf{x}^1), \pi(\mathbf{x}^2), \ldots, \pi(\mathbf{x}^n)) \sim \mathsf{RGG}(n, \mathcal{N}(\varepsilon), \mathcal{D}', \sigma, p').$$

All that is left to show is that with probability $1 - o(1)$ over $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^n \overset{\text{i.i.d.}}{\sim} \mathcal{D}$, it is the case that $\mathbf{gg}_\Omega(\pi(\mathbf{x}^1), \pi(\mathbf{x}^2), \ldots, \pi(\mathbf{x}^n)) = \mathbf{gg}_{\mathcal{N}(\varepsilon)}(\pi(\mathbf{x}^1), \pi(\mathbf{x}^2), \ldots, \pi(\mathbf{x}^n))$.

Observe that whenever $\mathbf{gg}_\Omega(\pi(\mathbf{x}^1), \pi(\mathbf{x}^2), \ldots, \pi(\mathbf{x}^n)) \neq \mathbf{gg}_{\mathcal{N}(\varepsilon)}(\pi(\mathbf{x}^1), \pi(\mathbf{x}^2), \ldots, \pi(\mathbf{x}^n))$, there exist some i, j such that

$$\mathbb{1}\Big[\mu(\mathbf{x}^i, \mathbf{x}^j) \le \tau\Big] \neq \mathbb{1}\Big[\mu(\pi(\mathbf{x}^i), \pi(\mathbf{x}^j)) \le \tau\Big].$$

However, by triangle inequality,

$$\Big|\mu(\mathbf{x}^i,\mathbf{x}^j)-\mu(\pi(\mathbf{x}^i),\pi(\mathbf{x}^j))\Big|\le \mu(\mathbf{x}^i,\pi(\mathbf{x}^i))+\mu(\mathbf{x}^j,\pi(\mathbf{x}^j))\le 2\varepsilon.$$

In particular, this means that $\mu(\mathbf{x}^i,\mathbf{x}^j)\in[\tau-2\varepsilon,\tau+2\varepsilon]$. As this happens with probability $o(n^{-2})$ for a fixed pair i,j, the union bound implies that this happens with probability $o(1)$ for some i,j, which finishes the proof that $\mathsf{TV}\Big(\mathsf{RGG}(n,\Omega,\mathcal{D},\sigma,p),\mathcal{Q}\Big)=o(1)$. Thus, it is enough to show that $\mathsf{TV}\Big(\mathcal{Q},\mathsf{G}(n,p)\Big)=1-o(1)$ under the given conditions. This follows immediately from $|\mathcal{N}(\varepsilon)|\le \exp\Big(Cnp\log 1/p\Big)$ as shown in (Bangachev and Bresler, 2023, Theorem 7.5). ■

Theorem 10 now immediately follows from the following proposition.

Theorem 45 *Consider any* $q\in[1,+\infty)\cup\{\infty\}, d\ge n^{\delta}, p\ge n^{-1+\epsilon}$. *For* $\varepsilon=\exp(-(\log nd)^4)$, *there exists an* ε*-net of* $(\mathbb{T}^d,L_q)$ *of size* $\exp(\tilde\Theta(d))$. *Furthermore,* $\mathbf{P}_{\mathbf{x},\mathbf{y}\overset{\text{i.i.d.}}{\sim}\mathbb{T}^d}[\|\mathbf{x}-\mathbf{y}\|_q\in[\tau_p^q-2\varepsilon,\tau_p^q+2\varepsilon]\le n^{-3}$.

Proof First, we will show the existence of a small ε net. Let $k=\lceil d/\varepsilon\rceil$ be an integer and consider the set $\mathcal{N}=\{i/k\in\mathbb{T}^1\;:\;0\le i\le 2k-1\}^d\subseteq\mathbb{T}^d$. This is a set of size $(2k)^d=\exp(\tilde\Theta(d))$. Furthermore, it is a ε-net for any L_q geometry for the following reason. Take $\mathbf{x}\in\mathbb{T}^d$ and let $\mathbf{u}=(u_1/k,u_2/k,\ldots,u_d/k)$ be the projection of $\mathbf{x}$ to $\mathcal{N}$. Then, for any $q\in[1,+\infty)\cup\{\infty\}$,

$$\|\mathbf{x}-\mathbf{u}\|_q\le\|\mathbf{x}-\mathbf{u}\|_1=\sum_{j=1}^d|x_u-u_i/k|\le d/k\le\varepsilon.$$

Now, we need to show that for each q, $\mathbf{P}_{\mathbf{x},\mathbf{y}\overset{\text{i.i.d.}}{\sim}\mathbb{T}^d}[\|\mathbf{x}-\mathbf{y}\|_q\in[\tau_p^q-2\varepsilon,\tau_p^q+2\varepsilon]\le n^{-3}$ holds. This is equivalent to showing that for $U_1,U_2,\ldots,U_d\overset{\text{i.i.d.}}{\sim}\mathsf{Unif}([0,1])$, it is the case that

$$\mathbf{P}\Bigg[\|(U_1,U_2,\ldots,U_d)\|_q\in[\tau_p^q-2\varepsilon,\tau_p^q+2\varepsilon]\Bigg]\le n^{-3}\text{ or equivalently}$$
$$\mathbf{P}\Bigg[\sum_{j=1}^d U_j^d\in[(\tau_p^q-2\varepsilon)^q,(\tau_p^q+2\varepsilon)^q]\Bigg]\le n^{-3}.$$

