

# Sandwiching Random Geometric Graphs and Erdős-Rényi with Applications: Sharp Thresholds, Robust Testing, and Enumeration

Kiril Bangachev\* Guy Bresler†

August 5, 2024

## Abstract

The distribution  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  is formed by sampling independent vectors  $\{V_i\}_{i=1}^n$  uniformly on  $\mathbb{S}^{d-1}$  and placing an edge between pairs of vertices  $i$  and  $j$  for which  $\langle V_i, V_j \rangle \geq \tau_d^p$ , where  $\tau_d^p$  is such that the expected density is  $p$ . Our main result is a poly-time implementable coupling between Erdős-Rényi and RGG such that  $\mathbb{G}(n, p(1 - \tilde{O}(\sqrt{np/d}))) \subseteq \text{RGG}(n, \mathbb{S}^{d-1}, p) \subseteq \mathbb{G}(n, p(1 + \tilde{O}(\sqrt{np/d})))$  edgewise with high probability when  $d \gg np$ . We apply the result to:

1) *Sharp Thresholds*: We show that for any monotone property having a sharp threshold with respect to the Erdős-Rényi distribution and critical probability  $p_n^c$ , random geometric graphs also exhibit a sharp threshold when  $d \gg np_n^c$ , thus partially answering a question of Perkins.

2) *Robust Testing*: The coupling shows that testing between  $\mathbb{G}(n, p)$  and  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  with  $\epsilon n^2 p$  adversarially corrupted edges for any constant  $\epsilon > 0$  is information-theoretically impossible when  $d \gg np$ . We match this lower bound with an efficient (constant degree SoS) spectral refutation algorithm when  $d \ll np$ .

3) *Enumeration*: We show that the number of geometric graphs in dimension  $d$  is at least  $\exp(dn \log^{-7} n)$ , recovering (up to the log factors) the sharp result of Sauermann.

**Keywords**— Random Geometric Graphs, Algorithmic Coupling, Robust Testing, Spectral Refutation, Sharp Thresholds, Enumeration.

---

\*Dept. of EECS, MIT. [kirilb@mit.edu](mailto:kirilb@mit.edu). Supported by a Siebel Scholarship.

†Dept. of EECS, MIT. [guy@mit.edu](mailto:guy@mit.edu). Supported by NSF Career Award CCF-1940205.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Our Angle . . . . .	1
1.2	Main Coupling Theorem . . . . .	2
1.3	Sharp Thresholds . . . . .	3
1.4	Robust Testing . . . . .	4
1.5	Enumeration . . . . .	5
1.6	Entropic and Spectral Interpretations of $d = \tilde{\Theta}(np)$ . . . . .	6
1.7	Challenges and Comparison to Prior Work . . . . .	7
1.8	Organization . . . . .	7
<b>2</b>	<b>Preliminaries</b>	<b>8</b>
<b>3</b>	<b>Overview of Main Ideas</b>	<b>9</b>
3.1	The Coupling Algorithm . . . . .	9
3.2	The Robust Testing Algorithm . . . . .	11
3.3	A Curious Representation: RGG as Recursively Planted Erdős-Rényi . . . . .	12
<b>4</b>	<b>Coupling Random Geometric Graphs and Erdős-Rényi</b>	<b>13</b>
4.1	The Flip-Orientation Map . . . . .	14
4.2	Proof of Lemma 3.1 . . . . .	15
4.3	Proof of Lemma 3.2 . . . . .	15
4.4	Deriving Corollary 1.1 from Theorem 1.1 . . . . .	19
<b>5</b>	<b>Robust Testing</b>	<b>20</b>
5.1	The SoS Program and Refutation . . . . .	20
5.2	Analysis of Corrupted Erdős-Rényi . . . . .	20
5.3	Analysis of Corrupted RGG . . . . .	21
5.4	Putting It All Together . . . . .	23
<b>6</b>	<b>Enumerating Geometric Graphs and the Support Size of RGG</b>	<b>23</b>
<b>7</b>	<b>Sharp Thresholds</b>	<b>24</b>
7.1	Proof of Theorem 1.2 . . . . .	24
7.2	RGG and the FKG Lattice Condition . . . . .	24
<b>8</b>	<b>Discussion and Open Problems</b>	<b>26</b>
8.1	Robust Testing With Fewer Corruptions . . . . .	26
8.2	Embedding Random Geometric Graphs . . . . .	26
8.3	A Reverse Coupling? . . . . .	27
8.4	Other Random Geometric Graph Distributions . . . . .	27
<b>A</b>	<b>Proof of Lemma 2.1</b>	<b>31</b>
<b>B</b>	<b>Simple Lemmas on Graph Counting</b>	<b>33</b>
<b>C</b>	<b>Random Geometric Graphs as Recursively Planted Erdős-Rényi</b>	<b>34</b>

# 1 Introduction

Random geometric graphs have emerged as a fruitful way to model dependence between edges in random networks in both theory and practice [Pen03, SKB08, BBDD<sup>+</sup>21, DdC22]. Their defining property is that edges are included based on similarity of feature vectors associated to the nodes. Geometric graphs accurately capture certain properties of social networks [MB73, MSLC01, HRH02, PJ09, ES16], biological networks [HRP08, RV23, SGN<sup>+</sup>21], and wireless communication networks [Gil61, Hae12].

A simple example is given by the *Spherical Random Geometric Graph*, popular in the mathematical community due to its simplicity and elegance.

**Definition 1** (Spherical Random Geometric Graph). A sample from the distribution  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  over graphs on  $n$  vertices  $[n] = \{1, \dots, n\}$  of dimension  $d$  with expected density  $p$  is generated as follows. First,  $n$  vectors  $V_1, \dots, V_n$  are drawn independently from the uniform distribution on the sphere  $\mathbb{S}^{d-1}$ . Then, an edge between  $i$  and  $j$  is formed if and only if  $\langle V_i, V_j \rangle \geq \tau_d^p$ . Here,  $\tau_d^p$  is chosen so that  $p = \mathbb{P}[\langle V_i, V_j \rangle \geq \tau_d^p]$ .

Much is known about the theoretical properties of random geometric graphs in the low-dimensional regime when  $d$  is constant or on the order of  $O(\text{polylog}(n))$  [Pen03]. A recent breakthrough work in this setting [LMSY23] shows that in certain dimensions  $d = \Theta(\log n)$ , random geometric graphs are sparse two-dimensional expanders, an object for which no randomized construction was previously known.

The interest in the high-dimensional regime when  $d \gg \text{polylog}(n)$  is much more recent.<sup>1</sup> A strikingly simple, but fundamental, question remains open after more than a decade of continuous work and improvements: *For what sequences of  $n, d, p$  are the edges of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  actually dependent?* When  $d$  is large enough (as a function of  $n$  and  $p$ ), the  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  distribution is only  $o(1)$  total variation distance away from  $\text{G}(n, p)$ . In other words, for large enough  $d$ , the edges are asymptotically independent [DGLU11]. The majority of the work on high-dimensional random geometric graphs has focused on determining the precise dimension for which this occurs [DGLU11, BDER14, BBN20, LMSY22, BB24b]. When  $d \ll (np)^3$ , a simple triangle-based statistic witnesses edge dependence. Conjecturally, when  $d \gg (np)^3$ ,  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  converges to  $\text{G}(n, p)$  in total variation distance.

Insofar as the purpose of the  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  distribution is to model edge dependence, the interesting regime is  $d \ll (np)^3$ , where the distribution is far from Erdős-Rényi and the edges are meaningfully dependent. Yet, very little is known in the “bulk”  $\text{polylog}(n) \ll d \ll (np)^3$  of this regime. This leads to the main motivation of the current paper:

*What can be said about  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  in the regime  $\text{polylog}(n) \ll d \ll (np)^3$ ?*

The higher the dimension, the weaker the dependence is between edges and, thus, one can expect behaviour more similar to Erdős-Rényi. One well-studied property in this regime is the spectrum of the adjacency matrix [LR23, LS23, BB24b]. When  $d \ll np$ , the centered adjacency matrix of a sample from  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  has an operator norm of  $\tilde{\Theta}(np/\sqrt{d})$  which is significantly larger than that of  $\text{G}(n, p)$  at  $\tilde{\Theta}(\sqrt{np})$ . Conversely, when  $d \gg np$ , the operator norm is  $\tilde{\Theta}(\sqrt{np})$ , which is on the same order of  $\text{G}(n, p)$ . A different result is that when  $d \gg \text{polylog}(n)$ , the clique number of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  is essentially the same as that of  $\text{G}(n, p)$  [DGLU11]. Aside from existence of a clique of a given size, the thresholds in  $p$  for other natural properties such as connectivity and Hamiltonicity are completely open when  $\text{polylog}(n) \ll d \ll (np)^3$ . Is the behaviour of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  with respect to those properties similar to the behaviour of the much better-understood  $\text{G}(n, p)$  distribution?

## 1.1 Our Angle

In order to understand which properties might be shared between the random geometric graph and Erdős-Rényi in the regime  $\log(n) \ll d \ll (np)^3$ , we aim to show a new style of comparison result between the two graph models. Of course, such a comparison should *not be in total variation*, because in this regime the

---

<sup>1</sup>We use  $a \gg b$  to denote  $a \geq \log(n)^c b$  for some fixed, implicit, constant  $c$ . Similarly for  $\ll$ .

two distributions have total variation distance  $1 - o(1)$ . Large total variation is a consequence of differing signed triangle counts, but what about other properties? It seems plausible that the distributions share many properties in addition to their clique number (which was studied in [DGLU11]).

Rather than studying properties individually, we aim to simultaneously address all *edge-monotone properties*. These include existence of a clique of a given size, but also Hamiltonicity, connectivity, existence of a perfect matching and many more. One of the most fundamental facts about  $G(n, p)$  is that *monotone properties* subject to very general criteria exhibit a sharp-threshold in  $p$  [FabJB99]. Yet, extending this result to non-product measures has proven to be a challenging task. Important progress has been made in the case of measures satisfying the FKG lattice condition [Gri06]. However, the RGG does not satisfy this condition (see Proposition 7.1). [GRK05] prove sharp thresholds for monotone properties for random geometric graphs but in the low-dimensional regime  $o(\log n)$ . As addressing monotone properties is one of our main goals, we aim for a comparison with Erdős-Rényi under which similar distributions assign similar probabilities to monotone events.

Besides exploring the similarities between  $G(n, p)$  and  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$ , we will also aim for a comparison that evinces *meaningful* differences between the two distributions. Real-world networks typically do not exactly match any natural generative model such as the one in Definition 1, and model misspecification can be captured by allowing for a small proportion of the edges to be adversarially corrupted. While the signed triangle count distinguishes the two graph models when  $d \ll (np)^3$ , this test turns out to be rather brittle. *When are there tests between  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  and  $G(n, p)$  that are robust to edge corruptions?*

The combination of *monotonicity* and *edge corruptions* leads us to the following notion of comparison with Erdős-Rényi: approximate stochastic dominance. Our goal will be to

*Couple  $H_- \sim G(n, p(1-\epsilon))$ ,  $H_+ \sim G(n, p(1+\epsilon))$ , and  $G \sim \text{RGG}(n, \mathbb{S}^{d-1}, p)$  for some vanishingly small  $\epsilon$  in such a way that  $H_- \subseteq G \subseteq H_+$  with high probability (where inclusion is with respect to the edge sets).*

Our main result is that such a coupling exists if and only if  $d \gg np$ . ([LMSY22] also exhibit such a coupling but in the weaker regime  $d \gg n^2 p^2$ . Furthermore, their result is not algorithmic while ours is. A more detailed comparison is made in Section 1.7.) This has several applications:

1. *Sharp Thresholds:* We use the *edge-monotonicity* of the coupling to compare monotone properties of  $G(n, p)$  and  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$ . We give the first proof of sharp thresholds for monotone properties of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  for a polynomially large range of dimensions where  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  is TV-distance  $1 - o(1)$  away from  $G(n, p)$ . Furthermore, we show that the critical probabilities are nearly identical with those of  $G(n, p)$  in this range.
2. *Robust Testing:* One interpretation of the approximate stochastic dominance is that a sample of  $G(n, p)$  is a sample of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  with  $O(\epsilon n^2 p)$  adversarially corrupted edges. This leads us to revisit the well-studied question of testing between  $G(n, p)$  and  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  in the previously unexplored setting with adversarial edge-corruptions. Our comparison result shows impossibility of robust testing for  $d \gg np$  and in the complementary regime we provide an efficient sum-of-squares algorithm.
3. *Enumeration:* The approximate stochastic dominance implies a lower bound on the support size of the  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  distribution. This allows us to use  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  as a randomized construction for counting the number of geometric graphs in dimension  $d$ .

We next state our main coupling result and then describe these applications in more detail.

## 1.2 Main Coupling Theorem

**Theorem 1.1** (Main Coupling Result). *Consider some  $n, d, p$  such that  $d = \Omega(\max(np, \log n))$  and  $p = \Omega(1/n), p \leq 1/2$ . There exists a polynomial-time algorithm which on input  $H \sim G(n, p)$ , outputs  $n$  vectors  $V_1 = V_1(H), V_2 = V_2(H), \dots, V_n = V_n(H)$  in  $\mathbb{S}^{d-1}$  with the following two properties:*

1. The marginal distribution of the vectors is independent uniform:  $V_1, V_2, \dots, V_n \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$ .
2. With high probability (over the input and internal randomness),  $\mathbf{1}[\langle V_i, V_j \rangle \geq \tau_d^p] = H_{ij}$  for all pairs  $i, j$  such that  $|\langle V_i, V_j \rangle - \tau_p^d| \geq \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}$  for some absolute constant  $c$ .

In other words, the algorithm described above couples  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  and  $\mathbf{G}(n, p)$  in such a way that they only differ at pairs  $(i, j)$  for which  $\langle V_i, V_j \rangle$  is very close to the threshold  $\tau_d^p$ . Elementary calculations shows that with high probability,  $G, H$  differ at only  $\tilde{O}\left(\binom{n}{2}p \times \sqrt{\frac{np}{d}}\right) = o\left(\binom{n}{2}p\right)$  edges.

**Remark 1.** [Theorem 1.1](#) shows that one can represent  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  as an Erdős-Rényi graph with few “defect edges.” By performing multiple rounds of our coupling algorithm, we can sequentially “fix” these defect edges. We obtain a curious novel representation of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  by recursively planting Erdős-Rényi graphs (of exponentially shrinking size) in  $\mathbf{G}(n, p)$ . When  $d = n^{1+\epsilon}p$ , the recursion only takes constantly many rounds. See [Section 3.3](#) for the representation result.

An immediate implication of [Theorem 1.1](#) is the following approximate stochastic dominance.

**Corollary 1.1** (Approximate Stochastic Dominance). *Consider some  $n, d, p$  such that  $d = \Omega(np \log^4 n, \log^5 n)$  and  $p = \Omega(1/n)$ . Then, there exists a coupling of  $G \sim \text{RGG}(n, \mathbb{S}^{d-1}, p)$ ,  $H_- \sim \mathbf{G}(n, p(1 - \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^2}{\sqrt{d}}))$  and  $H_+ \sim \mathbf{G}(n, p(1 + \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^2}{\sqrt{d}}))$ , such that  $H_- \subseteq G \subseteq H_+$  with probability  $1 - o(1)$ , where inclusion is defined as inclusion over the edge set.*

In [\[LMSY22\]](#), the authors prove [Corollary 1.1](#) under the weaker condition  $d \gg n^2 p^2$  with the purpose of describing the local structure of random geometric graphs. Our result, which holds for  $d \gg np$  is based on an entirely different proof technique. The difference between  $np$  and  $(np)^2$  is not merely quantitative. Qualitatively,  $d = \tilde{\Theta}(np)$  is of fundamental importance to spherical random geometric graphs as it is simultaneously a spectral and an entropic threshold, see [Section 1.6](#). This dual interpretation makes it key to our applications and using  $d \gg n^2 p^2$  would give highly suboptimal results for robust testing and enumeration.

