Bernoulli **29**(3), 2023, 2417–2441
https://doi.org/10.3150/22-BEJ1547

A probabilistic view of latent space graphs and phase transitions

SUQI LIU[a] and MIKLÓS Z. RÁCZ[b]

Department of Operations Research and Financial Engineering, Princeton University, Sherrerd Hall, Charlton Street, Princeton, NJ 08544, USA, [a]*suqil@princeton.edu,* [b]*mracz@princeton.edu*

We study random graphs with latent geometric structure, where the probability of each edge depends on the underlying random positions corresponding to the two endpoints. We consider the setting where this conditional probability is a general monotone increasing function of the inner product of two vectors; such a function can naturally be viewed as the cumulative distribution function of some independent random variable. A one-parameter family of random graphs, characterized by the variance of this random variable, that smoothly interpolates between a random dot product graph and an Erdős–Rényi random graph, is investigated. Focusing on the dense regime, we prove phase transitions of detecting geometry in these graphs, in terms of the dimension of the underlying geometric space and the variance parameter: When the dimension is high or the variance is large, the graph is similar to an Erdős–Rényi graph with the same edge density; in other parameter regimes, there is a computationally efficient signed triangle statistic that can distinguish them. The proofs make use of information-theoretic inequalities and concentration of measure phenomena.

Keywords: Random graph; random dot product graph; high-dimensional geometric structure; signed triangle

1. Introduction

Random graphs constructed with some latent geometric space are widely used to model a large variety of real-life networks. Random dot product graphs are a natural family of random graphs with simple latent geometry where the probability of connection depends on the inner product of two vectors in a Euclidean space [1,30]. For a graph on a set of vertices $V = [n] := \{1,2,\dots,n\}$, we write $i \sim j$ if and only if vertices i and j are connected by an undirected edge. Let $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n \in \mathbb{R}^d$ be independent identically distributed random vectors. In the general setting of random dot product graphs, conditioned on the latent vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, the event $i \sim j$ happens with probability $\sigma(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle)$, independently of everything else, where $\langle\cdot,\cdot\rangle$ denotes the inner product and $\sigma : \mathbb{R} \to [0,1]$ is usually a monotone increasing function called the *connection function*. That is,

$$\mathbb{P}(i \sim j \mid \boldsymbol{x}_1,\dots,\boldsymbol{x}_n) = \sigma(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle).$$

The distribution of the random graph is then specified by

$$\mathbb{P}(G) = \mathbb{E}\left[\prod_{i<j} \sigma(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle)^{a_{i,j}}(1-\sigma(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle))^{1-a_{i,j}}\right],$$

where $\boldsymbol{A} = [a_{i,j}]$ is the adjacency matrix of the graph G.

The properties of the connection function σ play a key role in the presence of geometry in the random graph. When σ is a threshold function, i.e., $\sigma(x) = \mathbb{1}\{x \ge t\}$, then such a graph is known simply as a (hard) *random dot product graph* [16]. At the other extreme, if σ is a constant, we have an Erdős–Rényi random graph, with geometry lost. In between, intuitively, when the connection function is “steep”, the edges are correlated through the latent geometric space; when the connection function is “flat”, the edges become less dependent of each other. The flatness of the connection function can also be understood as the level of noise: A connection function that is close to a constant implies large noise in a geometric graph. The trade-off between noise and dimensionality in detecting geometry in random graphs was first studied by the authors [16], where a particular one-parameter family of connection functions, consisting of step functions, were considered. An immediate question is how does this trade-off generalize to other connection functions (in particular, smooth ones) that are widely used in practice? We attempt to answer this question by studying a natural one-parameter family of smooth connection functions that interpolates between the two extremes.

1.1. A probabilistic view of the connection function

We focus on the case when the latent positions are independent standard normal random vectors: $\boldsymbol{x}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_d)$ for $1 \le i \le n$. We consider a broad class of connection functions bearing a probabilistic view. Observing that the connection function is usually monotone increasing between 0 and 1, it is natural to view it as the cumulative distribution function (CDF) of some random variable. Keeping this setup in mind, we seek a parametrization that characterizes the “flatness” of the function and interpolates between the random geometric graph and the Erdős–Rényi model. A natural parameter is the variance of this random variable. If the variance is large, the CDF changes slowly. On the other hand, if the variance goes to zero, the random variable converges to a constant and the CDF becomes a threshold function. We formulate the idea as follows.

Let $\mathcal{D}(0,1)$ be an arbitrary zero mean and unit variance distribution with the CDF denoted by $F : \mathbb{R} \to [0,1]$. Suppose the probability measure is absolutely continuous with respect to Lebesgue measure and let f be the probability density function (PDF). We assume that the PDF is strictly positive:

$$f(x) > 0, \quad \forall x \in \mathbb{R}. \tag{A0}$$

We also assume that f is continuously differentiable and the derivative f' is bounded:

$$\alpha := \sup_x |f'(x)| < \infty. \tag{A1}$$

For technical reasons, we further assume that the second order derivative f'' exists and for any fixed Gaussian random variable X,

$$\mathbb{E}[|f''(X)|] < \infty. \tag{A2}$$

We then create a one-parameter family of connection functions using F. The process can be viewed as scaling and translation of a random variable following $\mathcal{D}(0,1)$. The “flatness” of the connection function is parametrized by r, which measures the deviation of the random variable from a constant. However, since the inner product of two d-dimensional standard normal random vectors has variance equal to d, we account for it by letting the variance be $r^2 d$. We also need to match the marginal probability of an edge, or edge density, in the graph, which is achieved by choosing an appropriate mean $\mu_{p,d,r}$. Then, we can view the connection function as the CDF of the distribution $\mathcal{D}(\mu_{p,d,r}, r^2 d)$:

$$\sigma(x) := F\left(\frac{x - \mu_{p,d,r}}{r\sqrt{d}}\right), \tag{1}$$

where $\mu_{p,d,r}$ is determined by setting the edge density in the graph to be equal to p:

$$\mathbb{P}(i \sim j) = \mathbb{E}[\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle)] = \mathbb{E}\left[F\left(\frac{\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)\right] = p.$$

We denote the random graph constructed in this way by $\mathcal{G}(n,p,d,r)$.

The edge generating process for $\mathcal{G}(n,p,d,r)$ can also be viewed as adding noise to the hard random dot product graph. For each pair of vertices i and j, we draw an independent random variable $z_{i,j} \sim \mathcal{D}(\mu_{p,d,r}, r^2 d)$. Then, conditioned on the latent positions $\boldsymbol{x}_i, \boldsymbol{x}_j$, and the random variable $z_{i,j}$, vertices i and j are connected if and only if $\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle \geq z_{i,j}$. In other words, for a graph with adjacency matrix $\boldsymbol{A}$,

$$a_{i,j} = \mathbb{1}\{i \sim j\} = \mathbb{1}\{\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle \geq z_{i,j}\}.$$

In contrast, in the conventional definition of a hard random dot product graph, the connection between i and j is determined by comparing the inner product with a threshold that is constant [16].

Canonical examples of the connection function, which satisfy the assumptions above, include

- logistic (here $\alpha = \pi^2/(18\sqrt{3})$):

$$F(x) = \frac{1}{1 + \exp(-\pi x/\sqrt{3})},$$

- Gaussian (here $\alpha = 1/\sqrt{2e\pi}$):

$$F(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{y^2}{2}\right) dy.$$

The assumptions are also satisfied broadly by many more CDFs.

For a random graph with latent geometric structure, the only visible structure is the connectivity, while the underlying random vectors are not observed. It is natural to ask for a given graph, if geometry is reflected in the combinatorial structure. A canonical random graph model that does not possess geometry is the Erdős–Rényi graph $\mathcal{G}(n,p)$, in which the edges are generated independently with probability p. Therefore, in this work, we focus on how $\mathcal{G}(n,p,d,r)$ compares to $\mathcal{G}(n,p)$. The detection of geometry can then be understood through bounds on total variation distance between them which will be defined formally in Section 2. Roughly speaking, when the total variation distance is close to 0, there is no algorithm that can tell the difference.

1.2. Main result

Our main result is summarized as the following theorem, where $\mathrm{TV}(\cdot,\cdot)$ stands for total variation distance as defined in Definition 2.2(a).

Theorem 1.1. *Let $\mathcal{G}(n,p,d,r)$ be defined above with p fixed in $(0,1)$ and $r \geq 1$.*

(a) Assume (A1). If

$$\frac{n^3}{r^4 d} \to 0,$$

then $\mathrm{TV}(\mathcal{G}(n,p,d,r), \mathcal{G}(n,p)) \to 0$.

(b) Additionally, assume (A0) and (A2). Suppose that $d/\log^2 d \gg r^6$ or $r/\log^2 r \gg d^{1/6}$. If

$$\frac{n^3}{r^6 d} \to \infty,$$

then $\mathrm{TV}(\mathcal{G}(n,p,d,r), \mathcal{G}(n,p)) \to 1$.

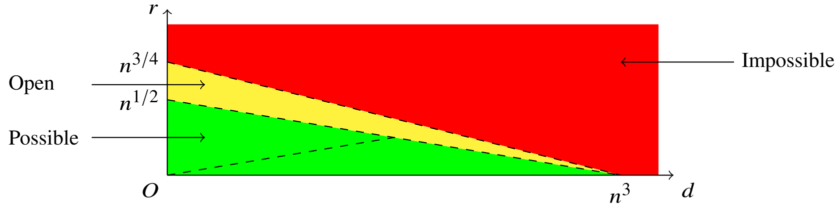

Figure 1. Phase diagram for detecting geometry in $\mathcal{G}(n, p, d, r)$.

Theorem 1.1 can be displayed graphically by a phase diagram in the space of the dimension d and the parameter r, as shown in Figure 1.

Remark 1.2. There is an intermediate regime that is not covered by the theorem. We believe the phase transitions happen at a certain power of r. We conjecture that Theorem 1.1(b) gives the actual threshold. In the proof of Theorem 1.1(a), a factor of r^2 might have been lost when applying Jensen's inequality to the KL divergence (see Section 3).

Remark 1.3. The noisy high-dimensional random geometric graphs studied in [16] do not fall into the family of graphs considered in this work since the connection functions there are not continuously differentiable, which is crucial in the proofs. Nevertheless, we can draw some comparisons between them. We see that the signed triangle statistic gives the same detection boundary in both models. This could be understood as despite the discrepancies, the parameters are both a linear measurement of the deviation from a constant.

Remark 1.4. The theorem is stated in the case $r \geq 1$, which is the regime where there is an interplay between r and d. When $r = o(1)$, by the data processing inequality, the total variation distance between $\mathcal{G}(n, p, d, r)$ and $\mathcal{G}(n, p)$ is upper bounded by the Wishart to GOE transition, which gives the condition $d \gg n^3$ for the impossibility of detecting geometry [16]. At the same time, by the calculation in [16, Section 5], the detection power for the signed triangle statistic when $r \to 0$ (the connection function becomes an indicator function) is $d \ll n^3$. Combining these, we have that the phase transition is at $d \asymp n^3$ regardless of the rate of r.

Remark 1.5. The main result focuses on the dense regime, with $p \in (0, 1)$ fixed. Although the dependency on p is explicit in some steps of the proofs, we do not expect it to be the precise characterization. Instead, we focus on making the assumptions on the connection function as weak as possible. Several important quantities cannot be made precise without further knowledge of the connection function. Nevertheless, it would be interesting to study, for particular families of functions, how the bounds would depend on p precisely, thus shedding light on the sparse regime. This is beyond the scope of the current work.

1.3. Related work

The study of random graphs generated from latent positions traces back to the work of Gilbert [12] in the early 1960s, illustrated by applications to communication networks. After several decades, a latent space model with large flexibility was proposed in [13] and applied to social network analysis, which popularized the modern study of latent space graphs in statistics. The inner product model was later generalized to a latent position graph model equipped with a continuous positive definite kernel in [26], where feature map estimation and universally consistent vertex classification were discussed. A recent work [18] studied a similar inner product model focusing on model fitting methods. We refer to [1] for a survey on such topics in random dot product graphs. Another family of random graphs with latent geometric structure is the random geometric graphs [20]. The soft variants wherein the connection probability depends on the distance of latent positions through a function were also investigated in [21].