As in the proof of Lemma 42, clearly $(\tau_p^q)\ge 1-(\log 1/p)/d\ge 1/2$. Furthermore, note that $(\tau_p^q)^q\le d$ as $\|(U_1,U_2,\ldots,U_d)\|_q^q\le d$ a.s. Now, we consider two cases:

Case 1) When $q=o(d/(\log d))$. Note that

$$(\tau_p^q+2\varepsilon)^q-(\tau_p^q+2\varepsilon)^q=(\tau_p^q)^q\Bigg((1+2\varepsilon/\tau_q^p)^q-(1-2\varepsilon/\tau_q^p)^q\Bigg).$$

Using that $q=o(d/\log d)=o(1/\varepsilon),(\tau_p^q)^q\le d,\tau_p^q\ge 1/2$, the last expression is of order $O(dq\varepsilon)=o(n^{-3})$. By Lemma 38, $\mathbf{P}\Bigg[\sum_{j=1}^d U_j^d\in[(\tau_p^q-2\varepsilon)^q,(\tau_p^q+2\varepsilon)^q]\Bigg]=o(n^{-3})$, as desired.

Case 2) When $q = \Omega(d/\log d)$. Using Lemma 43,

$$\mathbf{P}\Big[\sum_{j=1}^{d} U_j^d \in [(\tau_p^q - 2\varepsilon)^q, (\tau_p^q + 2\varepsilon)^q]\Big] \le \Big((\tau_p^q + 2\varepsilon)^q\Big)^{d/q} - \Big((\tau_p^q - 2\varepsilon)^q\Big)^{d/q}$$
$$= ((\tau_p^q)^q)^{d/q}\bigg(\Big(1 + 2\varepsilon/\tau_p^q\Big)^d - \Big(1 - 2\varepsilon/\tau_p^q\Big)^d\bigg)$$
$$\le d^{d/q} \times O(d\epsilon) = \exp(O((\log d)^2)) \times \exp(-(\log(nd))^4) \le n^{-3}.$$

■

Appendix G. Signed Counts in L_q Geometries

Our only rigorous progress towards signed subgraph tests in L_q geometries is the following.

Theorem 46 *The signed 4-cycle test cannot distinguish between* $H_0 : \mathsf{G}(n, 1/2)$ *and* $H_1 : \mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^q_{1/2}, 1/2)$ *in the following regimes:*

1. *When* $q = o(d/\log d)$ *and* $dq = \omega(n^2)$.
2. *When* $q = \Omega(d/\log d)$ *and* $d = \tilde{\omega}(n)$.

Proof The signed 4-cycle count corresponds to the second moment of γ :

$$\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma^q_{1/2},1/2)}[\mathsf{SW}_{C_4}(\mathbf{G})]$$
$$= \mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[(\mathbf{G}_{12} - 1/2)(\mathbf{G}_{23} - 1/2)(\mathbf{G}_{34} - 1/2)(\mathbf{G}_{41} - 1/2)]$$
$$= \mathbf{E}_{\mathbf{g}_1,\mathbf{g}_2,\mathbf{g}_3,\mathbf{g}_4 \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}(\mathbb{T}^d)}\Big[\big(\sigma(\mathbf{g}_1 - \mathbf{g}_2) - 1/2\big)\big(\sigma(\mathbf{g}_2 - \mathbf{g}_3) - 1/2\big)\times$$
$$\times \big(\sigma(\mathbf{g}_3 - \mathbf{g}_4) - 1/2\big)\big(\sigma(\mathbf{g}_4 - \mathbf{g}_1) - 1/2\big)\Big]$$
$$= \mathbf{E}_{\mathbf{h},\mathbf{z}_1,\mathbf{z}_2 \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}(\mathbb{T}^d)}\Big[\big(\sigma(\mathbf{z}_1) - 1/2\big)\big(\sigma(\mathbf{h} - \mathbf{z}_1) - 1/2\big)\big(\sigma(\mathbf{z}_2) - 1/2\big)\big(\sigma(\mathbf{h} - \mathbf{z}_2) - 1/2\big)\Big]$$
$$= \mathbf{E}\Big[\big(\sigma * \sigma(\mathbf{h}) - 1/4\big)^2\Big],$$

as desired. We used the substitution $\mathbf{z}_1 = \mathbf{g}_1 - \mathbf{g}_2, \mathbf{z}_2 = \mathbf{g}_1 - \mathbf{g}_4, \mathbf{h} = \mathbf{g}_1 - \mathbf{g}_3$. Recalling (53) and (68), we conclude that the signed count is of order $O(1/dq)$ in the regime $q = o(d/\log d))$ and of order $\tilde{O}(1/d^2)$ in the regime $q = \Omega(d/\log d)$. However, K_n has $\Theta(n^4)$ subgraphs isomorphic to C_4 and $\mathbf{Var}_{\mathbf{H}\sim\mathsf{G}(n,1/2)}[\mathsf{SC}_{C_4}(\mathbf{H})] = \Theta(n^4)$ by (45). Therefore, a necessary condition for detection via the signed 4-cycle test is $n^4\mathbf{E}_{\mathbf{G}\sim\mathsf{RGG}}[\mathsf{SW}_{C_4}(\mathbf{G})] = \omega(\sqrt{n^4})$. ■