**Remark 2.** We note that in general a relaxed stochastic dominance of the form in [Corollary 1.1](#) does not guarantee closeness in total variation even if one couples  $H_- \sim \mathbf{G}(n, p)$ ,  $H_+ \sim \mathbf{G}(n, p)$ , and  $G \sim \mathcal{D}$  such that  $H_- \subseteq G \subseteq H_+$  with high probability for some  $\mathcal{D}$ . Specifically, consider a distribution  $\mathcal{D}$  defined by the following coupling.  $H_-$  is drawn from  $\mathbf{G}(n, p)$ . Then,  $H_+$  is a copy of  $H_-$  except that a uniformly random non-edge is changed to an edge. One can check that whenever  $p = \Omega(1/n)$ , the distribution of  $H_+$  is only  $o(1)$  TV distance away from  $\mathbf{G}(n, p)$ . Finally,  $G$  is the graph among  $H_-, H_+$  with an even number of edges. Clearly,  $H_- \subseteq G \subseteq H_+$  w.h.p. but  $d_{\text{TV}}(\mathcal{D}, \mathbf{G}(n, p)) \geq 1/2 - o(1)$  as the distribution  $\mathcal{D}$  is supported on graphs with an even number of edges.

### 1.3 Sharp Thresholds

The question of whether monotone properties exhibit sharp thresholds is by far best understood for product distributions.<sup>2</sup> The seminal paper of [\[FabJB99\]](#) building on decades of work on the subject [\[Mar74, Rus81, Rus82, BKK<sup>+</sup>92, Tal94, FK96\]](#) establishes very broad conditions on vertex-transitive properties for which the Erdős-Rényi distribution exhibits sharp-thresholds. In non-product settings, much less is understood. Important progress has been made in the case of positive measures satisfying the FKG lattice property [\[Gri06, GG06, GG11\]](#). However, this property does not hold for random geometric graphs, which we check in [Proposition 7.1](#).

In the case of low-dimensional random geometric graphs (over the solid cube  $[0, 1]^d$ ), [\[GRK05\]](#) show that monotone properties have *additive* sharp-thresholds by proving that if  $G_1$  and  $G_2$  are independent samples with densities  $p(1 - o(1))$  and  $p$ , then  $G_1 \subseteq G_2$  with high probability after an appropriate vertex relabelling. Also in low dimension, but in a somewhat different setting where the vertices come from a Poisson process

---

<sup>2</sup>See [Section 2](#) for the formal definitions of sharp thresholds, critical probability, and critical window.

and hence the random geometric graph is of variable size, [BP13] gives necessary conditions for properties monotone in *both* edges and vertices that do not exhibit *multiplicative* sharp thresholds. Under certain compatibility of the vertex and edge thresholds, they can transfer the results to the fixed-size edge-monotone setting. However, a generic result for the fixed-size edge-monotone setting is only conjectured in [BP13]. The methods of [GRK05, BP13] do not seem applicable to the high-dimensional setting because both rely on the fact that a very fine (grid) subdivision of the space is well-covered by the latent vectors. In  $d$  dimensions, the subdivision has  $\exp(\Omega(d))$  cells, which is much larger than  $n$  when  $d = \omega(\log n)$ .

This suggests a natural question, posed by Perkins [Per24]: *Do random geometric graphs in higher dimensions exhibit sharp thresholds?* Of course, when the dimension is sufficiently large, the  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  distribution is asymptotically the same as  $\text{G}(n, p)$  and the answer is “yes”. But what about the case when  $d \ll n^3 p^3$  and the two distributions are provably far in total variation?

Our coupling Corollary 1.1 in the regime  $d \gg np$  between random geometric graph  $G$  and Erdős-Rényi  $H_-$  and  $H_+$  is monotone in the sense that  $H_- \subseteq G \subseteq H_+$ . This implies that whenever  $H_- \sim \text{G}(n, p(1 - \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^2}{\sqrt{d}}))$  satisfies an increasing property with high probability, so does  $G \sim \text{RGG}(n, \mathbb{S}^{d-1}, p)$ , and whenever  $G \sim \text{RGG}(n, \mathbb{S}^{d-1}, p)$  satisfies an increasing property, so does  $H_+ \sim \text{G}(n, p(1 + \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^2}{\sqrt{d}}))$ . This allows us to translate sharp-threshold phenomena from Erdős-Rényi to RGG. We formalize this argument in Section 7.1 and record it as the following theorem.

**Theorem 1.2.** *Suppose that  $\mathcal{P}_n$  is a monotone property with critical probability  $p_n^c$  in  $\text{G}(n, p)$ , which satisfies  $p_n^c = \Omega(1/n)$ . Suppose further that  $\mathcal{P}_n$  exhibits a sharp threshold with respect to  $\text{G}(n, p)$ . If  $d_n = \omega(\max(np_n^c \log^4 n, \log^5 n))$ , the property  $\mathcal{P}_n$  also exhibits a sharp threshold for  $\text{RGG}(n, d_n, p)$  and has critical probability  $p_n^c(1 + o_n(1))$ .*

## 1.4 Robust Testing

The problem of testing between Erdős-Rényi and a spherical random geometric graph has attracted significant attention in recent years [DGLU11, BDER14, BBN20, LMSY22, BB24b]. Formally, one observes a graph  $G$  and needs to decide  $H_0 : G \sim \text{G}(n, p)$  or  $H_1 : G \sim \text{RGG}(n, \mathbb{S}^{d-1}, p)$ . By comparing the number of signed triangles  $\sum_{i < j < k} (G_{ij} - p)(G_{jk} - p)(G_{ki} - p)$  to a threshold, one can succeed with high probability whenever  $d = \omega(n^3 p^3 \log^3(1/p))$  [BDER14, SW22]. The current state-of-the-art information-theoretic lower bounds are that when  $d \geq (\log n)^C n^3 p^2$ , the problem is information theoretically impossible as well as when  $p = \Theta(1/n)$  and  $d \geq (\log n)^C$  for some constant  $C$  [LMSY22]. Even ignoring polylog factors, there is still a  $1/p$  gap between the testing algorithm and the information-theoretic lower bound when  $1/n \ll p \ll 1/2$ . This gap was recently closed with respect to low-degree polynomial tests by [BB24b], i.e., counting signed triangles is optimal among all degree  $\log^{2-\epsilon} n$  polynomial tests.

The existing methods in the literature for both the upper and lower-bounds for feasibility of testing are extremely brittle – they fail even under a very small fraction of adversarially corrupted edges. On the algorithmic side, consider the well-studied test mentioned above based on comparing the number of signed triangles to a threshold. The expected number of signed triangles in  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  is  $\tilde{\Theta}(n^3 p^3 / \sqrt{d})$  [LMSY22] while in  $\text{G}(n, p)$  it is 0. Thus, planting a clique of size  $\tilde{O}(np/d^{1/6})$  in  $\text{G}(n, p)$  (which requires  $\tilde{O}(n^2 p^2 / d^{1/3}) = o(n^2 p)$  adversarial corruptions) makes the expected counts match and hence the test fails. One can argue similarly for other tests based on counting subgraphs. Regarding lower-bounds, all of the prior works in the setting without adversarial corruptions [BDER14, BBN20, LMSY22] are based on bounding total variation distance and strongly exploit distributional assumptions. We do not expect that the techniques in any of those papers apply to the robust testing setting.

The methods of the current work not only apply to, but also give tight upper and lower bounds, in a setting of a “maximum possible” fraction of corruptions. We work in the strong contamination model with edge corruptions, which has been extensively studied for the stochastic block model, e.g. [BMR21] for testing and [MS16, DdNS22, MRW24] for recovery.

**Problem 1** (Robust Testing of Geometry With Edge Corruptions). Let  $n, d \in \mathbb{N}$ ,  $p \in [0, 1]$  and  $\epsilon > 0$  be a fixed constant. One observes a graph  $H$  and needs to decide between the two hypotheses  $H_0 : H =$

$\mathcal{A}(\mathbf{G})$ ,  $\mathbf{G} \sim \mathsf{G}(n, p)$  and  $H_1 : H = \mathcal{A}(\mathbf{G})$ ,  $\mathbf{G} \sim \mathsf{RGG}(n, d, p)$ . Here,  $\mathcal{A} : \{0, 1\}^{\binom{n}{2}} \rightarrow \{0, 1\}^{\binom{n}{2}}$  is any unknown mapping such that  $\mathcal{A}(G)$  and  $G$  differ in at most  $\epsilon \binom{n}{2} p$  entries for any adjacency matrix  $G \in \{0, 1\}^{\binom{n}{2}}$ .

We restrict the adversary to changing at most  $\epsilon \binom{n}{2} p$  edges for a small constant  $\epsilon$  as otherwise [Problem 1](#) is trivially impossible. To see this, note that both  $\mathsf{RGG}(n, d, p)$  and  $\mathsf{G}(n, p)$  have with high probability at most  $1.5 \binom{n}{2} p$  edges. Hence, an adversary  $\mathcal{A}$  who simply deletes all edges of a sample from  $\mathsf{RGG}$  and then inserts Bernoulli  $p$  edges (in total  $3 \binom{n}{2} p$  w.h.p.) can transform  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  into a sample from  $\mathsf{G}(n, p)$ . Thus, [Problem 1](#) captures the setting of “maximal adversaries.”

Our first result is an information-theoretic lower bound which immediately follows from [Theorem 1.1](#). Whenever  $d \geq K(\epsilon) \max(np(\log n)^4, (\log n)^5)$  (where  $K(\epsilon)$  depends solely on  $\epsilon$ ), one can with high probability (adversarially) corrupt  $\epsilon \binom{n}{2} p$  edges of a sample from  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  and produce a sample from  $\mathsf{G}(n, p)$ . Hence, the robust testing problem is information-theoretically impossible when  $d \geq K(\epsilon) \max(np(\log n)^4, (\log n)^5)$ . When  $d = o((np)^3(\log 1/p)^3)$ ,  $d \geq K(\epsilon) \max(np(\log n)^4, (\log n)^5)$ , the problem is information-theoretically impossible even against a polynomial-time adversary. In this regime, an adversary can see a sample  $H$  and (correctly with high probability) test using the signed-triangle statistic [[LMSY22](#)] whether the sample is from  $\mathsf{G}(n, p)$  or  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$ . If  $H$  is from  $\mathsf{G}(n, p)$ , the adversary can use the efficient algorithm from [Theorem 1.1](#) and produce (with high-probability) a sample from  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  differing from  $H$  only at  $o(\binom{n}{2} p)$  edges. If  $H$  is from  $\mathsf{RGG}$  already, the adversary does not change edges. We record this in the following corollary.

**Corollary 1.2.** *Suppose that  $\epsilon \in (0, 1)$  is any fixed constant. Whenever  $d \geq K(\epsilon) \max(np(\log n)^4, (\log n)^5)$  the robust testing problem [Problem 1](#) is information-theoretically impossible with  $\epsilon \binom{n}{2} p$  corruptions. If, furthermore,  $d = o((np)^3(\log 1/p)^3)$ , the problem is information-theoretically impossible even against polynomial-time adversaries.*

It turns out that when  $d \ll np$ , there exists a polynomial time algorithm solving [Problem 1](#) with high-probability even against computationally unbounded adversaries. As already mentioned, it is unlikely that a subgraph-counting algorithm solves the robust testing problem. Instead, we require a more sophisticated sum-of-squares spectral refutation algorithm which uses the true embedding as a witness. To the best of our knowledge, [[MWZ24](#)] is the only other work which uses an embedding-based approach for testing for random geometric graphs, but their work is in low-dimension and the algorithm is not computationally efficient.

**Theorem 1.3.** *There is an absolute constant  $K_r > 0$  such that when  $d \leq K_r \min(np \log(1/p), \frac{n^2 p^2 \log(1/p)}{\log^2 n})$  and  $d = \omega(\log 1/p)$ , there exists a polynomial-time algorithm solving [Problem 1](#) with high probability for all  $\epsilon \leq K_r$ .*

We overview the algorithm in [Section 3.2](#) and give full detail in [Section 5](#). It is a constant degree SoS program for spectral refutation of a small operator norm of the centered adjacency matrix. The key insight, implicit in [[LS23](#)], is that the latent vectors witness a large second eigenvalue for a random geometric graph when  $d \ll np$ . Coincidentally, this is the complementary regime from the one where [Theorem 1.1](#) holds.

**Remark 3.** [Theorem 1.3](#) implies that when  $d \leq K_r \min(np \log(1/p), n^2 p^2 \log(1/p) \log^{-2} n)$ , there does not exist a coupling between  $H_- \sim \mathsf{G}(n, p(1 - o(1)), G \sim \mathsf{RGG}(n, \mathbb{S}^{d-1}, p), H_+ \sim \mathsf{G}(n, p(1 + o(1))$  such that  $H_- \subseteq G \subseteq H_+$  w.h.p. Otherwise, one can use this coupling as an adversary for [Problem 1](#) even against computationally-unbounded testers, but [Theorem 1.3](#) shows that such an adversary cannot exist. Hence, the dimension requirement  $d = \tilde{\Omega}(np)$  in [Theorem 1.1](#) and [Corollary 1.1](#) is tight (up to polylog factors).

## 1.5 Enumeration

A classic result in algebraic geometry due to Warren [[War68](#)] nearly immediately implies that the number of geometric graphs in  $d$  dimensions is at most  $(Cn/d)^{dn}$  for some explicit constant  $C$  when  $d = o(n)$  [[MM11](#), [Sau21](#)]. The main idea is that the edges of a geometric graph correspond to the sign patterns of  $\binom{n}{2}$

polynomials over the  $dn$  variables  $\{(v_i)_j\}_{1 \leq i \leq n, 1 \leq j \leq d}$ . The variables are the coordinates of the latent vectors and the polynomials are the squared pairwise  $\ell_2$  distance minus the radius squared.

Proving lower bounds, however, turns out to be a much more challenging task. In [MM11], it is shown that in 2 dimensions the number of geometric graphs (known also as unit disk graphs in this case) is  $n^{(2+o(1))n}$ . Using more sophisticated tools from algebraic geometry and topology, [Sau21] shows that the number of geometric graphs in  $d$  dimensions is  $n^{(d+o(1))n}$  when  $d = o(n)$ .

We recover this result up to a polylogarithmic factor in the exponent. Namely, using [Corollary 1.1](#), we demonstrate that  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  is a distribution over at least  $\exp(\tilde{\Omega}(n^2p))$  graphs when  $d = \max(np \log^4 n, \log^5 n)$  which yields the following statement.

**Theorem 1.4.** *The number of graphs which can be realized as intersection graphs of unit spheres in dimension  $d$  is at least  $\exp(nd \log^{-7} n)$ .*

Admittedly, our proof loses a non-trivial polylogarithmic factor. Yet, one advantage of our proof is its simplicity – it is based on a natural probabilistic construction, spherical random geometric graphs. By comparison, the techniques from algebraic geometry and algebraic topology in [Sau21] are significantly more involved, although certainly more systematic and sharper.

The proof idea in [Theorem 1.4](#) is rather simple: if  $\mathcal{M}$  is the support of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$ , a counting argument bounds the number of graphs that are edge-subgraphs of some graph  $H$  in  $\mathcal{M}$  and differ from  $H$  in  $o(\binom{n}{2}p)$  edges. Combined with [Corollary 1.1](#), this immediately gives an upper bound on the entropy of  $G(n, p(1 - o(1)))$  in terms of  $|\mathcal{M}|$ . As the latter is  $\exp(\tilde{\Omega}(n^2p))$ , we get a lower bound on  $|\mathcal{M}|$ .

