Detecting High-Dimensional Geometric Structure in Random Graphs

Henry Smith

Yale University

November 27, 2021

Abstract

The paper *Testing for High-Dimensional Geometry in Random Graphs* by Bubeck and colleagues considers the hypothesis testing problem of whether or not a random graph G on n vertices has some underlying high-dimensional structure. Under the "no structure" regime, the graph is distributed according to the Erdős–Rényi standard model, whereas under the alternative its edges are determined by n independent random vectors that are uniformly distributed on $\mathbb{S}^{d-1}$. The authors explore two test statistics, the number of triangles and signed triangles in G, and show that the latter is better suited for higher-dimensional problems. In our report, we present a measure-theoretic examination of two key results from the paper; these results study the distribution of the two aforementioned statistics under H_0 and H_1.

1 Introduction

Prior to digging into the main results we will explore, we present an overview of the problem considered in [1]. Bubeck and colleagues motivate their work by contemporary ambitions in the fields of computational biology and social network theory to understand structure in high-dimensional graphs. Particularly, they consider the question of whether latent geometrical structure exists altogether in such graphs. In the context of graph theory, the term 'structure' refers to the edges connecting vertices in a given graph G. Under a regime in which a graph has no geometric structure, the presence or absence of edges between vertices would be purely random. Alternatively, a graph with underlying structure would have some pattern to its edges. This paradigm of structure versus no structure lends itself to a hypothesis testing problem. In the remainder of this section, we discuss the hypotheses and test statistics presented in [1] that are relevant to our paper.

1.1 Hypothesis Testing

Bubeck and colleagues test the null hypothesis that a random graph G defined on n vertices has been generated according to the *Erdős–Rényi standard model*. That is, we have

$$H_0 : G \sim G(n, p). \tag{1}$$

Under this model, each pair of distinct vertices $i, j \in [n]$ is connected by an edge independently with probability p. This null hypothesis represents the aforementioned "no structure" paradigm since the presence or absence of each edge in G has no dependence on the remaining edges in G.

As for the alternative hypothesis, the authors consider random graph G on n vertices distributed as

$$H_1 : G \sim G(n, p, d), \tag{2}$$

where $G(n, p, d)$ is defined as follows. Let $X_1, \dots X_n \in \mathbb{R}^d$ be independent random vectors that are uniformly distributed on $\mathbb{S}^{d-1}$. Then distinct vertices $i, j \in [n]$ are connected by an edge if and only if $\langle X_i, X_j \rangle \geq t_{p,d}$, where $\langle \cdot, \cdot \rangle$ is the standard inner product in $\mathbb{R}^d$. Here $t_{p,d} \in [-1, 1]$ is a constant in $p \in (0, 1)$ and $d \in \mathbb{N}$ satisfying $\mathbb{P}(\langle X_i, X_j \rangle \geq t_{p,d}) = p$. Intuitively, vertices i and j in graph G are connected by an edge if and only if the corresponding vectors $X_i, X_j \in \mathbb{S}^{d-1}$ are "close enough". Note that $\langle X_i, X_j \rangle = \cos(\theta)$, where θ is the angle between X_i, X_j in $\mathbb{R}^d$. Thus, we can say that vertices i and j are connected by an edge if and only if the angle between random vectors X_i and X_j is sufficiently small.

1.2 The Uniform Distribution on $\mathbb{S}^{d-1}$

Our previous statement that $X_1, \dots, X_n \in \mathbb{R}^d$ are independent random vectors, uniformly-distributed on $\mathbb{S}^{d-1}$ is admittedly lacking in mathematical rigor. More specifically, each X_i is a random variable that maps from latent probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathbb{S}^{n-1}, \mathcal{B}(\mathbb{S}^{d-1}), Q)$. Here, Q is the image measure of $\mathbb{P}$ under X_i and is equal to the Lebesgue measure on $\mathbb{S}^{d-1}$.

For a more formal definition of the spherical measure Q, [2] proves that there exists a unique Borel measure σ_{d-1} on $\mathbb{S}^{d-1}$ such that for each $f \in \mathcal{M}^+(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, then $\mu(f) = (\rho_d \times \sigma_{d-1})(f)$.[1] Here, μ denotes the Lebesgue measure on $\mathbb{R}^d$ and ρ_d is the measure on $((0, \infty), \mathcal{B}(0, \infty))$ having density $\Delta(r) = r^{d-1}$ with respect to Lebesgue measure. That is, the d-dimensional Lebesgue measure may be expressed as the product measure of ρ_n defined on measurable space $((0, \infty), \mathcal{B}(0, \infty))$ and σ_{d-1} defined on measurable space $(\mathbb{S}^{d-1}, \mathcal{B}(\mathbb{S}^{d-1}))$. Using this Borel measure on $\mathbb{S}^{d-1}$, we can equivalently define $Q = \frac{\sigma}{\sigma(\mathbb{S}^{d-1})} = \frac{\Gamma(n/2)}{2\pi^{d/2}}\sigma$ to be our desired spherical measure [2].[2]

[1] This result is subsequently extended to $f \in \mathcal{L}^1(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu)$ by splitting f into f^+ and f^-.

[2] For further details about spherical measure, one should reference section 2.7, "Integration in Polar Coordination" of [2].