We believe that $1/dq$ and $1/d^2$ are the correct (up to log factors) orders of the signed 4-cycle count in the two regimes. Note that when $q = \infty$, the signed 4-cycle count is indeed $\Theta(d^{-2})$ by Corollary 27. Similarly, in L_2 geometry (admittedly over a different latent space such as $\{\pm 1\}^d$, but again with a hard threshold connection with density $1/2$), the signed 4-cycle count is $\tilde{\Theta}(1/d)$ (follows directly from (Bangachev and Bresler, 2023, Observation 2.12)). As this is the correct behaviour at both ends, we believe that it is also correct for all q, which leads to the following conjecture.

Conjecture 47 *The signed four-cycle test distinguishes w. h. p. between H_0 : $\mathsf{G}(n, 1/2)$ and H_1 : $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^q_{1/2}, 1/2)$ under (**A**) in the following regimes:*

1. *When $q = o(d/\log d)$ and $dq = \tilde{o}(n^2)$.*
2. *When $q = \Omega(d/\log d)$ and $d = \tilde{o}(n)$.*

Similarly, we conjecture the performance of the signed-triangle statistic by extrapolating from behaviour at $q = 2$ and $q = \infty$.

Conjecture 48 *The signed triangle test distinguishes w. h. p. between H_0 : $\mathsf{G}(n, 1/2)$ and H_1 : $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^q_{1/2}, 1/2)$ under (**A**) in the following regimes:*

1. *When $q = o(d/\log d)$ and $dq^3 = \tilde{o}(n^3)$.*
2. *When $q = \Omega(d/\log d)$ and $d = \tilde{o}(n^{3/4})$.*

These conjectures can be summarized with the following diagram.

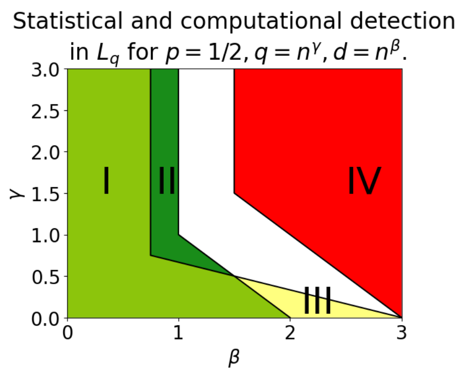

Figure 5: Visualizing Theorems 9 and 46 and Conjectures 47 and 48. $I + III$ is the conjectured region in which the signed triangle test solves (**P1**) for $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^q_{1/2}, 1/2)$ with high probability. Region $I + II$ is the conjectured region in which the signed four-cycle test succeeds with high probability. In IV, it is information theoretically impossible to solve (**P1**) with high probability. The last region is potentially suboptimal. Interestingly, if these conjectures are correct, the signed 4-cycle statistic is always at least as good as the entropic upper bound Theorem 10 but this is not the case for the signed 3-cycle statistic.

A Fourier-based Approach to Signed Cycle Counts. We end with a Fourier-based approach to computing the signed cycle counts for $\mathsf{RGG}(n, \mathbb{T}^d, \mathsf{Unif}, \sigma^q_p, p)$ (which extends to any random algebraic graph over $\mathbb{T}^d$ or a discrete torus).

We begin with some brief refresher on Fourier analysis over $\mathbb{T}^d$. Recall that we defined $\mathbb{T}^d$ as a product of d circles of circumference 2, or, equivalently, $\mathbb{T}^d = \mathbb{R}^d/\sim$, where $\mathbf{x} \sim \mathbf{y}$ if and only if $\mathbf{x} - \mathbf{y} \in 2\mathbb{Z}^d$. Similarly to the Boolean case, we will use the fact that any L_2-integrable function $f : \mathbb{T}^d \longrightarrow \mathbb{R}$ can be uniquely written as $f(\mathbf{x}) = \sum_{\mathbf{v} \in \mathbb{Z}^d} \widehat{f}(\mathbf{v}) \exp(i\pi\langle \mathbf{v}, \mathbf{x}\rangle)$. We make the following simple well-known observation. If f satisfies $f(\mathbf{x}) = f(-\mathbf{x})$ for all $\mathbf{x}$, then each coefficient $\widehat{f}(\mathbf{v})$ is real and, furthermore, $\widehat{f}(\mathbf{v}) = \widehat{f}(-\mathbf{v})$. Indeed, this follows by uniqueness as

$$\sum_{\mathbf{v}} \widehat{f}(\mathbf{v}) \exp(-i\pi\langle \mathbf{v}, \mathbf{x}\rangle) = f(-\mathbf{x}) = f(\mathbf{x}) = \overline{f(\mathbf{x})} = \sum_{\mathbf{v}} \overline{\widehat{f}(\mathbf{v}) \exp(i\pi\langle \mathbf{v}, \mathbf{x}\rangle)}$$
$$= \sum_{\mathbf{v}} \overline{\widehat{f}(\mathbf{v})} \exp(-i\pi\langle \mathbf{v}, \mathbf{x}\rangle).$$

