Phase transition on the maximal overlap of two independent random geometric graphs

Course project report for 6.S896

Hang Du*

December 11, 2023

Abstract

In this note, we consider the maximal overlap among vertex bijections between two independent d-dimensional random geometric graphs, both with n vertices and average degree $n^\theta, \theta \in (0,1)$. It is know that the maximal overlap compared with the size of a single graph undergoes a phase transition from $1-o(1)$ to $o(1)$ as d grows. We conjecture that the transition threshold is at $d \asymp \log n / \log\log n$, and we provide upper and lower bounds which are tight up to $\log\log n$ factors.

1 Introduction

The *random geometric graph* is a probabilistic model for graphs with geometric natures. We fix two integers n and d as the size and dimension for the graph, respectively. Let ρ be the uniform measure on the d-dimensional sphere $\mathbb{S}^d$, the random geometric is defined as follows.

Definition 1.1 (Random geometric graph with connecting threshold τ)**.** *Let $u_1, \dots, u_n \overset{i.i.d.}{\sim} \rho$, construct a graph G on $[n]$ by forming edges between pairs (i,j) if and only if $u_i \cdot u_j \geq \tau$. We denote the law of G as $\mathcal{G}(n,d,\tau)$.*

In this note, we will also fix a constant $\theta \in (0,1)$, and let $\tau_* = \tau_*(n,d,\theta)$ be such that $G \sim \mathcal{G}(n,d,\tau(n,d,\theta))$ has average degree n^θ. We will also write $\mathcal{G}$ for the shorthand of $\mathcal{G}(n,d,\tau_*)$. We make the following (heuristic) observations:

- When d is relatively small compared with n, $G \sim \mathcal{G}$ is basically a discretization of $\mathbb{S}^d$ (see e.g. [5] for the case $d = O(1)$). Hence in the low-dimensional regime, G possesses a strong geometric rigidity and there is essentially "no randomness" in $\mathcal{G}$.
- When d is large enough compared with n , it can be shown that $G \sim \mathcal{G}$ is indistinguishable with an Erdős-Rényi graph on n vertices with average degree n^θ (see [2] for the case when $d \gg n^3$, and see also [1, 6] for some later improvements). Thus, in sharp contrary to the low-dimensional regime, G is "purely random" without any geometric property that can be observed.

These observations indicate that there must be certain geometry-randomness phase transition of the random geometric graph as the dimension grows, and this phenomenon serves as one of the central topics in the study of (high-dimensional) random geometric graphs in the past decade. While significant progress have been made on determining the regime where randomness takes over everything [2, 1, 6],

*Department of Mathematics, MIT, hangdu@mit.edu.

to the author's best knowledge, there is few result on the threshold when the geometric rigidity starts to fade and randomness comes to play a substantial role. In this note, we make a step towards answering the latter question by studying the maximal overlap of two instances that are independently sampled from $\mathcal{G}$. We begin by the following definition.

Definition 1.2. *For two graphs G_1, G_2 on $[n]$ and a permutation $\pi \in \mathrm{S}_n$, define*

$$\mathrm{O}(G_1, G_2, \pi) = \sum_{1 \le i < j \le n} G^1_{i,j} G^2_{\pi(i),\pi(j)} \tag{1.1}$$

(where $(G^1_{i,jj})$ and $(G^2_{i,j})$ are the adjacency matrices for G_1 and G_2, respectively), and we further denote that

$$\Lambda(G_1, G_2) = \frac{\max_{\pi \in \mathrm{S}_n} \mathrm{O}(G_1, G_2, \pi)}{|E(G_1)| \wedge |E(G_2)|} \,. \tag{1.2}$$

Remark 1.3. *It is clear from the definition that $\Lambda(G_1, G_2) \le 1$. Heuristically speaking, if $\Lambda(G_1, G_2)$ is close to 1, then it implies that the "smaller" graph can be nearly embedded into the larger one, and furthermore, if these two graph have almost same sizes, then they look very similar in certain sense.*

It is clear from the definition that $\Lambda(G_1, G_2) \le 1$. In addition, if $\Lambda(G_1, G_2)$ is close to G and $|E(G_1)|, |E(G_2)|$ are close, then G_1, G_2 look very similar. One way to measure the geometric rigidity of a random geometric graph sampled from $\mathcal{G}$ is to look at the behavior $\Lambda(G_1, G_2)$ for two independent $G_1, G_2 \sim \mathcal{G}$. On the one hand, for low-dimensional case when $d = O(1)$, it can be deduced from [5, Lemma 2.1] that $\Lambda(G_1, G_2)$ is typically $1 - o(1)$. On the other hand, for high-dimensional case when G_1, G_2 are indistinguishable with a pair of independent Erdős-Rényi graphs, a straightforward union bound yields that $\Lambda(G_1, G_2) = o(1)$ with high probability (see also [3, 4] for a detailed study for the case of Erdős-Rényi graph). Heuristically speaking, the threshold for randomness come into play more or less corresponds to the threshold when $\Lambda(G_1, G_2)$ starts to be away from 1, so we turn to consider the transition of $\Lambda(G_1, G_2)$ as d grows.

In this note, we provide sufficient conditions on d in in terms of n such that with high probability $\Lambda(G_1, G_2) = 1 - o(1)$ or $o(1)$. In particular, our results determine the transition threshold of the dimension d for $\Lambda(G_1, G_2)$ going from $1 - o(1)$ to $o(1)$ up to a $\log \log n$ factor.

Theorem 1.4. *Let $d_0 = \log n / \log \log n$, then for any constant $\theta \in (0, 1)$, it holds that for some constant $\lambda = \lambda(\theta) > 0$,*

$$d \le \lambda d_0 \Rightarrow \mathbb{P}[\Lambda(G_1, G_2) = 1 - o(1)] = 1 - o(1) \,, \tag{1.3}$$

and

$$d \gg \log n \Rightarrow \mathbb{P}[\Lambda(G_1, G_2) = o(1)] = 1 - o(1) \,, \tag{1.4}$$

where the probability $\mathbb{P}$ is taken over $(G_1, G_2) \sim \mathcal{G}$.

We expect our lower bound (1.3) is tight. More precisely, we make the following conjecture:

Conjecture 1.5. *With the same notations, it holds that for some constant $\lambda' = \lambda'(\theta) > 0$,*

$$d \ge \lambda' d_0 \Rightarrow \mathbb{P}[\Lambda(G_1, G_2) = o(1)] = 1 - o(1) \,. \tag{1.5}$$

The note is organized as follows: in Section 2 we introduce several basic properties and estimates for random geometric graphs, in Section 3 and Section 4 we prove the lower bound (1.3) and the upper bound (1.4), respectively. We conclude this note by some further discussions on potential approaches towards solving the main conjecture.

2 Preliminaries on random geometric graphs

In this section, we give some basic properties on random geometric graphs which will be useful in later proofs. We begin by a handful estimation on the dot-product between uniformly chosen points on $\mathbb{S}^d$.