## 1.6 Entropic and Spectral Interpretations of $d = \tilde{\Theta}(np)$

The reader might have observed the omnipresence of the dimension  $d = \tilde{\Theta}(np)$  in the current work. This is a polynomially smaller dimension than the one required for  $d_{\text{TV}}(\text{RGG}(n, \mathbb{S}^{d-1}, p), G(n, p)) = o(1)$ . For the latter to hold,  $d = \tilde{\Omega}(n^3p^3)$  is needed as demonstrated by the signed triangle count test [BDER14, LMSY22] which is also conjectured to be optimal [BBN20, LMSY22, BB24b]. While  $d = \tilde{\Theta}(n^3p^3)$  is still rather mysterious, the critical dimension  $d = \tilde{\Theta}(np)$  in [Theorem 1.1](#) and [Corollary 1.1](#) has two natural interpretations:

1. **Entropic Threshold:** The maximum entropy of a graph distribution with expected density  $p$  is  $\binom{n}{2}H(p) = \tilde{\Theta}(n^2p)$  and this is achieved by  $G(n, p)$ . On the other hand, a simple heuristic<sup>3</sup> (made rigorous in [Section 6](#)) suggests that the effective support size of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  is  $\exp(\tilde{\Omega}(nd))$  and, hence, its entropy is  $\tilde{\Theta}(nd)$ . To match (up to lower order terms) the  $\tilde{\Theta}(n^2p)$  entropy of  $G(n, p)$ , one therefore needs  $d = \tilde{\Omega}(np)$ . In the regime  $d \gg np$ , we indeed use [Corollary 1.1](#) to show that the support size of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  is  $\exp(\tilde{\Omega}(n^2p))$ . On the other hand, when  $d \ll np$ , the entropy is  $\tilde{\Theta}(nd)$ . As in [Remark 7](#), this implies optimality of our [Corollary 1.1](#), an approximate stochastic dominance of the given form cannot exist when  $d \ll np$ . In [MZ24], the authors use a more sophisticated rate-distortion-theoretic version of this entropy argument to show that the when  $d \gg np$ , the latent vectors have a much larger entropy than the adjacency matrix. This yields the tight lower bound that when  $d \gg np$ , one cannot do a non-trivial recovery of the latent vectors from the adjacency matrix.
2. **Spectral Threshold:** Somewhat miraculously, as soon as  $d \ll np$ , there is a spectral separation between  $G(n, p)$  and  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$ : The second largest eigenvalue of  $G(n, p)$  is  $\tilde{\Theta}(\sqrt{np})$  whereas the second largest eigenvalue of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  is  $\tilde{\Theta}(np/\sqrt{d}) \gg \tilde{\Theta}(\sqrt{np})$  [LMSY23, LS23]. This spectral separation allows for efficient spectral and SDP algorithms, as in [LS23] (for reconstructing the latent vectors from the adjacency matrix) and [Section 5](#) (for robust testing).

We leave the open-ended question of explaining why the entropic and spectral thresholds coincide as an exciting direction for future work. While we are not aware of any reasons for why this is the case, it

---

<sup>3</sup> $\mathbb{S}^{d-1}$  has an  $\epsilon$ -net of size  $\exp(\tilde{\Omega}(d \log(1/\epsilon)))$ , discretizing the sphere. As there are  $n$  latent vectors  $V_1, V_2, \dots, V_n$ , each effectively from the  $\exp(\tilde{\Omega}(d))$ -discretization, and the latent vectors uniquely determine the geometric graph, the  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  distribution effectively has support size  $\exp(\tilde{\Omega}(nd))$ . Hence, its entropy is at most  $\tilde{\Theta}(nd)$ .

is fundamentally the reason why problems like latent vector embedding and robust testing for random geometric graphs do *not* have statistical-computational gaps. Entropy roughly controls the size of the support and, hence, the success of (inefficient) exhaustive search algorithms. The spectrum of a matrix, on the other hand, is one of the most fundamental linear-algebraic primitives and is used throughout theoretical computer science for the design of computationally efficient algorithms.

The fact that simple spectral and SDP approaches match brute-force algorithms for random geometric graphs is not a given, and indeed this is not the case for many other latent space graph distributions. Most famously, for the planted clique distribution [Jer92, Kuc95, AKS98] an exponential gap arises with strong supporting evidence in the Sum-of-Squares hierarchy [BHK<sup>+</sup>19]. For a random geometric graphs example, the planted dense cycle model also exhibits an information-computation gap [MWZ23, MWZ24]. On the other hand, there are also many models without an information-computation gap such as exact recovery in the 2-community symmetric stochastic block model [MNS15, ABH16]. In general, it is poorly understood when an information-computation gap for testing and recovery of latent space structure arises. In [EAM21] the authors give a general condition under which a gap does not exist, but their model is qualitatively different and does not capture our setting. In it, one only observes a predefined subset of the edges. The absence of a gap depends on the structure of this subset.

## 1.7 Challenges and Comparison to Prior Work

In [LMSY22], the authors prove Corollary 1.1 under the weaker condition  $d \geq n^2 p^2 \log^4 n$ . In their construction, they sample  $V_1, V_2, \dots, V_n$  sequentially. A caricature of their argument is that with high probability, when sampling  $V_1, V_2, \dots, V_{k-1}$  the intersection of any  $i \approx pk$   $p$ -caps<sup>4</sup> and  $k-i-1 \approx k(1-p)$   $p$ -anti-caps corresponding to  $V_1, V_2, \dots, V_{k-1}$  is sufficiently close to  $p^i(1-p)^{k-i-1}$ . This implies that the edges from  $k$  to  $[k-1]$  are roughly distributed as independent  $\text{Bern}(p)$  variables. They make this statement precise by quantifying “sufficiently close” via a concentration argument based on transportation inequalities.

While our proof of Corollary 1.1 is entirely different, it suggests what might be suboptimal in the approach of [LMSY22]. For typical  $V_1, \dots, V_{k-1}$ , it turns out that the volume of *most* intersections of  $i \approx pk$  caps and  $k-i-1 \approx k(1-p)$  anti-caps concentrates much better than the worst-case over choice of the  $i$  caps and  $k-i-1$  anti-caps due to a square-root CLT-type cancellation phenomenon. We use this idea in our proof to get  $d \geq \max(np, \log n)$ ; if we instead consider *all* intersections of caps and anti-caps, we also obtain a result for  $d \geq n^2 p^2 \text{polylog} n$ . While the square-root cancellation comes naturally in our martingale-based analysis, it is unclear whether it can be incorporated in the transportation-based approach of [LMSY22]. Beyond that, an advantage of our approach is that Theorem 1.1 gives an *efficient algorithm* producing the coupling. The result of [LMSY22] seems difficult to make algorithmically efficient as it requires keeping track of an intersection of caps and anti-caps and being able to sample and estimate the volume of this intersection.

The work [BB24b] also shows that in certain regimes random geometric graphs can be represented as  $G(n, p)$  with some modifications. In fact, in their representation, just as in Theorem 1.1 the non-Erdős-Rényi pairs  $(i, j)$  are exactly the ones for which  $\langle V_i, V_j \rangle$  are close to  $\tau_d^p$  (called *fragile pairs* in [BB24b]). While some of our ideas are related to and motivated by [BB24b], there is a fundamental barrier to their proof technique: It crucially relies on performing the Gram-Schmidt operation on the latent vectors<sup>5</sup>  $V_1, V_2, \dots, V_n$ , requiring  $d \geq n$ . This makes it inapplicable to the regime  $np \ll d \ll n$ . We circumvent this barrier in the current work by replacing Gram-Schmidt with a different transformation that requires only near-orthogonality of the latent vectors, which holds as soon as  $\text{polylog}(n) \ll d$ .

## 1.8 Organization

In Section 2 we introduce some basic preliminaries and notation. In Section 3, we overview our two main algorithmic results: the coupling algorithm in Theorem 1.1 and the spectral refutation algorithm in Theorem 1.3.

---

<sup>4</sup>For  $V \in \mathbb{S}^{d-1}$ , the associated  $p$ -cap is the set  $\{U \in \mathbb{S}^{d-1} : \langle U, V \rangle \geq \tau_d^p\}$  and the  $p$ -anti-cap is its complement.

<sup>5</sup>To be precise, Gram-Schmidt is performed on isotropic Gaussian vectors  $Z_1, Z_2, \dots, Z_n$ . This, however, makes little difference to the arguments in both [BB24b] and the current work, see Remark 8.

Additionally, in [Section 3.3](#) we address the representation mentioned in [Remark 1](#). We give the full details of [Theorem 1.1](#) and [Theorem 1.3](#) in [Section 4](#) and [Section 5](#). In [Section 6](#) we prove [Theorem 1.4](#). In [Section 7](#) we prove [Theorem 1.2](#) as well as the fact that  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  does not satisfy the FKG property. We end with several remarks and open questions regarding our results in [Section 8](#).

## 2 Preliminaries

**Graphs and Matrices.** We denote by  $\mathsf{G}(n, p)$  the Erdős-Renyi distribution over undirected graphs on  $n$  vertices, in which each edge appears independently with probability  $p$ . For a graph  $G$  on  $n$  vertices, its adjacency matrix  $A$  is an  $n \times n$  matrix for which  $A_{ij} = \mathbf{1}[(i, j) \text{ is an edge in } G]$ . The  $p$ -centered adjacency matrix of  $G$  is an  $n \times n$  matrix for which  $A_{ii} = 0$  for all  $i$ ,  $A_{ij} = 1 - p$  if  $(i, j)$  is an edge in  $G$ ,  $A_{i,j} = -p$  if  $(i, j)$  is not an edge. When we only say ‘‘centered’’ adjacency matrix of  $G$  where  $G \sim \mathcal{G}$ , the value of  $p$  is implied to be the expected marginal edge probability according to  $\mathcal{G}$ . The largest in absolute value eigenvalue of a matrix  $M$  is denoted by  $\lambda_{\max}(M)$ .

We will need the following well-known fact about the  $\mathsf{G}(n, p)$  distribution.

**Theorem 2.1** ([BGBK19, BGBK20, ADK21]). *Let  $A$  be the centered adjacency matrix of  $G \sim \mathsf{G}(n, p)$ . Then, for some absolute constant  $C$ , with high probability  $|\lambda_{\max}(A)| \leq C \max(\sqrt{np}, \sqrt{\log n})$ .*

**Spherical Coordinates.** Denote by  $\mathcal{D}_d$  the marginal distribution of a coordinate from the unit sphere. That is,  $\mathcal{D}_d$  is the distribution of  $\langle V, e_1 \rangle$  for  $V \sim \text{Unif}(\mathbb{S}^{d-1})$ . One can define the threshold in [Definition 1](#) by the equation  $p = \mathbb{P}_{X \sim \mathcal{D}_d}[X \geq \tau_d^p]$ . We will use the following simple facts. The intuition behind all of them is that for large enough  $d$ , the distribution  $\mathcal{D}_d$  is very close to  $\mathcal{N}(0, \frac{1}{d})$ .

**Lemma 2.1.** *Suppose that  $p \leq 1/2$  and  $d = \omega(\log(1/p))$ . Then, for some absolute constants  $C_A > 1, C_B > 1, C_D > 1$ , and  $C_i > 1$  for  $i \in \mathbb{N}$ , depending only on  $i$ , but not on  $d$  or  $p$ :*

$$A) \min\left(\frac{1}{C_A}\left(\frac{1}{2} - p\right)\frac{\sqrt{\log(1/p)}}{\sqrt{d}}, \frac{1}{2}\right) \leq \tau_d^p \leq C_A \frac{\sqrt{\log(1/p)}}{\sqrt{d}}.$$

$$B) \frac{\sqrt{\log 1/p}}{C_B \sqrt{d}} \leq \mathbb{E}_{X \sim \mathcal{D}_d}[X \mid X \geq \tau_d^p].$$

$$C) \mathbb{E}_{X \sim \mathcal{D}_d}[X^k \mid X \geq \tau_d^p] \leq C_k \frac{\log(1/p)^{k/2}}{d^{k/2}}.$$

$$D) \text{ Suppose that } \Delta \geq 0 \text{ and } \Delta \leq \frac{1}{C_D} \frac{(\log 1/p)^{-1/2}}{\sqrt{d}}. \text{ Then,}$$

$$\mathbb{P}_{X \sim \mathcal{D}_d}\left[X \in [\tau_d^p - \Delta, \tau_d^p + \Delta]\right] \leq C_D \Delta p \sqrt{d} \sqrt{\log 1/p}.$$

The proofs are simple and are deferred to [Appendix A](#). A) is from [BDER14]. D) appears in [BB24b] with slightly weaker dependence on  $d$  and  $p$ , so for completeness we also give a proof.

**Martingales.** In the analysis of our coupling algorithm for [Theorem 1.1](#) we will use the well-known inequality due to Freedman.

**Theorem 2.2** ([Fre75]). *Consider a sequence of martingale differences  $(\xi_i)_{i=1}^r$ , with respect to the filtration  $\{\mathcal{F}_i\}_{i=0}^r$ . That is,  $\xi_i$  is  $\mathcal{F}_i$ -measurable, bounded, and  $\mathbb{E}[\xi_i \mid \mathcal{F}_{i-1}] = 0$ . Suppose that*

1.  $\sum_{i=1}^r \mathbb{E}[\xi_i^2 \mid \mathcal{F}_{i-1}] \leq L$  for some deterministic  $L$ , and
2.  $|\xi_i| \leq a$  for all  $i$  for some deterministic value of  $a$ .

Then,  $\mathbb{P}\left[\left|\sum_{i=1}^r \xi_i\right| \geq z\right] \leq 2 \exp\left(-\frac{z^2/2}{L + az/3}\right)$ .

**Sharp Thresholds.** Let  $S$  be a finite set and consider a family of distributions  $(\mathcal{D}_p)_{p \in [0,1]}$  over  $\{0,1\}^S$  (we identify the power set of  $S$  with the set of indicator vectors  $\{0,1\}^S$ ) parametrized by the common marginal coordinate density  $p$ , for example  $\text{G}(n,p)$  or  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$ . Let  $\mathcal{P} \subseteq \{0,1\}^S$  be a given property (such as connectivity or Hamiltonicity of a graph). We say that  $\mathcal{P}$  is monotone increasing if for any  $A \in \mathcal{P}$  and  $B \subseteq \{0,1\}^S$  such that  $A \subseteq B$ , it is also true that  $B \in \mathcal{P}$ . If  $\mathcal{P}$  is a non-trivial monotone increasing property (that is, other than “no sets satisfy it” or “all sets satisfy it”), then  $\emptyset \notin \mathcal{P}$  and  $S \in \mathcal{P}$ . Hence,  $\mathbb{P}_{X \sim \mathcal{D}_0}[X \in \mathcal{P}] = 0$ ,  $\mathbb{P}_{X \sim \mathcal{D}_1}[X \in \mathcal{P}] = 1$ .

For natural distributions such as  $\text{G}(n,p)$  and  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$ , the function  $f : [0,1] \rightarrow [0,1]$  defined by  $f(p) = \mathbb{P}_{X \sim \mathcal{D}_p}[X \in \mathcal{P}]$  is continuous and increasing. As  $f(0) = 0$ ,  $f(1) = 1$ , there exists a *critical probability*  $p_c$  such that  $\mathbb{P}_{X \sim \mathcal{D}_{p_c}}[X \in \mathcal{P}] = 1/2$ , i.e.  $p_c = f^{-1}(1/2)$ . For a small constant  $\epsilon > 0$ , the *critical window* is denoted  $w_\epsilon$  and equals  $f^{-1}(1 - \epsilon) - f^{-1}(\epsilon)$ . Clearly,  $w_\epsilon > 0$ .

When  $w_\epsilon = o(p_c)$  for a given property  $\mathcal{P}$  for all fixed  $\epsilon > 0$ , we say that  $\mathcal{P}$  has a *multiplicative sharp threshold* with respect to  $(\mathcal{D}_p)_{p \in [0,1]}$ . This is the main setting we will be interested in. Common in the literature is also the weaker *additive sharp threshold* notion which only requires  $w_\epsilon = o_n(1)$ . When we say *sharp thresholds* in the current paper, we always mean the stronger multiplicative variant unless otherwise stated. All of the above can be analogously defined for monotone decreasing properties. See, for example, [O'D14, Gri06] for a more thorough discussion on sharp thresholds.