Finally, recall that $\widehat{f}(\mathbf{0}) = \int_{\mathbb{T}}^d f(\mathbf{x})d\mathbf{x}$. Now, σ_p^q is clearly L_2-integrable. Thus, for any signed k-cycle weight,

$$\begin{aligned}
&\mathbf{E}_{\mathbf{G}\sim \mathsf{RGG}(n,\mathbb{T}^d,\mathsf{Unif},\sigma_p^q,p)}\Big[\mathsf{SW}_{C_k}(\mathbf{G})\Big] \\
&= \mathbf{E}_{\mathbf{g}_1,\mathbf{g}_2,\cdots,\mathbf{g}_k \overset{\text{i.i.d.}}{\sim} \mathsf{Unif}(\mathbb{T}^d)}\left[\prod_{i=1}^k (\sigma(\mathbf{g}_i - \mathbf{g}_{i+1}) - p)\right] \\
&= \int_{(\mathbb{T}^d)^k} \prod_{i=1}^k \sum_{\mathbf{v}\in\mathbb{Z}^d\setminus\mathbf{0}} \widehat{\sigma}(\mathbf{v}) \exp(i\pi\langle \mathbf{v}, \mathbf{g}_i - \mathbf{g}_{i+1}\rangle) d\mathbf{g}_1 d\mathbf{g}_2\cdots \mathbf{g}_k, \\
&= \sum_{\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k\in\mathbb{Z}^d\setminus\mathbf{0}} \int_{(\mathbb{T}^d)^k} \widehat{\sigma}(\mathbf{v}_1)\widehat{\sigma}(\mathbf{v}_1)\cdots\widehat{\sigma}(\mathbf{v}_k) \exp(-i\pi \sum_{i=1}^k \langle \mathbf{v}_i, \mathbf{g}_i - \mathbf{g}_{i+1}\rangle) \\
&= \sum_{\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_k\in\mathbb{Z}^d\setminus\mathbf{0}} \int_{(\mathbb{T}^d)^k} \widehat{\sigma}(\mathbf{v}_1)\widehat{\sigma}(\mathbf{v}_1)\cdots\widehat{\sigma}(\mathbf{v}_k) \exp(-i\pi \sum_{i=1}^k \langle \mathbf{g}_i, \mathbf{v}_i - \mathbf{v}_{i-1}\rangle) \\
&= \sum_{\mathbf{v}\in\mathbb{Z}^d\setminus\mathbf{0}} \widehat{\sigma}(\mathbf{v})^k,
\end{aligned} \tag{71}$$

where the last line follows from the simple observation that if $\mathbf{v}_i \neq \mathbf{v}_{i-1}$ for some i, the integral vanishes. It must be noted, however, that even if one manages to compute a signed cycle count, there still remains the obstacle of computing its variance.

Appendix H. Random Algebraic Graphs Over the Hypercube: Theorem 12

H.1. Preliminaries

We begin with some preliminaries on Boolean Fourier analysis. Any function $f : \{\pm 1\}^d \longrightarrow \mathbb{R}$ can be written uniquely as $f(\mathbf{x}) = \sum_{S\subseteq[d]} \widehat{f}(S)\omega_S(\mathbf{x})$, where $\omega_S(\mathbf{x}) := \prod_{i\in S} x_i$ is the Walsh polynomial O'Donnell (2014). The influence $\mathbf{Inf}_i[f]$ of variable i is defined as

$$\mathbf{Inf}_i[f] = \sum_{i\in S} \widehat{f}(S)^2 = \mathbf{E}_{\mathbf{x}\sim\mathsf{Unif}(\{\pm1\}^d)}\Big[(f(\mathbf{x}) - f(\mathbf{x}^{\oplus i}))^2/4\Big], \tag{72}$$

where $\mathbf{x}^{\oplus i}$ is $\mathbf{x}$ with the i-th coordinate flipped. We denote $\overrightarrow{\mathbf{Inf}}[f]$ as the vector in $\mathbb{R}^d_{\geq 0}$ with i'th coordinate equal to $\mathbf{Inf}_i[f]$. In particular, $\|\overrightarrow{\mathbf{Inf}}[f]\|_1 = \sum_{i=1}^d \mathbf{Inf}_i[f] = \mathbf{Inf}[f]$, which is the total influence, and $\|\overrightarrow{\mathbf{Inf}}[f]\|_\infty = \mathbf{MaxInf}[f]$, which is the max influence. Also, $\|\overrightarrow{\mathbf{Inf}}[f]\|_2^2 = \sum_i \mathbf{Inf}_i^2[f]$, which is the quantity of interest in Theorem 12.