The study of random geometric graphs in high dimensions originates from the pioneering work of Devroye, György, Lugosi, and Udina [8], where they used a multivariate central limit theorem to show that the graph becomes similar to an Erdős–Rényi graph when the dimension grows. Subsequent work of Bubeck, Ding, Eldan, and Rácz [5] determined that the dimension threshold is $d \asymp n^3$ at which the phase transition of losing geometry happens in dense random geometric graphs. In prior work, we generalized this phase transition phenomena to a noisy setting and studied the trade-off between noise and dimensionality for the first time [16]; this paper is the closest to the current one, see Remark 1.3 above for a discussion. An excellent recent survey [9] provides a detailed summary of progress and discussions of open questions on these problems (see also [23]).

Lying underneath the loss of geometry in random graphs is the Wishart to GOE transition in high dimensions, and a line of work explores this direction [5,6,15,24]. In particular, the phase transition was shown for log-concave measures using entropy-based methods in [6]. This was further extended to an anisotropic setting [10]. In a recent work, masked Wishart matrices were considered, and phase transitions were proven to matching orders in various types of combinatorial masks [3].

1.4. Organization

The rest of the paper is organized as follows. In Section 2, we introduce several notations and key facts used throughout the paper. Section 3 consists of the proof of Theorem 1.1(a), where several information-theoretic inequalities are used. Detecting geometry using the signed triangle statistic is presented in Section 4, where the main body consists of estimating the expectation of a signed triangle in $\mathcal{G}(n,p,d,r)$ in two different parameter regimes. Applying Chebyshev's inequality with these estimates concludes the proof of Theorem 1.1(b).

2. Notations and preliminaries

A graph $G=(V,E)$ is a tuple consisting of a set of vertices $V=[n]:=\{1,2,\dots,n\}$ and a set of edges $E\subset\binom{[n]}{2}$, where the collection of all subsets of a set S with cardinality k is denoted by $\binom{S}{k}$. We use $\|\cdot\|$ to denote the Euclidean norm of a vector. For a random variable $X\in\mathcal{X}$ and a measurable function $f:\mathcal{X}\to\mathbb{R}$, $\|f(X)\|_p$ denotes the L^p-norm of $f\in L^p(\mathcal{X})$.

Our proofs build upon various inequalities involving f-divergences, which are briefed here. The definitions and properties stated in this section can be found in most standard texts on information theory (see, e.g., [22, Chapter 6]).

To begin with, we state the definition of an f-divergence.[1]

[1] Note that the f in f-divergence should not be confused with the probability density function f in the description of $\mathcal{G}(n,p,d,r)$. In general, f is overloaded in this section, but the meaning of f will always be clear from the context.

Definition 2.1 (f-divergence). Let $\mathcal{P}$ and $\mathcal{Q}$ be probability measures on the measurable space $(\Omega, \mathcal{F})$. Suppose that $\mathcal{P}$ is absolutely continuous with respect to $\mathcal{Q}$. For a convex function f such that $f(1) = 0$, the f-divergence of $\mathcal{P}$ and $\mathcal{Q}$ is defined as

$$\mathrm{D}_f(\mathcal{P} \,\|\, \mathcal{Q}) := \mathbb{E}_{\mathcal{Q}}\left[f\left(\frac{d\mathcal{P}}{d\mathcal{Q}}\right)\right] = \int_\Omega f\left(\frac{d\mathcal{P}}{d\mathcal{Q}}\right) d\mathcal{Q},$$

where $\frac{d\mathcal{P}}{d\mathcal{Q}}$ is the Radon–Nikodym derivative of $\mathcal{P}$ with respect to $\mathcal{Q}$.

Different choices of f lead to the following f-divergences encountered in the proofs.

Definition 2.2. Let $\mathcal{P}$ and $\mathcal{Q}$ be probability measures on the measurable space $(\Omega, \mathcal{F})$ such that $\mathcal{P}$ is absolutely continuous with respect to $\mathcal{Q}$.

(a) *Total variation distance* (corresponding to $f(x) = \frac{1}{2}|x-1|$):

$$\mathrm{TV}(\mathcal{P}, \mathcal{Q}) := \sup_{A \in \mathcal{F}} |\mathcal{P}(A) - \mathcal{Q}(A)| = \frac{1}{2}\int_\Omega \left|\frac{d\mathcal{P}}{d\mathcal{Q}} - 1\right| d\mathcal{Q}.$$

(b) *Kullback–Leibler (KL) divergence* (corresponding to $f(x) = x \log x$):

$$\mathrm{KL}(\mathcal{P} \,\|\, \mathcal{Q}) := \mathbb{E}_{\mathcal{P}}\left[\log \frac{d\mathcal{P}}{d\mathcal{Q}}\right] = \int_\Omega \frac{d\mathcal{P}}{d\mathcal{Q}} \log \frac{d\mathcal{P}}{d\mathcal{Q}}\, d\mathcal{Q}.$$

(c) *(Pearson) χ^2-divergence* (corresponding to $f(x) = (x-1)^2$):

$$\chi^2(\mathcal{P} \,\|\, \mathcal{Q}) := \mathbb{E}_{\mathcal{Q}}\left[\left(\frac{d\mathcal{P}}{d\mathcal{Q}} - 1\right)^2\right] = \mathbb{E}_{\mathcal{Q}}\left[\left(\frac{d\mathcal{P}}{d\mathcal{Q}}\right)^2\right] - 1.$$

The f-divergences defined above are connected through the following inequalities (see [11]).

Proposition 2.3 (Pinsker's inequality). *Let $\mathcal{P}$ and $\mathcal{Q}$ be probability measures on the measurable space $(\Omega, \mathcal{F})$ such that $\mathcal{P}$ is absolutely continuous with respect to $\mathcal{Q}$. Then*

$$\mathrm{TV}(\mathcal{P}, \mathcal{Q}) \le \sqrt{\frac{1}{2}\mathrm{KL}(\mathcal{P} \,\|\, \mathcal{Q})}.$$

Proposition 2.4. *Let $\mathcal{P}$ and $\mathcal{Q}$ be probability measures on the measurable space $(\Omega, \mathcal{F})$ such that $\mathcal{P}$ is absolutely continuous with respect to $\mathcal{Q}$. Then*

$$\mathrm{KL}(\mathcal{P} \,\|\, \mathcal{Q}) \le \log(1 + \chi^2(\mathcal{P} \,\|\, \mathcal{Q})).$$

For KL divergence, a useful property is the chain rule, stated as the following proposition.

Proposition 2.5 (Chain rule). *For joint distributions $\mathcal{P}_{X,Y} = \mathcal{P}_{X|Y}\mathcal{P}_Y$ and $\mathcal{Q}_{X,Y} = \mathcal{Q}_{X|Y}\mathcal{Q}_Y$, the chain rule for KL divergence reads*

$$\mathrm{KL}(\mathcal{P}_{X,Y} \,\|\, \mathcal{Q}_{X,Y}) = \mathrm{KL}(\mathcal{P}_Y \,\|\, \mathcal{Q}_Y) + \mathbb{E}_{\mathcal{P}_Y} \mathrm{KL}(\mathcal{P}_{X|Y} \,\|\, \mathcal{Q}_{X|Y}).$$

We show the following lemma for Lebesgue integrable functions with bounded derivatives.

Lemma 2.6. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is an integrable function on $\mathbb{R}$ that is continuously differentiable. If f' is bounded, then f is bounded as well. Further, let $M := \int_{\mathbb{R}} |f(x)|\, dx < +\infty$ and $\alpha := \sup_x |f'(x)|$; then $\sup_x |f(x)| \le 2\sqrt{M\alpha}$.*

Proof. For all $a \in \mathbb{R}$ and $b > 0$, consider the integral

$$\int_a^{a+b} |f(x)|\, dx \le \int_{\mathbb{R}} |f(x)|\, dx = M.$$

Let $f_m := \inf_{x\in[a,a+b]} |f(x)|$. Since $f_m \le |f(x)|$ for $x \in [a, a + b]$,

$$b f_m = \int_a^{a+b} f_m\, dx \le \int_a^{a+b} |f(x)|\, dx \le M,$$

which gives $f_m \le M/b$.

Since f is continuous, there exists a number $c \in [a, a + b]$ such that $|f(c)| = f_m$. The mean value theorem gives that for $\xi \in [a, a + b]$,

$$f(a) \le f_m + |f'(\xi)| b \le \frac{M}{b} + \alpha b.$$

Applying the proof to $-f$, one could get

$$|f(a)| \le \frac{M}{b} + \alpha b.$$

If $\alpha > 0$, by choosing $b = \sqrt{M/\alpha}$, the claim directly follows.

If $\alpha = 0$, then f is constant on $\mathbb{R}$. Since f is integrable, $f(x) = 0$ for all $x \in \mathbb{R}$. Thus, the claim is also true. □

Applying Lemma 2.6 to a probability density function, we have the following corollary.

Corollary 2.7. *Let $f : \mathbb{R} \to \mathbb{R}_+$ be a probability density function that is continuously differentiable. If $\sup_x |f'(x)| \le \alpha$, then $\sup_x |f(x)| \le 2\sqrt{\alpha}$.*

For standard normal random variables, Stein's lemma (also known as Gaussian integration by parts) provides a powerful tool and is frequently used in the proofs. The lemma and Stein's method built upon it are broadly used in probability and statistics (see, e.g., [14, Example 13.13], also [7]). We state it as the following proposition.

Proposition 2.8 (Stein's lemma [25, Lemma 1]). *Let Y be a $\mathcal{N}(0, 1)$ real random variable and let $g : \mathbb{R} \to \mathbb{R}$ be an indefinite integral of the Lebesgue measurable function g', essentially the derivative of g. Suppose also that $\mathbb{E}[|g'(Y)|] < \infty$. Then,*

$$\mathbb{E}[g'(Y)] = \mathbb{E}[Y g(Y)].$$

For a continuously differentiable function of a standard normal random vector, the following proposition provides a sharp bound for the variance, known as the Gaussian Poincaré inequality (see [2, Theorem 3.20]).

Proposition 2.9 (Gaussian Poincaré inequality). *Suppose $\boldsymbol{x} = (x_1, \ldots, x_d)$ is a vector of i.i.d. standard Gaussian random variables. Let $f : \mathbb{R}^d \to \mathbb{R}$ be any continuously differentiable function. Then,*

$$\mathbb{V}\mathrm{ar}[f(\boldsymbol{x})] \leq \mathbb{E}[\|\nabla f(\boldsymbol{x})\|^2].$$

We also make frequent use of the properties of sub-exponential random variables. We state the definition in terms of the moment generating function and the equivalent tail bound, which appear in most texts (see, e.g., [28, Definition 2.7 and Proposition 2.9]).

Definition 2.10. A random variable X is *sub-exponential* if there are nonnegative parameters (a, b) such that

$$\log \mathbb{E}[e^{t(X-\mathbb{E}[X])}] \leq \frac{a^2 t^2}{2} \qquad \text{for all } |t| < \frac{1}{b}.$$

Proposition 2.11. *Suppose X is sub-exponential with parameters (a, b). Then,*

$$\mathbb{P}(X - \mathbb{E}[X] \geq t) \leq \begin{cases} \exp\left(-\frac{t^2}{2a^2}\right) & \text{if } 0 \leq t \leq \frac{a^2}{b}, \\ \exp\left(-\frac{t}{2b}\right) & \text{for } t > \frac{a^2}{b}. \end{cases}$$

Equivalently,

$$\mathbb{P}(X - \mathbb{E}[X] \geq t) \leq \exp\left(-\frac{1}{2} \min\left\{\frac{t^2}{a^2}, \frac{t}{b}\right\}\right).$$

The following concentration lemma regarding the normal distribution is used in various proofs.