Proposition 2.1. *For any fixed $u \in \mathbb{S}^d$ and $v \sim \rho$, $u \cdot v$ has probability density*

$$\psi(x) = \frac{\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2})\sqrt{\pi}}(1-x^2)^{\frac{d-2}{2}} \triangleq C_d(1-x^2)^{\frac{d-2}{2}},\ x \in [-1,1]\,. \tag{2.1}$$

Furthermore, for any $0 \le r \le 2$, it holds that

$$\frac{C_d}{2ed}\left(r^2 - r^4/4\right)^{\frac{d}{2}} \le \mathbb{P}[\mathrm{d}(u,v) \le r] \le \frac{C_d}{2} r^2 \left(r^2 - r^4/2\right)^{\frac{d-2}{2}}\ . \tag{2.2}$$

Proof. (2.1) follows from standard calculations (see, e.g., [1, Lemma 5.1]), and given this we have (note that $\mathrm{d}(u,v) \le r \iff u \cdot v \ge 1 - r^2/2$),

$$\mathbb{P}[\mathrm{d}(u,v) \le r] = \int_{1-r^2/2}^{1} C_d(1-x^2)^{\frac{d-2}{2}}\,\mathrm{d}x \le \frac{C_d r^2}{2}(1-(1-r^2/2)^2)^{\frac{d-2}{2}} = \frac{C_d r^2}{2}\cdot(r^2-r^4/4)^{\frac{d-2}{2}}\ ,$$

which gives the upper bound in (2.2). On the other hand, take $r' = \frac{r^2 - r^4/4}{d}$, we have that

$$\mathbb{P}[\mathrm{d}(u,v) \le r] \ge \int_{1-r^2/2}^{1-r^2/2+r'} C_d(1-x^2)^{\frac{d-2}{2}}\,\mathrm{d}x \ge \frac{C_d(r^2-r^4/4)}{2d}\left(1-\frac{1}{d}\right)^{\frac{d-2}{2}}\cdot(r^2-r^4/4)^{\frac{d-2}{2}}\ ,$$

which is bounded below by $\frac{C_d}{2ed}\cdot(r^2-r^4/4)^{\frac{d}{2}}$. This verifies (2.2) can completes the proof. ☐

The next lemma consider the degree of a vertex in the random geometric graph.

Lemma 2.2. *For any $\tau \in [-1,1]$ and $v \in [n]$, let $G \sim \mathcal{G}(n,d,\tau)$, then the degree d_v of v in G has the binomial distribution $\mathbf{B}(n,p)$, where $p = \mathbb{P}[\mathrm{d}(u,v) \le \sqrt{2-2\tau}]$. Furthermore, for any $\delta > 0$, it holds that*

$$\mathbb{P}[|d_v - np| \ge K] \le 2\exp\left(-\frac{K^2}{2(np+K)}\right)\,. \tag{2.3}$$

Proof. The first claim follows immediately from the fact that the vectors $u_1, \dots, u_n$ generate G are independent, while (2.3) is a standard estimate on the tail probability of binomial variables. ☐

3 The lower bound

In this section, we assume that $d \le \theta d_0/10$ and prove (1.3). We begin with a lemma relating the connecting radius and the size of neighborhood in the random geometric graph.

Lemma 3.1. *For any constant $\theta' \in (0,1)$, let $r_{\theta'}$ be such that $\mathbb{P}_{u,v\sim\rho^{\otimes 2}}[\mathrm{d}(u,v) \le r_{\theta'}] = n^{-1+\theta'}$, then it holds that $r_{\theta'} = n^{\frac{-1+\theta'+o(1)}{d}}$.*

Proof. By the definition of $r_{\theta'}$ and the estimations given in Proposition 2.1 (together with the fact that $C_d/2 = n^{o(1)}, C_d/2ed = n^{o(1)}$ when $d \ll d_0$), we have

$$n^{\frac{-2+2\theta'+o(1)}{d}} \le r_\theta^2 - r_\theta^4/4 \le r_\theta^2 \le n^{-\frac{2+2\theta'+o(1)}{d}}\ ,$$

which implies $r_\theta = n^{\frac{-1+\theta'+o(1)}{d}}$, as desired. ☐

The core of the proof of (1.3) lies in the following proposition, which is largely inspired by [5].

Proposition 3.2. *For $d \le \lambda d_0$, let $u_1, \dots, u_n, v_1, \dots, v_n \overset{i.i.d.}{\sim} \rho$. Then with high probability, there is $\pi \in \mathrm{S}_n$ such that for except $o(n)$ many $i \in [n]$, it holds*

$$\mathrm{d}(u_i, v_{\pi(i)}) \le n^{-\frac{-1+\theta/3+o(1)}{d}} . \tag{3.1}$$

With Proposition 3.2 in hand, we may finish the proof of (1.3) as follows.

Proof of the lower bound (1.3). Recall the definition of $r_{\theta'}, \theta' \in (0,1)$ as in Lemma 3.1. Let $u_1, \dots, u_n \overset{\text{i.i.d.}}{\sim} \rho$, and let G_1 be the random geometric graph generated by forming edges between pairs (i,j) which satisfy $\mathrm{d}(u_i, u_j) \le r_\theta$. We note that $G_1 \sim \mathcal{G}$ by the definition of τ_* and $\mathcal{G} = \mathcal{G}(n, d, \tau_*)$. Moreover, let G_1' be the graph on $[n]$ constructed by forming edges between the pairs (i,j) such that $\mathrm{d}(u_i, u_j) \le r_\theta - r_{\theta/2}$. It is clear that G_1' is a subgraph of G_1. In addition, denoting $p_* = \mathbb{P}[\mathrm{d}(u,v) \le r_\theta] = n^{-1+\theta}$ and $p_{**} = \mathbb{P}[\mathrm{d}(u,v) \le r_\theta - r_{\theta/2}]$ (where $u, v \sim \rho^{\otimes 2}$) as the edge density of G_1 and G_1', respectively, then we note that

$$\begin{aligned} p_* - p_{**} =\ & \mathbb{P}[r_\theta - r_{\theta/2} < \mathrm{d}(u,v) \le r_*] \\ \overset{(2.1)}{=}\ & \int_{1-r_\theta^2/2}^{1-(r_\theta - r_{\theta/2})^2/2} C(1-x^2)^{\frac{d-2}{2}}\,\mathrm{d}x \le C_d r_\theta r_{\theta/2} (r_\theta^2 - r_\theta^4/4)^{\frac{d-2}{2}} \\ \overset{(2.2)}{\le}\ & 2ed \cdot \frac{r_{\theta/2}}{r_\theta} \cdot \mathbb{P}[\mathrm{d}(u,v) \le r_\theta] \overset{\text{Lemma 3.1}}{\le} dn^{-\frac{\theta/2+o(1)}{d}} p_* , \end{aligned}$$

which is $o(p_*)$ since $dn^{-\theta/2d} \ll 1$ by our assumption that $d \le \theta d_0/10$. This suggests that $p' = (1-o(1))p$. Therefore, both G_1 and G_1' has average degree $(1+o(1))np_*$.