H.2. The Proof of Theorem 12

Throughout, we make the following assumption, without which the statement of Theorem 12 is trivial (as it gives an upper bound of a total variation by a number larger than 1).

$$\frac{n\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2}{p(1-p)} = o(1) \tag{73}$$

Write σ in the standard Fourier basis as $\sigma(\mathbf{g}) = p + \sum_{\emptyset\subsetneq S\subsetneq[d]} \widehat{\sigma}(S)\omega_S(\mathbf{g})$, where $\widehat{\sigma}(\emptyset) = \mathbf{E}[\sigma] = p$. Then, $\gamma(\mathbf{g}) = \sigma * \sigma(\mathbf{g}) - p^2 = \sum_{\emptyset\subsetneq S\subseteq[d]} \widehat{\sigma}(S)^2\omega_S(\mathbf{g})$. In particular, this means that for any $i \in [d]$, and $\mathbf{h}_{-i} \in \{\pm 1\}^{d-1}$, we have

$$\gamma(\mathbf{g})|_{\mathbf{g}_{-i}=\mathbf{h}_{-i}} = \sum_{i\notin S}\widehat{\sigma}(S)^2\omega_S(\mathbf{h}_{-i}) + \mathbf{g}_i\sum_{i\in S}\widehat{\sigma}(S)^2\omega_{S\setminus\{i\}}(\mathbf{h}_{-i}).$$

It follows that

$$D_i\gamma(\mathbf{h}_{-i}) = 2\sum_{i\in S}\widehat{\sigma}(S)^2\omega_{S\setminus\{i\}}(\mathbf{h}_{-i}) \le 2\sum_{i\in S}\widehat{\sigma}(S)^2 = 2\mathbf{Inf}_i[\sigma]$$

and $\mathbf{Var}_i[\gamma(\mathbf{h}_{-i})] = \Bigg(\sum_{i\in S}\widehat{\sigma}(S)^2\omega_{S\setminus\{i\}}(\mathbf{h}_{-i})\Bigg)^2 \le \Bigg(\sum_{i\in S}\widehat{\sigma}(S)^2\Bigg)^2 = \mathbf{Inf}_i^2[\sigma]$. Therefore, by Lemma 21,

$$\|\gamma\|_k \le C\Bigg(\sqrt{k}\sqrt{\sum_{i=1}^d \mathbf{Inf}_i^2[\sigma]} + k\times\mathbf{MaxInf}[\sigma]\Bigg) = C(\sqrt{k}\times\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2 + k\times\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_\infty.)$$

This implies

$$\|\gamma\|_k^k \le (2C)^k\sqrt{k}^k\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2^k + (2C)^k k^k\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_\infty^k.$$

Plugging this into (12) and using $\mathbf{E}_{\mathbf{g}}[\gamma(\mathbf{g})] = 0$, we obtain the following bound. The computation is analogous to (54).

$$\begin{aligned}
\mathsf{KL}\Big(\mathsf{RAG}(n,\{\pm1\}^d,\sigma,p)\|\mathsf{G}(n,p)\Big) &\le \sum_{k=0}^{n-1}\log\left(\mathbf{E}_{\mathbf{g}}\left[\left(1+\frac{\gamma(\mathbf{g})}{p(1-p)}\right)^k\right]\right)\\
&\le \sum_{k=0}^{n-1}\log\left(1+\sum_{t=2}^k\binom{k}{t}\frac{\mathbf{E}[|\gamma|^t]}{p^t(1-p)^t}\right)\\
&\le n\sum_{k=2}^n\binom{n}{k}\frac{\mathbf{E}[|\gamma|^k]}{p^k(1-p)^k}\\
&\le n\sum_{k\ge2}\binom{n}{k}(2C)^k\sqrt{k}^k\frac{\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2^k}{p^k(1-p)^k}\\
&+ n\sum_{k\ge2}\binom{n}{k}(2C)^k k^k\frac{\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_\infty^k}{p^k(1-p)^k}.
\end{aligned}$$

We now handle the two sums separately. We will use the inequality $\binom{n}{k} \le (ne/k)^k$.

Sum depending on L_2 norm.

$$n\sum_{k\ge2}\binom{n}{k}(2C)^k\sqrt{k}^k\frac{\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2^k}{p^k(1-p)^k} \le n\sum_{k\ge2}\left(\frac{2eCn\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2}{\sqrt{k}p(1-p)}\right)^k. \tag{74}$$

We will show exponential decay in the summands. That is, for all $k \geq 2$,

$$\left(\frac{2eCn\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2}{\sqrt{k}p(1-p)}\right)^k \geq 2\left(\frac{2eCn\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2}{\sqrt{k+1}p(1-p)}\right)^{k+1}.$$

This is equivalent to $\sqrt{k+1} \geq C'\frac{n\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2}{p(1-p)}$ for some absolute constant C'. The latter inequality clearly holds for all $k \geq 2$ by (73). Since there is exponential decay, the term for $k = 2$ is dominant and, thus, the entire expression is of order $O\left(\frac{n^3\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2^2}{p^2(1-p)^2}\right)$.

Sum depending on L_∞ norm. Using the same reasoning, the expression can be bounded by

$$n\sum_{k\geq 2}\left(\frac{2eCn\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_\infty}{p(1-p)}\right)^k. \tag{75}$$

Again, whenever $\frac{n\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_\infty}{p(1-p)} = o(1)$, we have exponential decay. This, however, clearly is the case by (73) as $\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_\infty \leq \|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2$. Thus, the term for $k = 2$ is dominant, so the L_∞ contribution is bounded by $O\left(\frac{n^3\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_\infty^2}{p^2(1-p)^2}\right)$. Combining with the L_2 contribution, the statement follows as $\|\overrightarrow{\mathbf{Inf}}[\sigma]\|_\infty^2 \leq \|\overrightarrow{\mathbf{Inf}}[\sigma]\|_2^2$.