## 3 Overview of Main Ideas

### 3.1 The Coupling Algorithm

**Motivating Our Construction.** Suppose that we are given the adjacency matrix  $(H_{ji})_{1 \leq i < j \leq n} \stackrel{\text{i.i.d.}}{\sim} \text{Bern}(p)$  of  $H \sim \text{G}(n,p)$ . Our goal is to couple  $H$  with  $n$  vectors  $V_1, V_2, \dots, V_n$  on the sphere such that:

1.  $V_1, V_2, \dots, V_n \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$ .
2. Conditioned on  $(H_{ji})_{1 \leq i < j \leq n}$ , for most pairs  $(j,i)$ , it is the case that  $\mathbf{1}[\langle V_i, V_j \rangle \geq \tau_d^p] = H_{ji}$ .

While property 1 is easy to achieve on its own, property 2 is much less trivial. One needs to approximately embed the adjacency matrix of  $H$  as a geometric graph on the sphere. Even if the adjacency matrix already corresponds to a sample from  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$ , this step is non-trivial and requires some embedding algorithm, for example the spectral approach of [LMSY22]. And even if one succeeds in this embedding step, it is unclear how one could enforce the distributional assumption in property 1.

Our approach is to build up a coupling incrementally, starting with vectors satisfying property 1 and maintaining this as an invariant. That is, the first step is to sample  $n$  vectors  $V_1^{(0)}, V_2^{(0)}, \dots, V_n^{(0)} \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$ . We then incrementally transform them so that 1) they remain uniform on the sphere and 2) they produce a geometric graph that is close to the input  $H$ .

We start with the edge  $(2,1)$ . If  $H_{21} = \mathbf{1}[\langle V_1^{(0)}, V_2^{(0)} \rangle \geq \tau_d^p]$ , we do not need to do anything since property 2 is already satisfied for this edge. Otherwise, we transform  $V_2^{(0)}$  into  $V_2^{(1)}$  such that  $\mathbf{1}[\langle V_1^{(0)}, V_2^{(1)} \rangle \geq \tau_d^p] = 1 - \mathbf{1}[\langle V_1^{(0)}, V_2^{(0)} \rangle \geq \tau_d^p]$ , i.e., presence of edge  $(2,1)$  in our RGG is flipped. One way to do this transformation is as follows: we keep unchanged the direction of  $V_2^{(0)}$  in the space orthogonal to  $V_1^{(0)}$  and flip  $V_2^{(0)}$  in the direction of  $V_1^{(0)}$  via a measure-preserving (with respect to the  $\text{Unif}(\mathbb{S}^{d-1})$  distribution) map so that the inner product is on the correct side of  $\tau_d^p$ . We call this procedure a *flip orientation* of  $V_2^{(0)}$  along  $V_1^{(0)}$  so that the edge between them is set to  $H_{21}$  and we write  $V_2^{(1)} = \rho_{V_1^{(0)}}^{H_{21}}(V_2^{(0)})$ . In particular, in this notation  $\rho_a^b(z) = z$  whenever  $b = \mathbf{1}[\langle a, z \rangle \geq \tau_d^p]$ .

Having coupled  $H_{21}$  with  $V_1^{(0)}, V_2^{(1)}$ , we proceed to flip  $V_3^{(0)}$  first around  $V_1^{(0)}$  and then around  $V_2^{(1)}$  and so on. In short, the algorithm we use to prove [Theorem 1.1](#) is as follows.

**Algorithm 1.** First, sample  $V_1^{(0)}, V_2^{(0)}, \dots, V_n^{(0)} \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$ . Then, flip sequentially:

$$\begin{aligned} V_1^{(0)}, \quad \tilde{V}_1 &:= V_1^{(0)} \\ V_2^{(0)} &\xrightarrow{\rho_{\tilde{V}_1}^{H_{21}}} V_2^{(1)}, \quad \tilde{V}_2 := V_2^{(1)} \\ V_3^{(0)} &\xrightarrow{\rho_{\tilde{V}_1}^{H_{31}}} V_3^{(1)} \xrightarrow{\rho_{\tilde{V}_2}^{H_{32}}} V_3^{(2)}, \quad \tilde{V}_3 := V_3^{(2)} \\ &\vdots \\ V_n^{(0)} &\xrightarrow{\rho_{\tilde{V}_1}^{H_{n1}}} V_n^{(1)} \xrightarrow{\rho_{\tilde{V}_2}^{H_{n2}}} \cdots \xrightarrow{\rho_{\tilde{V}_{n-1}}^{H_{n,n-1}}} V_n^{(n-1)}, \quad \tilde{V}_n := V_n^{(n-1)}. \end{aligned}$$

The final output vectors are denoted by  $(\tilde{V}_i)_{i=1}^n$  where  $\tilde{V}_i := V_i^{(i-1)}$ . Importantly, observe that *flips occur with respect to the final vectors*. In general, the flip orientation transformation is succinctly expressed as

$$V_j^{(i)} = \rho_{\tilde{V}_i}^{H_{ji}}(V_j^{(i-1)}),$$

meaning that we flip  $V_j^{(i-1)}$  in the direction of  $\tilde{V}_i$  so that  $\mathbf{1}[\langle V_j^{(i)}, \tilde{V}_i \rangle \geq \tau_d^p] = H_{ji}$ .

**Analysis Sketch.** We need to show that the two properties claimed in [Theorem 1.1](#) hold.

Independence follows directly from the ‘‘measure-reserving’’ property of  $\rho_a^b(\cdot)$ , which we formalize in [Section 4.1](#). For technical reasons, the following slightly stronger independence statement will be useful.

**Lemma 3.1.** *For any choice of  $s \in \{1, 2, \dots, n\}$  and  $i_s \in \{0, 1, 2, \dots, s-1\}$ ,*

$$\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{s-1}, V_s^{(i_s)} \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1}).$$

The main idea behind the proof is that before any operation is performed on  $V_s^{(0)}$ , it is uniform on the sphere and independent from  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{s-1}$ . Since the transformation  $V_s^{(0)} \longrightarrow V_s^{(1)}$  is measure-preserving for any choice of  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{s-1}$ , the distribution of  $V_s^{(1)}$  remains uniform on the sphere and independent from  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{s-1}$ . We proceed by induction.

We also need to prove that with high probability, all pairs  $i, j$  such that  $|\langle \tilde{V}_i, \tilde{V}_j \rangle - \tau_d^p| \geq \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}$  satisfy  $\mathbf{1}[\langle \tilde{V}_i, \tilde{V}_j \rangle \geq \tau_d^p] = H_{ji}$ . By construction of the flip orientation map  $\rho$ ,  $\mathbf{1}[\langle \tilde{V}_i, V_j^{(i)} \rangle \geq \tau_d^p] = H_{ji}$ , because  $V_j^{(i)}$  has just been oriented relative to  $\tilde{V}_i$  to achieve this. We will show that the final  $\tilde{V}_j = V_j^{(j-1)}$  has not changed too significantly from  $V_j^{(i)}$  in the  $\tilde{V}_i$  direction:

**Lemma 3.2.** *For any  $1 \leq i < j \leq n$ ,*

$$\mathbb{P}\left[|\langle \tilde{V}_i, \tilde{V}_j \rangle - \langle \tilde{V}_i, V_j^{(i)} \rangle|\right] \geq \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d} \leq \frac{1}{n^{100}}. \quad (1)$$

To give intuition about the quantity  $\tilde{\Theta}(\sqrt{np}/d)$ , we write  $\langle \tilde{V}_i, \tilde{V}_j \rangle - \langle \tilde{V}_i, V_j^{(i)} \rangle$  as a telescoping sum, recalling that  $\tilde{V}_j = V_j^{(j-1)}$ :

$$\langle \tilde{V}_i, \tilde{V}_j \rangle - \langle \tilde{V}_i, V_j^{(i)} \rangle = \sum_{u=i+1}^{j-1} \langle \tilde{V}_i, V_j^{(u)} - V_j^{(u-1)} \rangle = \sum_{u=i+1}^{j-1} \langle \tilde{V}_i, \rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)} \rangle. \quad (2)$$

Each of the terms is equal to 0 with probability  $1-p$ . That is because  $\rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)} = 0$  whenever  $\mathbf{1}[\langle \tilde{V}_u, V_j^{(u-1)} \rangle \geq \tau_d^p] = H_{ji}$  and the two sides are independent  $\text{Bern}(p)$  random variables (the left-hand side because  $\tilde{V}_u, V_j^{(u-1)}$  are independent by [Lemma 3.1](#)). Hence, we expect only  $\tilde{O}(np)$  non-zero terms.

Whenever  $\rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)}$  is non-zero, its norm is on the order of  $\tilde{O}(1/\sqrt{d})$  with high probability. This is the case because  $\tilde{V}_i, V_j^{(u)}$  and  $\tilde{V}_i, V_j^{(u-1)}$  are pairwise independent vectors on the unit sphere by Lemma 3.1 and  $V_j$  changes in direction  $\tilde{V}_u$  when  $\rho_{\tilde{V}_u}^{H_{ju}}(\cdot)$  is applied. Moreover,  $\rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)}$  is in the span of  $\tilde{V}_u, V_j^{(u-1)}$ . Both directions are independent of  $\tilde{V}_i$  and, hence, only have a  $\tilde{O}(1/\sqrt{d})$  inner product with  $\tilde{V}_i$ . Altogether, this means that each  $\langle \tilde{V}_i, \rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)} \rangle$  is of order  $\tilde{O}(1/d)$ . As there are  $\tilde{O}(np)$  non-zero terms, the sum is of order  $\tilde{O}(np/d)$  with high probability.

However, we need to prove that the sum is of order  $\tilde{O}(\sqrt{np}/d)$ . The key insight is that a CLT-type square-root cancellation occurs in the sum Eq. (2). We show this via the classic martingale inequality Theorem 2.2. The technical challenge is that Eq. (2) is *not actually a sum of martingale differences* for the following ‘‘martingale violation’’ reasons:

- (V1) Whenever  $\rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)} \neq 0$ , it has a non-trivial component in the direction of  $V_j^{(u-1)}$ . The reason is that while the map  $V \mapsto \rho_{\tilde{V}_u}^{H_{ju}}(V)$  preserves the direction of  $V$  on the space orthogonal to  $\tilde{V}_u$ , it does not preserve the norm of  $V$  in this orthogonal space. This is because  $\rho_{\tilde{V}_u}^{H_{ju}}(V)$  must also have unit norm and, thus, a rescaling occurs. As the difference  $\rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)}$  is not independent of  $V_j^{(u-1)}$ , one can easily show that the martingale property is violated.
- (V2) Whether  $\rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)} = 0$  holds depends on the projection of  $V_j^{(u-1)}$  on  $\tilde{V}_u$ . Conditioned on  $H_{ji}, H_{j,i+1}, \dots, H_{j,u-1}$  (which need to appear in the underlying filtration),  $V_j^{(u-1)}$  and  $\tilde{V}_i$  are not independent. Hence, the projections of  $\tilde{V}_u$  on  $V_j^{(u-1)}, \tilde{V}_i$  are not independent. So,  $\langle \tilde{V}_i, \rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)} \rangle$  does not have a zero conditional expectation.

We deal with these issues by manually subtracting off the component of  $\rho_{\tilde{V}_u}^{H_{ju}}(V_j^{(u-1)}) - V_j^{(u-1)}$  in the direction of  $V_j^{(u-1)}$  and the component of  $V_j^{(u-1)}$  in the  $\tilde{V}_i$  direction. We show that these lead to lower-order contributions and what remains is a true sum of martingale differences.

**Remark 4.** On a high-level ‘‘flipping  $V_j^{(i-1)}$  in the direction of  $\tilde{V}_i$ ’’ is analogous to the Gram-Schmidt procedure of [BB24b] which aims to isolate the component of  $V_j$  in the direction of  $V_i$ . In [BB24b], this is achieved by forming an orthonormal basis  $e_1, e_2, \dots, e_n$  such that  $V_i \approx e_i$ . In fact, one can modify their argument and derive the equivalent statement of Theorem 1.1 when  $d \gg n$  (by flipping measure-preservingly the  $Z_{ji}$  values around  $\phi_d^p$  in their notation). Nevertheless, when  $d < n$ , one cannot form an orthonormal basis. A ‘‘nearly orthonormal’’ set is not sufficient for the arguments in [BB24b] as they heavily rely on the fact that the components of a Gaussian vector in orthogonal directions are independent. Hence, we need a different approach for the regime  $np \ll d \ll n$  that is central to our applications.

### 3.2 The Robust Testing Algorithm

**Motivation.** The motivation behind our testing algorithm comes from the interpretation of  $d = \tilde{\Theta}(np)$  as a spectral threshold. Specifically, when  $d \ll np$ , the centered adjacency matrix of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  typically has  $d$  large eigenvalues of order  $\tilde{\Theta}(np/\sqrt{d})$ . This is larger than the typical second eigenvalue of a sample from  $\mathbb{G}(n, p)$  which is only  $\sqrt{np}$ . Something more turns out to be true: the eigenvectors corresponding to these  $d$  eigenvalues are very close to the latent vectors  $V_1, V_2, \dots, V_n$ , a key insight in [LS23]. Formally, the  $d$  rows of the matrix  $V = (V_1, V_2, \dots, V_n)$  are close to the  $d$  eigenvectors with large eigenvalues. In [LS23], the authors use this fact to estimate the true Gram matrix  $VV^T$  from the adjacency matrix. Our end goal is different – to construct a spectral refutation algorithm. Formally, we give a certificate robust to  $\epsilon \binom{n}{2} p$  edge corruptions that the  $p$ -centered adjacency matrix of  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  has a large operator norm.

**The case of no corruptions.** Suppose for now that we wanted to use the spectrum to test between  $\mathsf{G}(n, p)$  and  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  in the case of no corruptions when  $d \ll np$ .<sup>6</sup> For the uncorrupted centered adjacency matrix  $A$  of a sample from  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$ , we know that there exists some matrix  $Y \in \mathbb{R}^{d \times n}$  such that  $\|Y_{:,i}\|_2 = 1$  for each  $i$  and  $\langle Y^T Y, A \rangle$  is large: One such matrix is  $Y = V$ , the matrix of latent vectors. As discussed above,  $V$  is close to the top  $d$  eigenvectors of  $A$ , each corresponding to eigenvalue  $\lambda = \tilde{\Theta}(np/\sqrt{d})$ . Furthermore,  $\|V\|_F^2 = n$ . Hence, the objective value  $\langle V^T V, A \rangle$  is on the order of  $\lambda \times \frac{\|V\|_F^2}{d} \times d = \tilde{\Theta}(n^2 p / \sqrt{d})$ . To illustrate the main ideas, assume that it is exactly  $n^2 p / \sqrt{d}$ .

Consider now the centered adjacency matrix  $A$  of a  $\mathsf{G}(n, p)$  sample. Its largest eigenvalue is  $\tilde{O}(\sqrt{np})$ , so

$$\langle Y^T Y, A \rangle \leq \langle Y^T Y, \tilde{O}(\sqrt{np}) \times I \rangle \leq \tilde{O}(\sqrt{np}) \times \text{tr}(Y^T Y) = \tilde{O}(n\sqrt{np}).$$

Again, for simplicity, assume that it is  $n\sqrt{np}$ . Note that  $n\sqrt{np} \ll n^2 p / \sqrt{d}$  when  $d \ll np$ .

We will therefore search (via a suitable SoS program) for a matrix  $Y \in \mathbb{R}^{d \times n}$  such that  $\langle YY^T, A \rangle$  is larger than some threshold. If we succeed, then our graph ought to be a sample from  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$ .