Lemma 2.12. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuously differentiable function. If the norm of the gradient satisfies $\|\nabla f(\boldsymbol{x})\| \leq \|\boldsymbol{x}\|$, then for $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)$, $f(\boldsymbol{x})$ is sub-exponential with parameters $(2\sqrt{d}, 1)$. In particular, the tails of $f(\boldsymbol{x})$ satisfy*

$$\mathbb{P}(|f(\boldsymbol{x}) - \mathbb{E}[f(\boldsymbol{x})]| \geq t) \leq \begin{cases} 2\exp\left(-\frac{t^2}{8d}\right) & \text{if } 0 \leq t \leq 4d, \\ 2\exp\left(-\frac{t}{2}\right) & \text{for } t > 4d. \end{cases}$$

Proof. Without loss of generality, we may assume $\mathbb{E}[f(\boldsymbol{x})] = 0$.

For a nonnegative random variable Z, the *entropy* of Z is defined as

$$\mathrm{Ent}(Z) := \mathbb{E}[\varphi(Z)] - \varphi(\mathbb{E}[Z]),$$

where $\varphi(x) = x \log x$ (see, e.g., [2]; note that this notion of entropy is not to be confused with the Shannon entropy).

By the Gaussian logarithmic Sobolev inequality (see, e.g., [2, Theorem 5.4]),

$$\mathrm{Ent}(e^{tf(\boldsymbol{x})}) \leq 2\,\mathbb{E}[\|\nabla e^{tf(\boldsymbol{x})/2}\|^2] = \frac{t^2}{2}\,\mathbb{E}[e^{tf(\boldsymbol{x})}\|\nabla f(\boldsymbol{x})\|^2] \leq \frac{t^2}{2}\,\mathbb{E}[\|\boldsymbol{x}\|^2 e^{tf(\boldsymbol{x})}].$$

The duality formula of the entropy (see [2, Remark 4.4]) implies that for any random variable W such that $\mathbb{E}[e^W] < \infty$, the entropy of e^{tZ} for a random variable Z satisfies

$$\mathbb{E}[We^{tZ}] \leq \mathbb{E}[e^{tZ}] \log \mathbb{E}[e^W] + \mathrm{Ent}(e^{tZ}).$$

Applying the inequality with $Z = f(\boldsymbol{x})$ and $W = \frac{e-1}{2e}\|\boldsymbol{x}\|^2$, we have that

$$\mathbb{E}[\|\boldsymbol{x}\|^2 e^{tf(\boldsymbol{x})}] \le \frac{2e}{e-1}\,\mathbb{E}[e^{tf(\boldsymbol{x})}] \log \mathbb{E}\left[\exp\left(\frac{e-1}{2e}\|\boldsymbol{x}\|^2\right)\right] + \operatorname{Ent}(e^{tf(\boldsymbol{x})}).$$

Since $\boldsymbol{x} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_d)$, $\|\boldsymbol{x}\|^2$ has a chi-squared distribution with d degrees of freedom. By the moment generating function of the chi-squared distribution, we have that

$$\log \mathbb{E}\left[\exp\left(\frac{e-1}{2e}\|\boldsymbol{x}\|^2\right)\right] = \log\left(1 - 2 \cdot \frac{e-1}{2e}\right)^{-d/2} = \frac{d}{2}.$$

Hence, by putting together the previous displays, we have that

$$\operatorname{Ent}(e^{tf(\boldsymbol{x})}) \le \frac{t^2}{2}\left(\frac{ed}{e-1}\,\mathbb{E}[e^{tf(\boldsymbol{x})}] + \operatorname{Ent}(e^{tf(\boldsymbol{x})})\right) \le \frac{t^2}{2}\left(2d\,\mathbb{E}[e^{tf(\boldsymbol{x})}] + \operatorname{Ent}(e^{tf(\boldsymbol{x})})\right).$$

By rearranging this inequality, we have that

$$\operatorname{Ent}(e^{tf(\boldsymbol{x})}) \le \frac{dt^2}{1 - t^2/2}\,\mathbb{E}[e^{tf(\boldsymbol{x})}] \le 2dt^2\,\mathbb{E}[e^{tf(\boldsymbol{x})}],$$

where the second inequality holds for $|t| \le 1$. Write $M(t) := \mathbb{E}[e^{tf(\boldsymbol{x})}]$, and note that $\operatorname{Ent}(e^{tf(\boldsymbol{x})}) = tM'(t) - M(t)\log M(t)$. Therefore, the inequality in the previous display becomes

$$tM'(t) - M(t)\log M(t) \le 2dt^2 M(t).$$

Solving this equation exactly, and noting that $M(0) = 1$, gives that

$$\log \mathbb{E}[e^{tf(\boldsymbol{x})}] \le 2dt^2.$$

By Definition 2.10, the claim is hence proved.

Further by Proposition 2.11, the tail bound directly follows. □

Remark 2.13. This lemma can also be derived from an exponential Poincaré inequality [27, problem 3.16], as pointed out by Ramon van Handel (personal communication).

The following proposition characterizes the tail behavior of the inner product of two independent d-dimensional standard normal random vectors. The sub-exponential tails in the proposition can be derived from Lemma 2.12; however, since the function is explicit, we prove it directly using the moment generating function, resulting in slightly different constants.

Proposition 2.14. *For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ independently distributed as $\mathcal{N}(\mathbf{0}, \boldsymbol{I}_d)$, the inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is sub-exponential with parameters $(\sqrt{2d}, \sqrt{2})$.*

Proof. Consider two independent random variables $X, Y \sim \mathcal{N}(0,1)$. The moment generating function of their product satisfies

$$\mathbb{E}[e^{tXY}] = \iint \frac{1}{2\pi} e^{txy} e^{-(x^2+y^2)/2}\,dx\,dy = \begin{vmatrix} 1 & -t \\ -t & 1 \end{vmatrix}^{-1/2} = \frac{1}{\sqrt{1-t^2}}.$$

Since $(1-x)^{-1/2} \le e^x$ for $0 \le x \le 1/2$, we have for $t^2 \le 1/2$ that $\mathbb{E}[e^{tXY}] \le e^{t^2}$.

Consequently, for two independent standard normal random vectors $\boldsymbol{x}, \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)$, we have that

$$\mathbb{E}[e^{t\langle \boldsymbol{x},\boldsymbol{y}\rangle}] = \mathbb{E}[e^{t\sum_{i=1}^d x_i y_i}] = \prod_{i=1}^d \mathbb{E}[e^{t x_i y_i}] \le e^{dt^2}$$

for $t^2 \le 1/2$. □

We state a lemma concerning the concentration of the inner product of two independent random vectors uniformly distributed on the unit sphere $\mathbb{S}^{d-1}$.

Lemma 2.15. *For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ independently uniformly distributed on the unit sphere $\mathbb{S}^{d-1}$, when $t \ge 1$ and $d \ge 2$,*

$$\mathbb{P}\left(|\langle \boldsymbol{x},\boldsymbol{y}\rangle| \ge \frac{t}{\sqrt{d}}\right) \le 2\exp\left(-\frac{t^2}{4}\right).$$

Proof. By rotation invariance on the d-dimensional sphere, we can fix $\boldsymbol{y} = \boldsymbol{e}_1 := (1,0,\dots,0)$, the first vector of the standard basis. Let $\boldsymbol{z} \in \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)$, then $\boldsymbol{z}/\|\boldsymbol{z}\|$ is a uniform random point in $\mathbb{S}^{d-1}$. Therefore, we have that

$$\mathbb{P}\left(|\langle \boldsymbol{x},\boldsymbol{y}\rangle| \ge \frac{t}{\sqrt{d}}\right) = \mathbb{P}\left(\frac{|z_1|}{\|\boldsymbol{z}\|} \ge \frac{t}{\sqrt{d}}\right) = \mathbb{P}\left(\frac{z_1^2}{\sum_{i=1}^d z_i^2} \ge \frac{t^2}{d}\right).$$

Since z_i's are i.i.d. standard normal random variables, $z_1^2/\sum_{i=1}^d z_i^2$ has a $\mathrm{Beta}(\frac{1}{2}, \frac{d-1}{2})$ distribution. Hence, by the density function of a beta distribution, we have that

$$\mathbb{P}\left(\frac{z_1^2}{\sum_{i=1}^d z_i^2} \ge \frac{t^2}{d}\right) = \frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{d-1}{2})} \int_{\frac{t^2}{d}}^1 z^{-1/2}(1-z)^{(d-1)/2-1}\, dz.$$

By Wendel's double inequality (see [29, equation (7)], also [16, equation (3.15)]), we have that

$$\frac{\Gamma(\frac{d}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{d-1}{2})} \le \sqrt{\frac{d-1}{2\pi}}.$$

Additionally,

$$\begin{aligned}\int_{\frac{t^2}{d}}^1 z^{-1/2}(1-z)^{(d-1)/2-1}\, dz &\le \frac{\sqrt{d}}{t}\int_{\frac{t^2}{d}}^1 (1-z)^{(d-1)/2-1}\, dz = \frac{\sqrt{d}}{t}\left(-\frac{2}{d-1}(1-z)^{(d-1)/2}\right)\Bigg|_{\frac{t^2}{d}}^1 \\ &= \frac{2\sqrt{d}}{t(d-1)}\left(1-\frac{t^2}{d}\right)^{(d-1)/2} \le \frac{2\sqrt{d}}{t(d-1)}\exp\left(-\frac{(d-1)t^2}{2d}\right).\end{aligned}$$

Putting the previous displays together, we have for $t \ge 1$ and $d \ge 2$ that

$$\mathbb{P}\left(|\langle \boldsymbol{x},\boldsymbol{y}\rangle| \ge \frac{t}{\sqrt{d}}\right) \le \frac{2}{\sqrt{\pi}}\exp\left(-\frac{t^2}{4}\right).$$

The claim directly follows. □

Remark 2.16. The proof of Lemma 2.15 makes use of the explicit density function of a Beta distribution. The tails of Beta random variables have also been studied in recent years [19,31], where sub-Gaussian and Bernstein-type bounds are given respectively in different parameter regimes.

Remark 2.17. We have seen in Proposition 2.14 that the inner product of two d-dimensional standard normal random vectors has sub-exponential tails, and since the moment generating function does not exist when $t > 1$, the exponential rate cannot be improved. In comparison, Lemma 2.15 gives sub-Gaussian tails when the random vectors are uniformly distributed on a sphere of the same dimension, which decays much faster. In other words, the inner product of independent high dimensional random vectors concentrates better on a sphere than in a Gaussian space after proper normalization.

3. Impossibility of detecting geometry

In this section, we show that $\mathcal{G}(n,p,d,r)$ and $\mathcal{G}(n,p)$ are indistinguishable when the dimension d or the parameter r is large, thus proving Theorem 1.1(a).

We first introduce several notations used in the proofs. We denote the adjacency matrix of $\mathcal{G}(n,p,d,r)$ by $\boldsymbol{A}$. Let $\boldsymbol{B} \in \mathbb{R}^{n\times n}$ be a symmetric Bernoulli ensemble, that is, $\{b_{i,j}\}_{1\le i<j\le n}$ are independent Bernoulli random variables with parameter p. We also use the following shorthand notations in this section. For the matrices $\boldsymbol{A},\boldsymbol{B} \in \mathbb{R}^{n\times n}$, we denote their principal minor of order k by $\boldsymbol{A}_k$ and $\boldsymbol{B}_k$. The bold lower case letter $\boldsymbol{a}_k$ denotes the last row of $\boldsymbol{A}_k$. The k by d matrix consisting of the first k rows of the matrix $\boldsymbol{X} \in \mathbb{R}^{n\times d}$ is denoted by $\boldsymbol{X}_k$, and $\boldsymbol{x}_k$ denotes the kth row of $\boldsymbol{X}$.