Now let $v_1, \dots, v_n \overset{\text{i.i.d.}}{\sim} \rho$, and let $G_2 \sim \mathcal{G}$ be generated from $v_1, \dots, v_n$. Define $\mathcal{S}_1$ as the event that $u_1, \dots, u_n, v_1, \dots, v_n$ satisfies the condition in Proposition 3.2, and define $\mathcal{S}_2$ as the event that each vertex in G_1, G_1' and G_2 has degree $(1+o(1))np_*$. From Proposition 3.2 and Lemma 2.2 together with a union bound, $\mathbb{P}[\mathcal{S}_1 \cap \mathcal{S}_2] = 1 - o(1)$. On the event $\mathcal{S}_1 \cap \mathcal{S}_2$, let $\pi \in \mathrm{S}_n$ and $I \subset [n]$ be the witness of $\mathcal{S}_1$, i.e. $|I| = n - o(n)$ and $\mathrm{d}(u_i, v_{\pi(i)}) \le n^{\frac{-1+\theta/3+o(1)}{d}}$. Note that for $i, j \in I$ and $(i,j) \in E(G_1')$, it holds that (recall that $r_{\theta'} = n^{\frac{-1+\theta'+o(1)}{d}}$ for $\theta' \in (0,1)$ by Lemma 3.1)

$$\mathrm{d}(v_{\pi(i)}, v_{\pi(j)}) \le \mathrm{d}(u_i, v_{\pi(i)}) + \mathrm{d}(u_i, u_j) + \mathrm{d}(u_j, v_{\pi(j)}) \le r_\theta - r_{\theta/2} + 2n^{\frac{-1+\theta/3+o(1)}{d}} \le r_\theta ,$$

and thus $(\pi(i), \pi(j)) \in E(G_2)$. This implies that $O(G_1, G_2, \pi)$ is at least the number of edges in the induced subgraph of G_1' on I. Under the event $\mathcal{S}_2$, we have the induced subgraph has at least

$$|E(G_1')| - (1+o(1))np_* |[n] \setminus I| \ge (1/2 - o(1))n^{1+\theta} = (1+o(1))|E(G_1)| ,$$

which implies that

$$\Lambda(G_1, G_2) \ge \frac{|O(G_1, G_2, \pi)}{|E(G_1)|} \ge 1 - o(1) ,$$

and thus $\mathbb{P}[\Lambda(G_1, G_2) \ge 1 - o(1)] \ge \mathbb{P}[\mathcal{S}_1 \cap \mathcal{S}_2] = 1 - o(1)$, as desired. □

Now we turn to the proof of Proposition 3.2. For a set of some sets $A_1, \dots, A_N \subset [n]$, we say $(a_1, \dots, a_N)$ is a system of distinct representatives (SDR) of $(A_1, \dots, A_n)$, if $a_i \in A_i$ for any $1 \le i \le N$ and $a_1, \dots, a_N$ are distinct. Recall Hall's theorem for the existence of SDR: $(A_1, , \dots, A_N)$ has a SDR if and only if for any $1 \le k \le N$ and any $1 \le i_1 < \cdots < i_k \le N$, it holds that $|A_{i_1} \cup \cdots \cup A_{i_N}| \ge k$.

Proof of Proposition 3.2. Let $u_1, \ldots, u_n, v_1, \ldots, v_n \overset{\text{i.i.d.}}{\sim} \rho$ and $N = \lfloor n - n/\log n \rfloor = (1 - o(1))N$. For each $1 \le i \le N$, we define

$$A_i = \{j \in [n] : \mathrm{d}(u_i, v_j) \le r_{\theta/3}\} ,$$

and for each $1 \le j \le N$, we define $N_j = |\{i : j \in A_i\}|$. It is clear that for each $1 \le i \le N$, $|A_i| \sim \mathbf{B}(n, n^{-1+\theta/3})$ and for each $1 \le j \le n$, $N_j \sim \mathbf{B}(N, n^{-1+\theta/3})$. Then it follows from Lemma 2.2 together with a union bound that, with high probability,

$$|A_i| \ge n^{\theta/3} - n^{\theta/6} \log n, \forall 1 \le i \le N ,$$

and similarly (note that $N_j \sim \mathbf{B}(N, p)$)

$$N_j \le N n^{-1+\theta/3} + n^{\theta/6} \log n \le n^{\theta/3} - n^{\theta/3}/\log n + n^{\theta/6} \log n \, , \forall 1 \le j \le n \, .$$

Therefore, it holds with high probability that

$$\min_{1 \le i \le N} |A_i| \ge \max_{1 \le j \le n} N_j \, . \tag{3.2}$$

Now we verify Hall's condition under (3.2): for any $1 \le k \le N$ and any $1 \le i_1 < \cdots < i_k \le N$, let $B = A_{i_1} \cup \cdots \cup A_{i_k}$. We have that

$$k \min_{1 \le i \le N} |A_i| \le |\{(i,j) : i \in \{i_1, \ldots, i_k\}, j \in [n], \mathrm{d}(u_i, v_j) \le r_{\theta/3}\} \le B \max_{1 \le j \le n} N_j \, ,$$

and thus $|B| \ge k$ by (3.2). This verifies the Hall's condition and hence $(A_1, \ldots, A_d)$ has a SDR $(j_1, \ldots, j_N)$. Let $\pi \in \mathrm{S}_n$ satisfies $\pi(i) = j_i, 1 \le i \le N$, then by our definition of A_i we see $\mathrm{d}(u_i, v_{\pi(i)}) \le r_{\theta/3}$ for any $1 \le i \le N = (1 - o(1))n$. Combined with Lemma 3.1, this completes the proof. □

4 The upper bound

This section devotes for proving the upper bound (1.4). We fix some $d \gg \log n$ together with an arbitrary constant $\delta \in (0, 1)$, then it suffices to show that

$$\mathbb{P}[\Lambda(G_1, G_2) \ge 2\delta] \le \delta + o(1) \, . \tag{4.1}$$

Since given this is true, we can get the desired result by sending $\delta \to 0$.

4.1 Random sparsification

The starting point is to try to show that for each fixed $\pi \in \mathrm{S}_n$, $\mathbb{P}[\mathrm{O}(G_1, G_2, \pi) \ge \delta n^{1+\theta}] = o(1/n!)$. Once this is true, the desired result follows by a union bound over $\pi \in \mathrm{S}_n$. Unfortunately, the distribution of $\mathrm{O}(G_1, G_2, \pi)$ is far from clear, and even after some relaxation it is still difficult to analysis. This is basically due to denseness of G_1, G_2, which causes complicated correlations inside $\mathrm{O}(G_1, G_2.\pi)$. On the other hand, it turns out that the sparse random geometric graphs (where the average degree is $O(1)$) are much more tractable. For example, many tools in the sparse regime have been developed in [6], and this note is also hugely inspired by this seminal work. In light of this, our first step is to do a random "sparsification" of G_1.