**Handling corruptions.** Observe that the objective value  $\langle Y^T Y, A \rangle$  for  $Y = V$  changes by very little when an adversary corrupts an edge  $A_{ij}$ . The reason is that  $(V^T V)_{ij} = \langle V_i, V_j \rangle$ , which with high probability is only on the order of  $\tilde{O}(1/\sqrt{d})$ . Hence, after  $\epsilon \binom{n}{2} p$  corruptions, the objective value can change by at most  $\tilde{\Theta}(\epsilon n^2 p / \sqrt{d})$ . Again, for simplicity, assume that the change is at most  $\epsilon n^2 p / \sqrt{d}$ . This, however, is less than the difference between  $n^2 p / \sqrt{d}$  (the uncorrupted RGG value) and  $n\sqrt{np}$  (the uncorrupted  $\mathsf{G}(n, p)$  value)! Hence, imposing that  $|\langle Y_i, Y_j \rangle| = \tilde{O}(1/\sqrt{d})$ , the SoS program will distinguish between the two graph models.

To get the full statement of [Theorem 1.3](#), we need to be more careful due to the log factors that we suppressed everywhere. Otherwise, we may end up with an algorithm that only works for  $\epsilon = \log n^{-C}$  for some absolute constant  $C > 0$ . We deal with this by imposing a more fine-grained bound on the empirical distribution of  $\{|\langle Y_i, Y_j \rangle|\}_{1 \leq i < j \leq n}$  instead of  $|\langle Y_i, Y_j \rangle| = \tilde{O}(1/\sqrt{d})$ .

**Remark 5.** The idea of maximizing a quadratic form against the centered adjacency has been used for robust statistics in the stochastic block model, e.g. [\[GV16\]](#). What is new in our work is the realization that the latent vectors witness a large operator norm in the case of random geometric graphs. To the best of our knowledge, novel is also the use of a constraint on the tails of the empirical distribution of the entries of a given matrix  $(\{|\langle Y_i, Y_j \rangle|\}_{1 \leq i < j \leq n})$ .

### 3.3 A Curious Representation: RGG as Recursively Planted Erdős-Rényi

We now describe a curious representation result, which furthers the view of RGG as Erdős-Rényi with few planted edges arising in [Corollary 1.1](#) and [\[BB24b\]](#). Most strikingly, when  $d = n^{1+\epsilon} p$  for some absolute constant  $\epsilon$ , we show that one can represent  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  by starting with a sample from  $\mathsf{G}(n, p)$  and performing constantly many rounds of planting independent density 1/2 Erdős-Rényi graphs. The sizes of these planted Erdős-Rényi graphs decrease exponentially in  $n$ . The entire “geometry” of the  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  distribution is, thus, hidden in the locations of the planted Erdős-Rényi graphs.

Concretely, in the current [Theorem 1.1](#), we allow for some amount of “defect edges”, i.e. the ones for which  $|\langle V_i, V_j \rangle - \tau_d^p| \leq \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}$ . One may try to correct these edges after the termination of [Algorithm 1](#). Indeed, recall from [Section 3.1](#) that one starts with arbitrary vectors  $V_1^{(0)}, V_2^{(0)}, \dots, V_n^{(0)} \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$  and transforms them into  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_n$ . However,  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_n \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$ . Thus, one may try to apply the algorithm again and correct the defect edges. That is, sequentially apply  $V_j^{(i)} = \rho_{\tilde{V}_i}^{H_{ji}}(\tilde{V}_j)$  for all  $i, j$  such that  $|\langle \tilde{V}_i, \tilde{V}_j \rangle - \tau_d^p| \leq \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}$ .

---

<sup>6</sup>Of course, this regime is suboptimal in the case of no adversarial corruptions as the signed triangle count succeeds all the way up to  $d = \tilde{\Theta}(n^3 p^3)$  [\[BDER14, LMSY22\]](#).

Intuitively, this should fix most of the disagreements between  $H$  and  $G_{ji} = \mathbb{1}[\langle V_i, V_j \rangle \geq \tau_d^p]$ . The reason is that each  $\tilde{V}_i$  will only be changed  $\tilde{O}(p\sqrt{np/d})$  times as (with high probability there are at most  $\tilde{O}(p\sqrt{np/d})$  values of  $j$  for which  $|\langle \tilde{V}_i, \tilde{V}_j \rangle - \tau_d^p| \leq \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}$ ). Furthermore, one can easily show that the changes will be of smaller order than the ones [Section 3.1](#) (see [Appendix C](#)). Altogether, running through the same martingale-based analysis, one should expect to fix all edges except the ones for which  $|\langle \tilde{V}_i, \tilde{V}_j \rangle - \tau_d^p| \lll (\max(\frac{1}{\sqrt{d}} \times (np/d)^{5/4}), \frac{1}{d}(np/d)^{1/2}))$ . This will vastly reduce the number of defect edges. After several more iterations, one can hope to completely remove all defect edges.

Of course, this should not be possible since  $\mathsf{G}(n, p)$  and  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  are *far in total variation* in the regime  $np \ll d \ll n^3 p^3$ . The issue with the argument above is that the independence is lost since we have already used  $H_{ji}$  once to flip  $V_j^{(i)} = \rho_{\tilde{V}_i}^{H_{ji}}(V_j^{(i-1)})$  in [Algorithm 1](#). Hence, the distributional [Lemma 3.1](#) does not hold any more and the resulting sample will not be a sample from  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$ .

A simple way to resolve this issue is to draw fresh values of  $H_{ji}^1$  for the second round of flippings. Then, draw fresh values for a third round and so on until no defect edges are left. This algorithm yields the following representation of  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$ . We give more detail in [Appendix C](#).

**Theorem 3.1** (Description of Random Geometric Graphs as Recursively Planted Erdős-Rényi). *Suppose that  $p = \Omega(1/n)$  and  $d = \kappa \times np \log^6 n$  for some  $\kappa > 1$ . Then, there exist some  $T = O(\frac{\log d}{\log \kappa} + \frac{\log n}{\log d})$  and a set of random functions  $F_1, F_2, \dots, F_T$ , where each  $F_i$  takes an  $n$ -vertex graph as an input and outputs a subset of  $[n] \times [n]$  with the following property. Consider the algorithm:*

1. Draw  $H^0 \sim \mathsf{G}(n, p)$  and set  $G^0 = H^0$ .
2. For  $t = 1, 2, \dots, T$  :
  - (a) Draw a fresh  $H^t \sim \mathsf{G}(n, 1/2)$ .
  - (b) Compute the set of pairs of vertices  $F_t(G^{t-1}) \in [n] \times [n]$ .
  - (c) Define  $G^t$  by  $G^t|_{F_t(G^{t-1})} = H^t$  and  $G^t|_{F_t(G^{t-1})^c} = G^{t-1}$ .

Then,  $d_{\text{TV}}(\mathcal{L}(G^T), \mathsf{RGG}(n, \mathbb{S}^{d-1}, p)) = o(1)$ . Furthermore, with high probability,  $F_t(G^{t-1}) \subseteq F_{t+1}(G^t)$  for each  $t$  and there exists some  $T_1 = \Theta(\frac{\log d}{\log \kappa})$  such that

$$|F_t(G^{t-1})| \leq \begin{cases} C_D n^2 p \sqrt{\log 1/p} \times \left(\frac{np \log^6 n}{d}\right)^{\frac{3^{t-1}}{2^t}} & \text{for } t \leq T_1 \\ C_D n^{3/2} p^{1/2} \sqrt{\log 1/p} \times \left(\frac{\log^2 n}{\sqrt{d}}\right)^{t-T_1} & \text{for } t > T_1. \end{cases}$$

One should think of  $F_t(G^{t-1})$  as the set of defect edges after  $t$  rounds of the coupling algorithm. Note that we always resolve the defect edges by planting an Erdős-Rényi graph on them. Hence, the above observation produces a description of  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  as a recursively planted Erdős-Rényi. Each next planted Erdős-Rényi is (doubly-)exponentially smaller.

When  $\kappa = n^\epsilon$  for some constant  $\epsilon > 0$ , in the above statement  $T = O(1/\epsilon)$ . Hence, to construct a random geometric graph in dimension  $n^{1+\epsilon} p$ , one only needs to recursively plant constantly many times an Erdős-Rényi graph in an Erdős-Rényi graph according to the functions  $F_i$ .

## 4 Coupling Random Geometric Graphs and Erdős-Rényi

In this section, we make the intuition in [Section 3.1](#) rigorous. We first define the map  $\rho_a^b(z)$  in [Section 4.1](#) and analyse some of its properties. Then, in [Section 4.2](#) we prove [Lemma 3.1](#) and in [Section 4.3](#) we prove [Lemma 3.2](#). In [Section 4.4](#) we deduce [Corollary 1.1](#) from [Theorem 1.1](#).

## 4.1 The Flip-Orientation Map

Define the unique monotone map  $\phi : [-1, 1] \rightarrow [-1, 1]$  such that  $\phi(\tau_d^p) = \tau_d^p$ ,  $\phi(1) = -1$ ,  $\phi(-1) = 1$  and  $\phi : [-1, \tau_d^p] \rightarrow [\tau_d^p, 1]$ ,  $\phi : [\tau_d^p, 1] \rightarrow [-1, \tau_d^p]$  is measure-preserving with respect to  $\mathcal{D}_d|_{(-1, \tau_d^p]}$  and  $\mathcal{D}_d|_{[\tau_d^p, 1]}$ . In particular,  $\phi \circ \phi = \text{identity}$ .

For  $b \in \{0, 1\}$ , define the map  $\rho_a^b(z)$  which (possibly) flips  $z \in \mathbb{S}^{d-1}$  around  $a \in \mathbb{S}^{d-1}$  so that  $b = \mathbf{1}[\langle a, z \rangle \geq \tau_d^p]$ . The vector  $z$  is transformed in the  $a$ -direction so that the projection on  $a$  changes to  $\phi(\langle a, z \rangle)$  instead of  $\langle a, z \rangle$  if  $\mathbf{1}[\langle \rho_a^b(z), a \rangle \geq \tau_d^p] \neq b$ . Formally,

$$\rho_a^b(z) := \begin{cases} z & \text{if } \mathbf{1}[\langle a, z \rangle \geq \tau_d^p] = b, \\ (z - \langle a, z \rangle a) \sqrt{\frac{1-\phi^2(\langle a, z \rangle)}{1-\langle a, z \rangle^2}} + \phi(\langle a, z \rangle) a & \text{if } \mathbf{1}[\langle a, z \rangle \geq \tau_d^p] \neq b. \end{cases} \quad (3)$$

To simplify notation, we will denote

$$\begin{aligned} \sqrt{\frac{1-\phi^2(\langle a, z \rangle)}{1-\langle a, z \rangle^2}} &:= 1 + \psi(\langle a, z \rangle), \\ s^b(\langle a, z \rangle) &:= \mathbf{1}[\mathbf{1}[\langle a, z \rangle \geq \tau_d^p] \neq b], \text{ and} \\ \kappa(\langle a, z \rangle) &:= (-\langle a, z \rangle - \psi(\langle a, z \rangle) \langle a, z \rangle + \phi(\langle a, z \rangle)). \end{aligned}$$

Thus,

$$\begin{aligned} \rho_a^b(z) &= z + \mathbf{1}[\mathbf{1}[\langle a, z \rangle \geq \tau_d^p] \neq b] \times (\psi(\langle a, z \rangle) z + \kappa(\langle a, z \rangle) a) \\ &= z + s^b(\langle a, z \rangle) \psi(\langle a, z \rangle) z + s^b(\langle a, z \rangle) \kappa(\langle a, z \rangle) a. \end{aligned} \quad (4)$$

We make several simple but key observations.

**Observation 4.1.** Let  $a \in \mathbb{S}^{d-1}$ . If  $b \sim \text{Bern}(p)$ ,  $z \sim \text{Unif}(\mathbb{S}^{d-1})$ , and  $a, b$ , and  $z$  are independent, then  $\rho_a^b(Z) \sim \text{Unif}(\mathbb{S}^{d-1})$ .

*Proof.* First, we will show that the projection  $\langle \rho_a^b(z), a \rangle$  of  $\rho_a^b(z)$  on  $a$  has distribution  $\mathcal{D}_d$ . By the independence of  $a$  and  $z$ , the distribution of the projection of  $z$  on  $a$  is  $\mathcal{D}_d$ , which is the same as the mixture  $(1-p)\mathcal{D}_d|_{[-1, \tau_d^p]} + p\mathcal{D}_d|_{[\tau_d^p, 1]}$ . As  $\phi$  measure-preservingly maps  $\mathcal{D}_d|_{[-1, \tau_d^p]}$  and  $\mathcal{D}_d|_{[\tau_d^p, 1]}$  to each other,  $\langle a, \rho_a^1(z) \rangle \sim \mathcal{D}_d|_{[\tau_d^p, 1]}$  and  $\langle a, \rho_a^0(z) \rangle \sim \mathcal{D}_d|_{[-1, \tau_d^p]}$ . As  $b$  is independent of  $a, z$ ,  $\langle a, \rho_a^b(z) \rangle$  has distribution  $(1-p)\mathcal{D}_d|_{[-1, \tau_d^p]} + p\mathcal{D}_d|_{[\tau_d^p, 1]} = \mathcal{D}_d$ . The distribution of  $\rho_a^b(z)$  in the  $a$  direction is correct.

The distribution in the orthogonal space to  $a$  has the same direction as  $z$  (which is uniform in this orthogonal space). The norm is uniquely determined by  $\langle \rho_a^b(z), a \rangle$  as  $\rho_a^b(z)$  is unit norm. By the previous paragraph,  $\langle \rho_a^b(z), a \rangle$  has the correct distribution, hence so does the norm and so too the distribution in the orthogonal space.  $\square$

**Observation 4.2.** If  $b \sim \text{Bern}(p)$ ,  $z \sim \text{Unif}(\mathbb{S}^{d-1})$  and  $a, b, z$  are independent, then  $\mathbb{E}[s^b(\langle a, z \rangle)] = p$ .

*Proof.*  $b$  and  $\mathbf{1}[\langle a, z \rangle \geq \tau_d^p]$  are i.i.d.  $\text{Bern}(p)$ .  $\square$

**Proposition 4.1.** Suppose that  $|\langle a, z \rangle| \leq \frac{100C_A \sqrt{\log n}}{\sqrt{d}}$  where  $C_A$  is the constant from [Lemma 2.1](#). Then, for some absolute constant  $C$ :

1.  $|\phi(\langle a, z \rangle)| \leq \frac{C \sqrt{\log n}}{\sqrt{d}}$ ,
2.  $|\psi(\langle a, z \rangle)| \leq \frac{C \log n}{d}$ , and
3.  $|\kappa(\langle a, z \rangle)| \leq \frac{C \sqrt{\log n}}{\sqrt{d}}$ .

*Proof.* For item 1, if  $\langle a, z \rangle \geq \tau_d^p$ , then clearly  $|\phi(\langle a, z \rangle)| \leq |\langle a, z \rangle|$ . Else,  $|\phi(\langle a, z \rangle)| \leq \left| \phi\left(-\frac{100C_A\sqrt{\log n}}{\sqrt{d}}\right) \right|$ . Note that

$$\mathbb{P}_{x \sim \mathcal{D}_d |_{[-1, \tau_d^p]}}[x \leq -\frac{100C_A\sqrt{\log n}}{\sqrt{d}}] = \mathbb{P}_{x \sim \mathcal{D}_d}[x \leq -\frac{100C_A\sqrt{\log n}}{\sqrt{d}}]/(1-p) \geq n^{-K},$$

where  $K$  is a constant independent of  $n, d$ . Hence,  $\mathbb{P}_{x \sim \mathcal{D}_d |_{[\tau_d^p, 1]}}[x \geq \phi(-\frac{100C_A\sqrt{\log n}}{\sqrt{d}})] \geq n^{-K}$ , so  $\mathbb{P}_{x \sim \mathcal{D}_d}[x \geq \phi(-\frac{100C_A\sqrt{\log n}}{\sqrt{d}})] \geq pn^{-K} \geq n^{-K-2}$ . This is enough by Lemma 2.1.