Pinsker's inequality (Proposition 2.3) gives

$$\mathrm{TV}(\mathcal{G}(n,p,d,r),\mathcal{G}(n,p)) \le \sqrt{\frac{1}{2}\mathrm{KL}(\mathcal{G}(n,p,d,r)\parallel \mathcal{G}(n,p))}.$$

By the chain rule of KL divergence (Proposition 2.5) we have that

$$\begin{aligned}\mathrm{KL}(\mathcal{G}(n,p,d,r)\parallel \mathcal{G}(n,p)) = \mathrm{KL}(\boldsymbol{A}\parallel \boldsymbol{B}) &= \sum_{k=0}^{n-1}\mathbb{E}_{\boldsymbol{A}_k}\mathrm{KL}(\boldsymbol{a}_{k+1}\mid \boldsymbol{A}_k \parallel \boldsymbol{b}_{k+1}\mid \boldsymbol{B}_k=\boldsymbol{A}_k)\\ &= \sum_{k=0}^{n-1}\mathbb{E}_{\boldsymbol{A}_k}\mathrm{KL}(\boldsymbol{a}_{k+1}\mid \boldsymbol{A}_k \parallel \boldsymbol{b}_{k+1}),\end{aligned}\tag{2}$$

where the last equality is due to the independence of $\boldsymbol{b}_{k+1}$ and $\boldsymbol{B}_k$. By convexity of KL divergence (see, e.g., [16, Proposition 3.4]), applying Jensen's inequality gives that

$$\begin{aligned}\mathbb{E}_{\boldsymbol{A}_k}\mathrm{KL}(\boldsymbol{a}_{k+1}\mid \boldsymbol{A}_k \parallel \boldsymbol{b}_{k+1}) &\le \mathbb{E}_{\boldsymbol{A}_k,\boldsymbol{X}_k}\mathrm{KL}(\boldsymbol{a}_{k+1}\mid \boldsymbol{A}_k,\boldsymbol{X}_k \parallel \boldsymbol{b}_{k+1})\\ &= \mathbb{E}_{\boldsymbol{X}_k}\mathrm{KL}(\boldsymbol{a}_{k+1}\mid \boldsymbol{X}_k \parallel \boldsymbol{b}_{k+1}),\end{aligned}$$

where the last equality is because $\boldsymbol{a}_{k+1}$ and $\boldsymbol{A}_k$ are conditionally independent given $\boldsymbol{X}_k$. The KL divergence is bounded from above by the χ^2 divergence (see Proposition 2.4):

$$\begin{aligned}\mathbb{E}_{\boldsymbol{X}_k}\mathrm{KL}(\boldsymbol{a}_{k+1}\mid \boldsymbol{X}_k \parallel \boldsymbol{b}_{k+1}) &\le \mathbb{E}_{\boldsymbol{X}_k}[\log(1+\chi^2(\boldsymbol{a}_{k+1}\mid \boldsymbol{X}_k \parallel \boldsymbol{b}_{k+1}))]\\ &\le \log(1+\mathbb{E}_{\boldsymbol{X}_k}\chi^2(\boldsymbol{a}_{k+1}\mid \boldsymbol{X}_k \parallel \boldsymbol{b}_{k+1}))\\ &= \log \mathbb{E}_{\boldsymbol{X}_k,\boldsymbol{b}_{k+1}}\left[\left(\frac{\mathbb{P}_{\mathcal{G}(n,p,d,r)}(\boldsymbol{b}_{k+1}\mid \boldsymbol{X}_k)}{\mathbb{P}_{\mathcal{G}(n,p)}(\boldsymbol{b}_{k+1})}\right)^2\right],\end{aligned}\tag{3}$$

where the second line is by Jensen's inequality.

By the definition of $\mathcal{G}(n,p,d,r)$, we have that

$$\mathbb{P}_{\mathcal{G}(n,p,d,r)}(\boldsymbol{b}_{k+1} \mid \boldsymbol{X}_k) = \mathbb{E}_{\boldsymbol{x}_{k+1}}\left[\prod_{1\le i\le k} \sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle)^{b_{i,k+1}}(1-\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle))^{1-b_{i,k+1}}\right].$$

Following an idea similar to the second moment method [4], we can write the square as the product of two expectations of independent copies and then apply Fubini's theorem:

$$\begin{aligned}\mathbb{P}_{\mathcal{G}(n,p,d,r)}(\boldsymbol{b}_{k+1} \mid \boldsymbol{X}_k)^2 &= \mathbb{E}_{\boldsymbol{x}_{k+1}}\left[\prod_{1\le i\le k} \sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle)^{b_{i,k+1}}(1-\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle))^{1-b_{i,k+1}}\right] \\ &\quad \times \mathbb{E}_{\boldsymbol{x}'_{k+1}}\left[\prod_{1\le i\le k} \sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}'_{k+1}\rangle)^{b_{i,k+1}}(1-\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}'_{k+1}\rangle))^{1-b_{i,k+1}}\right] \\ &= \mathbb{E}_{\boldsymbol{x}_{k+1},\boldsymbol{x}'_{k+1}}\left[\prod_{1\le i\le k}\Big(\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle)\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}'_{k+1}\rangle)\Big)^{b_{i,k+1}}\right. \\ &\quad \left.\times\Big((1-\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle))(1-\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}'_{k+1}\rangle))\Big)^{1-b_{i,k+1}}\right],\end{aligned}$$

where the last equality is by the independence of $\boldsymbol{x}_{k+1}$ and $\boldsymbol{x}'_{k+1}$. Therefore, by interchanging the expectations and taking out the product by independence, we obtain that

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{X}_k,\boldsymbol{b}_{k+1}}\left[\left(\frac{\mathbb{P}_{\mathcal{G}(n,p,d,r)}(\boldsymbol{b}_{k+1} \mid \boldsymbol{X}_k)}{\mathbb{P}_{\mathcal{G}(n,p)}(\boldsymbol{b}_{k+1})}\right)^2\right] \\ &\quad = \mathbb{E}_{\boldsymbol{x}_{k+1},\boldsymbol{x}'_{k+1}}\left[\prod_{1\le i\le k}\mathbb{E}_{\boldsymbol{x}_i,b_{i,k+1}}\left[\left(\frac{1}{p^2}\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle)\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}'_{k+1}\rangle)\right)^{b_{i,k+1}}\right.\right. \\ &\quad\quad \left.\left.\times\left(\frac{1}{(1-p)^2}(1-\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle))(1-\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}'_{k+1}\rangle))\right)^{1-b_{i,k+1}}\right]\right].\end{aligned}$$

Since each entry of $\boldsymbol{B}$ is an independent Bernoulli random variable with parameter p, we can compute the inner expectation over $b_{i,k+1}$ directly, obtaining that

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{X}_k,\boldsymbol{b}_{k+1}}\left[\left(\frac{\mathbb{P}_{\mathcal{G}(n,p,d,r)}(\boldsymbol{b}_{k+1} \mid \boldsymbol{X}_k)}{\mathbb{P}_{\mathcal{G}(n,p)}(\boldsymbol{b}_{k+1})}\right)^2\right] \\ &\quad = \mathbb{E}_{\boldsymbol{x}_{k+1},\boldsymbol{x}'_{k+1}}\left[\prod_{1\le i\le k}\left(1+\frac{1}{p(1-p)}\mathbb{E}_{\boldsymbol{x}_i}[(\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}_{k+1}\rangle)-p)(\sigma(\langle \boldsymbol{x}_i, \boldsymbol{x}'_{k+1}\rangle)-p)]\right)\right] \qquad (4) \\ &\quad = \mathbb{E}_{\boldsymbol{x}_{k+1},\boldsymbol{x}'_{k+1}}\left[\left(1+\frac{1}{p(1-p)}\mathbb{E}_{\boldsymbol{x}_1}[(\sigma(\langle \boldsymbol{x}_1, \boldsymbol{x}_{k+1}\rangle)-p)(\sigma(\langle \boldsymbol{x}_1, \boldsymbol{x}'_{k+1}\rangle)-p)]\right)^k\right],\end{aligned}$$

where the last equality holds since all $\boldsymbol{x}_i$'s are identically distributed.

Define

$$\gamma(\boldsymbol{x},\boldsymbol{y}) := \mathbb{E}_{\boldsymbol{z}\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{I}_d)}[(\sigma(\langle \boldsymbol{x},\boldsymbol{z}\rangle)-p)(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle)-p)]. \qquad (5)$$

We show the following lemma concerning $\gamma(\boldsymbol{x},\boldsymbol{y})$.

Lemma 3.1. *Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ be independent standard normal random vectors. Recall the definition of σ, as well as the assumption (A1).*

(a) The mean of $\gamma(\boldsymbol{x}, \boldsymbol{y})$ satisfies

$$0 \le \mathbb{E}[\gamma(\boldsymbol{x}, \boldsymbol{y})] \le \frac{\alpha^2}{r^4 d}.$$

(b) The variance of $\gamma(\boldsymbol{x}, \boldsymbol{y})$ is upper bounded by

$$\mathbb{V}\mathrm{ar}[\gamma(\boldsymbol{x}, \boldsymbol{y})] \le \frac{68\alpha^2}{r^4 d}.$$

(c) Let $L := \sqrt{34}\alpha/(r^2 d)$. Then, $\gamma(\boldsymbol{x}, \boldsymbol{y})/L$ is sub-exponential with parameters $(2\sqrt{2d}, 1)$, that is,

$$\log \mathbb{E}\left[\exp\left(\frac{t}{L}(\gamma(\boldsymbol{x}, \boldsymbol{y}) - \mathbb{E}[\gamma(\boldsymbol{x}, \boldsymbol{y})])\right)\right] \le 4dt^2 \quad \textit{for all } |t| \le 1.$$

Proof of Lemma 3.1(a). Let

$$\eta(\boldsymbol{x}) := \mathbb{E}_{\boldsymbol{z}\sim\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)}[\sigma(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)] \qquad \text{and} \qquad \xi(\boldsymbol{x}, \boldsymbol{y}) := \mathbb{E}_{\boldsymbol{z}\sim\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)}[\sigma(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)\sigma(\langle \boldsymbol{y}, \boldsymbol{z}\rangle)].$$

By the construction of σ, we have that

$$\mathbb{E}_{\boldsymbol{x}\sim\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)}[\eta(\boldsymbol{x})] = \mathbb{E}[\sigma(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)] = p. \tag{6}$$

Next, we bound the variance of $\eta(\boldsymbol{x})$. Observe that

$$\frac{\partial \eta(\boldsymbol{x})}{\partial x_i} = \mathbb{E}_{\boldsymbol{z}\sim\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)}\left[\frac{\partial \sigma(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)}{\partial x_i}\right] = \mathbb{E}_{\boldsymbol{z}\sim\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)}[z_i \sigma'(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)].$$

Thus, by Stein's lemma we have that

$$\frac{\partial \eta(\boldsymbol{x})}{\partial x_i} = \mathbb{E}_{\boldsymbol{z}_{-i}}[\mathbb{E}_{z_i}[z_i \sigma'(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)]] = \mathbb{E}_{\boldsymbol{z}_{-i}}\left[\mathbb{E}_{z_i}\left[\frac{\partial \sigma'(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)}{\partial z_i}\right]\right] = x_i\, \mathbb{E}_{\boldsymbol{z}}[\sigma''(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)],$$

where $\boldsymbol{z}_{-i}$ denotes the rest of $\boldsymbol{z}$ except for z_i. Hence, by the definition of σ in (1),

$$\begin{aligned}
\|\nabla \eta(\boldsymbol{x})\|^2 &= \sum_{i=1}^d \left(\frac{\partial \eta(\boldsymbol{x})}{\partial x_i}\right)^2 = \frac{1}{r^4 d^2}\sum_{i=1}^d x_i^2\, \mathbb{E}_{\boldsymbol{z}}\left[f'\left(\frac{\langle \boldsymbol{x}, \boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)\right]^2 \\
&= \frac{\|\boldsymbol{x}\|^2}{r^4 d^2}\mathbb{E}_{\boldsymbol{z}}\left[f'\left(\frac{\langle \boldsymbol{x}, \boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)\right]^2 \le \frac{\|\boldsymbol{x}\|^2}{r^4 d^2}\mathbb{E}_{\boldsymbol{z}}\left[f'\left(\frac{\langle \boldsymbol{x}, \boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)^2\right] \le \frac{\alpha^2\|\boldsymbol{x}\|^2}{r^4 d^2},
\end{aligned}$$

where we used the assumption (A1) in the last inequality. The Gaussian Poincaré inequality (Proposition 2.9) thus gives

$$\mathbb{V}\mathrm{ar}[\eta(\boldsymbol{x})] \le \mathbb{E}[\|\nabla \eta(\boldsymbol{x})\|^2] \le \frac{\alpha^2}{r^4 d^2}\mathbb{E}[\|\boldsymbol{x}\|^2] = \frac{\alpha^2}{r^4 d}. \tag{7}$$