1.3 Test Statistics

Having established the distribution of G under both H_0 and H_1, we present two statistics to test H_0 against H_1. Define the *adjacency matrix* $A \in \mathbb{R}^{n\times n}$ of graph G such that

$$A_{i,j} = \begin{cases} 1 & \text{there exists an edge between vertices } i \text{ and } j \\ 0 & \text{otherwise} \end{cases}$$

As Bubeck and collaborators note in their work, the number of triangles present in random graph G may be used to uncover its latent geometric structure.[3] Accordingly, the authors consider the *triangle statistic*

$$T(G) := \sum_{\{i,j,k\}\in\binom{[n]}{3}} A_{i,j}A_{i,k}A_{j,k} \tag{3}$$

as well as the *signed triangle statistic*

$$\tau(G) := \sum_{\{i,j,k\}\in\binom{[n]}{3}} (A_{i,j}-p)(A_{i,k}-p)(A_{j,k}-p) \tag{4}$$

for testing the null hypothesis. Clearly, $T(G)$ is doing no more than counting the number of triangles present in G. The signed triangle statistic, on the other hand, reduces the variance of T under the null hypothesis H_0. In particular, a key result of [1] is that $\mathrm{Var}(\tau(G(n,p)))$ is on the order of n^3, whereas $\mathrm{Var}(T(G(n,p)))$ is on the order of n^4. This fact is then used in Theorem 2 to prove that the signed triangle statistic is asymptotically powerful so long as $d \ll n^3$. The meaning of "asymptotically powerful" is considered in section 3 of our report.

2 Expected Number of Triangles in $G(n,p,d)$

The first result in [1] that we consider is Lemma 1. Particularly, this result establishes a lower bound on the value of $\mathbb{E}T(G(n,p,d))$, the expected value of the triangle statistic for graph G under H_1.

In order to prove this result, the authors define the event

$$E := \left\{ \langle X_1, X_2\rangle \geq t_{p,d}, \langle X_1, X_3\rangle \geq t_{p,d}, \langle X_2, X_3\rangle \geq t_{p,d} \right\}$$

for independent random vectors X_1, X_2, X_3 uniformly distributed on $\mathbb{S}^{n-1}$.

Note that the map $f(\omega) := \langle X_i(\omega), X_j(\omega)\rangle$ is, in fact, $\mathcal{F}\setminus\mathcal{B}([-1,1])$-measurable. To understand why this is the case, recall that each of $X_i, X_j : (\Omega,\mathcal{F},\mathbb{P}) \to (\mathbb{S}^{n-1}, \mathcal{B}(\mathbb{S}^{n-1}))$ is measurable, and so $g : (\Omega,\mathcal{F},\mathbb{P}) \to (\mathbb{S}^{n-1}\times\mathbb{S}^{n-1}, \mathcal{B}(\mathbb{S}^{n-1})\otimes\mathcal{B}(\mathbb{S}^{n-1}))$ defined $g(\omega) = (X_i(\omega), X_j(\omega))$ is measurable. Further, the inner product mapping $h : (\mathbb{S}^{n-1}\times\mathbb{S}^{n-1}, \mathcal{B}(\mathbb{S}^{n-1})) \to ([-1,1], \mathcal{B}([-1,1]))$ defined $h(x,y) = \langle x, y\rangle$ is continuous. All together, this gives us that $f = h\circ g$, the composition of a continuous map with a measurable map is indeed $\mathcal{F}\setminus\mathcal{B}([-1,1])$-measurable. This is important because it gives us that $\{\langle X_i, X_j\rangle \geq t_{p,d}\} \in \mathcal{F}$, and so $E \in \mathcal{F}$.

To understand why we define this event E, recall from (3) that

$$T(G(n,p,d)) = \sum_{\{i,j,k\}\in\binom{[n]}{3}} A_{i,j}A_{i,k}A_{j,k},$$

[3] See p. 5 for a rigorous justification of this statement.

and so

$$\begin{aligned}
\mathbb{E}T(G(n,p,d)) &= \mathbb{E}\left(\sum_{\{i,j,k\}\in\binom{[n]}{3}} A_{i,j}A_{i,k}A_{j,k}\right) \\
&= \sum_{\{i,j,k\}\in\binom{[n]}{3}} \mathbb{E}(A_{i,j}A_{i,k}A_{j,k}) && \text{linearity} \\
&= \sum_{\{i,j,k\}\in\binom{[n]}{3}} \mathbb{E}(\mathbb{I}_E) \\
&= \sum_{\{i,j,k\}\in\binom{[n]}{3}} \mathbb{P}(E) && \mathbb{I}_E \text{ a simple function} \\
&= \binom{n}{3}\mathbb{P}(E).
\end{aligned}$$

Note that the third equality holds because the X_i's are identically distributed, and so random variables X_1, X_2, X_3 are chosen for convenience in the definition of E. The above derivation tells us that in order to bound $\mathbb{E}T(G(n,p,d))$, we must simply bound $\mathbb{P}(E)$. This task is the subject of the following lemma:

Lemma 1. *There exists a universal constant $C > 0$ such that whenever $p < \frac{1}{4}$ we have that*

$$\mathbb{P}(E) \geq C^{-1}p^3\frac{(log\frac{1}{p})^{3/2}}{\sqrt{d}}. \tag{5}$$

Moreover, for every fixed $0 < p < 1$, there exists a constant $C_p > 0$ such that for all $d \geq C_p^{-1}$

$$\mathbb{P}(E) \geq p^3\left(1 + \frac{C_p}{\sqrt{d}}\right). \tag{6}$$

As a preliminary result, we establish the one-dimensional marginal distribution of a random vector uniformly distributed on $\mathbb{S}^{n-1}$. Particularly, section 2 of [5] calculates that this one-dimensional marginal has density

$$f_d(x) = \frac{\Gamma(d/2)}{\Gamma((d-1)/2)\sqrt{\pi}}(1-x^2)^{(d-3)/2} \quad x \in [-1,1]$$

with respect to Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. For convenience, we let

$$\int f(x)dx$$

denote the Lebesgue integral of $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proof of Lemma 1. The first case that Bubeck and colleagues treat is that of $d \leq \frac{1}{4}\log(1/p)$.‘