Definition 4.1. *Fix a constant M such that $M > 100(1-\theta)/\delta^2$ and let $s = Mn^{-\theta}$. Let $G_1 \sim \mathcal{G}(d, \tau_*)$, sample a random subgraph $H = H(G_1, s)$ of G_1 by keeping each edge of G_1 in H with probability s, independently. Note that H has average degree $M = O(1)$ by our choice.*

After the sparsifying G_1 to H, we claim that (4.1) reduces to show that

$$\mathbb{P}[\Lambda(H, G_1) \geq 2\delta + o(1)] = \delta + o(1), , \tag{4.2}$$

where $\Lambda(H, G_2) = \frac{\max_{\pi \in \mathrm{S}_n} \mathrm{O}(H, G_2, \pi)}{|E(H)| \wedge |E(G_2)|}$ as in Definition 1.2. This is because that we have with high probability, $|E(G_1)| \wedge |E(G_2)| = (1 + o(1))n^{1+\theta}/2$ and $|E(H)| \wedge |E(G_2)| = (1 + o(1))Mn/2$, which implies that

$$\mathbb{P}[\Lambda(G_1, G_2) \geq 2\delta] \leq \mathbb{P}[\max_{\pi \in \mathrm{S}_n} \mathrm{O}(G_1, G_2, \pi) \geq (1 + o(1))n^{1+\theta/2}/2] + o(1)\,,$$

and

$$\mathbb{P}[\Lambda(H, G_2) \geq 2\delta + o(1)] \geq \mathbb{P}[\max_{\pi \in \mathrm{S}_n} \mathrm{O}(H, G_2, \pi)] - o(1)\,.$$

In addition, for any (G_1, G_2) such that $\max_{\pi \in \mathrm{S}_n} \mathrm{O}(G_1, G_2, \pi) = \mathrm{O}(G_1, G_2, \pi^*) \geq (1 + o(1))n^{1+\theta}/2$, we have that

$$\max_{\pi \in \mathrm{S}_n} \mathrm{O}(H, G_2, \pi) \geq \mathrm{O}(H, G_2, \pi^*) \sim \mathbf{B}(\mathrm{O}(G_1, G_2, \pi^*), s)\,,$$

which is no less than $(1 + o(1))sn^{1+\theta}/2 = (1 + o(1))Mn/2$ with high probability. Therefore,

$$\begin{aligned}
\mathbb{P}[\Lambda(G_1, G_2) \geq 2\delta] &\leq \mathbb{P}[\max_{\pi \in \mathrm{S}_n} \mathrm{O}(G_1, G_2, \pi) \geq (1 + o(1))n^{1+\theta}/2] + o(1) \\
&\leq (1 + o(1))\mathbb{P}[\max_{\pi \in \mathrm{S}_n} O(H, G_2, \pi) \geq (1 + o(1))Mn/2] + o(1) \\
&\leq (1 + o(1))\mathbb{P}[\Lambda(H, G_2) \geq 2\delta + o(1)] + o(1)\,,
\end{aligned}$$

verifying the claim.

4.2 Truncations

Now we turn to the proof of (4.2). To this end, we need to introduce some appropriate truncations on both H and G_2. We begin by the desiring properties we would like for H.

Definition 4.2. *We say a vertex $v \in [n]$ is good in H, if the following hold:*

- *The 3-neighborhood of v is a tree;*
- *For any u in the 3-neighborhood of v, the degree of v in H is at most $2M$.*

Otherwise v is said to be bad. Moreover, we define $\mathcal{H}$ for the event that $|E(H)| = (1 + o(1))Mn/2$, and the sum of degrees of bad vertices is no more than $\delta Mn/3$.

By standard facts on sparse random graphs, we can conclude the following lemma, for which we omit the proof here.

Lemma 4.3. *For $H = H(s, G_1)$ defined as in Definition 4.1, $\mathcal{H}$ happens with probability $1 - o(1)$.*

We then turn to a slight modification of the random geometric graph model $\mathcal{G}$.

Definition 4.4. *Denote $\mu = \rho^{\otimes n}$, and for $(u_1, \ldots, u_n) \sim \mu$, let $\mathcal{U}$ be the event that $|u_i \cdot u_j| \leq \gamma$ for all $1 \leq i \neq j \leq n$, where $\gamma = (\log n/d)^{1/4}$. Furthermore, let $\mathcal{G}'$ be the law of the graph constructed by forming edges between pairs (i, j) such that $u_i \cdot u_j \geq \tau_*$ where $(u_1, \ldots, u_n)$ is sampled from $\mu[\cdot \mid \mathcal{U}]$.*

We remark that such a truncation will facilitate our analysis on the joint probability of the form $\mathbb{P}_{v\sim\rho}[v \cdot u_{i_1} \geq \tau_*, \ldots, v \cdot u_{i_K} \geq \tau_*]$, see Lemma 4.9 for details. The following lemma shows that $\mathcal{G}$ and $\mathcal{G}'$ are indeed very close to each other.

Lemma 4.5. *It holds that* $\mu[\mathcal{U}] = 1 - n^{-\omega(1)}$ *and* $\mathrm{TV}(\mathcal{G}, \mathcal{G}') = 1 - n^{-\omega(1)}$.

Proof. From (2.1) we have for any $i \neq j$,

$$\mu[|u_i \cdot u_j| \geq \gamma] \leq 2C_d\big(1 - \sqrt{\log n/d}\big)^{\frac{d-2}{2}} \leq \exp\Big(-\Omega\big(\sqrt{d \log n}\big)\Big)\,,$$

which is $n^{-\omega(1)}$ since $d \gg \log n$. Then we conclude that $\mu[\mathcal{U}^c] \leq n^2 \cdot n^{-\omega(1)} = n^{-\omega(1)}$ from a union bound. This proves that $\mu[\mathcal{U}] = 1 - n^{-\omega(1)}$, and thus by the data processing inequality, $\mathrm{TV}(\mathcal{G}, \mathcal{G}') \leq \mathrm{TV}(\mu, \mu[\cdot \mid \mathcal{U}]) = n^{-\omega(1)}$, as desired. $\square$

Now it is straightforward to see that

$$\mathbb{P}[\Lambda(H, G_2) \leq 2\delta + o(1)] \leq \mathbb{P}[H \notin \mathcal{H}] + \mathbb{P}[H \in \mathcal{H}, \Lambda(H, G') \leq 2\delta + o(1)] + \mathrm{TV}(\mathcal{G}, \mathcal{G}')\,,$$

where in the second term G' is sampled from $\mathcal{G}'$. In light of Lemma 4.3 and Lemma 4.5, (4.2) reduces to prove that for any $H \in \mathcal{H}$, we have $\mathbb{P}[\Lambda(H, G) \leq 2\delta + o(1)] \leq \delta + o(1)$, where the probability is taken over the random graph $G \sim \mathcal{G}'$. In addition, since $|E(H)| = (1 + o(1))Mn/2$ under $\mathcal{H}$, we see it remains to show that for any $H \in \mathcal{H}$,

$$\mathbb{P}_{G \sim \mathcal{G}'}[\max_{\pi \in \mathrm{S}_n} \mathrm{O}(H, G, \pi) \leq (1 + o(1))\delta Mn] \leq \delta + o(1)\,. \tag{4.3}$$

4.3 Domination by binomial variable

Henceforth we fix a graph $H \in \mathcal{H}$ and show that (4.3) is true. For any $\pi \in \mathrm{S}_n$ and $1 \leq k \leq n$, we denote $\mathcal{F}_{k-1}^\pi$ for the σ-field generated by the random variables $G_{\pi(i),\pi(j)}$ where $(i, j) \in E(H)$ and $i < j < k$. We write

$$\mathrm{O}(H, G_2, \pi) = \sum_{(i,j) \in E(H)} G_{\pi(i),\pi(j)} = \sum_{k=1}^{n} O_k(\pi)\,,$$

where $O_k(\pi) = \sum_{j<k,(j,k)\in E(H)} G_{\pi(j),\pi(k)}$, then it is clear that $O_j, 1 \leq j \leq k$ are all measurable with respect to $\mathcal{F}_k^\pi$. The key ingredient we will need for controlling $\Lambda(H, G) = \max_{\pi \in \mathrm{S}_n} \mathrm{O}(H, G, \pi)$ is the following stochastic domination relation.