Additionally, by interchanging the order of expectations we have that

$$\mathbb{E}[\xi(\boldsymbol{x}, \boldsymbol{y})] = \mathbb{E}[\mathbb{E}_{\boldsymbol{z}}[\sigma(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)\sigma(\langle \boldsymbol{y}, \boldsymbol{z}\rangle)]] = \mathbb{E}[\mathbb{E}_{\boldsymbol{x}}[\sigma(\langle \boldsymbol{x}, \boldsymbol{z}\rangle)]\,\mathbb{E}_{\boldsymbol{y}}[\sigma(\langle \boldsymbol{y}, \boldsymbol{z}\rangle)]] = \mathbb{E}[\eta(\boldsymbol{z})^2]. \tag{8}$$

Finally, expanding the product in the definition of $\gamma(\boldsymbol{x},\boldsymbol{y})$ and putting together the expressions in (6), (7), and (8), we obtain that

$$\mathbb{E}[\gamma(\boldsymbol{x},\boldsymbol{y})] = \mathbb{E}[\xi(\boldsymbol{x},\boldsymbol{y})] - p\,\mathbb{E}[\eta(\boldsymbol{x})] - p\,\mathbb{E}[\eta(\boldsymbol{y})] + p^2 = \mathbb{E}[\xi(\boldsymbol{x},\boldsymbol{y})] - p^2 = \mathbb{V}\mathrm{ar}[\eta(\boldsymbol{z})] \le \frac{\alpha^2}{r^4 d}. \tag{9}$$

The nonnegativity of $\mathbb{E}[\gamma(\boldsymbol{x},\boldsymbol{y})]$ directly follows from it being a variance. □

Recall the assumption (A1), and note that Corollary 2.7 thus implies that f is bounded and

$$\sup_x |f(x)| \le 2\sqrt{\alpha}. \tag{10}$$

Proof of Lemma 3.1(b) and Lemma 3.1(c). Taking the partial derivative with respect to x_i, we have that

$$\begin{aligned}\frac{\partial \gamma(\boldsymbol{x},\boldsymbol{y})}{\partial x_i} &= \mathbb{E}_{\boldsymbol{z}\sim\mathcal{N}(0,\boldsymbol{I}_d)}\left[\frac{\partial(\sigma(\langle \boldsymbol{x},\boldsymbol{z}\rangle) - p)}{\partial x_i}(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p)\right] \\ &= \mathbb{E}_{\boldsymbol{z}\sim\mathcal{N}(0,\boldsymbol{I}_d)}[z_i \sigma'(\langle \boldsymbol{x},\boldsymbol{z}\rangle)(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p)].\end{aligned}$$

By Stein's lemma, we have that

$$\begin{aligned}\frac{\partial \gamma(\boldsymbol{x},\boldsymbol{y})}{\partial x_i} &= \mathbb{E}_{\boldsymbol{z}_{-i}}[\mathbb{E}_{z_i}[z_i \sigma'(\langle \boldsymbol{x},\boldsymbol{z}\rangle)(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p)]] = \mathbb{E}_{\boldsymbol{z}_{-i}}\left[\mathbb{E}_{z_i}\left[\frac{\partial}{\partial z_i}(\sigma'(\langle \boldsymbol{x},\boldsymbol{z}\rangle)(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p))\right]\right] \\ &= y_i\,\mathbb{E}_{\boldsymbol{z}}[\sigma'(\langle \boldsymbol{x},\boldsymbol{z}\rangle)\sigma'(\langle \boldsymbol{y},\boldsymbol{z}\rangle)] + x_i\,\mathbb{E}_{\boldsymbol{z}}[\sigma''(\langle \boldsymbol{x},\boldsymbol{z}\rangle)(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p)].\end{aligned}$$

Then, by the elementary inequality $(a+b)^2 \le 2(a^2+b^2)$ and Jensen's inequality, we have that

$$\begin{aligned}\left(\frac{\partial \gamma(\boldsymbol{x},\boldsymbol{y})}{\partial x_i}\right)^2 &\le \frac{2y_i^2}{r^4 d^2}\mathbb{E}_{\boldsymbol{z}}\left[f\left(\frac{\langle \boldsymbol{x},\boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right) f\left(\frac{\langle \boldsymbol{y},\boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)\right]^2 \\ &\quad + \frac{2x_i^2}{r^4 d^2}\mathbb{E}_{\boldsymbol{z}}\left[f'\left(\frac{\langle \boldsymbol{x},\boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p)\right]^2 \\ &\le \frac{2y_i^2}{r^4 d^2}\mathbb{E}_{\boldsymbol{z}}\left[f\left(\frac{\langle \boldsymbol{x},\boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)^2 f\left(\frac{\langle \boldsymbol{y},\boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)^2\right] \\ &\quad + \frac{2x_i^2}{r^4 d^2}\mathbb{E}_{\boldsymbol{z}}\left[f'\left(\frac{\langle \boldsymbol{x},\boldsymbol{z}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)^2(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p)^2\right] \\ &\le \frac{2\alpha^2(x_i^2 + 16y_i^2)}{r^4 d^2},\end{aligned}$$

where in the last inequality we used (A1) and (10). Hence, we obtain that

$$\|\nabla\gamma(\boldsymbol{x},\boldsymbol{y})\|^2 = \sum_{i=1}^d \left(\frac{\partial \gamma(\boldsymbol{x},\boldsymbol{y})}{\partial x_i}\right)^2 + \sum_{i=1}^d \left(\frac{\partial \gamma(\boldsymbol{x},\boldsymbol{y})}{\partial y_i}\right)^2 \le \frac{34\alpha^2}{r^4 d^2}(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2). \tag{11}$$

Taking expectation on both sides of the above display yields

$$\mathbb{E}[\|\nabla\gamma(\boldsymbol{x},\boldsymbol{y})\|^2] \le \frac{34\alpha^2}{r^4 d^2}\mathbb{E}[\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2] = \frac{68\alpha^2}{r^4 d}.$$

By the Gaussian Poincaré inequality we thus have that

$$\mathbb{V}\mathrm{ar}[\gamma(\boldsymbol{y},\boldsymbol{z})] \le \mathbb{E}[\|\nabla\gamma(\boldsymbol{y},\boldsymbol{z})\|^2] \le \frac{68\alpha^2}{r^4 d}.$$

Lemma 3.1(b) is hence proved.

By viewing $(\boldsymbol{x},\boldsymbol{y})$ as a $(2d)$-dimensional vector, (11) exactly gives the upper bound of the norm of the gradient in terms of the norm of the vector. Thus, by applying Lemma 2.12, the sub-exponential tails of $\gamma(\boldsymbol{x},\boldsymbol{y})$ in Lemma 3.1(c) directly follow. □

With Lemma 3.1 in place, we return to bounding the KL divergence from above. Using the definition of $\gamma(\boldsymbol{x},\boldsymbol{y})$, we can express (4) with $\gamma(\boldsymbol{x},\boldsymbol{y})$ as

$$\begin{aligned}\mathbb{E}_{\boldsymbol{X}_k,\boldsymbol{b}_{k+1}}\left[\left(\frac{\mathbb{P}_{\mathcal{G}(n,p,d,r)}(\boldsymbol{b}_{k+1}\mid \boldsymbol{X}_k)}{\mathbb{P}_{\mathcal{G}(n,p)}(\boldsymbol{b}_{k+1})}\right)^2\right] &= \mathbb{E}\left[\left(1+\frac{1}{p(1-p)}\gamma(\boldsymbol{x},\boldsymbol{y})\right)^k\right]\\ &\le \mathbb{E}\left[\exp\left(\frac{k}{p(1-p)}\gamma(\boldsymbol{x},\boldsymbol{y})\right)\right],\end{aligned}$$

where we use the fact that $1+x\le\exp(x)$. Using Lemma 3.1(c) with $t = Lk/(p(1-p))$, we have, for $r^2 d/k \ge \sqrt{34}\alpha/(p(1-p))$, that

$$\mathbb{E}\left[\exp\left(\frac{k}{p(1-p)}(\gamma(\boldsymbol{x},\boldsymbol{y})-\mathbb{E}[\gamma(\boldsymbol{x},\boldsymbol{y})])\right)\right] \le \exp\left(\frac{136\alpha^2}{p^2(1-p)^2}\cdot\frac{k^2}{r^4 d}\right).$$

Recall that we assume that $n^3/(r^4 d)\to 0$, which implies that $r^2 d/n\to\infty$, so the inequality in the display above holds eventually (uniformly for all $k\le n$). Combined with Lemma 3.1(a), we thus have that

$$\begin{aligned}\mathbb{E}\left[\exp\left(\frac{k}{p(1-p)}\gamma(\boldsymbol{x},\boldsymbol{y})\right)\right] &= \exp\left(\frac{k}{p(1-p)}\mathbb{E}[\gamma(\boldsymbol{x},\boldsymbol{y})]\right)\mathbb{E}\left[\exp\left(\frac{k}{p(1-p)}(\gamma(\boldsymbol{x},\boldsymbol{y})-\mathbb{E}[\gamma(\boldsymbol{x},\boldsymbol{y})])\right)\right]\\ &\le \exp\left(\frac{\alpha^2}{p(1-p)}\cdot\frac{k}{r^4 d}+\frac{136\alpha^2}{p^2(1-p)^2}\cdot\frac{k^2}{r^4 d}\right).\end{aligned}$$

Putting the previous displays together and inserting the upper bound into (3), we obtain that

$$\mathbb{E}_{\boldsymbol{X}_k}\mathrm{KL}(\boldsymbol{a}_{k+1}\mid \boldsymbol{X}_k\,\|\,\boldsymbol{b}_{k+1}) \le \frac{\alpha^2}{p(1-p)}\cdot\frac{k}{r^4 d}+\frac{136\alpha^2}{p^2(1-p)^2}\cdot\frac{k^2}{r^4 d}.$$

Plugging the above display into (2), we conclude that

$$\begin{aligned}\mathrm{KL}(\mathcal{G}(n,p,d,r)\,\|\,\mathcal{G}(n,p)) = \mathrm{KL}(\boldsymbol{A}\,\|\,\boldsymbol{B}) &\le \sum_{k=0}^{n-1}\left(\frac{\alpha^2}{p(1-p)}\cdot\frac{k}{r^4 d}+\frac{136\alpha^2}{p^2(1-p)^2}\cdot\frac{k^2}{r^4 d}\right)\\ &\le \frac{\alpha^2}{2p(1-p)}\cdot\frac{n^2}{r^4 d}+\frac{68\alpha^2}{3p^2(1-p)^2}\cdot\frac{n^3}{r^4 d}.\end{aligned}$$

The asymptotic in Theorem 1.1(a) directly follows.

4. Detecting geometry using the signed triangle statistic

In this section, we show that the geometric structure in $\mathcal{G}(n,p,d,r)$ can be detected in certain parameter regimes of d and r using the signed triangle statistic proposed in [5], thus proving Theorem 1.1(b).