The authors begin by letting $\angle(\cdot,\cdot)$ denote the geodesic distance between two vectors on $\mathbb{S}^{d-1}$. We note that the geodesic distance on $\mathbb{S}^{d-1}$ is the same as the great circle distance; that is, the shortest distance between any two vectors on $\mathbb{S}^{d-1}$ is along a great circle. From [5], we know that the great circle distance between $x, y \in \mathbb{S}^{n-1}$ is $\arccos(\langle x, y\rangle)$.

We then define $g(\theta) := \mathbb{P}(\angle(X_1, X_2) < \theta)$. Just as in our previous argument for E, we have $\{\angle(X_1, X_2) < \theta\} \in \mathcal{F}$ since $h(z) = \arccos(z)$ is continuous on $z \in [-1, 1]$, and so $\arccos(\langle X_i, X_j\rangle)$ is measurable as it is the composition of a continuous function and a measurable function.

Since the coordinate system in $\mathbb{R}^d$ is unspecified, then without loss of generality we may assume $X_1 = e_1$, where e_1 denotes the first standard basis vector in $\mathbb{R}^d$. Moreover, let φ denote the angle between $X_1 = e_1$ and X_2 in $\mathbb{S}^{d-1}$. Then we may rewrite $g(\theta)$ in terms of this angle φ:

$$\begin{aligned} g(\theta) &= \mathbb{P}(\angle(X_1, X_2) < \theta) \\ &= \mathbb{P}(\arccos(\langle X_1, X_2\rangle) < \theta) \\ &= \mathbb{P}(\arccos(\cos(\varphi)) < \theta) \\ &= \mathbb{P}(\varphi < \theta) && \text{for } \varphi \in [0, \pi/2]. \end{aligned}$$

The event that the angle φ between $X_1 = e_1$ and X_2 is less than θ defines a *hyperspherical cap* of $\mathbb{S}^{d-1}$. Note that we assume (without loss of generality) $\varphi \in [0, \pi/2]$ because otherwise the complementary spherical cap $\{\varphi < \theta \leq \pi\}$ has angle less than $\pi/2$, in which case we would consider this cap instead. Thus, $g(\theta)$ is no more than the spherical measure of this hyperspherical cap. Equivalently, $g(\theta)$ is the surface area of the hyperspherical cap normalized by the total surface area of $\mathbb{S}^{d-1}$. And so from [4], which computes the area of a hyperspherical cap of $\mathbb{S}^{n-1}$, we deduce that

$$g(\theta) = \frac{(d-1)\pi^{(d-1)/2}}{\Gamma\left(\frac{d+1}{2}\right)} \int_0^\theta \sin(x)^{d-2} dx.$$

With a straightforward change of variables $u = 2x$, we compute

$$\begin{aligned} g(\theta/2) &= \frac{(d-1)\pi^{(d-1)/2}}{\Gamma\left(\frac{d+1}{2}\right)} \int_0^{\theta/2} \sin(x)^{d-2} dx \\ &= \frac{(d-1)\pi^{(d-1)/2}}{\Gamma\left(\frac{d+1}{2}\right)} \int_0^{\theta} \frac{1}{2} \sin(u/2)^{d-2} du. \end{aligned}$$

And since $\sin(u/2) \geq \sin(u)$ for every $u \in [0, \pi]$, then we get that for all $\theta \in [0, \pi)$:

$$\begin{aligned} g(\theta/2) &\geq \frac{(d-1)\pi^{(d-1)/2}}{\Gamma\left(\frac{d+1}{2}\right)} \int_0^{\theta} \frac{1}{2} \left(\frac{\sin(u)}{2}\right)^{d-2} du && \text{monotonicity} \\ &= \frac{(d-1)\pi^{(d-1)/2}}{\Gamma\left(\frac{d+1}{2}\right)} \int_0^{\theta} 2^{1-d} \sin(u)^{d-2} du \\ &\geq \frac{(d-1)\pi^{(d-1)/2}}{\Gamma\left(\frac{d+1}{2}\right)} \int_0^{\theta} 2^{-d} \sin(u)^{d-2} du \\ &= 2^{-d} g(\theta). \end{aligned} \tag{7}$$

Having given this inequality for $g(\theta/2)$, let us momentarily consider the event E. First we note that, by the definition of $\angle(\cdot, \cdot)$,

$$\begin{aligned} E &= \{\arccos(\langle X_1, X_2\rangle) \geq \arccos(t_{p,d}), \arccos(\langle X_1, X_3\rangle) \geq \arccos(t_{p,d}), \arccos(\langle X_2, X_3\rangle) \geq \arccos(t_{p,d})\} \\ &= \{\angle(X_1, X_2) \geq \arccos(t_{p,d}), \angle(X_1, X_3) \geq \arccos(t_{p,d}), \angle(X_2, X_3) \geq \arccos(t_{p,d})\}. \end{aligned}$$