For item 2,

$$\sqrt{\frac{1 - \phi^2(\langle a, z \rangle)}{1 - \langle a, z \rangle^2}} \in \left[ \sqrt{1 - 2(\phi^2(\langle a, z \rangle) + \langle a, z \rangle^2)}, \sqrt{1 + 2(\phi^2(\langle a, z \rangle) + \langle a, z \rangle^2)} \right].$$

As  $2(\phi^2(\langle a, z \rangle) + \langle a, z \rangle^2) \leq C_1 \log n/d$  for some constant  $C_1$ , we have that  $\left| \sqrt{\frac{1 - \phi^2(\langle a, z \rangle)}{1 - \langle a, z \rangle^2}} - 1 \right| = O(\log n/d)$ .

Item 3 Follows from the definition of  $\kappa$  and items 1 and 2.  $\square$

## 4.2 Proof of Lemma 3.1

*Proof of Lemma 3.1.* The proof will proceed by induction on  $(s, i_s)$  in the lexicographic order. Suppose that we have proven the statement for some  $(s, \ell)$ . We consider two cases.

**Case 1:** The next lexicographic term is  $(s+1, 0)$ , i.e.  $\ell = s-1$ . Hence,  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_s \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$  by induction. By definition,  $V_{s+1}^{(0)}$  is independent of these vectors and also uniform on  $\mathbb{S}^{d-1}$ .

**Case 2:** Now, suppose that  $\ell < s-1$ . Hence,  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{s-1}, V_s^{(\ell)} \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$ . We need to prove the statement for  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{s-1}, V_s^{(\ell+1)} = \rho_{\tilde{V}_{\ell+1}}^{H_{s,\ell+1}}(V_s^{(\ell)})$ . This follows from Observation 4.1 since  $H_{s,\ell+1}, V_s^{(\ell)}$  are independent of  $\{\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{s-1}\}$  (and  $\ell \leq s-2$ ).  $\square$

Using Lemma 2.1, we immediately conclude the following statement.

**Corollary 4.3.** *With high probability  $1 - \frac{1}{n^{100}}$ , for all  $1 \leq i < j \leq n$ ,  $0 \leq u \leq i-1$ ,  $0 \leq v \leq j-1$ ,*

$$|\langle V_i^{(u)}, V_j^{(v)} \rangle| \leq \frac{100C_A\sqrt{\log n}}{\sqrt{d}},$$

and all of the remaining bounds in Proposition 4.1 also hold.

## 4.3 Proof of Lemma 3.2

**Step 1: Expanding the telescoping sum.** We begin by expanding the terms of Eq. (2) using the expression in Eq. (4). The goal is to isolate the ‘‘martingale violations’’ discussed in (V1) and (V2).

$$\begin{aligned}
& \sum_{u=i+1}^{j-1} \langle \tilde{V}_i, V_j^{(u)} - V_j^{(u-1)} \rangle \\
&= \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \left( \psi(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_i, V_j^{(u-1)} \rangle + \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, \tilde{V}_i \rangle \right) \\
&= \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \psi(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_i, V_j^{(u-1)} \rangle \quad (\text{Violations 1}) \\
&\quad + \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, V_j^{(u-1)} \rangle \langle V_j^{(u-1)}, \tilde{V}_i \rangle \quad (\text{Violations 2}) \\
&\quad + \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp} V_j^{(u-1)} \rangle. \quad (\text{Martingale Part})
\end{aligned}$$

Above,  $(\tilde{V}_i)^{\perp} V_j^{(u-1)}$  denotes the part of  $\tilde{V}_i$  perpendicular to  $V_j^{(u-1)}$ .

We have decomposed the sum into three parts. The first one corresponds to martingale violations in **(V1)**, the second to the violations in **(V2)**. These will be bounded in Steps 2 through 4 below. The last part is a true martingale and will be analyzed in Steps 5 and 6.

To bound the martingale violations, we first show that with high probability there are few nonzero terms and then show that each of these terms is small.

## Step 2: The number of non-zero terms.

**Lemma 4.4.** *With high probability  $1 - 1/n^{100}$ , the following inequality holds for all  $i, j$  simultaneously:*

$$\sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \leq 200C_A \max(np, \log n).$$

*Proof.* We first show that  $\left\{ M_u = s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) - p \right\}_{u=i+1}^{j-1}$  is a martingale difference sequence with respect to the filtration

$$\mathcal{F}_u := \sigma \left\{ \tilde{V}_i, \tilde{V}_{i+1}, \dots, \tilde{V}_u, H_{j,i+1}, H_{j,i+2}, \dots, H_{j,u}, V_j^{(i)} \right\}. \quad (5)$$

Measurability with respect to this filtration holds because  $V_j^{(u-1)}$  is a deterministic function of  $V_j^{(i)}, \tilde{V}_i, \dots, \tilde{V}_{u-1}$  and  $H_{j,i+1}, H_{j,i+2}, \dots, H_{j,u-1}$ . If  $H_{ju}, \tilde{V}_u$  are independent of  $\mathcal{F}_{u-1}$ , the zero expectation follows immediately from [Observation 4.2](#) as  $\mathbb{E} [s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) - p | \mathcal{F}_{u-1}] = 0$ . The independence of  $H_{ju}, \tilde{V}_u$  holds because the sigma algebra  $\mathcal{F}_{u-1}$  is coarser than the sigma algebra

$$\sigma \{ \tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{u-1}, H_{j,0}, H_{j,1}, \dots, H_{j,u-1}, V_j^{(0)} \}.$$

However,  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{u-1}, \tilde{V}_u, V_j^{(0)}$  are independent by [Lemma 3.1](#). And also  $H_{j,0}, \dots, H_{j,u}$  are independent from  $\tilde{V}_i, \tilde{V}_{i+1}, \dots, \tilde{V}_{u-1}, \tilde{V}_u, V_j^{(0)}$ . Hence, the collection

$$\{ \tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_{u-1}, \tilde{V}_u, V_j^{(0)}, H_{j,0}, H_{j,1}, \dots, H_{j,u-1}, H_{j,u} \}$$

is jointly independent and so  $H_{ju}, \tilde{V}_u$  are independent of the sigma algebra generated by  $\{\tilde{V}_i, \tilde{V}_{i+1}, \dots, \tilde{V}_{u-1}, H_{j,0}, H_{j,1}, \dots, H_{j,u-1}, V_j^{(0)}\}$ , which makes  $V_j^{(i-1)}$  measurable.

Going back to the problem, we have the martingale difference sum

$$Q = \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) - p.$$

Each term is bounded by 1 and has conditional variance  $p - p^2 \leq p$ . By [Theorem 2.2](#) for  $L = np$  and  $a = 1$ , the inequality  $\mathbb{P}[|Q| \geq z] \leq 2 \exp(-\frac{z^2/2}{np+z/3})$  holds. Take  $z = 100 \max(\sqrt{np \log n}, \log n)$ . The number of non-zero terms is  $Q + np \leq np + 100C_A \max(\sqrt{np \log n}, \log n) \leq 200C_A \max(np, \log n)$ .  $\square$

**Step 3: Martingale violations (V1).** We now deal with the term [\(Violations 1\)](#). Recall from [\(V1\)](#) that this term appears due to the rescaling in the direction orthogonal to  $\tilde{V}_u$  when flipping  $V_j^{(u-1)}$  with respect to  $\tilde{V}_u$ . For each  $u$ , by [Proposition 4.1](#) and [Corollary 4.3](#), with high probability  $1 - 1/n^{100}$

$$|\psi(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_i, V_j^{(u-1)} \rangle| \leq C' \frac{\log^{3/2} n}{d^{3/2}}$$

for some absolute constant  $C'$ . Combining with [Lemma 4.4](#), with probability  $1 - O(1/n^{100})$ ,

$$\begin{aligned} & \left| \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \psi(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_i, V_j^{(u-1)} \rangle \right| \\ & \leq \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) |\psi(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_i, V_j^{(u-1)} \rangle| \leq C' \max(np, \log n) \frac{\log^{3/2} n}{d^{3/2}} \end{aligned}$$

When  $d = \Omega(\max(np, \log n))$ , we have  $d^{-3/2} \max(np, \log n) \log^{3/2} n = O(d^{-1} \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2})$ .

**Step 4: Martingale violations (V2).** Next, we consider the term in [\(Violations 2\)](#). As in [\(V2\)](#), it appears due to fact that  $V_j^{(u-1)}$  and  $\tilde{V}_i$  are not independent. By [Proposition 4.1](#) and [Corollary 4.3](#), with high probability  $1 - 1/n^{100}$ ,

$$|\kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, V_j^{(u-1)} \rangle \langle V_j^{(u-1)}, \tilde{V}_i \rangle| \leq C' \frac{\log^{3/2} n}{d^{3/2}}.$$

Using [Lemma 4.4](#) we conclude that with probability  $1 - O(1/n^{100})$ ,

$$\left| \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, V_j^{(u-1)} \rangle \langle V_j^{(u-1)}, \tilde{V}_i \rangle \right| \leq \frac{C' \max(np, \log n) \log^{3/2} n}{d^{3/2}}.$$

Again, when  $d = \Omega(\max(np, \log n))$ , this expression is of order  $O(\max(\sqrt{np}, \sqrt{\log n}) \frac{(\log n)^{3/2}}{d})$ .

**Step 5: Truncating the martingale differences.** The remaining sum

$$\sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle$$

is a sum of martingale differences. However, we need to truncate it so that we have a small uniform bound to apply Freedman's inequality [Theorem 2.2](#). To this end, we write

$$\begin{aligned}
& \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \\
&= \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \times \\
&\quad \times \mathbb{1} \left[ |\langle \tilde{V}_u, V_j^{(u-1)} \rangle| \leq \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \right] \mathbb{1} \left[ |\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle| \leq \frac{200C_A \sqrt{\log n}}{\sqrt{d}} \right] \\
&\quad + \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) (V_j^{(u-1)}) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \\
&\quad \times \mathbb{1} \left[ |\langle \tilde{V}_u, V_j^{(u-1)} \rangle| > \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \text{ or } |\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle| > \frac{200C_A \sqrt{\log n}}{\sqrt{d}} \right]
\end{aligned}$$

Observe that

$$\begin{aligned}
|\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle| &= |\langle \tilde{V}_u, \tilde{V}_i \rangle - \langle \tilde{V}_u, V_j^{(u-1)} \rangle \langle V_j^{(u-1)}, \tilde{V}_i \rangle| \\
&\leq |\langle \tilde{V}_u, \tilde{V}_i \rangle| + |\langle \tilde{V}_u, V_j^{(u-1)} \rangle \langle V_j^{(u-1)}, \tilde{V}_i \rangle| \leq |\langle \tilde{V}_u, \tilde{V}_i \rangle| + |\langle \tilde{V}_u, V_j^{(u-1)} \rangle|.
\end{aligned}$$

Hence, by [Corollary 4.3](#),

$$\mathbb{P} \left[ |\langle \tilde{V}_u, V_j^{(u-1)} \rangle| > \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \text{ or } |\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle| > \frac{200C_A \sqrt{\log n}}{\sqrt{d}} \right] \leq 1 - \frac{1}{n^{100}}.$$

Thus, with probability  $1 - O(1/n^{100})$ , the second sum is 0 for all  $i, j$ .

**Step 6: Martingale analysis.** All that is left to bound is the sum

$$\begin{aligned}
& \sum_{u=i+1}^{j-1} s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \times \\
&\quad \times \mathbb{1} \left[ |\langle \tilde{V}_u, V_j^{(u-1)} \rangle| \leq \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \right] \mathbb{1} \left[ |\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle| \leq \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \right]. \tag{6}
\end{aligned}$$

Let  $K_u$  be the term corresponding to  $u$  above. We claim that  $\{K_u\}_{u=i+1}^{j-1}$  is a martingale difference sequence with respect to the filtration in [Eq. \(5\)](#). Measurability follows in the same as way as for  $\{M_u\}_{u=i+1}^{j-1}$  defined in [Lemma 4.4](#). We also need to show the martingale property.

**Step 6.1. Martingale property.** For brevity of notation, denote

$$\begin{aligned}
\mathbb{1}(u, j) &:= \mathbb{1} \left[ |\langle \tilde{V}_u, V_j^{(u-1)} \rangle| \leq \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \right] \\
\mathbb{1}(u, i, \perp j) &:= \mathbb{1} \left[ |\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle| \leq \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \right].
\end{aligned}$$

We need to show that

$$\mathbb{E} \left[ s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \mathbb{1}(u, j) \mathbb{1}(u, i, \perp j) \middle| \mathcal{F}_{u-1} \right] = 0.$$

This follows immediately from the fact that  $\tilde{V}_u$  is independent of  $\mathcal{F}_{u-1}, H_{ju}$ . Hence,

$$\begin{aligned} & \mathbb{E} \left[ s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \mathbf{1}(u, j) \mathbf{1}(u, i, \perp j) \middle| \mathcal{F}_{u-1} \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \times \right. \right. \\ &\quad \times \mathbf{1}(u, j) \mathbf{1}(u, i, \perp j) \left. \middle| \mathcal{F}_{u-1}, H_{ju}, \langle \tilde{V}_u, V_j^{(u-1)} \rangle \right] \left. \middle| \mathcal{F}_{u-1} \right] \\ &= \mathbb{E} \left[ s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \mathbf{1}(u, j) \times \right. \\ &\quad \times \mathbb{E} \left[ \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \mathbf{1}(u, i, \perp j) \middle| \mathcal{F}_{u-1}, H_{ju}, \langle \tilde{V}_u, V_j^{(u-1)} \rangle \right] \left. \middle| \mathcal{F}_{u-1} \right] \end{aligned}$$

We used the fact that  $s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle), \mathbf{1}(u, j)$  are functions of only  $H_{ju}, \langle \tilde{V}_u, V_j^{(u-1)} \rangle$ . Now, the inner expectation is equal to 0 as the map

$$V + \langle \tilde{V}_u, V_j^{(u-1)} \rangle V_j^{(u-1)} \longrightarrow -V + \langle \tilde{V}_u, V_j^{(u-1)} \rangle V_j^{(u-1)},$$

where  $V$  is the component of  $\tilde{V}_u$  orthogonal to  $V_j^{(u-1)}$ , is measure-preserving under the sigma algebra  $\sigma\{\mathcal{F}_{u-1}, H_{ju}, \langle \tilde{V}_u, V_j^{(u-1)} \rangle\}$  and acts by

$$\begin{aligned} & \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \mathbf{1} \left[ |\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle| \leq \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \right] \\ & \longrightarrow -\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \mathbf{1} \left[ |\langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle| \leq \frac{100C_A \sqrt{\log n}}{\sqrt{d}} \right]. \end{aligned}$$

**Step 6.2. Martingale concentration.** We end by applying [Theorem 2.2](#). Recall that

$$K_u = s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle \mathbf{1}(u, j) \mathbf{1}(u, i, \perp j)$$

is bounded deterministically by  $C'' \frac{\log n}{d}$  for some absolute constant  $C''$  due to the indicators and the bound on  $\kappa$  from [Proposition 4.1](#). The conditional second moment is at most  $(C'')^2 p \times \frac{\log^2 n}{d^2}$ :

$$\begin{aligned} & \mathbb{E}[K_u^2 | \mathcal{F}_{u-1}] \\ &= \mathbb{E} \left[ \left( s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \right)^2 \kappa(\langle \tilde{V}_u, V_j^{(u-1)} \rangle)^2 \langle \tilde{V}_u, (\tilde{V}_i)^{\perp V_j^{(u-1)}} \rangle^2 \mathbf{1}(u, j) \mathbf{1}(u, i, \perp j) \middle| \mathcal{F}_{u-1} \right] \\ &\leq \mathbb{E} \left[ \left( s^{H_{ju}}(\langle \tilde{V}_u, V_j^{(u-1)} \rangle) \right)^2 (C'')^2 \frac{\log^2 n}{d^2} \middle| \mathcal{F}_{u-1} \right] \leq (C'')^2 p \times \frac{\log^2 n}{d^2}. \end{aligned}$$

Altogether, by [Theorem 2.2](#),

$$\begin{aligned} \mathbb{P} \left[ \sum_{u=i+1}^{j-1} K_u \geq z \right] &\leq \exp \left( - \frac{z^2/2}{(C'')^2 (j-i-1) pd^{-2} \log^2 n + z C'' d^{-1} \log n / 3} \right) \\ &= \exp \left( - \frac{z^2}{O(np \log^2 n / d^2) + O(z \log n / d)} \right). \end{aligned}$$

Taking  $z = \frac{c \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}$  for a large enough  $c$  makes this last expression  $O(1/n^{100})$ .