Let $\boldsymbol{A}$ be the adjacency matrix of a sample graph G with edge probability p. The signed triangle over vertices $\{i,j,k\}$ is defined as

$$\tau_{\{i,j,k\}}(G) := (a_{i,j}-p)(a_{j,k}-p)(a_{k,i}-p).$$

Further, the signed triangle statistic is the sum of all possible signed triangles in the graph:

$$\tau(G) := \sum_{\{i,j,k\}\in\binom{[n]}{3}} \tau_{\{i,j,k\}}(G).$$

In $\mathcal{G}(n,p)$, due to the independence of edges, calculation of the mean and the variance of the signed triangle statistic is straightforward, which has been done in [5]. We state the results here as the following lemma.

Lemma 4.1. *The signed triangle statistic in $\mathcal{G}(n,p)$ satisfies*

$$\mathbb{E}[\tau(\mathcal{G}(n,p))] = 0 \qquad \textit{and} \qquad \mathbb{V}\mathrm{ar}[\tau(\mathcal{G}(n,p))] \le n^3.$$

The main goal of this section is to estimate the mean and the variance of the signed triangle statistic in $\mathcal{G}(n,p,d,r)$. The analysis of the mean is quite delicate due to the generality of the connection function. We have two parameters d and r that affect the detectability of geometry acting in different ways. Informally, when the variance parameter r is large, the connection function becomes "flat" and the edges in the graph are less dependent on the latent space. When the dimension d is large, the inner products of the latent vectors become more independent hence the edges. The challenge here is to implement the intuition in the two regimes while retaining the correct dependency of the other parameter. The estimation is divided into two parts accordingly.

4.1. Estimating the mean in the large variance regime

We start our discussion with the estimate for the mean in the regime when r is large, specifically $r/\log^2 r \gg d^{1/6}$. Before diving into the details, we first present a concentration result which provides a tail bound for the remainder of a linear approximation of σ.

Lemma 4.2. *Suppose $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d$ are independent standard normal random vectors. Denote $\eta := \mathbb{E}[\sigma'(\langle \boldsymbol{x},\boldsymbol{y}\rangle)]$ and define the remainder of a linear approximation of σ as*

$$g(x) := \sigma(x) - p - \eta x.$$

Then, the tails of $g(\langle \boldsymbol{x},\boldsymbol{y}\rangle)$ satisfy that for $t \ge 6$,

$$\mathbb{P}\left(|g(\langle \boldsymbol{x},\boldsymbol{y}\rangle)| \ge \frac{3\alpha t}{2r^2}\right) \le \exp\left(-\sqrt{\frac{t}{2e}}\right).$$

Proof. Let $\boldsymbol{x}', \boldsymbol{y}' \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_d)$ be independent copies of $\boldsymbol{x}, \boldsymbol{y}$. For ease of notation, denote $Z := \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ and $Z' := \langle \boldsymbol{x}', \boldsymbol{y}' \rangle$. Then, by definition,

$$p = \mathbb{E}[\sigma(Z')] \quad \text{and} \quad \eta = \mathbb{E}[\sigma'(Z')].$$

Plugging them into $g(x)$ and by the triangle inequality, we have that

$$\begin{aligned}
|g(Z)| &= |\sigma(Z) - \mathbb{E}[\sigma(Z')] - \mathbb{E}[\sigma'(Z')]Z| = |\mathbb{E}_{Z'}[\sigma(Z) - \sigma(Z') - \sigma'(Z')(Z - Z')] - \mathbb{E}[\sigma'(Z')Z']| \\
&\le \underbrace{|\mathbb{E}_{Z'}[\sigma(Z) - \sigma(Z') - \sigma'(Z')(Z - Z')]|}_{V_1} + \underbrace{|\mathbb{E}[\sigma'(Z')Z']|}_{V_2}.
\end{aligned}$$

We first bound V_1 and V_2 from above and then utilize Gaussian hypercontractivity.

By Taylor's theorem,

$$\sigma(Z) - \sigma(Z') - \sigma'(Z')(Z - Z') = \frac{\sigma''(\xi)}{2}(Z - Z')^2$$

for some ξ between Z' and Z. Then, we have that

$$\begin{aligned}
V_1 &\le \mathbb{E}_{Z'}[|\sigma(Z) - \sigma(Z') - \sigma'(Z')(Z - Z')|] = \mathbb{E}_{Z'}\left[\frac{|\sigma''(\xi)|}{2}(Z - Z')^2\right] = \mathbb{E}_{Z'}\left[\frac{|f'(\xi)|}{2r^2 d}(Z - Z')^2\right] \\
&\le \frac{\alpha}{2r^2 d}\mathbb{E}_{Z'}[(Z - Z')^2],
\end{aligned}$$

where we used the assumption (A1) in the last inequality.

Since $Z' = \langle \boldsymbol{x}', \boldsymbol{y}' \rangle = \sum_{i=1}^d x_i' y_i'$, we have that $\mathbb{E}[Z'] = 0$ and $\mathbb{E}[Z'^2] = d$. Therefore,

$$\mathbb{E}_{Z'}[(Z - Z')^2] = \mathbb{E}_{Z'}[Z^2 - 2ZZ' + Z'^2] = Z^2 + d.$$

Hence, we obtain that

$$V_1 \le \frac{\alpha}{2r^2 d}(Z^2 + d) = \frac{\alpha}{2r^2 d}Z^2 + \frac{\alpha}{2r^2}.$$

Turning to V_2, by the triangle inequality we have that

$$V_2 = |\mathbb{E}[\sigma'(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)\langle \boldsymbol{x}, \boldsymbol{y} \rangle]| = \left|\sum_{i=1}^d \mathbb{E}[\sigma'(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)x_i y_i]\right| \le \sum_{i=1}^d |\mathbb{E}[\sigma'(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)x_i y_i]|,$$

Thus, by Stein's lemma and Jensen's inequality, we have that

$$V_2 \le \sum_{i=1}^d \left|\mathbb{E}\left[y_i \frac{\partial \sigma'(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)}{\partial x_i}\right]\right| = \sum_{i=1}^d |\mathbb{E}[y_i^2 \sigma''(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)]| \le \sum_{i=1}^d \mathbb{E}[y_i^2 |\sigma''(\langle \boldsymbol{x}, \boldsymbol{y} \rangle)|] \le \frac{\alpha}{r^2 d}\mathbb{E}[\|\boldsymbol{y}\|^2] = \frac{\alpha}{r^2}.$$

Therefore, by the triangle inequality, we have for any $q \ge 1$ that

$$\|g(Z)\|_q \le \|V_1\|_q + \|V_2\|_q \le \frac{\alpha}{2r^2 d}\|Z^2\|_q + \frac{3\alpha}{2r^2}.$$

Since $Z^2 = (\sum_{i=1}^d x_i y_i)^2$ is a fourth order polynomial of independent standard normal random variables, by Gaussian hypercontractivity (see [2, Corollary 5.21] for the univariate case and [14, Theorem 6.7] for a general argument), we have that

$$\|Z^2\|_q \le (q-1)^2 \|Z^2\|_2.$$

Since

$$\mathbb{E}[Z^4] = \mathbb{E}\left[\left(\sum_{i=1}^d x_i y_i\right)^4\right] = 3\sum_{i\neq j} \mathbb{E}[x_i^2]\mathbb{E}[x_j^2]\mathbb{E}[y_i^2]\mathbb{E}[y_j^2] + \sum_i \mathbb{E}[x_i^4]\mathbb{E}[y_i^4] = 3d^2 + 6d \le 9d^2,$$

we obtain that $\|Z^2\|_2 \le 3d$ and so

$$\|g(Z)\|_q \le \frac{3\alpha}{2r^2}(q^2 - 2q + 2).$$

By Markov's inequality,

$$\begin{aligned}\mathbb{P}\left(|g(Z)| \ge \frac{3\alpha}{2r^2}t\right) &\le \left(\frac{3\alpha}{2r^2}t\right)^{-q} \mathbb{E}[|g(Z)|^q] \\ &\le t^{-q}(q^2 - 2q + 2)^q = \exp(-q\log t + q\log(q^2 - 2q + 2)).\end{aligned}$$

For $t \ge 3$, by choosing $q = \sqrt{t/e - 1} + 1$, we have that

$$\mathbb{P}\left(|g(Z)| \ge \frac{3\alpha t}{2r^2}\right) \le \exp\left(-\left(\sqrt{\frac{t}{e} - 1} + 1\right)\right).$$

Hence, for $t \ge 6$,

$$\mathbb{P}\left(|g(Z)| \ge \frac{3\alpha t}{2r^2}\right) \le \exp\left(-\sqrt{\frac{t}{2e}}\right). \qquad \square$$

Let

$$\lambda := \mathbb{E}\left[f\left(\frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle - \mu_{p,d,r}}{r\sqrt{d}}\right)\right]. \tag{12}$$

Then, by the construction of σ,

$$\mathbb{E}[\sigma'(\langle \boldsymbol{x}, \boldsymbol{y}\rangle)] = \frac{\lambda}{r\sqrt{d}}.$$

We also have

$$0 \le \lambda \le \sup_x |f(x)| \le 2\sqrt{\alpha}.$$

In the following lemma, we show that λ is bounded away from 0 uniformly for all r and d.

Lemma 4.3. *Let λ be defined in (12). For a fixed $p \in (0,1)$, there exists a constant $C_p > 0$ that does not depend on r and d such that $\lambda \ge C_p$.*

Proof. Let

$$Z := \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle - \mu_{p,d,r}}{r\sqrt{d}}.$$

By Proposition 2.14, we have

$$\mathbb{P}\left(Z - \mathbb{E}[Z] \geq \frac{t}{r\sqrt{d}}\right) \leq \exp\left(-\frac{1}{2}\min\left\{\frac{t^2}{2d}, \frac{t}{\sqrt{2}}\right\}\right),$$

which gives

$$\mathbb{P}(Z - \mathbb{E}[Z] \geq t) \leq \exp\left(-\frac{1}{2}\min\left\{\frac{t^2 r^2}{2}, \frac{tr\sqrt{d}}{\sqrt{2}}\right\}\right).$$

Since $r \geq 1$ and $d \geq 1$, we have for $t \geq 1$,

$$\mathbb{P}(Z - \mathbb{E}[Z] \geq t) \leq \exp\left(-\frac{t}{4}\right). \tag{13}$$

The lower tail directly follows from $Z - \mathbb{E}[Z]$ being symmetric about zero.

Without loss of generality, we may assume $p \in (0, 1/2]$. (If $p \in (1/2, 1)$, we can consider the connection function $1 - F(x)$, which shares the same properties with $F(x)$.) Since F is strictly monotonic when $F \in (0,1)$, the inverse F^{-1} exists in $(0,1)$. Let $z_{p/2} := F^{-1}(p/2)$. We first show by contradiction that $\mathbb{E}[Z] > z_{p/2} - 4\log\frac{4}{p}$. Suppose that it does not hold, that is, $\mathbb{E}[Z] \leq z_{p/2} - 4\log\frac{4}{p}$. Then, by the monotonicity of F,

$$\begin{aligned} p = \mathbb{E}[F(Z)] &= \mathbb{E}[F(Z)\,\mathbb{1}\{Z < z_{p/2}\}] + \mathbb{E}[F(Z)\,\mathbb{1}\{Z \geq z_{p/2}\}] \leq \frac{p}{2} + \mathbb{P}(Z \geq z_{p/2}) \\ &\leq \frac{p}{2} + \mathbb{P}\left(Z \geq \mathbb{E}[Z] + 4\log\frac{4}{p}\right), \end{aligned}$$

where the last inequality holds since $\{Z \geq z_{p/2}\} \subset \{Z \geq \mathbb{E}[Z] + 4\log\frac{4}{p}\}$. Using (13), we have that

$$p \leq \frac{p}{2} + \frac{p}{4} = \frac{3}{4}p,$$

which is a contradiction since $p > 0$. Hence, $\mathbb{E}[Z] > z_{p/2} - 4\log\frac{4}{p}$.