Recall that the geodesic distance on $\mathbb{S}^{n-1}$ defines a proper metric on $\mathbb{S}^{n-1}$, and so we may invoke the triangle inequality to say that

$$\begin{aligned} &\angle(X_1, X_2) < \frac{1}{2}\arccos(t_{p,d}), \angle(X_1, X_3) < \frac{1}{2}\arccos(t_{p,d}) \\ \implies &\angle(X_2, X_3) \leq \angle(X_2, X_1) + \angle(X_1, X_3) \leq \frac{1}{2}\arccos(t_{p,d}) + \frac{1}{2}\arccos(t_{p,d}) = \arccos(t_{p,d}). \end{aligned}$$

Putting together the two previous statements, we deduce

$$E \supseteq \{\angle(X_1, X_2) < \frac{1}{2}\arccos(t_{p,d}), \angle(X_1, X_3) < \frac{1}{2}\arccos(t_{p,d})\}.$$

Since X_2 and X_3 are independent random variables, then any function of these random variables are independent. Particularly, we have $\angle(X_1, X_2) \perp \angle(X_1, X_3)$. Wrapping up this portion of the proof, we have

$$\begin{aligned}
\mathbb{P}(E) &\geq \mathbb{P}\{\angle(X_1, X_2) < \frac{1}{2}\arccos(t_{p,d}), \angle(X_1, X_3) < \frac{1}{2}\arccos(t_{p,d})\} && \text{monotonicity} \\
&= \mathbb{P}\{\angle(X_1, X_2) < \frac{1}{2}\arccos(t_{p,d})\}^2 && \angle(X_1, X_2) \perp \angle(X_1, X_3) \\
&= g\left(\frac{1}{2}\arccos(t_{p,d})\right)^2 \\
&\geq 2^{-2d} g(\arccos(t_{p,d})) && \text{from (5)} \\
&= p^2 2^{-2d} && \mathbb{P}(\langle X_i, X_j\rangle \geq t_{p,d}) = p
\end{aligned}$$

Because we assumed that $d \leq \frac{1}{4}\log(1/p)$, then

$$p^2 2^{-2d} \geq p^2 2^{\frac{1}{2}\log(1/p)} \geq p^3\left(1 + c\left(\log(1/p)\right)^{3/2}\right)$$

for some constant $c > 0$. And since $\mathbb{P}(E) \geq p^3\left(1 + c\left(\log(1/p)\right)^{3/2}\right) \geq cp^3\left(\log(1/p)\right)^{3/2}$, taking $C = c$ satisfies (5), the first inequality in Lemma 1. Similarly, $\mathbb{P}(E) \geq p^3\left(1 + c\left(\log(1/p)\right)^{3/2}\right)$ means that $C_p = c\log(1/p)^{3/2}$ satisfies (6). And so we have proven that Lemma 1 holds when $d \leq \frac{1}{4}\log(1/p)$.

Now, for the remainder of the proof, we assume $d \geq \frac{1}{4}\log(1/p)$. We begin with the case that $p < 1/2$.

Define the events $E_{i,j} = \{\langle X_i, X_j\rangle \geq t_{p,d}\} \in \mathcal{F}$ as well as $E_{i,j}(x) = \{\langle X_i, X_j\rangle = x\} \in \mathcal{F}$.

We consider the mapping $T : (\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}, \mathcal{B}(\mathbb{S}^{d-1}) \otimes \mathcal{B}(\mathbb{S}^{d-1}), Q \times Q) \to ([-1,1], \mathcal{B}([-1,1]), V)$ defined $T(x, y) := \langle x, y\rangle$, where V is the image measure of $Q \times Q$ under T. We claim that there exists a conditional probability of $Q \times Q$ under T, which is unique $Q \times Q$-almost surely. To understand why this is the case, recall that the following are sufficient conditions for the existence of a conditional probability:

1. $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ is a complete, separable metric space (a *Polish space*)
2. The sigma-field from which we are mapping is $\mathcal{B}(\mathbb{S}^{d-1} \times \mathbb{S}^{d-1})$
3. V is the image measure of $Q \times Q$ under map T
4. $g_r(T)$ is $\left(\mathcal{B}(\mathbb{S}^{d-1}) \otimes \mathcal{B}(\mathbb{S}^{d-1})\right) \otimes \mathcal{B}([-1,1])$-measurable

Notice that the third condition is immediately true as a consequence of how we defined our mapping T. Further, $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ endowed with the Euclidean metric $\|\cdot\|_2$ is clearly a metric space. $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ is separable because it is a subset of separable metric space $(\mathbb{R}^d \times \mathbb{R}^d, \|\cdot\|_2)$, and it is complete because it is a closed subset of complete metric space $(\mathbb{R}^d \times \mathbb{R}^d, \|\cdot\|_2)$. Thus, $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ is indeed a Polish space. For the second condition, we note that the Borel sigma-field on $\mathbb{R}$ satisfies:

$$\underbrace{\mathcal{B}(\mathbb{R}) \otimes \ldots \otimes \mathcal{B}(\mathbb{R})}_{k\text{-times}} = \mathcal{B}(\mathbb{R}^k).$$

Using this property, we have $\mathcal{B}(\mathbb{S}^{d-1}) \otimes \mathcal{B}(\mathbb{S}^{d-1}) = \mathcal{B}(\mathbb{S}^{d-1} \times \mathbb{S}^{d-1})$, and so T indeed maps from the Borel sigma-field on $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$. Add proof of condition 4. Having established each of the sufficient conditions, we deduce that a conditional probability $\{(Q \times Q)_t\}$ mapping from $([-1,1], \mathcal{B}([-1,1]), V)$ to $(\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}, \mathcal{B}(\mathbb{S}^{d-1}) \otimes \mathcal{B}(\mathbb{S}^{d-1}), Q \times Q)$ indeed exists.