Proposition 4.6. *For any $\pi \in \mathrm{S}_n$, any $1 \leq k \leq n$ which is a good vertex in H, and any realization of $\mathcal{F}_{k-1}^\pi$, it holds that the random variable O_k conditioned on $\mathcal{F}_{k-1}^\pi$ is stochastically dominated by $\delta M/3 + 2MI_k$, where $I_k \sim \mathbf{B}(1, n^{-10/\delta})$ is a Bernoulli indicator independent with $\mathcal{F}_{k-1}^\pi$.*

Proof of (4.3) *assuming Proposition 4.6.* For each fixed $\pi \in \mathrm{S}_n$, it is clear that

$$\sum_{k \text{ is bad in } H} O_k(\pi) \leq \sum_{k \text{ is bad in } H} d_H(k) \leq \delta Mn/3$$

by the condition $\mathcal{H}$. In addition, we conclude from Proposition 4.6 that

$$\sum_{k \text{ is good in } H} O_k(\pi)$$

is stochastically dominated by $\delta Mn/3 + 2MX$, where $X \sim \mathbf{B}(n, n^{-10/\delta})$ is a binomial variable. From the Chernoff bound for binomial variables, we conclude that

$$\mathbb{P}[\mathrm{O}(H, G, \pi) \geq (1 + o(1))\delta Mn] \leq \mathbb{P}[X \geq \delta n/8] \leq \exp\left(-\frac{\delta n}{8} \cdot \left(\log\left(\frac{\delta n/8}{n^{1-\delta/10}}\right) - 1\right)\right)\,,$$

which is upper-bounded by $\exp(-n \log n) \ll 1/n!$. The desired result then follows by taking a union bound over $\pi \in \mathrm{S}_n$. $\square$

The remaining of this section dedicates to proving Proposition 4.6. Without loss of generality, we may assume that $\pi = \mathrm{id}$, and we simply write O_k for $\mathrm{O}_k(\mathrm{id})$ and $\mathcal{F}_{k-1}$ for $\mathcal{F}^{\mathrm{id}}_{k-1}$, $1 \le k \le n$. We fix some $1 \le k \le n$ which is good in H as well as some particular realization of $\mathcal{F}_{k-1}$, and we define

$$\mathcal{C} = \{(i,j) \in E(H) : i < j < k, G_{i,j} = 1\}\,, \quad \mathcal{D} = \{(i,j) \in E(H) : i < j < k, G_{i,j} = 0\}\,.$$

Assume that G is generated from $u_1, \dots, u_n \in \mathbb{S}^d$. Then the joint distribution of $u_1, \dots, u_n$ conditioned on $\mathcal{F}_{k-1}$ is given by

$$\mu_k := \mu[\cdot \mid u_i \cdot u_j \ge \tau_*, \forall (i,j) \in \mathcal{C}, u_i \cdot u_j < \tau_*, \forall (i,j) \in \mathcal{D}, \mathcal{U}]\,, \tag{4.4}$$

where $\mathcal{U}$ denotes the event that $|u_i \cdot u_j| \le \gamma$ for any $i \ne j$ as in Definition 4.4. Let H_k be the induced subgraph of H on $[k]$, and let $i_1, \dots, i_D$ be the neighbors of k in H_k. Then, we note that the distribution of O_k given $\mathcal{F}_{k-1}$ is the distribution of

$$\sum_{t=1}^{D} \mathbf{1}(u_{i_t} \cdot u_k \ge \tau_*)$$

under μ_k. At this point, while μ_k is a measure conditioned on an event with sparse correlation structure, the joint probability on all coordinates $(u_1, \dots, u_n)$ is still difficult to analyze. To tackle this, we will further fix the positions of u_i for vertices i that are far from k in H, and restrict the correlation to a local neighborhood of k.

We denote T for the 3-neighborhood of k in H_k, then T is a tree rooted at k with maximal degree no more than $2M$ by the assumption that $H \in \mathcal{H}$. Moreover, for $1 \le t \le D$, we let T_t be the subtree of T rooted at i_t, and let $N(T_t)$ be the set of non-leaf vertices in T_t. In addition, we define

$$\mathcal{C}_t = \{(i,j) \in E(T_t) : G_{i,j} = 1\}\,, \quad \mathcal{D}_t = \{(i,j) \in E(T_t) : G_{i,j} = 0\}\,,$$

and we say a realization of $(u_i)_{i \notin N(T_t)}$ is compatible with $\mathcal{F}_{k-1}$, if the conditioning in (4.4) is not violated by $(u_i)_{i \notin N(T_t)}$. With these notations, we see that the joint law on $(u_i)_{i \in N(T_t)}$ of μ_k conditioned on a compatible realization $(u_i)_{i \notin N(T_t)}$ can be expressed as

$$\widetilde{\mu}_t := \widetilde{\mu}\big[\cdot \mid u_i \cdot u_j \ge \tau_*, \forall (i,j) \in \mathcal{C}_t, u_i \cdot u_j < \tau_*, \forall (i,j) \in \mathcal{D}_t, |u_i \cdot u_j| \le \gamma, \forall i \in N(T_t), j \in [n]\big]\,, \tag{4.5}$$

where $\widetilde{\mu}$ is the product measure of uniform distributions on $(u_i)_{i \in N(T_t)}$.

For a probability measure ν on $\mathbb{S}^d$ with probability density $\nu(x)$ with respect to ρ, we define

$$F(\nu) = \sup_{x \in \mathbb{S}^d} \mathbb{P}_{y \sim \nu}[x \cdot y \ge \tau_*] = \sup_{x \in \mathbb{S}^d} \int_{y \cdot x \ge \tau_*} \nu(y)\, \mathrm{d}\rho(y)\,. \tag{4.6}$$

We claim that it suffices to show the following control on $F(\widetilde{\nu}_t)$ for $\widetilde{\nu}_t$ be the marginal of u_{i_t} under $\widetilde{\mu}_t$.

Proposition 4.7. *For any realization of $\mathcal{F}_{k-1}$, any $1 \le t \le D$, and any compatible specification of $(u_i)_{i \notin N(T_t)}$, let $\widetilde{\mu}_t$ be defined as in (4.5) and let $\widetilde{\nu}_t$ be the marginal of u_{i_t} under $\widetilde{\mu}_t$, then $F(\widetilde{\nu}_t) \le n^{o(1)} p$.*

Proof of Proposition 4.6 assuming Proposition 4.7. For any distinct $t_1, \dots, t_k \in \{i_1, \dots, i_D\}$, it follows from the total probability formula that

$$\mu_k[u_{t_j} \cdot u_k \ge \tau_* \mid u_{t_l} \cdot u_k \ge \tau_*, l < j] = \mathbb{E}_{(u_i)_{i \notin N(T_{t_j})}} \widetilde{\nu}_{t_j}[u_{t_j} \cdot u_k \ge \tau_*] \le \mathbb{E}_{(u_i)_{i \notin N(T_{t_j})}} F(\widetilde{\nu}_{t_j})\,,$$

where the expectation over $(u_i)_{i \notin N(T_{t_j})}$ is taken under the conditional measure $\mu_k[\cdot \mid u_{t_l} \cdot u_k \ge \tau_*, l < j]$, and thus the realizations of $(u_i)_{i \notin N(T_{t_j})}$ are almost surely compatible. By Proposition 4.7, we see $F(\widetilde{\nu}_t) \le n^{o(1)} p = n^{-1+\theta+o(1)}$ almost surely and thus

$$\mu_k[u_{t_j} \cdot u_k \ge \tau_* \mid u_{t_l} \cdot u_k \ge \tau_*, l < j] \le n^{-1+\theta+o(1)}\,.$$