Similarly, we show that $\mathbb{E}[Z] < z_{7p/4} + 4\log 3$, where $z_{7p/4} := F^{-1}(7p/4)$. Again suppose $\mathbb{E}[Z] \geq z_{7p/4} + 4\log 3$. Then, by the monotonicity of F,

$$p = \mathbb{E}[F(Z)] \geq \mathbb{E}[F(Z)\,\mathbb{1}\{Z \geq z_{7p/4}\}] \geq \frac{7p}{4}\mathbb{P}(Z \geq z_{7p/4}) \geq \frac{7p}{4}\mathbb{P}(Z \geq \mathbb{E}[Z] - 4\log 3),$$

where the last inequality holds since $\{Z \geq z_{7p/4}\} \supset \{Z \geq \mathbb{E}[Z] - 4\log 3\}$. Therefore, by the lower tail bound,

$$p \geq \frac{7p}{4}(1 - \mathbb{P}(Z < \mathbb{E}[Z] - 4\log 3)) \geq \frac{7p}{4}\left(1 - \frac{1}{3}\right) = \frac{7}{6}p,$$

which is a contradiction since $p > 0$. Hence, $\mathbb{E}[Z] < z_{7p/4} + 4\log 3$.

Consider the interval $I := [z_{p/2} - 4\log\frac{4}{p} - 4, z_{7p/4} + 4\log 3 + 4]$. The interval is bounded since F is a CDF. Since f is a continuous function and strictly positive in the closed interval by assumption (A0), using the extreme value theorem, we have that $\inf_{x\in I} f(x) \geq f(\xi) = \epsilon_p > 0$ for some $\xi \in I$. Therefore,

$$\begin{aligned}\mathbb{E}[f(Z)] &\geq \mathbb{E}[f(Z)\,\mathbb{1}\{Z \in I\}] \geq \epsilon_p\,\mathbb{P}(Z \in I)\\ &= \epsilon_p\left(1 - \mathbb{P}\left(Z < z_{p/2} - 4\log\frac{4}{p} - 4\right) - \mathbb{P}(Z > z_{7p/4} + 4\log 3 + 4)\right)\\ &\geq \epsilon_p(1 - \mathbb{P}(Z \leq \mathbb{E}[Z] - 4) - \mathbb{P}(Z \geq \mathbb{E}[Z] + 4)) \geq \left(1 - \frac{2}{e}\right)\epsilon_p.\end{aligned}$$

The claim directly follows. □

With the help of Lemma 4.2 and Lemma 4.3, we can show the following lower bound for the signed triangle statistic in $\mathcal{G}(n,p,d,r)$ by expanding the signed triangle and then bounding each term individually. The proof is postponed to the appendix [17] due to limit of space.

Lemma 4.4. *For $r/\log^2 r \gg d^{1/6}$, there exist constants $C_\alpha, C_p > 0$ such that for $r \geq C_\alpha$,*

$$\mathbb{E}[\tau(\mathcal{G}(n,p,d,r))] \geq \frac{C_p n^3}{r^3\sqrt{d}}.$$

4.2. Estimating the mean in the high dimension regime

In this part, we focus on the case when $d/\log^2 d \gg r^6$, which complements the parameter regime discussed in the previous subsection.

We start by bounding the probability of the following two events in $\mathcal{G}(n,p,d,r)$:

$$\Lambda := \{1 \sim 2, 1 \sim 3\} \qquad \text{and} \qquad \Delta := \{1 \sim 2, 2 \sim 3, 3 \sim 1\}.$$

Since

$$\mathbb{P}(\Lambda) = \mathbb{E}[a_{1,2}a_{1,3}] = \mathbb{E}[\sigma(\langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle)\sigma(\langle \boldsymbol{x}_1, \boldsymbol{x}_3\rangle)],$$

we directly have the following lemma as a consequence of (9).

Lemma 4.5. *Let Λ be defined above. Then,*

$$\mathbb{P}(\Lambda) \leq p^2 + \frac{\alpha^2}{r^4 d}.$$

The probability of a triangle in $\mathcal{G}(n,p,d,r)$ is given by

$$\mathbb{P}(\Delta) = \mathbb{E}[a_{1,2}a_{2,3}a_{3,1}] = \mathbb{E}[\sigma(\langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle)\sigma(\langle \boldsymbol{x}_2, \boldsymbol{x}_3\rangle)\sigma(\langle \boldsymbol{x}_3, \boldsymbol{x}_1\rangle)].$$

We have the following estimate for the lower bound of the probability of a triangle.

Lemma 4.6. *Suppose $d/\log^2 d \gg r^6$. There exists a constant $C_\alpha > 0$ such that for $d \geq C_\alpha$,*

$$\mathbb{P}(\Delta) \geq p^3 + \frac{\lambda^3}{4r^3\sqrt{d}}.$$

Due to limit of space, the proof of Lemma 4.6 is provided in the appendix [17].

Lemma 4.7. *Assume $d/\log^2 d \gg r^6$. Then, there exist constants $C_p, C_\alpha \geq 0$, for $d \geq C_\alpha$,*

$$\mathbb{E}[\tau_{\{1,2,3\}}] \geq \frac{C_p}{r^3\sqrt{d}}.$$

Proof. The expected signed triangle can be written as

$$\begin{aligned}\mathbb{E}[\tau_{\{1,2,3\}}] &= \mathbb{E}[(a_{1,2}-p)(a_{2,3}-p)(a_{3,1}-p)] \\ &= \mathbb{E}[a_{1,2}a_{2,3}a_{3,1}] - p(\mathbb{E}[a_{1,2}a_{2,3}] + \mathbb{E}[a_{1,2}a_{3,1}] + \mathbb{E}[a_{2,3}a_{3,1}]) \\ &\quad + p^2(\mathbb{E}[a_{1,2}] + \mathbb{E}[a_{2,3}] + \mathbb{E}[a_{3,1}]) - p^3 \\ &= \mathbb{P}(\Delta) - 3p\,\mathbb{P}(\Lambda) + 2p^2.\end{aligned}$$

By Lemma 4.5 and Lemma 4.6, we have that for $d \geq C_\alpha$,

$$\mathbb{E}[\tau_{\{1,2,3\}}] \geq \frac{\lambda^3}{4r^3\sqrt{d}} - \frac{3p\alpha^2}{r^4 d}.$$

Additionally, with Lemma 4.3, the claim follows directly. □

The expected signed triangle statistic in $\mathcal{G}(n,p,d,r)$ satisfies

$$\mathbb{E}[\tau(\mathcal{G}(n,p,d,r))] := \mathbb{E}\Bigg[\sum_{\{i,j,k\}\in\binom{[n]}{3}} \tau_{\{i,j,k\}}\Bigg] = \sum_{\{i,j,k\}\in\binom{[n]}{3}} \mathbb{E}[\tau_{\{i,j,k\}}] = \binom{n}{3}\mathbb{E}[\tau_{\{1,2,3\}}].$$

Putting them together, we have the following lemma.

Lemma 4.8. *When $d/\log^2 d \gg r^6$, there exist constants $C_p, C_\alpha > 0$ such that for $d \geq C_\alpha$,*

$$\mathbb{E}[\tau(\mathcal{G}(n,p,d,r))] \geq \frac{C_p n^3}{r^3\sqrt{d}}.$$

4.3. Estimating the variance

The variance of the signed triangle statistic in $\mathcal{G}(n,p,d,r)$ can be written as

$$\begin{aligned}\mathbb{V}\mathrm{ar}[\tau(\mathcal{G}(n,p,d,r))] &= \sum_{\{i,j,k\},\{i',j',k'\}\subset[n]} V_{\{i,j,k\},\{i',j',k'\}} \\ &= \binom{n}{3} V_{\{1,2,3\},\{1,2,3\}} + \binom{n}{4}\binom{4}{2} V_{\{1,2,3\},\{1,2,4\}} \\ &\quad + \binom{n}{5}\binom{5}{1}\binom{4}{2} V_{\{1,2,3\},\{1,4,5\}} + \binom{n}{6}\binom{6}{3} V_{\{1,2,3\},\{4,5,6\}},\end{aligned}$$

where $V_{\{i,j,k\},\{i',j',k'\}}$ is the covariance of two signed triangles defined by

$$V_{\{i,j,k\},\{i',j',k'\}} := \mathbb{E}_{\mathcal{G}(n,p,d,r)}[\tau_{\{i,j,k\}}\tau_{\{i',j',k'\}}] - \mathbb{E}_{\mathcal{G}(n,p,d,r)}[\tau_{\{i,j,k\}}]^2.$$

Since two triangles that do not share a vertex are independent,

$$V_{\{1,2,3\},\{4,5,6\}} = \mathbb{E}[\mathbb{E}[\tau_{\{1,2,3\}} \mid \boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3]\mathbb{E}[\tau_{\{4,5,6\}} \mid \boldsymbol{x}_4,\boldsymbol{x}_5,\boldsymbol{x}_6]] - \mathbb{E}[\tau_{\{1,2,3\}}]\mathbb{E}[\tau_{\{4,5,6\}}] = 0.$$

For a signed triangle, $\mathbb{E}_{\mathcal{G}(n,p,d,r)}[\tau^2_{\{1,2,3\}}] \le 1$. Then, we have that

$$V_{\{1,2,3\},\{1,2,3\}} \le \mathbb{E}_{\mathcal{G}(n,p,d,r)}[\tau^2_{\{1,2,3\}}] \le 1.$$

Before proceeding to bounding the other two covariances, we present the following lemma which directly follows from the results of previous parts.

Lemma 4.9. *Let $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{I}_d)$ be independent standard normal random vectors. Then, there exists a constant C_α such that*

$$\mathbb{E}[\mathbb{E}_{\boldsymbol{z}}[(\sigma(\langle \boldsymbol{x},\boldsymbol{z}\rangle) - p)(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p)]^2] \le \frac{C_\alpha}{r^4 d}.$$

Proof. From the definition in (5),

$$\mathbb{E}_{\boldsymbol{z}}[(\sigma(\langle \boldsymbol{x},\boldsymbol{z}\rangle) - p)(\sigma(\langle \boldsymbol{y},\boldsymbol{z}\rangle) - p)] = \gamma(\boldsymbol{x},\boldsymbol{y}).$$

By Lemma 3.1(a) and Lemma 3.1(b), we have that

$$\mathbb{E}[\gamma(\boldsymbol{x},\boldsymbol{y})^2] = \mathbb{V}\mathrm{ar}[\gamma(\boldsymbol{x},\boldsymbol{y})] + \mathbb{E}[\gamma(\boldsymbol{x},\boldsymbol{y})]^2 \le \frac{68\alpha^2}{r^4 d} + \frac{\alpha^4}{r^8 d^2} \le \frac{68\alpha^2 + \alpha^4}{r^4 d},$$

where we used that $r, d \ge 1$. □

Lemma 4.10. *There exists a constant $C_\alpha > 0$ such that*

$$\mathbb{E}[\tau_{\{1,2,3\}}\tau_{\{1,2,4\}}] \le \frac{C_\alpha}{r^4 d}.$$

Proof. For simplicity of notation, denote

$$\bar{\sigma}_{i,j} := \sigma(\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle) - p.$$

Then,

$$\begin{aligned}\mathbb{E}[\tau_{\{1,2,3\}}\tau_{\{1,2,4\}}] &= \mathbb{E}[\bar{\sigma}^2_{1,2}\bar{\sigma}_{1,3}\bar{\sigma}_{2,3}\bar{\sigma}_{1,4}\bar{\sigma}_{2,4}] \le \mathbb{E}[\bar{\sigma}_{1,3}\bar{\sigma}_{2,3}\bar{\sigma}_{1,4}\bar{\sigma}_{2,4}] \\ &= \mathbb{E}[\mathbb{E}_{\boldsymbol{x}_3}[\bar{\sigma}_{1,3}\bar{\sigma}_{2,3}]\mathbb{E}_{\boldsymbol{x}_4}[\bar{\sigma}_{1,4}\bar{\sigma}_{2,4}]] = \mathbb{E}[\mathbb{E}_{\boldsymbol{x}_3}[\bar{\sigma}_{1,3}\bar{\sigma}_{2,3}]^2].\end{aligned}$$

The claim then follows directly from Lemma 4.9. □

By Lemma 4.10 we thus have that

$$V_{\{1,2,3\},\{1,2,4\}} \le \mathbb{E}[\tau_{\{1,2,3\}}\tau_{\{1,2,4\}}] \le \frac{C_\alpha}{r^4 d}.$$

Lemma 4.11. *There exists a constant $C_\alpha > 0$ such that*

$$\mathbb{E}[\tau_{\{1,2,3\}}\tau_{\{1,4,5\}}] \le \frac{C_\alpha}{r^4 d}.$$

Proof. By the definition of the signed triangle,

$$\begin{aligned}\mathbb{E}[\tau_{\{1,2,3\}}\tau_{\{1,4,5\}}] &= \mathbb{E}[\bar{\sigma}_{1,2}\bar{\sigma}_{2,3}\bar{\sigma}_{3,1}\bar{\sigma}_{1,4}\bar{\sigma}_{4,5}\bar{\sigma}_{5,1}] = \mathbb{E}[\mathbb{E}_{\boldsymbol{x}_2,\boldsymbol{x}_3}[\bar{\sigma}_{1,2}\bar{\sigma}_{2,3}\bar{\sigma}_{3,1}]\,\mathbb{E}_{\boldsymbol{x}_4,\boldsymbol{x}_5}[\bar{\sigma}_{1,4}\bar{\sigma}_{4,5}\bar{\sigma}_{5,1}]] \\ &= \mathbb{E}[\mathbb{E}_{\boldsymbol{x}_2,\boldsymbol{x}_3}[\bar{\sigma}_{1,2}\bar{\sigma}_{2,3}\bar{\sigma}_{3,1}]^2],\end{aligned}$$

where the last equality holds since $(\boldsymbol{x}_2, \boldsymbol{x}_3)$ and $(\boldsymbol{x}_4, \boldsymbol{x}_5)$ are identically distributed.