#### 4.4 Deriving [Corollary 1.1](#) from [Theorem 1.1](#)

*Proof of [Corollary 1.1](#).* Suppose that one is given the coupling in [Theorem 1.1](#) of  $H \sim \mathsf{G}(n, p)$  with vectors  $\tilde{V}_1, \tilde{V}_2, \dots, \tilde{V}_n \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$  such that the second condition holds. Thus, with high probability, all edges of the random geometric graph  $G_+$  defined by

$$(G_+)_{ij} = \mathbf{1} \left[ \langle \tilde{V}_i, \tilde{V}_j \rangle \geq \tau_d^p - c \max(\sqrt{np}, \sqrt{\log n}) \frac{(\log n)^{3/2}}{d} \right]$$

are also edges in  $H$  defined by  $H_{ji}$ . Note that  $G_+$  is a sample from  $\text{RGG}(n, \mathbb{S}^{d-1}, p_+)$  where

$$p_+ := \mathbb{P}_{U, V \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})} [\langle U, V \rangle \geq \tau_d^p - c \max(\sqrt{np}, \sqrt{\log n}) \frac{(\log n)^{3/2}}{d}].$$

From [Lemma 2.1](#),  $p_+ = p\left(1 + O\left(\max(\sqrt{np}, \sqrt{\log n}) \frac{(\log n)^2}{\sqrt{d}}\right)\right)$ . Similarly, we consider

$$(G_-)_{ij} = \mathbf{1}\left[\langle \tilde{V}_i, \tilde{V}_j \rangle \geq \tau_d^p + \max(\sqrt{np}, \sqrt{\log n}) \frac{(\log n)^{3/2}}{d}\right].$$

Hence, we have a coupling of  $H \sim \mathsf{G}(n, p)$ ,  $G_+ \sim \text{RGG}\left(n, d, p\left(1 + O\left(\max(\sqrt{np}, \sqrt{\log n}) \frac{(\log n)^2}{\sqrt{d}}\right)\right)\right)$  and  $G_- \sim \text{RGG}\left(n, d, p\left(1 - O\left(\max(\sqrt{np}, \sqrt{\log n}) \frac{(\log n)^2}{\sqrt{d}}\right)\right)\right)$  such that  $G_- \subseteq H \subseteq G_+$  with high probability. This is clearly enough.  $\square$

## 5 Robust Testing

### 5.1 The SoS Program and Refutation

We now give the details for the SoS program. Consider the following axiom set over  $n$   $d$ -dimensional variables  $(Y_i)_{i=1}^n$  stacked together in a  $n \times d$  matrix  $Y$  and indicator variables  $w = (w_{ij})_{1 \leq i < j \leq n}$ :

$$\mathcal{A}(Y, w) := \begin{cases} \langle Y_i, Y_i \rangle = 1 \forall i, \\ w_{i,j}^2 = w_{i,j} \forall i \neq j, \\ (1 - w_{i,j})\langle Y_i, Y_j \rangle^2 \leq (\tau_d^p)^2 \forall i \neq j, \\ \sum_{i,j} w_{i,j} \langle Y_i, Y_j \rangle^2 \leq \frac{2C_2 n^2 p \log(1/p)}{d}, \end{cases} \quad (7)$$

where  $C_2$  is the constant from [Lemma 2.1](#). One should interpret the condition imposed by variables  $w_{ij}$  as a tail bound on the empirical distribution of  $\{\langle Y_i, Y_j \rangle^2\}_{1 \leq i < j \leq n}$ . The constraint essentially reads  $\sum_{i,j : \langle Y_i, Y_j \rangle^2 \geq (\tau_d^p)^2} \langle Y_i, Y_j \rangle^2 \leq \frac{2C_2 n^2 p \log(1/p)}{\sqrt{d}}$ , which holds w.h.p. when one takes  $Y_i = V_i$ .

**Spectral Refutation Algorithm.** On input a  $p$ -centered adjacency matrix  $A = A_G$ , try to prove in degree 12 SoS that  $\left(\sum_{i,j} A_{i,j} \langle Y_i, Y_j \rangle\right)^2 \leq \frac{n^4 p^2 \log(1/p)}{8C_B^2 d}$ , where  $C_B$  is the constant from [Lemma 2.1](#). If the algorithm fails, report that  $G$  is a corrupted  $\text{RGG}(n, d, p)$ . If it succeeds, report that it is a corruption of a graph with small operator norm of its  $p$ -centered adjacency matrix (e.g.  $\mathsf{G}(n, p)$  for concreteness).

### 5.2 Analysis of Corrupted Erdős-Rényi

Let  $A$  be the centered adjacency matrix of a corrupted  $\mathsf{G}(n, p)$ . Let  $A = U + S$ , where  $U$  is the uncorrupted adjacency matrix and  $S$  is defined by  $S = A - U$ . The entries of  $S$  are in  $\{-1, 0, 1\}$  and at most  $\epsilon p \binom{n}{2}$  of them are non-zero ( $S$  is supported on corrupted entries). Hence,

$$\begin{aligned} \left(\sum_{i,j} A_{i,j} \langle Y_i, Y_j \rangle\right)^2 &= \left(\sum_{i,j} U_{i,j} \langle Y_i, Y_j \rangle + \sum_{i,j} S_{i,j} \langle Y_i, Y_j \rangle\right)^2 \\ &= \left(\sum_{i,j} U_{i,j} \langle Y_i, Y_j \rangle + \sum_{i,j} S_{i,j} w_{i,j} \langle Y_i, Y_j \rangle + \sum_{i,j} S_{i,j} (1 - w_{i,j}) \langle Y_i, Y_j \rangle\right)^2 \\ &\leq 4 \left(\sum_{i,j} U_{i,j} \langle Y_i, Y_j \rangle\right)^2 + 4 \left(\sum_{i,j} S_{i,j} w_{i,j} \langle Y_i, Y_j \rangle\right)^2 + 4 \left(\sum_{i,j} S_{i,j} (1 - w_{i,j}) \langle Y_i, Y_j \rangle\right)^2, \end{aligned} \quad (8)$$

where so far all inequalities are in SoS. We bound each of the three terms separately.

**Case 1) Uncorrupted value:** Note that  $\sum_{i,j} U_{i,j} \langle Y_i, Y_j \rangle = \langle U, Y^T Y \rangle$ . Since  $U$  is the signed adjacency matrix of  $G(n, p)$ , with high probability  $\lambda_{\max}(U) \leq C \max(\sqrt{np}, \log n)$  for some absolute constant  $C$  by [Theorem 2.1](#). Hence,  $C \max(\sqrt{np}, \log n)I - U \succeq 0$  and, thus, in SoS,

$$\langle U, YY^T \rangle \leq \langle C \max(\sqrt{np}, \log n)I, YY^T \rangle = C \max(\sqrt{np}, \log n) \times \text{tr}(YY^T) = C \max(\sqrt{np}, \log n) \times n.$$

Similarly,  $\langle U, YY^T \rangle \geq -C \max(\sqrt{np}, \log n) \times n$ . Thus, in SoS,  $\left( \sum_{i,j} U_{i,j} \langle Y_i, Y_j \rangle \right)^2 \leq C^2 \max(n^3 p, n^2 \log^2 n)$ .

**Case 2)** Corrupted value with large entries. By Cauchy-Schwartz inequality,

$$\begin{aligned} & \left( \sum_{i,j} S_{i,j} w_{i,j} \langle Y_i, Y_j \rangle \right)^2 \\ & \leq \left( \sum_{i,j} S_{i,j}^2 \right) \left( \sum_{i,j} w_{i,j}^2 \langle Y_i, Y_j \rangle^2 \right) \\ & \leq \left( \sum_{i,j : (ij) \text{ corrupted}} 1 \right) \left( \sum_{i,j} w_{i,j} \langle Y_i, Y_j \rangle^2 \right) \\ & \leq \epsilon p n^2 \times \frac{4C_2^2 n^2 p \log(1/p)}{d}. \end{aligned}$$

**Case 3)** Corrupted value with small entries. Again, by Cauchy-Schwartz,

$$\begin{aligned} & \left( \sum_{i,j} S_{i,j} (1 - w_{i,j}) \langle Y_i, Y_j \rangle \right)^2 \\ & = \left( \sum_{i,j} S_{i,j}^3 (1 - w_{i,j}) \langle Y_i, Y_j \rangle \right)^2 \\ & \leq \left( \sum_{i,j} S_{i,j}^2 \right) \left( S_{i,j}^4 \sum_{i,j} (1 - w_{i,j})^2 \langle Y_i, Y_j \rangle^2 \right) \\ & \leq \epsilon p n^2 \times \sum_{1 \leq i < j \leq n} S_{i,j}^4 (\tau_d^p)^2 \\ & \leq \epsilon p n^2 \times \epsilon p n^2 \times (\tau_d^p)^2 \leq \frac{C_A^2 \epsilon^2 n^4 p^2 \log(1/p)}{d}, \end{aligned}$$

using [Lemma 2.1](#). Altogether, for some absolute constant  $C'$ , when  $A$  is an  $\epsilon$ -corruption of  $G(n, p)$ , with high probability there is a degree 12 proof in SoS that

$$\left( \sum_{i,j} A_{i,j} \langle Y_i, Y_j \rangle \right)^2 \leq C' \left( \max(n^3 p, n^2 \log^2 n) + \frac{\epsilon n^4 p^2 \log(1/p)}{d} \right). \quad (9)$$

### 5.3 Analysis of Corrupted RGG

We will show that with high probability over  $\text{RGG}(n, d, p)$ , for no  $\epsilon$ -corruption  $A$  can one prove  $\left( \sum_{i,j} A_{i,j} \langle Y_i, Y_j \rangle \right)^2 \leq \frac{n^4 p^2 \log(1/p)}{d}$ . We simply need to exhibit a pair  $(Y, w)$  that satisfies the axioms  $\mathcal{A}(Y, w)$  and  $\left( \sum_{i,j} A_{i,j} \langle Y_i, Y_j \rangle \right)^2 \geq \frac{n^4 p^2 \log(1/p)}{8C_B^2 d}$ .

Let  $Y_i = V_i, 1 \leq i \leq n$ , where this is the true embedding that produced  $A$ . Let  $w_{i,j} = \mathbb{1}[\langle V_i, V_j \rangle^2 \geq (\tau_d^p)^2]$ .

**Feasibility.** Clearly, the first three axioms are satisfied. For the last one,

$$\begin{aligned}
& \mathbb{E}_{V_1, V_2, \dots, V_n \sim_{iid} \text{Unif}(\mathbb{S}^{d-1})} \left[ \sum_{i \neq j} \mathbb{1}[\langle V_i, V_j \rangle^2 \geq (\tau_d^p)^2] \times \langle V_i, V_j \rangle^2 \right] \\
& \leq n^2 \mathbb{E}_{V_1, V_2 \sim_{iid} \text{Unif}(\mathbb{S}^{d-1})} \left[ \mathbb{1}[\langle V_1, V_2 \rangle^2 \geq (\tau_d^p)^2] \times \langle V_1, V_2 \rangle^2 \right] \\
& = n^2 \mathbb{E}_{V_1, V_2 \sim_{iid} \text{Unif}(\mathbb{S}^{d-1})} \left[ \langle V_1, V_2 \rangle^2 | \langle V_1, V_2 \rangle^2 \geq (\tau_d^p)^2 \right] \times \mathbb{P}[\langle V_1, V_2 \rangle^2 \geq (\tau_d^p)^2] \\
& \leq n^2 p \times \frac{C_2^2 \log(1/p)}{d}
\end{aligned}$$

by [Lemma 2.1](#). Furthermore, the value concentrates extremely well as the collection of random variables  $\{\mathbb{1}[\langle V_i, V_j \rangle^2 \geq (\tau_d^p)^2] \times \langle V_i, V_j \rangle^2\}_{1 \leq i < j \leq n}$  is pairwise independent! (Note that  $\langle V_1, V_2 \rangle$  and  $\langle V_2, V_3 \rangle$  are independent.) Hence,

$$\begin{aligned}
& \mathbb{V}_{V_1, V_2, \dots, V_n \sim_{iid} \text{Unif}(\mathbb{S}^{d-1})} \left[ \sum_{i \neq j} \mathbb{1}[\langle V_i, V_j \rangle^2 \geq (\tau_d^p)^2] \times \langle V_i, V_j \rangle^2 \right] \\
& \leq n^2 \mathbb{V}_{V_1, V_2 \sim_{iid} \text{Unif}(\mathbb{S}^{d-1})} \left[ \mathbb{1}[\langle V_1, V_2 \rangle^2 \geq (\tau_d^p)^2] \times \langle V_1, V_2 \rangle^2 \right] \\
& \leq n^2 \mathbb{E}_{V_1, V_2 \sim_{iid} \text{Unif}(\mathbb{S}^{d-1})} \left[ \langle V_1, V_2 \rangle^4 | \langle V_1, V_2 \rangle^2 \geq (\tau_d^p)^2 \right] \times \mathbb{P}[\langle V_1, V_2 \rangle^2 \geq (\tau_d^p)^2] \\
& \leq n^2 p \times \frac{C_4 \log(1/p)^2}{d^2} = o\left(\left(n^2 p \times \frac{C_2^2 \log(1/p)}{d}\right)^2\right).
\end{aligned}$$

Hence,  $\sum_{i \neq j} \mathbb{1}[\langle V_i, V_j \rangle^2 \geq (\tau_d^p)^2] \times \langle V_i, V_j \rangle^2 \leq \frac{2C_2^2 n^2 p \log(1/p)}{d}$  w.h.p. by Chebyshev's inequality.

**Objective Value.** Finally, we need to bound  $\sum_{i,j} A_{i,j} \langle V_i, V_j \rangle$ . As in the Erdős-Rényi case,

$$\begin{aligned}
\left| \sum_{i,j} S_{i,j} \langle V_i, V_j \rangle \right| & \leq 2 \sqrt{\left( \sum_{i,j} S_{i,j} w_{i,j} \langle V_i, V_j \rangle \right)^2 + \left( \sum_{i,j} S_{i,j} (1 - w_{i,j}) \langle Y_i, Y_j \rangle \right)^2} \\
& \leq 2 \sqrt{\epsilon p n^2 \times \frac{4C_2^2 n^2 p \log(1/p)}{d} + \frac{C_A^2 \epsilon^2 n^4 p^2 \log(1/p)}{d}} \leq \frac{C'' \sqrt{\epsilon} n^2 p \sqrt{\log(1/p)}}{\sqrt{d}}.
\end{aligned}$$

Hence, we only need to show that  $\sum_{i,j} U_{i,j} \langle V_i, V_j \rangle$  is large. Again, as the different summands are pairwise independent and this value concentrates well, we only consider the expectation:

$$\begin{aligned}
& \mathbb{E} \left[ \sum_{i,j} U_{i,j} \langle V_i, V_j \rangle \right] \\
& = n(n-1) \mathbb{E} [\mathbb{1}[\langle V_i, V_j \rangle \geq \tau_d^p] \langle V_i, V_j \rangle] \\
& = n(n-1) \mathbb{E} [\langle V_i, V_j \rangle | \langle V_i, V_j \rangle \geq \tau_d^p] \times \mathbb{P}[\langle V_i, V_j \rangle \geq \tau_d^p] \\
& \geq n(n-1)p \mathbb{E} [\langle V_i, V_j \rangle | \langle V_i, V_j \rangle \geq \tau_d^p] \geq \frac{n^2 p \sqrt{\log(1/p)}}{2C_B \sqrt{d}},
\end{aligned}$$

the last inequality by [Lemma 2.1](#). Hence,  $\sum_{i,j} A_{i,j} \langle V_i, V_j \rangle \geq \frac{n^2 p \sqrt{\log(1/p)}}{2C_B \sqrt{d}} - \frac{C'' \sqrt{\epsilon} n^2 p \sqrt{\log(1/p)}}{\sqrt{d}}$  and so  $\left( \sum_{i,j} A_{i,j} \langle V_i, V_j \rangle \right)^2 \geq \frac{n^4 p^2 \log(1/p)}{8C_B^2 d}$  for all  $\epsilon \leq K'_r$  for some absolute constant  $K'_r$ .