Therefore, it follows from the multiplicative rule that

$$\mu_k[u_{t_j} \cdot u_k \geq \tau_*, 1 \leq j \leq k] \leq (n^{-1+\theta+o(1)})^k \,.$$

Given with this, we see from a union bound that

$$\widetilde{\mu}_k[O_k \geq \delta M/3] \leq \binom{2M}{\delta M/3}(n^{-1+\theta+o(1)})^{\delta M/3} \leq n^{-10/\delta} \,,$$

where the second inequality follows since M is chosen as a constant greater than $100(1-\theta)/\delta^2$. In addition, it trivially holds that $O_k \leq 2M$, and thus $O_k \leq \delta M/3 + 2MX$ for $X \sim \mathbf{B}(1, n^{-10/\delta})$. $\square$

4.4 Compute the marginal via belief propagation

Now we are left with Proposition 4.7. We fix a realization of $\mathcal{F}_{k-1}$ and $1t \in \{1, \ldots, D\}$, as well as a compatible realization $(u_i)_{i \notin N(T_t)}$. For simplicity, we denote

$$\mathcal{A}_t = \{u_i \cdot u_j \geq \tau_*, \forall (i,j) \in \mathcal{C}_t\}, \mathcal{B}_t = \{u_i \cdot u_j < \tau_*\}, \mathcal{U}_t = \{u_i \cdot u_j \geq -\gamma, \forall i \in N(T_i), j \in [n]\} \,.$$

Intuitively, the main impact from the conditioning comes from $\mathcal{A}_t$ since both $\mathcal{B}_t$ and $\mathcal{U}_t$ are typical events under $\widetilde{\mu}$, and indeed this will be made rigorous in Lemma 4.10 below. At this point, let us first focus on the measure $\widehat{\mu}_t[\cdot] = \widetilde{\mu}[\cdot \mid \mathcal{A}_t]$, which is the bulk part of the analysis. Let $\widehat{\nu}_t$ be its marginal on u_{i_t}, we first derive a nice control on the density $\widehat{\nu}(x)$ of $\widehat{\nu}_t$ with respect to ρ.

Let $F_t \subset T_t$ be the subgraph generated by edges $(i,j) \in \mathcal{C}_t$. Then F_t is a forest, and we let $T_t' \subset F_t$ be the component in F_t that contains i_t. Crucially, we note $\widehat{\mu}_t$ is the uniform measure on the solution space for a constraint-satisfying problem on $(\mathbb{S}^d)^{N(T_t)}$ given by $u_i \cdot u_j \geq \tau_*, \forall (i,j) \in E(F_t)$. Such a CSP has factor graph F_t and so $\widehat{\mu}_t$ decomposes into products of marginals over components of F_t. Therefore, $\widehat{\nu}_t$ is just the marginal of the uniform measure over the solution space of the CSP given by $u_i \cdot u_j \geq \tau_*, \forall (i,j) \in E(T_t')$. In addition, since T_t' is a tree, we have that $\widehat{\nu}_t$ can be explicitly computed by the belief propagation algorithm. Since here T_t' has depth at most 2, so we can write down the explicit expression for the belief propagation.

For clearance, we redefine the notations as follows: let T_t' have root $\emptyset$, and $\emptyset$ has sons $1, \ldots, D$, and i has sons $(i,1), \ldots, (i, D_i)$ for $1 \leq i \leq D$. Under this context, we fix $u_{i,j} \in \mathbb{S}^d, 1 \leq i \leq D, 1 \leq j \leq D_i$ such that $|u_{i,j} \cdot u_{i',j'}| \leq \gamma$ for all $(i,j) \neq (i',j')$ and $u_\emptyset, u_i, 1 \leq i \leq D$ are flexible. We restate our goal as follows:

Proposition 4.8. *Given that*

$$\widehat{\mu}[\cdot] = \widetilde{\mu}[\cdot \mid u_\emptyset \cdot u_i \geq \tau_*, 1 \leq i \leq D, u_i \cdot u_{(i,j)} \geq \tau_*, 1 \leq i \leq D, 1 \leq j \leq D_i]$$

where $\widetilde{\mu}$ is the product measure of uniform distributions of $u_\emptyset, u_i, 1 \leq i \leq D$, then the probability density $\widehat{\nu}_\emptyset(x)$ of the marginal of $u_\emptyset$ under $\widehat{\mu}$ satisfies that $F(\widehat{\nu}_t) \leq n^{o(1)}p$.

We now run belief propagation algorithm on T_t' to compute $\widehat{\nu}_t$ as follows:

Step 1 For $1 \leq i \leq D$ and $1 \leq j \leq D_i$, define $\nu_{(i,j)\to i}(x) = p^{-1}\mathbf{1}(x \cdot u_{i,j} \geq \tau_*), \forall x \in \mathbb{S}^d$.

Step 2 For $1 \leq i \leq D$, let $\widehat{\nu}_i(x) \propto \prod_{j=1}^{D_i} \nu_{(i,j)\to i}(x), \forall x \in \mathbb{S}^d$, such that $\widetilde{\nu}_i$ is a probability density on S^d.

Step 3 For $1 \leq i \leq D$, let $\nu_{i\to\emptyset}(x) = p^{-1}\mathbb{P}_{y\sim\widehat{\nu}_i}[x \cdot y \geq \tau_*], \forall x \in \mathbb{S}^d$.

Step 4 Let $\widehat{\nu}_\emptyset(x) \propto \prod_{i=1}^{D} \nu_{i\to\emptyset}(x), \forall x \in \mathbb{S}^d$, such that ν is a probability density on $\mathbb{S}^d$.

Then it follows that the output of Step 4 is just the probability density of the marginal distribution of $u_\emptyset$, see [6, Section 7] for more details.

We now look more carefully into each step of the belief propagation. Step 1 is straightforward. Plugging the expression in Step 1 into Step 2, we see that for each $1 \le i \le D$,

$$\widehat{\nu}_i(x) \propto \mathbf{1}(x \cdot u_{i,j} \ge \tau_*, \forall 1 \le j \le D_i)\,,$$

and so

$$\widehat{\nu}_i(x) = \frac{\mathbf{1}(x \cdot u_{i,j} \ge \tau_*, \forall 1 \le j \le D_i)}{\mathbb{P}_{u\sim\rho}[u \cdot u_{i,j} \ge \tau_*, 1 \le j \le D_i]}\,.$$

Therefore, in Step 3 we have that for each $1 \le i \le D$,

$$\nu_{i\to\emptyset}(x) = \mathbb{P}_{y\sim\widehat{\nu}_i}[x \cdot y \ge \tau_*] = \frac{\mathbb{P}_{u\sim\rho}[x \cdot u \ge \tau_*, u \cdot u_{i,j} \ge \tau_*, \forall 1 \le j \le D_i]}{p \cdot \mathbb{P}_{u\sim\rho}[u \cdot u_{i,j} \ge \tau_*, \forall 1 \le j \le D_i]}\,. \tag{4.7}$$

And finally, we have

$$\widehat{\nu}_\emptyset(x) = \frac{\prod_{i=1}^D \nu_{i\to\emptyset}(x)}{\int_{\mathbb{S}^d} \prod_{i=1}^D \nu_{i\to\emptyset}(y)\,\mathrm{d}\rho(y)}\,. \tag{4.8}$$

We have the following simple but useful estimate.