Note that, as part of their argument, Bubeck and colleagues provide a specific definition for the conditional probability. In particular, they claim that for $F \in \mathcal{F}$, the conditional probability of $Q \times Q$ given T is

$$\mathbb{P}(F|T(X_i, X_j) = x) = \lim_{\varepsilon \to \infty} \frac{\mathbb{P}\left(F \cap \{\langle X_i, X_j \rangle \in [x - \varepsilon, x + \varepsilon]\}\right)}{\mathbb{P}\{\langle X_i, X_j \rangle \in [x - \varepsilon, x + \varepsilon]\}} \quad x \in [-1, 1].$$

Here, we have measure $\mathbb{P}$ rather than $Q \times Q$ because, although the conditional probability was defined on the probability measure $Q \times Q$ under map T, Q is defined to be the image measure of $\mathbb{P}$ under each random variable X_i. While one could establish that this is, in fact, the conditional probability distribution of $Q \times Q$ given T, the remainder of the argument does not rely on this particular definition. Accordingly, we prefer our approach of establishing the existence of the conditional probability without providing superfluous details.

As a consequence of the existence of the conditional probability of $Q \times Q$ under T, we consider

$$\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}) - \mathbb{P}(E_{1,3}, E_{2,3}).$$

We rewrite the second term in the above expression in terms of the conditional probability as follows:

$$\mathbb{P}(E_{1,3}, E_{2,3}) = \mathbb{P}(E_{1,3}, E_{2,3}|\Omega) = \frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|\langle X_1, X_2 \rangle \geq 0) + \frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|\langle X_1, X_2 \rangle < 0).$$

The first equality uses the definition of the conditional probability and the second invokes linearity of expectation.[4]

We now make what seems like, at first glance, an abstruse claim. Particularly, we suggest that $f(x) := \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(x))$ is a monotonically-increasing function in $x \in [-1, 1]$. To understand why this is the case, we consider $\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(x))$ for some fixed x. Once again, since the coordinate system in $\mathbb{R}^d$ is unspecified, then, without loss of generality, we may let $X_1 = e_1$ and $X_2 = xe_1 + \sqrt{1 - x^2}e_2$ so that $\langle X_1, X_2 \rangle = x$. Now, conditioning on $E_{1,2}(x)$, each of $\{\langle X_1, X_3 \rangle \geq t_{p,d}\} = \{X_{3,1} \geq t_{p,d}\}$ and $\{\langle X_2, X_3 \rangle \geq t_{p,d}\} = \{xX_{3,1} + \sqrt{1 - x^2}X_{3,2} \geq t_{p,d}\}$ defines a hyperspherical cap of $\mathbb{S}^{d-1}$. In particular, $\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(x))$ is equal to the spherical measure of the intersection of these two hyperspherical caps. Notice that $\langle X_1, X_2 \rangle = x \Rightarrow cos(\varphi) = x$, where φ is the angle formed between X_1 and X_2. Therefore, as x increases from -1 to 1, the angle φ between X_1 and X_2 decreases from π to 0, and so the region of $\mathbb{S}^{d-1}$ on which the hyperspherical caps intersect has greater measure. In particular, when $x = 1$, then $\cos(\varphi) = 1 \Rightarrow \varphi = 0 \Rightarrow X_1 = X_2$, in which case we have $\{\langle X_1, X_3 \rangle \geq t_{p,d}\} = \{X_{3,1} \geq t_{p,d}\} = \{\langle X_2, X_3 \rangle \geq t_{p,d}\}$, meaning that the two hyperspherical caps coincide exactly.[5]

Having established the increasing property of $f(x) = \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(x))$ on $x \in [-1, 1]$, we now bound

$$\begin{aligned}
&\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}) - \mathbb{P}(E_{1,3}, E_{2,3}) \\
&= \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}) - \frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|\langle X_1, X_2 \rangle \geq 0) - \frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|\langle X_1, X_2 \rangle < 0) \\
&\geq \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}) - \frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|\langle X_1, X_2 \rangle \geq 0) - \frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(0)) \\
&= \left(\frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}) - \frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|\langle X_1, X_2 \rangle \geq 0)\right) + \left(\frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}) - \frac{1}{2}\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(0))\right) \\
&\geq \frac{1}{2}\left(\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}) - \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(0))\right) \\
&\geq \frac{1}{2}\left(\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(t_{p,d})) - \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(0))\right). \qquad (8)
\end{aligned}$$

Consider the random variable $Z_1 = \langle X_1, X_3 \rangle$. Using a similar trick to as above, we define the coordinate axes of $\mathbb{R}^n$ in terms of X_1, X_3 so that we can say something about the distribution of Z_1. In particular,

[4] In an elementary probability course, this is referred to as the law of total expectation.

[5] For more about the measure of the intersection of hyperspherical caps, see [3].

without loss of generality, we may take $X_1 = e_1$, $X_3 = z_1 e_1 + \sqrt{1-z_1^2} e_2$ so that $Z_1 = \langle X_1, X_3 \rangle = z_1$, the first coordinate of the vector X_2. Thus, from the preliminary result to Lemma 1, we know that Z_1 has density f_d with respect to Lebesgue measure. And so we invoke Bayes' rule to write

$$\begin{aligned}
&\mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}(t_{p,d})) - \mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}(0)) \\
&= \int_{z_1 \geq t_{p,d}} \left\{ \mathbb{P}(E_{2,3} | Z_1 = z_1, E_{1,2}(t_{p,d})) - \mathbb{P}(E_{2,3} | Z_1 = z_1, E_{1,2}(0)) \right\} f_d(z_1) dz_1. \qquad (9)
\end{aligned}$$

That is, we disintegrate the conditional probability $\mathbb{P}(\cdot | E_{1,2}(x))$ as the measure over the one-dimensional marginal distribution of X_i and the conditional probability $\mathbb{P}(\cdot | Z_1 = z_1, E_{1,2}(x))$.