H.3. Comparison of Theorem 12 with Bangachev and Bresler (2023)

In Bangachev and Bresler (2023), the authors prove the following theorem in the same setup.

Theorem 49 (Bangachev and Bresler (2023)) *Consider a dimension $d \in \mathbb{N}$, connection $\sigma : \{\pm 1\}^d \longrightarrow [0,1]$ with expectation p, and absolute constant $m \in \mathbb{N}$. There exists a constant K_m depending only on m, but not on σ, d, n, p, with the following property. Suppose that $n \in \mathbb{N}$ is such that $nK_m < d$. For $1 \leq i \leq d$, let $B_i = \max\left\{|\widehat{\sigma}(S)|\binom{d}{i}^{1/2} : |S| = i\right\}$. Denote also*

$$C_m = \sum_{i=m+1}^{\frac{d}{2en}} B_i^2 + \sum_{i=d-\frac{d}{2en}}^{d-m-1} B_i^2 \quad \textit{and} \quad D = \sum_{\frac{d}{2en}\leq j\leq d-\frac{d}{2en}} B_i^2.$$

If the following conditions additionally hold

- $d \geq K_m \times n \times \left(\frac{C_m}{p(1-p)}\right)^{\frac{2}{m+1}}$,
- $d \geq K_m \times n \times \left(\frac{B_u^2}{p(1-p)}\right)^{\frac{2}{u}}$ *for all* $2 \leq u \leq m$,
- $d \geq K_m \times n \times \left(\frac{B_{d-u}^2}{p(1-p)}\right)^{\frac{2}{u}}$ *for all* $2 \leq u \leq m$,

then

$$\begin{aligned}&\mathsf{TV}\Big(\mathsf{RAG}(n,\{\pm 1\}^d,p,\sigma)\|\mathsf{G}(n,p)\Big)^2\\&\le K_m\times\frac{n^3}{p^2(1-p)^2}\times\left(\sum_{i=1}^m\frac{B_i^4}{d^i}+\sum_{i=d-m}^d\frac{B_i^4}{d^i}+\frac{C_m^2}{d^{m+1}}+D^2\times\exp\left(-\frac{d}{2en}\right)\right).\end{aligned}$$

We make several remarks on the comparison between those two theorems along several lines.

- *Simplicity.* Theorem 12 is much easier to apply than Theorem 49 and its proof is substantially shorter and less involved. In addition, it gives a bound based on influences which are a much more standard quantity in Boolean analysis than the values B_i, C_m, D in Theorem 49.
- *Applicability.* Furthermore, Theorem 12 can be applied in setting when $d = o(n)$ as opposed to Theorem 49. Thus, for example in Bangachev and Bresler (2023) prove the first part of Corollary 13 only when $d = \Omega(n)$, that is $r = O(\sqrt{n})$.
- *Sharpness.* Still, in many cases Theorem 49 is much stronger. For example, consider the double threshold connection $\sigma(\mathbf{g}) = \mathbb{1}\big[|\sum_{i=1}^d g_i| \ge \chi_d\big]$, where χ_d is chosen so that $\mathbf{E}[\sigma] = 1/2$. Then, Theorem 49 implies that $\mathsf{TV}\Big(\mathsf{RAG}(n,\{\pm1\}^d,\sigma,1/2),\mathsf{G}(n,1/2)\Big) = o(1)$ whenever $d = \omega(n^{3/2})$ (Bangachev and Bresler, 2023, Corollary 4.10). However, Theorem 12 only implies this for $d = \omega(n^3)$. The reason Theorem 12 is much weaker in this setting is that the expression $\sum_{i=1}^d \mathbf{Inf}_i^2[\sigma]$ puts a much larger weight on levels close to d. Indeed, note that for $S \subset [d]$, the Fourier coefficient $\widehat{\sigma}(S)$ contributes to $|S|$ of the terms $\mathbf{Inf}_i^2[\sigma]$, but it only contributes once to the expression $\sum_{i=d-m}^d \frac{B_i^4}{d^i}$ from Theorem 49.

Appendix I. Anticoncentration of Convolutions and the Proof of Lemma 39

Suppose that X is a real-valued random variable with density which is absolutely continuous with respect to the Lebesgue density on $\mathbb{R}$. Denote by $M(X) \in \mathbb{R}_+ \cup \{+\infty\}$ the maximum value of the density of X. We will use the following fact from Bobkov and Chistyakov (2014).[13]

Theorem 50 *Suppose that $Y_1, Y_2, \ldots, Y_d$ are independent real random variables with densities absolutely continuous with respect to the Lebesgue measure. Then,*

$$M^{-2}(Y_1 + Y_2 + \cdots + Y_d) \ge \frac{1}{e}\sum_{i=1}^d M^{-2}(Y_i).$$

In particular, when $Y_1, Y_2, \ldots, Y_d$ are iid, this implies that $M(Y_1 + Y_2 + \cdots + Y_d) \le \sqrt{\frac{e}{d}}M(Y_1)$. As already mentioned in Section E, in the setup of Lemma 39, $M(U^q) = +\infty$ when $q > 1$ and, thus, we need to generalize Theorem 50.