Lemma 4.9. *For any $K = O(1)$ and $v_1, \dots, v_k \in \mathbb{S}^d$ with $|v_i \cdot v_j| \le o(1), \forall 1 \le i \ne j \le K$, it holds that*

$$\mathbb{P}_{u\sim\rho}[u \cdot v_j \ge \tau_*, \forall 1 \le j \le K] = n^{o(1)} p^K\,.$$

Proof. By rotation symmetry and the condition, we may assume that

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_K \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ o(1) & 1-o(1) & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & 0 \\ o(1) & o(1) & o(1) & \cdots & 1-o(1) & \cdots & 0 \end{pmatrix}$$

First, recall $\gamma = (\log n/d)^{1/4} = o(1)$, from a simple union bound it holds that

$$\mathbb{P}_{u\sim\rho}[u \cdot u_j \ge \tau_*, \forall 1 \le j \le K] = \mathbb{P}_{u\sim\rho}[\tau_* \le u \cdot v_j \le \gamma, \forall 1 \le j \le K] + n^{-\omega(1)}\,.$$

Now let $u = (u_1, \dots, u_{d+1}) \sim \rho$, we note that conditioned on the first m coordinates, $(u_{m+1}, \dots, u_{d+1})$ is just uniformly chosen from the $(d-m)$-dimensional sphere with radius $(1 - u_1^2 - \cdots - u_m^2)^{1/2}$. With the observation in mind, it can be easily deduced from 2.1 that for each $0 \le m \le K-1$, we have

$$\mathbb{P}_{u\sim\rho}[\tau_* \le u \cdot v_{m+1} \le \gamma \mid \tau_* \le u \cdot v_i \le \gamma, i \le m] = n^{o(1)} p\,,$$

and hence the result follows from the multiplicative rule. □

Now we can easily prove Proposition 4.8.

Proof of Proposition 4.8. From our choice of $u_{i,j}, 1 \le i \le D, 1 \le j \le D_i$, we have that $u_{i,j} \cdot u_{i,j'}| \le \gamma = o(1)$ for all $i, j \ne j'$. Therefore, we have the denominator in (4.7) is $n^{o(1)} p^{D_i+1}$. In addition, we define

$$B = \{x \in \mathbb{S}^d : |x \cdot u_{i,j}| \le \gamma, \forall 1 \le i \le D, 1 \le j \le D_i\}\,,$$

then it is clear that $\rho(B) = 1 - n^{-\omega(1)}$ by a union bound. Be lemma 4.9 we have for $x \in B$, the numerator in (4.7) is also $n^{o(1)} p^{D_i+1}$. Hence for $x \in B$ it holds $\nu_{i\to\emptyset}(x) = n^{o(1)}$. For $x \in \mathbb{S}^d \setminus N$, it trivially holds that $\nu_{i\to\emptyset}(x) \le p^{-1} \le n$. Given this, it follows that

$$F(\widehat{\nu}_\emptyset) = \sup_{x\in\mathbb{S}^d} \int_{y\cdot x\ge\tau_*} \widetilde{\nu}_\emptyset(y)\,\mathrm{d}\rho(y) \le p \cdot n^{o(1)} + n^{-\omega(1)} \cdot n = p n^{o(1)}\,. \qquad \square$$

Finally, we relate $\widehat{\mu}_t = \widetilde{\mu}[\cdot \mid \mathcal{A}_t]$ with $\widetilde{\mu}_t = \widetilde{\mu}[\cdot \mid \mathcal{A}_t, \mathcal{B}_t, \mathcal{U}_t]$, the probability measure we really care.

Lemma 4.10. *It holds that $\widetilde{\mu}[\mathcal{B}_t, \mathcal{U}_t \mid \mathcal{A}_t] \geq 1 - o(1)$, and hence $\widetilde{\mu}_t[\cdot] \leq (1 + o(1))\widehat{\mu}_t[\cdot]$.*

Proof. We only sketch the proof of the first claim. It suffices to show that $\widetilde{\mu}[\mathcal{B}_t^c \mid \mathcal{A}_t]$ and $\widetilde{\mu}[\mathcal{U}_t^c \mid \mathcal{A}_t]$ are both $o(1)$.

For the former one, we show that for any $(i, j) \in \mathcal{D}_t$, $\widetilde{\mu}[u_i \cdot u_j \geq \tau_* \mid \mathcal{A}_t]$. If either of i, j equals to i_t, then the result just follows from $F(\widehat{\nu}_t) \leq n^{o(1)}p$ as shown in Proposition 4.8. Otherwise, one of u_i, u_j is fixed, say u_j. It can be shown by analyzing a similar belief propagation that this reduces to some probability of the form

$$\frac{\mathbb{P}[u_i \cdot u_j \geq \tau_*, u_i \cdot u_{j_t} \geq \tau_*, 1 \leq t \leq K]}{\mathbb{P}[u_i \cdot u_{j_t} \geq \tau_*, \forall 1 \leq t \leq K]},$$

which can be done via Lemma 4.9.

For the latter one, we note that $\widetilde{\mu}[\mathcal{U}_t^c \mid \mathcal{A}_t] \leq \widetilde{\mu}[\mathcal{U}_t^c]/\widetilde{\mu}[\mathcal{A}_t]$, and $\widetilde{\mu}[\mathcal{U}_t^c] \leq n^{-\omega(1)}$ by a union bound. It can be shown via Lemma 4.9 that $\widetilde{\mu}[\mathcal{A}_t] \geq n^{O(1)}$, and thus the result follows.

Now given with the first claim, we have

$$\widetilde{\mu}_t[\cdot] = \frac{\widetilde{\mu}[\cdot, \mathcal{A}_t, \mathcal{B}_t, \mathcal{U}_t]}{\widetilde{\mu}[\mathcal{A}_t, \mathcal{B}_t, \mathcal{U}_t]} \leq \frac{\widetilde{\mu}[\cdot, \mathcal{A}_t]}{\widetilde{\mu}[\mathcal{A}_t]\widetilde{\mu}[\mathcal{B}_t, \mathcal{U}_t \mid \mathcal{A}_t]} = \frac{\widetilde{\mu}[\cdot \mid \mathcal{A}_t]}{\widetilde{\mu}[\mathcal{B}_t, \mathcal{U}_t \mid \mathcal{A}_t]} = (1 + o(1))\widehat{\mu}[\cdot]\,. \qquad \square$$

Proof of Proposition 4.7. It follows from the lemma that $F(\widetilde{\nu}_t) \leq (1 + o(1))F(\widehat{\nu}_t)$, which is $n^{o(1)}p$ as shown in Proposition 4.8. This proves Proposition 4.7 and thus finishes the proof of (1.4). $\square$

5 Further discussions

5.1 The power of rotation is futile

In this section, we discuss a potential approach for improving the lower bound proved in Section 3. Following the idea of the proof in Section 3, we note that there is some additional power of rotations on $\mathbb{S}^d$ which we did not exploit. Indeed, once we can prove that for typical G_1 and G_2 generated from $u_1, \ldots, u_n \in \mathbb{S}^d$ and $v_1, \ldots, v_n \in \mathbb{S}^d$, there exists $T \in \mathrm{O}(d)$ (where $\mathrm{O}(d)$ denotes the d-dimensional orthogonal group, the group of isometries on $\mathbb{S}^d$), such that for most $i \in [n]$, $\mathrm{d}(Tu_i, v_{\pi(i)}) \leq r$ for some appropriate r, then $\Lambda(G_1, G_2) = 1 - o(1)$. One may wonder that with the additional freedom of rotations, can we substantially improve the previous lower bound? This section suggests that the answer to this is no, the power of rotations does not help us break the conjectured threshold d_0.