Now, conditioning on $Z_1 = z_1$, we claim that $Z_2 := \left\langle X_3, \frac{\text{Proj}_{X_1^\perp} X_2}{\left| \text{Proj}_{X_1^\perp} X_2 \right|} \right\rangle$ is distributed as $\sqrt{1-z_1^2} Z'$, where Z' has density f_{d-1} with respect to Lebesgue measure. Here, $\text{Proj}_{X_1^\perp} X_2$ denotes the orthogonal projection of X_2 onto subspace $X_1^\perp$. To see why this statement is true, recall from above that we took $X_1 = e_1$ and $X_3 = z_1 e_1 + \sqrt{1-z_1^2} e_2$. And so we have that

$$\begin{aligned}
Z_2 &= \left\langle X_3, \frac{\text{Proj}_{X_1^\perp} X_2}{\left| \text{Proj}_{X_1^\perp} X_2 \right|} \right\rangle = \left\langle X_3, \frac{X_2 - \langle X_1, X_2 \rangle X_1}{\left| \text{Proj}_{X_1^\perp} X_2 \right|} \right\rangle \\
&= \frac{1}{\left| \text{Proj}_{X_1^\perp} X_2 \right|} \left(\langle X_2, X_3 \rangle - \langle X_1, X_2 \rangle \langle X_1, X_3 \rangle \right) \\
&= \frac{1}{\left| \text{Proj}_{X_1^\perp} X_2 \right|} \left(\left(z_1 X_{2,1} + \sqrt{1-z_1^2} X_{2,2} \right) - z_1 X_{2,1} \right) \\
&= \frac{1}{\left| \text{Proj}_{X_1^\perp} X_2 \right|} \sqrt{1-z_1^2} X_{2,2} \\
&= \frac{1}{\sqrt{1 - \langle X_1, X_2 \rangle^2}} \sqrt{1-z_1^2} X_{2,2} \\
&= \frac{1}{\sqrt{1 - X_{2,1}^2}} \sqrt{1-z_1^2} X_{2,2}.
\end{aligned}$$

Note that $\frac{X_{2,2}}{\sqrt{1-X_{2,1}^2}} \sim f_{d-1}$, and so Z_2 indeed has the same distribution as $\sqrt{1-z_1^2} Z'$.

Using this result, we rewrite $E_{2,3}$ conditioned on $Z_1 = z_1$ as $\left\{ \left| \text{Proj}_{X_1^\perp} X_2 \right| Z_2 + \langle X_3, \text{Proj}_{X_1} X_2 \rangle \geq t_{p,d} \right\}$:

$$\begin{aligned}
&\mathbb{P}(E_{2,3} | Z_1 = z_1, E_{1,2}(t_{p,d})) \\
&= \mathbb{P}\left(\left(\sqrt{1 - \langle X_1, X_2 \rangle^2} \right) \sqrt{1-z_1^2} Z' + \langle X_1, X_2 \rangle z_1 \geq t_{p,d} | E_{1,2}(t_{p,d}) \right)
\end{aligned}$$

And since $E_{1,2}(t_{p,d}) = \{ \langle X_1, X_2 \rangle \geq t_{p,d} \}$, then we have

$$\begin{aligned}
&\mathbb{P}\left(\left(\sqrt{1 - \langle X_1, X_2 \rangle^2} \right) \sqrt{1-z_1^2} Z' + \langle X_1, X_2 \rangle z_1 \geq t_{p,d} | E_{1,2}(t_{p,d}) \right) \\
&= \mathbb{P}\left(\sqrt{1-t_{p,d}^2} \sqrt{1-z_1^2} Z' + t_{p,d} z_1 \geq t_{p,d} \right) \\
&= \mathbb{P}\left(\sqrt{1-t_{p,d}^2} Z' \geq \sqrt{\frac{1-z_1}{1+z_1}} t_{p,d} \right). \qquad (10)
\end{aligned}$$

Notice that $f(z_1) = \sqrt{\frac{1-z_1}{1+z_1}}$ is a decreasing function in $z_1 \in (-1, 1]$, and so we get that $\mathbb{P}(E_{2,3}|Z_1 = z_1, E_{1,2}(t_{p,d}))$, and thus the right-hand side of (8), is an increasing function in z_1. Moreover, conditioning on $E_{1,2}(0)$, we see that $\langle X_2, X_3\rangle = \langle X_3, \text{Proj}_{X_1^\perp} X_2\rangle = Z_2$, since $\text{Proj}_{X_1} X_2 = \langle X_1, X_2\rangle X_1 = 0$. That is, conditioning on $Z_1 = z_1$ and $E_{1,2}(0)$, $E_{2,3}$ is distributed according to $\sqrt{1-z_1^2}Z'$. Recall from above that Z' has density f_{d-1} with respect to Lebesgue measure. This means that $\mathbb{P}(E_{2,3}|Z_1 = z_1, E_{1,2}(0)) = \mathbb{P}(\sqrt{1-z_1^2}Z' \geq t_{p,d})$ is the right tail probability of $\sqrt{1-z_1^2}Z'$ and so is a decreasing function in z_1.