Lemma 51 *Suppose that X is a real-valued random variable with the following property. There exists another random variable Y such that*

1. $\mathsf{TV}(X, Y) = 1 - p \in [0, 1)$, *and*
2. *The density of Y is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and $M(Y) = m < \infty$.*

13. The result in Bobkov and Chistyakov (2014) is more general and holds for random variables taking values in any $\mathbb{R}^a$.

Let d be an integer and let $X_1, X_2, \ldots, X_d$ be independent copies of X. Then, there exists a random variable Z_d on $\mathbb{R}$ such that

1. $\mathsf{TV}(X_1 + X_2 + \cdots + X_d, Z_d) \leq \exp(-dp/8)$, *and*
2. *The density of Z_d is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and $M(Z_d) \leq \sqrt{2e}\frac{m}{\sqrt{p^3 d}}$.*

Proof We first introduce two notational conventions.

If $(\mathcal{D}_i)_{i=1}^n$ are probability distributions and $\boldsymbol{p} \in \mathbb{R}^n_{\geq 0}$ is a vector with weights with sum to 1, we define the mixture $\sum_{i=1}^n p_i \mathcal{D}_i$ as follows. First, one takes $B \in [n]$ such that $\mathbf{P}[B = i] = p_i$. Then, one draws $Z \sim \mathcal{D}_B$ independently from B.

If $\mathcal{D}, \mathcal{F}$ are real-valued probability distributions, denote by $\mathcal{D} * \mathcal{F}$ the distribution of $Y_\mathcal{D} + Y_\mathcal{F}$, where $Y_\mathcal{D}, Y_\mathcal{F}$ are independent and $Y_\mathcal{D} \sim \mathcal{D}, Y_\mathcal{F} \sim \mathcal{F}$.

We will use the following trivial identity.

$$\left(\sum_{i=1}^n p_i \mathcal{D}_i\right) * \left(\sum_{j=1}^m q_j \mathcal{F}_j\right) = \sum_{1 \leq i \leq n, 1 \leq j \leq m} p_i q_j \mathcal{D}_i * \mathcal{F}_j.$$

Now, we go back to Lemma 51. Consider such a random variable X and let Y be its corresponding random variable from the statement of the lemma. Consider an optimal coupling (X', Y') of X and Y such that $X' = Y'$ with probability p. Denote by $\mathcal{D}_{\neq}$ the distribution of $X'|X' \neq Y'$ and by $\mathcal{D}_{=}$ the distribution of $X'|X' = Y'$, which is the same as the distribution of $Y'|X' = Y'$. Since Y is absolutely continuous with respect to the Lebesgue measure, so is $Y'|X' = Y'$. Furthermore, the maximum value of the density of $\mathcal{D}_{=}$ is at most mp^{-1} as m is the maximum value of the density of Y and $\mathbf{P}[X' = Y'] = p$.

In particular, note that the distribution $\mathcal{D}$ of X is the mixture $(1-p)\mathcal{D}_{\neq} + p\mathcal{D}_{=}$, where $\mathcal{D}_{=}$ is absolutely continuous with respect to the Lebesgue measure and its density is bounded by mp^{-1}. Therefore, the distribution of $X_1 + X_2 + \cdots + X_d$ is the mixture

$$\begin{aligned}
&\sum_{k=0}^d \binom{d}{k} p^k (1-p)^{d-k} (\mathcal{D}_{=})^{*k} * (\mathcal{D}_{\neq})^{*(d-k)} \\
&= \sum_{k < dp/2} \binom{d}{k} p^k (1-p)^{d-k} (\mathcal{D}_{=})^{*k} * (\mathcal{D}_{\neq})^{*(d-k)} \\
&+ \sum_{k \geq dp/2} \binom{d}{k} p^k (1-p)^{d-k} (\mathcal{D}_{=})^{*k} * (\mathcal{D}_{\neq})^{*(d-k)}.
\end{aligned} \tag{76}$$

We now show the following two facts. First, the weight on summands $k < dp/2$ is at most $\exp(-dp/8)$, which means that $X_1 + X_2 + \cdots + X_d$ is $\exp(-dp/8)$-close to the mixture $\sum_{k \geq dp/2} \binom{d}{k} p^k (1-p)^{d-k} (\mathcal{D}_{=})^{*k} * (\mathcal{D}_{\neq})^{*(d-k)}$. On the other hand, the latter mixture is absolutely continuous with respect to the Lebesgue measure and has density bounded by $\sqrt{2e}\frac{m}{\sqrt{p^3 d}}$. We begin with the first part.