Assume $d_0 \ll d \ll \log n$. Following the arguments in the proof of (1.3) in Section 3, in order to prove $\Lambda(G_1, G_2) = 1 - o(1)$ with high probability, we need to show that for $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ uniformly and independently drawn from $\mathbb{S}^d$, typically there exists $T \in \mathrm{O}(d)$ and $\pi \in \mathrm{S}_n$ such that

$$\mathrm{d}(Tu_i, v_{\pi(i)}) = o(r_\theta/d)\,, \text{ for } n - o(n) \text{ many } i \in [n]\,. \tag{5.1}$$

However, we will show that typically this cannot happen, and this suggests the previous approach cannot be extended to any dimension higher than d_0.

Let ε be such that $\mathbb{P}[\mathrm{d}(u, v) \leq \varepsilon] = n^{-2}$, then it follows that $\varepsilon = n^{-(2+o(1))/n}$ by similar arguments as in Lemma 3.1, and so $\varepsilon \gg r_\theta/d$ by our choice of d. Moreover, we take a minimal $\varepsilon/2$-net $\mathcal{N}$ of $\mathrm{O}(d)$ with respect to the operator norm $\|\cdot\|_{\mathrm{op}}$, then from [8, Propsition 6]we have $|\mathcal{N}| \leq (c_0\varepsilon)^{-d^2/2}$ for some universal constant $c_0 > 0$. If there exists $T \in \mathrm{O}(d)$ and $\pi \in \mathrm{S}_n$ such that (5.1) holds, then there exists $T_0 \in \mathcal{N}$, such that $\mathrm{d}(T_0u_i, v_{\pi(i)}) \leq \varepsilon$ for at least half of $i \in [n]$. Denote $B_{T_0,i} = \{v \in \mathbb{S}^d, \mathrm{d}(T_0u_i, v) \leq \varepsilon\}$, then the aforementioned event implies that $B_{T_0,i} \cap \{v_1, \ldots, v_n\} \neq \emptyset$ for at least half of $i \in [n]$. It can be shown by a union bound that with high probability for $u_1,\dots,u_n$, $\mathrm{d}(u_i,u_j)>2\varepsilon,\forall i\neq j\in[n]$, and so $B_{T_0,i}, i=1,\dots,n$ are pairwise disjoint. We fix any choice of $u_1,\dots,u_n$ satisfying this condition, then for each fixed $T_0\in\mathcal{N}$, the events $\{B_{T_0,i}\cap\{v_1,\dots,v_n\}\neq\emptyset\}$ are negative-correlated with each other, and so we have

$$\begin{aligned}
&\mathbb{P}\big[\exists T\in \mathrm{O}(d),\pi\in \mathrm{S}_n \text{ s.t. } \mathrm{d}(T_0u_i,v_{\pi(i)})\leq r_\theta/d \text{ for at least half of } i\in[n]\big]\\
\leq&\sum_{T_0\in\mathcal{N}}\mathbb{P}\big[B_{T_0,i}\cap\{v_1,\dots,v_n\}\neq\emptyset \text{ for at least half of } i\in[n]\big]\\
\leq&(c_0\varepsilon)^{-d^2/2}\cdot\binom{n}{n/2}\cdot(2n^{-2})^{n/2}=n^{O(\log n)-\Omega(n)}=o(1)\,,
\end{aligned}$$

where in the last inequality, we first employ a union bound on the half of $i\in[n]$ satisfying $B_{T_0,i}\cap\{v_1,\dots,v_n\}\neq\emptyset$, then use the fact that these events are negative-correlated together with a simple estimate $\mathbb{P}[B_{T_0,i}\cap\{v_1,\dots,v_n\}\neq\emptyset]\leq 2n^{-2}$ followed by Poisson approximation. This proves that the additional power of rotations does not essentially improve the previous lower bound.

5.2 Breaking the $\log n$ upper bound

Finally, we discuss how it is possible to improve our proof for upper bound to some case down to $\log n$. The key ingredient for our proof in Section 4 is the analysis of a belief propagation algorithm on a tree with depth 2 (see Proposition 4.8). Here the fundamental reason that a choice of depth 2 suffices is that the connecting threshold $\tau_*=o(1)$ when $d\gg\log n$, which implies the random walk on $\mathbb{S}^d$ with transition kernal $P(x,y)=p_*^{-1}\mathbf{1}(x\cdot y\geq\tau_*)$ mixes very well in 2 or 3 steps.

For smaller d, we have $\tau_*=\Omega(1)$ or even $\tau_*=1-o(1)$, then it is necessary to consider trees with large depth in order for analogues of Proposition 4.8. Intuitively, the depth D_* should satisfies that given the position of all the descendants in D_*-generations, the posterior distribution of the root is somehow close to the uniform. As a lower bound, D_* must be as large as the mixing time of the corresponding random walk on $\mathbb{S}^d$. The latter is studied in [7], and fortunately the lower bound is still far from the threshold that the local neighborhood in H becomes loopy provided that $d\geq d_0$.

However, given a large D_*, it is difficult to write down the explicit expression for each step in the belief propagation algorithm and some new idea is needed for analyzing the belief propagation.

Acknowledgments

The author would like to thank Sidhanth Mohanty for introducing the problem, and thank Byron Chin, Nike Sun and Sidhanth Mohanty for helpful discussions.

References

[1] M. Brennan G. Bresler and D. Nagaraj, Phase Transitions for Detecting Latent Geometry in Random Graphs, *Probability Theory and Related Fields*, Vol. 178, 2020, pp. 1215-1289.

[2] S. Bubeck, J. Ding, R. Eldan and M. Rácz, Testing for high-dimensional geometry in random graphs, *Random Structures and Algorithms*, vol. 49, no. 3, 2016, pp. 503-532.

[3] J. Ding, H. Du and S. Gong, A polynomial-time approximation scheme for the maximal overlap of two independent Erdős-Rényi graphs, Preprint, arXiv:2210.07823.

[4] H. Du, S. Gong and R. Huang, The Algorithmic Phase Transition of Random Graph Alignment Problem, Preprint, arXiv:2307.06590.

[5] A. Goel, S. Rai, and B. Krishnamachari, Monotone properties of random geometric graphs have sharp thresholds. *Annals of Applied Probability*, vol. 15, no. 4, 2005, pp. 2535–2552.

[6] S. Liu, S. Mohanty, T. Schramm, and E. Yang, Testing thresholds for high-dimensional sparse random geometric graphs, *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, June 2022, pp. 672–677.

[7] S. Li and T. Schramm, Spectral clustering in the Gaussian mixture block model. Preprint, arXiv:2305.00979.

[8] S. J. Szarek, Nets of Grassmann manifold and orthogonal group, *Proceeding of research workshop on Banach Space Theory*, 1981. pp. 169-186.