Putting everything together, we may bound the integral (9) below by each of the two conditional probabilities at $Z_1 = t_{p,d}$. That is, we have

$$\begin{aligned}
&\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(t_{p,d})) - \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(0)) \\
&= \int_{z_1 \geq t_{p,d}} \{\mathbb{P}(E_{2,3}|Z_1 = z_1, E_{1,2}(t_{p,d})) - \mathbb{P}(E_{2,3}|Z_1 = z_1, E_{1,2}(0))\} f_d(z_1)dz_1 \\
&\geq \int_{z_1 \geq t_{p,d}} \{\mathbb{P}(E_{2,3}|Z_1 = t_{p,d}, E_{1,2}(t_{p,d})) - \mathbb{P}(E_{2,3}|Z_1 = t_{p,d}, E_{1,2}(0))\} f_d(z_1)dz_1 && \text{monotonicity} \\
&= \{\mathbb{P}(E_{2,3}|Z_1 = t_{p,d}, E_{1,2}(t_{p,d})) - \mathbb{P}(E_{2,3}|Z_1 = t_{p,d}, E_{1,2}(0))\} \left(\int_{z_1 \geq t_{p,d}} f_d(z_1)dz_1\right)
\end{aligned}$$

Since $\mathbb{P}(Z_1 \geq t_{p,d}) = \mathbb{P}(\langle X_1, X_3\rangle \geq t_{p,d}) = p$, then

$$\begin{aligned}
&\{\mathbb{P}(E_{2,3}|Z_1 = t_{p,d}, E_{1,2}(t_{p,d})) - \mathbb{P}(E_{2,3}|Z_1 = t_{p,d}, E_{1,2}(0))\} \left(\int_{z_1 \geq t_{p,d}} f_d(z_1)dz_1\right) \\
&= p\{\mathbb{P}(E_{2,3}|Z_1 = t_{p,d}, E_{1,2}(t_{p,d})) - \mathbb{P}(E_{2,3}|Z_1 = t_{p,d}, E_{1,2}(0))\} \\
&= p\left\{\mathbb{P}\left(\sqrt{1-t_{p,d}^2}Z' \geq \sqrt{\frac{1-t_{p,d}}{1+t_{p,d}}}t_{p,d}\right) - \mathbb{P}\left(\sqrt{1-t_{p,d}^2}Z' \geq t_{p,d}\right)\right\} && \text{from (10)} \\
&= p\left\{\mathbb{P}\left(Z' \geq \frac{t_{p,d}}{1+t_{p,d}}\right) - \mathbb{P}\left(Z' \geq \frac{t_{p,d}}{\sqrt{1-t_{p,d}^2}}\right)\right\} \\
&= p\mathbb{P}\left(\frac{t_{p,d}}{1+t_{p,d}} \leq Z' \leq \frac{t_{p,d}}{\sqrt{1-t_{p,d}^2}}\right)
\end{aligned}$$

And since Z' has density f_{d-1} with respect to Lebesgue measure, then

$$\begin{aligned}
&\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(t_{p,d})) - \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(0)) \\
&\geq p\int_{\frac{t_{p,d}}{1+t_{p,d}}}^{\frac{t_{p,d}}{\sqrt{1-t_{p,d}^2}}} f_{d-1}(x)dx \\
&\geq p\int_{\frac{t_{p,d}}{1+t_{p,d}}}^{t_{p,d}} f_{d-1}(x)dx. && \text{monotonicity}
\end{aligned}$$

Now, by a previous technical lemma proved in [1], we have

$$f_{d-1}(z) \geq c'dpt_{p,d} \quad \text{for all } 0 \leq z \leq t_{p,d}$$

a bound on f_{d-1} for some absolute constant $c' > 0$.[6] And so again by monotonicity of the [Lebesgue] integral

$$\mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(t_{p,d})) - \mathbb{P}(E_{1,3}, E_{2,3}|E_{1,2}(0))$$

[6]For a formal derivation, see Lemma 2 on pp. 509-510 in [1].

$$
\begin{aligned}
&\geq p \int_{\frac{t_{p,d}}{1+t_{p,d}}}^{t_{p,d}} c' dp t_{p,d} \, dx \\
&= \left(c' \left(t_{p,d} - \frac{t_{p,d}}{1+t_{p,d}} \right) \right) dp^2 t_{p,d} \\
&= \left(\frac{c'}{1+t_{p,d}} \right) dp^2 t_{p,d}^3.
\end{aligned}
$$

From the same technical lemma in [1], we take $c = \frac{c'}{1+t_{p,d}} > 0$ to be an absolute constant so that

$$
cdp^2 t_{p,d}^3 \geq c' \left(\frac{1}{2} - p^3 \right) p^2 \left(\frac{(\log(1/p))^{3/2}}{\sqrt{d}} \wedge d \right)
$$

Moreover, because we assumed that $d \geq \frac{1}{4}\log(1/p) \Rightarrow d^{3/2} \geq \left(\frac{1}{4}\right)^{3/2} \log(1/p)^{3/2} \Rightarrow d \geq \frac{\left(\frac{1}{4}\right)^{3/2} \log(1/p)^{3/2}}{\sqrt{d}}$, then we achieve the bound

$$
\begin{aligned}
\mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}(t_{p,d})) - \mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}(0)) &\geq c' \left(\frac{1}{2} - p^3 \right) p^2 \left(\frac{(\log(1/p))^{3/2}}{\sqrt{d}} \wedge d \right) \\
&\geq \frac{1}{16} c' \left(\frac{1}{2} - p^3 \right) p^2 \frac{\log(1/p))^{3/2}}{\sqrt{d}}. \qquad (11)
\end{aligned}
$$