Lemma 52 $\displaystyle\sum_{k < dp/2} \binom{d}{k} p^k (1-p)^{d-k} \leq \exp(-dp/8).$

Proof This is a trivial application of Chernoff bounds. Let $V_1, V_2, \ldots, V_d$ be iid $\mathsf{Bern}(p)$ random variables. Then,

$$\sum_{k<dp/2} \binom{d}{k} p^k (1-p)^{d-k} = \mathbf{P}\left[\sum_{i=1}^{d} V_i < dp(1-1/2)\right] \le \exp(-dp/8).$$

■

Lemma 53 *For each $k \ge dp$, $M\Big((\mathcal{D}_=)^{*k} * (\mathcal{D}_{\neq})^{*(d-k)}\Big) \le \sqrt{2e}\frac{m}{\sqrt{dp^3}}$.*

Proof Note that the density of $\mathcal{D}_=$ is at most mp^{-1} as discussed. Therefore, by Theorem 50, we immediately obtain

$$M((\mathcal{D}_=)^{*k}) \le \sqrt{\frac{e}{k}} mp^{-1} \le \sqrt{\frac{2e}{dp}} mp^{-1}.$$

This is enough since $M\Big((\mathcal{D}_=)^{*k} * (\mathcal{D}_{\neq})^{*(d-k)}\Big) \le M((\mathcal{D}_=)^{*k}) \le \sqrt{2e}\dfrac{m}{\sqrt{dp^3}}$. ■

Now let $\mathcal{D}_<$ be an arbitrary random variable on $\mathbb{R}$ with maximal density at most $\sqrt{2e}\frac{m}{p\sqrt{dp}}$. Consider Z distributed according to

$$Z \sim \left(\sum_{k\ge dp/2} \binom{d}{k} p^k (1-p)^{d-k}\right) \mathcal{D}_< + \sum_{k<dp/2} \binom{d}{k} p^k (1-p)^{d-k} (\mathcal{D}_=)^{*k} * (\mathcal{D}_{\neq})^{*(d-k)}.$$

Lemma 52 implies $\mathsf{TV}(X_1 + X_2 + \cdots + X_d, Z) \le \exp(-dp/8)$. Lemma 53 implies that $M(Z) \le \sqrt{2e}\frac{m}{\sqrt{dp^3}}$. ■

An immediate corollary of Lemma 51 is Lemma 8 which we use to prove Lemma 39.

Proof [Proof of Lemma 8] Let $\Omega = \{x \ge 0 \ : \ f(x) \le m\}$. Clearly, $\int_\Omega f(x)dx = p$. Let Y be the real-valued random variable with density $f(x)$ for $x \in \Omega$ and density equal to m on $[-(1-p)/m, 0]$. Then, the density of Y is bounded by m and $\mathsf{TV}(Y, X) = 1 - p$ (as the two densities agree on Ω which has measure p). Now, we simply find the random variable Z_d given by Lemma 51 and observe that for an optimal coupling of $Z_d, X_1 + X_2 + \cdots + X_d$, we have

$$\begin{aligned}
&\mathbf{P}\big[X_1 + X_2 + \cdots + X_d \in [a,b]\big] \\
&\le \mathbf{P}[X_1 + X_2 + \cdots + X_d \neq Z_d] + \mathbf{P}[Z_d \in [a,b]] \le \exp(-dp/8) + M(Z_d)(b-a),
\end{aligned}$$

from which the claim follows. ■

Proof [Proof of Lemma 38] We apply Corollary 8 as follows. Consider the random variable U^q, where $U \sim \mathsf{Unif}([0,1])$. The CDF $\phi(x)$ of U^q for $x \in [0,1]$ is

$$\phi(x) = \mathbf{P}[U^q \le x] = \mathbf{P}[U \le x^{1/q}] = x^{1/q}.$$

Thus, the density $h(x)$ of U^q is $h(x) = (x^{1/q})' = \frac{1}{q}x^{1/q-1}\mathbb{1}\big[x \in (0,1]\big]$. Now, observe that $h(x) \le \frac{1}{q}(1/2)^{1/q-1} \le \frac{2}{q}$ for $x \in [1/2,1]$ and also

$$p \coloneqq \mathbf{P}\big[U^q \in [1/2,1]\big] = 1 - \mathbf{P}\big[U^q \le 1/2\big] = 1-(1/2)^{1/q} \ge \frac{1}{2q}.$$

Thus, applying Corollary 8 with $m = \frac{1}{q}(1/2)^{1/q-1} \le \frac{2}{q}, p \ge \frac{1}{2q}$ gives the result. ■

Proof [Proof of Lemma 39] Using Lemma 38,

$$\psi(\ell) = \mathbf{P}\big[U_1^q + U_2^q + \cdots + U_{d-1}^q \in [\tau^q, \tau^q - \ell^q]\big] \le \exp(-\Omega(d/q)) + O\left(\sqrt{\frac{q}{d}}\ell^q\right).$$

Using that $d/q = \omega(\log d)$ and integrating over $[0,1]$, we conclude

$$\begin{aligned}\int_0^1 \psi(\ell)d\ell &= \exp(-\Omega(d/q)) + O\left(\sqrt{\frac{q}{d}}\int_0^1 \ell^q d\ell\right)\\ &= \exp(-\Omega(d/q)) + O\left(\sqrt{\frac{1}{qd}}\right) = O\left(\sqrt{\frac{1}{qd}}\right).\end{aligned}$$

Similarly,

$$\begin{aligned}\int_0^1 \psi^2(\ell)d\ell &= \int_0^1 \left(\exp(-\Omega(d/q)) + O\left(\sqrt{\frac{q}{d}}\ell^q\right)\right)^2 d\ell\\ &= O(\exp(-\Omega(d/q))) + O\left(\frac{q}{d}\int_0^1 \ell^{2q}d\ell\right) = O\left(\frac{1}{d}\right).\end{aligned}$$

■