By Jensen's inequality,

$$\begin{aligned}\mathbb{E}_{\boldsymbol{x}_2,\boldsymbol{x}_3}[\bar{\sigma}_{1,2}\bar{\sigma}_{2,3}\bar{\sigma}_{3,1}]^2 &= \mathbb{E}_{\boldsymbol{x}_2}[\bar{\sigma}_{1,2}\,\mathbb{E}_{\boldsymbol{x}_3}[\bar{\sigma}_{2,3}\bar{\sigma}_{3,1}]]^2 \\ &\le \mathbb{E}_{\boldsymbol{x}_2}[\bar{\sigma}_{1,2}^2\,\mathbb{E}_{\boldsymbol{x}_3}[\bar{\sigma}_{2,3}\bar{\sigma}_{3,1}]^2] \le \mathbb{E}_{\boldsymbol{x}_2}[\mathbb{E}_{\boldsymbol{x}_3}[\bar{\sigma}_{2,3}\bar{\sigma}_{3,1}]^2].\end{aligned}$$

Therefore, by Lemma 4.9,

$$\mathbb{E}[\tau_{\{1,2,3\}}\tau_{\{1,4,5\}}] \le \mathbb{E}[\mathbb{E}_{\boldsymbol{x}_3}[\bar{\sigma}_{2,3}\bar{\sigma}_{3,1}]^2] \le \frac{C_\alpha}{r^4 d}.$$ □

By Lemma 4.11 we thus have that

$$V_{\{1,2,3\},\{1,4,5\}} \le \mathbb{E}[\tau_{\{1,2,3\}}\tau_{\{1,4,5\}}] \le \frac{C_\alpha}{r^4 d}.$$

Putting these bounds together, we have the following lemma.

Lemma 4.12. *There exists a constant $C_\alpha > 0$ such that*

$$\mathbb{V}\mathrm{ar}[\tau(\mathcal{G}(n,p,d,r))] \le n^3 + \frac{C_\alpha n^5}{r^4 d}.$$

4.4. Concluding the proof

Combining Lemma 4.4, Lemma 4.8, and Lemma 4.12, we have for $d, r \ge C_\alpha$ that when $d/\log^2 d \gg r^6$ or $r/\log^2 r \gg d^{1/6}$,

$$|\mathbb{E}[\tau(\mathcal{G}(n,p,d,r))] - \mathbb{E}[\tau(\mathcal{G}(n,p))]| \ge \frac{C_p n^3}{r^3\sqrt{d}}$$

and

$$\max\{\mathbb{V}\mathrm{ar}[\tau(\mathcal{G}(n,p,d,r)], \mathbb{V}\mathrm{ar}[\tau(\mathcal{G}(n,p))]\} \le n^3 + \frac{C_\alpha n^5}{r^4 d}.$$

Therefore, by Chebyshev's inequality,

$$\mathrm{TV}(\mathcal{G}(n,p,d,r), \mathcal{G}(n,p)) \ge 1 - \left(\frac{C'_p r^6 d}{n^3} + \frac{C'_{p,\alpha} r^2}{n}\right)$$

for some constants $C'_p, C'_{p,\alpha} > 0$. Notice that $r^6 d/n^3 \to 0$ implies $r^2/n \to 0$. Theorem 1.1(b) is hence proved.

Acknowledgments

The authors would like to thank Ramon van Handel for insightful comments on several results and Jiacheng Zhang for suggesting the proofs of Lemma 2.6 and Lemma 4.3. We also would like to thank the editor and anonymous reviewers for refereeing the article and their helpful comments.

Funding

The authors were supported in part by NSF grant DMS-1811724.

Supplementary Material

Supplement to "A probabilistic view of latent space graphs and phase transitions" (DOI: 10.3150/22-BEJ1547SUPP; .pdf). This supplement contains the proofs omitted from Section 4.

References

[1] Athreya, A., Fishkind, D.E., Tang, M., Priebe, C.E., Park, Y., Vogelstein, J.T., Levin, K., Lyzinski, V., Qin, Y. and Sussman, D.L. (2017). Statistical inference on random dot product graphs: A survey. *J. Mach. Learn. Res.* **18** Paper No. 226, 92. MR3827114

[2] Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford: Oxford Univ. Press. With a foreword by Michel Ledoux. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001

[3] Brennan, M., Bresler, G. and Huang, B. (2021). De Finetti-Style Results for Wishart Matrices: Combinatorial Structure and Phase Transitions. arXiv preprint arXiv:2103.14011.

[4] Brennan, M., Bresler, G. and Nagaraj, D. (2020). Phase transitions for detecting latent geometry in random graphs. *Probab. Theory Related Fields* **178** 1215–1289. MR4168398 https://doi.org/10.1007/s00440-020-00998-3

[5] Bubeck, S., Ding, J., Eldan, R. and Rácz, M.Z. (2016). Testing for high-dimensional geometry in random graphs. *Random Structures Algorithms* **49** 503–532. MR3545825 https://doi.org/10.1002/rsa.20633

[6] Bubeck, S. and Ganguly, S. (2018). Entropic CLT and phase transition in high-dimensional Wishart matrices. *Int. Math. Res. Not. IMRN* **2** 588–606. MR3801440 https://doi.org/10.1093/imrn/rnw243

[7] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. *Probability and Its Applications (New York)*. Heidelberg: Springer. MR2732624 https://doi.org/10.1007/978-3-642-15007-4

[8] Devroye, L., György, A., Lugosi, G. and Udina, F. (2011). High-dimensional random geometric graphs and their clique number. *Electron. J. Probab.* **16** 2481–2508. MR2861682 https://doi.org/10.1214/EJP.v16-967

[9] Duchemin, Q. and De Castro, Y. (2021). Random Geometric Graph: Some recent developments and perspectives. Preprint.

[10] Eldan, R. and Mikulincer, D. (2020). Information and dimensionality of anisotropic random geometric graphs. In *Geometric Aspects of Functional Analysis. Vol. I*. *Lecture Notes in Math.* **2256** 273–324. Cham: Springer. MR4175752 https://doi.org/10.1007/978-3-030-36020-7_13

[11] Gibbs, A.L. and Su, F.E. (2002). On choosing and bounding probability metrics. *Int. Stat. Rev.* **70** 419–435.

[12] Gilbert, E.N. (1961). Random plane networks. *J. Soc. Indust. Appl. Math.* **9** 533–543. MR0132566

[13] Hoff, P.D., Raftery, A.E. and Handcock, M.S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* **97** 1090–1098. MR1951262 https://doi.org/10.1198/016214502388618906

[14] Janson, S. (1997). *Gaussian Hilbert Spaces*. *Cambridge Tracts in Mathematics* **129**. Cambridge: Cambridge Univ. Press. MR1474726 https://doi.org/10.1017/CBO9780511526169

[15] Jiang, T. and Li, D. (2015). Approximation of rectangular beta-Laguerre ensembles and large deviations. *J. Theoret. Probab.* **28** 804–847. MR3413957 https://doi.org/10.1007/s10959-013-0519-7

[16] Liu, S. and Rácz, M.Z. (2021). Phase transition in noisy high-dimensional random geometric graphs. arXiv preprint arXiv:2103.15249.

[17] Liu, S. and Rácz, M.Z. (2023). Supplement to "A probabilistic view of latent space graphs and phase transitions." https://doi.org/10.3150/22-BEJ1547SUPP

[18] Ma, Z., Ma, Z. and Yuan, H. (2020). Universal latent space model fitting for large networks with edge covariates. *J. Mach. Learn. Res.* **21** Paper No. 4, 67. MR4071187 https://doi.org/10.1109/tnnls.2020.3010690

[19] Marchal, O. and Arbel, J. (2017). On the sub-Gaussianity of the beta and Dirichlet distributions. *Electron. Commun. Probab.* **22** Paper No. 54, 14. MR3718704 https://doi.org/10.1214/17-ECP92

[20] Penrose, M. (2003). *Random Geometric Graphs*. *Oxford Studies in Probability* **5**. Oxford: Oxford Univ. Press. MR1986198 https://doi.org/10.1093/acprof:oso/9780198506263.001.0001

[21] Penrose, M.D. (2016). Connectivity of soft random geometric graphs. *Ann. Appl. Probab.* **26** 986–1028. MR3476631 https://doi.org/10.1214/15-AAP1110

[22] Polyanskiy, Y. and Wu, Y. (2012-2017). Lecture notes on Information Theory.

[23] Rácz, M.Z. and Bubeck, S. (2017). Basic models and questions in statistical network analysis. *Stat. Surv.* **11** 1–47. MR3696007 https://doi.org/10.1214/17-SS117

[24] Rácz, M.Z. and Richey, J. (2019). A smooth transition from Wishart to GOE. *J. Theoret. Probab.* **32** 898–906. MR3959632 https://doi.org/10.1007/s10959-018-0808-2

[25] Stein, C.M. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135–1151. MR0630098

[26] Tang, M., Sussman, D.L. and Priebe, C.E. (2013). Universally consistent vertex classification for latent positions graphs. *Ann. Statist.* **41** 1406–1430. MR3113816 https://doi.org/10.1214/13-AOS1112

[27] van Handel, R. (2016). Probability in High Dimension. APC 550 Lecture Notes (Princeton University).

[28] Wainwright, M.J. (2019). *High-Dimensional Statistics: A Non-asymptotic Viewpoint*. *Cambridge Series in Statistical and Probabilistic Mathematics* **48**. Cambridge: Cambridge Univ. Press. MR3967104 https://doi.org/10.1017/9781108627771

[29] Wendel, J.G. (1948). Note on the gamma function. *Amer. Math. Monthly* **55** 563–564. MR0029448 https://doi.org/10.2307/2304460

[30] Young, S.J. and Scheinerman, E.R. (2007). Random dot product graph models for social networks. In *Algorithms and Models for the Web-Graph*. *Lecture Notes in Computer Science* **4863** 138–149. Berlin: Springer. MR2504912 https://doi.org/10.1007/978-3-540-77004-6_11

[31] Zhang, A.R. and Zhou, Y. (2020). On the non-asymptotic and sharp lower tail bounds of random variables. *Stat* **9** e314, 11. MR4193419 https://doi.org/10.1002/sta4.314

Received October 2021 and revised September 2022