To wrap up this case of the proof, we return to $\mathbb{P}(E) = \mathbb{P}(E_{1,2}, E_{1,3}, E_{2,3})$. Since $p \in (0,1)$, then the event $E_{1,2}$ has strictly positive measure with respect to $\mathbb{P}$. This means that $\mathbb{P}(F|E_{1,2})$, $F \in \mathcal{F}$ respects the traditional notion of conditional probability from an elementary statistics course: $\mathbb{P}(F|E_{1,2}) = \frac{\mathbb{P}(F \cap E_{1,2})}{\mathbb{P}(E_{1,2})}$. Specifically, this allows us to express

$$
\begin{aligned}
\mathbb{P}(E) &= \mathbb{P}(E_{1,2}, E_{1,3}, E_{2,3}) \\
&= \mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}) \mathbb{P}(E_{1,2}) \\
&= \{\mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}) - \mathbb{P}(E_{1,3}, E_{2,3})\} \mathbb{P}(E_{1,2}) + \mathbb{P}(E_{1,3}, E_{2,3}) \mathbb{P}(E_{1,2}) \\
&\geq \frac{1}{2} \{\mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}(t_{p,d})) - \mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}(0))\} \mathbb{P}(E_{1,2}) + \mathbb{P}(E_{1,3}, E_{2,3}) \mathbb{P}(E_{1,2}) \qquad \text{from (8)}
\end{aligned}
$$

With our previous lower bound (11) on $\mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}(t_{p,d})) - \mathbb{P}(E_{1,3}, E_{2,3} | E_{1,2}(0))$,

$$
\begin{aligned}
\mathbb{P}(E) &\geq \left\{ \frac{1}{32} c' \left(\frac{1}{2} - p^3 \right) p^2 \frac{\log(1/p))^{3/2}}{\sqrt{d}} \right\} \mathbb{P}(E_{1,2}) + \mathbb{P}(E_{1,3}, E_{2,3}) \mathbb{P}(E_{1,2}) \\
&= \left\{ \frac{1}{32} c' \left(\frac{1}{2} - p^3 \right) p^2 \frac{\log(1/p))^{3/2}}{\sqrt{d}} \right\} p + \mathbb{P}(E_{1,3}, E_{2,3}) p
\end{aligned}
$$

Since $\langle X_1, X_3 \rangle, \langle X_2, X_3 \rangle$ are functions of independent random variables X_1, X_2, then they are also independent. Finally, we conclude

$$
\begin{aligned}
\mathbb{P}(E) &\geq \left\{ \frac{1}{32} c' \left(\frac{1}{2} - p^3 \right) p^2 \frac{\log(1/p))^{3/2}}{\sqrt{d}} \right\} p + \mathbb{P}(E_{1,3}) \mathbb{P}(E_{2,3}) p \\
&= \left\{ \frac{1}{32} c' \left(\frac{1}{2} - p^3 \right) p^2 \frac{\log(1/p))^{3/2}}{\sqrt{d}} \right\} p + p^3 \\
&= p^3 \left(1 + \frac{1}{32} c' \left(\frac{1}{2} - p \right)^3 \frac{\log(1/p)^{3/2}}{\sqrt{d}} \right). \qquad (12)
\end{aligned}
$$

This, in turn, implies that $\mathbb{P}(E) \geq p^3 \frac{1}{32} c' \left(\frac{1}{2} - p\right)^3 \frac{\log(1/p)^{3/2}}{\sqrt{d}} \geq \left(\frac{c'}{256}\right) p^3 \frac{\log(1/p)^{3/2}}{\sqrt{d}}$ for $p < \frac{1}{4}$. And so taking $C = \frac{256}{c'} > 0$, we have that (5) holds. This concludes the proof of (5), since we have shown that the desired inequality holds for all cases in which $p < \frac{1}{4}$.

From (12), we also see that $C_p = \frac{1}{32}c'(\frac{1}{2} - p)^3 \log(1/p)^{3/2}$ is a positive constant for all $p < \frac{1}{2}$ such that $\mathbb{P}(E) \geq p^3 \left(1 + \frac{C_p}{\sqrt{d}}\right)$ whenever $d \geq C_p^{-1}$. Thus, this value of C_p satisfies inequality (6) for every $p < \frac{1}{2}$.

□

3 Total Variation Distance Between $\tau(G(n,p))$ and $\tau(G(n,p,d))$

4 Conclusion

References

[1] S. Bubeck, J. Ding, R. Eldan, and M. Z. Rácz, *Testing for high-dimensional geometry in random graphs*, Random Structures & Algorithms, 49 (2016), pp. 503–532.

[2] G. B. Folland, *Real analysis: modern techniques and their applications*, vol. 40, John Wiley & Sons, 1999.

[3] Y. Lee and W. C. Kim, *Concise formulas for the surface area of the intersection of two hyperspherical caps*, KAIST Technical Report, (2014).

[4] L. S, *Concise formulas for the area and volume of a hyperspherical cap*, Asian Journal of Mathematics Statistics, 4 (2011).

[5] S. Sodin, *Tail-sensitive gaussian asymptotics for marginals of concentrated measures in high dimension*, arXiv preprint math/0501382, (2005).