## 5.4 Putting It All Together

By the previous two sections, as long as  $\frac{n^4 p^2 \log(1/p)}{8C_B^2 d} \geq C'(\max(n^3 p, n^2 \log^2 n) + \frac{\epsilon n^4 p^2 \log(1/p)}{d})$ , the test succeeds with high probability. This occurs if  $\epsilon \leq K_r$  and  $d \leq K_r \min(np \log(1/p), n^2 p^2 \log(1/p) \log^{-2} n)$  for some absolute constant  $C$ .

**Remark 6.** The only property needed about the  $G(n, p)$  distribution in [Theorem 1.3](#) is that for the centered adjacency matrix  $U$  of a sample from  $G(n, p)$ ,  $|\lambda_{\max}(U)| \leq c_1 \frac{n^3 p^2 \log(1/p)}{d}$ . Hence, our result holds for robust testing between  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  and *any* distribution  $\mathcal{G}$  for which the  $p$ -centered adjacency matrix of a sample from  $\mathcal{G}$  has operator norm at most  $c_1 \frac{n^3 p^2 \log(1/p)}{d}$  for some absolute constant  $c_1$ .

## 6 Enumerating Geometric Graphs and the Support Size of RGG

*Proof of Theorem 1.4.* Suppose that we are given some  $d \geq \log^7 n$  (otherwise, the theorem statement follows immediately from the 2-dimensional case by [\[MM11\]](#)).

Take  $p = d/(n \log^{6.5} n) = \Omega(1/n)$ . Now, [Theorem 1.1](#) implies that there exists a coupling of  $H \sim G(n, p)$  and  $V_1, V_2, \dots, V_n \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$  such that with high probability:

1.  $\mathbb{1}\left[\langle V_i, V_j \rangle \geq \tau_d^p - c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}\right] \geq H_{i,j}$  for all  $i, j$ .
2.  $\mathbb{1}\left[\langle V_i, V_j \rangle \geq \tau_d^p + c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}\right] \leq H_{i,j}$  for all  $i, j$ .

Thus, the random geometric graph  $G$  defined by  $G_{i,j} = \mathbb{1}\left[\langle V_i, V_j \rangle \geq \tau_d^p - c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}\right]$  has  $H$  as an edge-subgraph.  $G$  is a sample from  $\text{RGG}(n, \mathbb{S}^{d-1}, p(1 + \delta))$ , where

$$p\delta = \mathbb{P}_{X \sim \mathcal{D}_d} \left[ X \in \left[ \tau_d^p - c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}, \tau_d^p \right] \right].$$

By Part D) of [Lemma 2.1](#),  $\delta = O\left(\frac{\max(\sqrt{np}, \log n) \log n^{3/2} \log(1/p)^{1/2}}{\sqrt{d}}\right) = o(1)$ .

Furthermore, as  $\mathbb{1}\left[\langle V_i, V_j \rangle \geq \tau_d^p + c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}\right] \leq H_{i,j}$  for all  $i, j$ , it is the case that  $G \setminus H$  is a set of edges which belongs to the edge-set defined by

$$\left\{ (i, j) : \langle V_i, V_j \rangle \in \left[ \tau_d^p - c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}, \tau_d^p + c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d} \right] \right\}. \quad (10)$$

By Part D) of [Lemma 2.1](#),  $q := \mathbb{P}\left[\langle V_i, V_j \rangle \in \left[ \tau_d^p - c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}, \tau_d^p + c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d} \right]\right] = O\left(p \frac{\max(\sqrt{np}, \log n) \log n^{3/2} \log(1/p)^{1/2}}{\sqrt{d}}\right) = o(p)$ . Hence, in expectation,  $G \setminus H$  is contained in a set of size  $o(n^2 p)$ . Note that the events  $\{\langle V_i, V_j \rangle \in [\tau_d^p - c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}, \tau_d^p + c \frac{\max(\sqrt{np}, \log n) \log n^{3/2}}{d}]\}_{i \neq j}$  are pairwise independent. Hence, by the second moment method, with high probability the set in [Eq. \(10\)](#) is of cardinality at most  $2 \max(q, 1/n^{3/2}) n^2 = o(pn^2)$ . We reach the following conclusion.

**Corollary 6.1.** *For  $G \sim \text{RGG}(n, \mathbb{S}^{d-1}, p(1 + \delta))$ ,  $H \sim G(n, p)$ , coupled as above, with high probability,  $H$  is an edge-subgraph of  $G$  and  $|E(G)| - |E(H)| \leq \epsilon n^2 p$  for some  $\epsilon = o(1)$ .*

This is enough to conclude that  $\text{RGG}(n, \mathbb{S}^{d-1}, p(1 + \delta))$  has support of size  $\exp(\Omega(n^2 p \log(1/p))) = \exp(\Omega(nd \log n^{-7}))$  using the following two nearly trivial observations proved in [Appendix B](#).

**Observation 6.2** ([\[BB23\]](#), Implicit in [Theorem 7.5.](#)). *Suppose that  $p \leq 1/2, p = \Omega(1/n)$ . Then, any set  $A$  such that  $P_{H \sim G(n, p)}[H \in A] \geq 1/2$  satisfies  $|A| \geq \exp(\frac{1}{6} n^2 p \log(1/p))$ .*

**Observation 6.3.** Suppose that  $p \leq 1/2$ ,  $p = \Omega(1/n)$  and  $\epsilon < 1/100$ . Let  $G$  be any graph on at most  $n^2p$  edges. Then, the number of edge-subgraphs  $H$  of  $G$  such that  $|E(G)| - |E(H)| \leq \epsilon n^2p$  is at most  $\exp(2n^2p\epsilon \log(1/\epsilon))$ .

With probability more than  $1/2$ , it is the case that  $H$  is an edge-subgraph of  $G$  by Corollary 6.1 and  $G$  has at most  $2\binom{n}{2}p \leq n^2p$  edges (again, edges of random geometric graphs are pairwise-independent). Let  $B$  be the support of  $\text{RGG}(n, \mathbb{S}^{d-1}, p(1 + \delta))$  restricted to graphs on at most  $n^2p$  edges. Hence, the number of edge-subgraphs of graphs in  $B$  which differ in at most  $\epsilon n^2p$  edges is at most  $|B| \times \exp(2n^2p\epsilon \log(1/\epsilon)) = |B| \times \exp(o(n^2p \log(1/p)))$  by Observation 6.3. By Observation 6.2,  $|B| \times \exp(o(n^2p \log(1/p))) \geq \exp(\frac{1}{6}n^2p \log(1/p))$ . Hence,  $|B| \geq \exp(\frac{1}{7}n^2p \log(1/p)) = \exp(\Omega(n^2d \log^{-7} n))$ .  $\square$

**Remark 7** (On The Support Size of The RGG distribution). The argument above shows that whenever  $d \geq \max(np, \log^2 n) \log(n)^3 \log(1/p)$ , the RGG distribution is supported on  $\exp(\Omega(n^2p \log(1/p)))$  graphs. Whenever  $d \leq np \log(1/p) \log(n)^{-1}$ , it is supported on  $\exp(o(n^2p \log(1/p)))$  as the number of geometric graphs in dimension  $d$  is  $\exp(o(nd \log n))$  [MM11, Sau21]. In particular, just like our spectral refutation algorithm shows (see Remark 3), this means that a coupling of the form  $H_- \subseteq G \subseteq H_+$  cannot exist when  $d \leq np \log(1/p) \log(n)^{-1}$ .

## 7 Sharp Thresholds

### 7.1 Proof of Theorem 1.2

*Proof of Theorem 1.2.* Consider some monotone property  $\mathcal{P}_n$  such that  $\mathsf{G}(n, p)$  exhibits sharp thresholds for  $\mathcal{P}_n$ . For any constant  $\epsilon > 0$ , let the critical window  $\omega_\epsilon$  be of length  $o_n(p)$ . Let  $f : [0, 1] \rightarrow [0, 1]$  be the function defined by  $f(p) = \mathbb{P}_{\mathbf{G} \sim \mathsf{G}(n, p)}[\mathbf{G} \in \mathcal{P}]$ . Let  $p_n^c = f^{-1}(1/2)$ . Similarly, let  $g(p) = \mathbb{P}_{\mathbf{G} \sim \text{RGG}(n, \mathbb{S}^{d-1}, p)}[\mathbf{G} \in \mathcal{P}]$ . Let  $q_n^c = g^{-1}(1/2)$ .

Take  $\delta \in (0, 0.1)$  to be any absolute constant. Since  $\mathsf{G}(n, p)$  exhibits sharp thresholds with respect to  $\mathcal{P}$ ,  $f^{-1}(\epsilon(1 - \delta)) = p(1 + o_n(1))$ . Now, by Corollary 1.1, one can couple with high probability  $1 - \psi$  for some  $\psi = o(1) = o(\epsilon\delta)$ , the graphs  $\mathbf{G}_- \sim \text{RGG}(n, \mathbb{S}^{d-1}, f^{-1}(\epsilon(1 - \delta)) \times (1 - \frac{c \max(\sqrt{np_n^c}, \sqrt{\log n})(\log^2 n)}{\sqrt{d}}))$  and  $\mathbf{H}_- \sim \mathsf{G}(n, f^{-1}(\epsilon(1 - \delta)))$ . Hence, as  $\mathcal{P}_n$  is monotone,

$$\mathbb{P}[\mathbf{G}_- \in \mathcal{P}_n] \leq \mathbb{P}[\mathbf{G}_- \not\subseteq \mathbf{H}_-] + \mathbb{P}[\mathbf{G}_- \subseteq \mathbf{H}_-, \mathbf{H}_- \in \mathcal{P}_n] \leq \psi + \epsilon(1 - \delta) \leq \epsilon.$$

Hence,  $g^{-1}(\epsilon) \geq f^{-1}(\epsilon(1 - \delta)) \times (1 - \frac{c \max(\sqrt{np_n^c}, \sqrt{\log n})(\log^2 n)}{\sqrt{d}}) = f^{-1}(\epsilon(1 - \delta))(1 - o_n(1))$ . As this holds for any  $\delta > 0$ ,  $g^{-1}(\epsilon) \geq f^{-1}(\epsilon(1 - o_n(1)))(1 - o_n(1))$ . Similarly,  $g^{-1}(1 - \epsilon) \leq f^{-1}(1 - \epsilon(1 - o_n(1)))(1 + o_n(1))$ .

Altogether, this means that

$$g(1 - \epsilon) - g(\epsilon) = o(f^{-1}(1 - \epsilon(1 - o_n(1)))) + o(f^{-1}(\epsilon(1 - o_n(1)))) + f^{-1}(1 - \epsilon(1 - o_n(1))) - f^{-1}(\epsilon(1 - o_n(1))) = o(p_n^c).$$

Taking  $\epsilon \rightarrow 1/2$  in

$$f^{-1}((1 - \epsilon)(1 - o_n(1)))(1 - o_n(1)) \leq g^{-1}(1 - \epsilon) \leq f^{-1}(1 - \epsilon(1 - o_n(1)))(1 + o_n(1))$$

also implies that  $q_n^c = p_n^c(1 + o(1))$ .  $\square$

### 7.2 RGG and the FKG Lattice Condition

A measure  $\mu$  over  $\{0, 1\}^N$  is strictly positive if  $\mu(\omega) > 0$  for each  $\omega \in \{0, 1\}^N$ . A strictly positive measure satisfies the FKG lattice condition if for any two  $\omega_1, \omega_2 \in \{0, 1\}^N$ , it holds that  $\mu(\omega_1 \vee \omega_2)\mu(\omega_1 \wedge \omega_2) \geq \mu(\omega_1)\mu(\omega_2)$  [GG06]. We denote pointwise maximum with  $\vee$  and pointwise minimum with  $\wedge$ .

**Proposition 7.1.** Suppose that  $d \geq 3$ . Then, the measure corresponding to  $\text{RGG}(3, \mathbb{S}^{d-1}, 1/2)$  is strictly positive but does not satisfy the FKG lattice property.

*Proof.* Denote by  $\mu$  the measure corresponding to  $\text{RGG}(3, \mathbb{S}^{d-1}, 1/2)$ . Due to symmetry, note that  $\mu(\omega)$  only depends on  $|\omega|$ . For simplicity of notation, we will also write  $\mu(|\omega|)$  for  $\mu(\omega)$ .

We will prove that for some  $a \in (0, 1/8)$ , it is the case that

$$\mu(3) = 1/8 + a, \mu(2) = 1/8 - a, \mu(1) = 1/8 + a, \mu(0) = 1/8 - a. \quad (11)$$

This is enough to show that the FKG lattice condition does not hold. Indeed, take  $\omega_1 = \{(1, 2)\}, \omega_2 = \{(2, 3)\}$ . Then,  $\omega_1 \vee \omega_2 = \{(1, 2), (2, 3)\}, \omega_1 \wedge \omega_2 = \emptyset$ . For the FKG lattice property to hold, it must be the case that  $\mu(2)\mu(0) \geq \mu(1)^2$ . This, however, is false as the left hand-side is  $(1/8 - a)^2$  while the right hand-side is  $(1/8 + a)^2$ .

Going back to Eq. (11). Let  $\mu(3) = 1/8 + a$ , where  $a$  is for now some number that we will later prove to be strictly positive. First, note that

$$\mu(3) + \mu(2) = \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{13} = 1, \mathbf{G}_{23} = 1] + \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{13} = 1, \mathbf{G}_{23} = 0] = \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{13} = 1].$$

The last expression equals  $1/4$  as every two edges in  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  are pairwise independent. Hence,  $\mu(2) = 1/4 - \mu(3) = 1/8 - a$ . Similarly,  $\mu(1) = 1/4 - \mu(2) = 1/8 + a$  as

$$\mu(2) + \mu(1) = \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{13} = 1, \mathbf{G}_{23} = 0] + \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{13} = 0, \mathbf{G}_{23} = 0] = \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{23} = 0],$$

and  $\mu(0) = 1/4 - \mu(1) = 1/8 - a$  as

$$\mu(1) + \mu(0) = \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{13} = 0, \mathbf{G}_{23} = 0] + \mathbb{P}[\mathbf{G}_{12} = 0, \mathbf{G}_{13} = 0, \mathbf{G}_{23} = 0] = \mathbb{P}[\mathbf{G}_{13} = 0, \mathbf{G}_{23} = 0].$$

Thus, all that is left to show is that  $a \geq 0$ . We accomplish this by bounding  $\mu(3)$ .

$$\begin{aligned} \mu(3) &= \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{13} = 1, \mathbf{G}_{23} = 1] \\ &= \mathbb{P}[\mathbf{G}_{23} = 1 | \mathbf{G}_{12} = 1, \mathbf{G}_{13} = 1] \times \mathbb{P}[\mathbf{G}_{12} = 1, \mathbf{G}_{13} = 1] \\ &= \frac{1}{4} \mathbb{P}[\mathbf{G}_{23} = 1 | \mathbf{G}_{12} = 1, \mathbf{G}_{13} = 1] \\ &= \frac{1}{4} \mathbb{P}[\langle V_2, V_3 \rangle \geq 0 | \langle V_1, V_2 \rangle \geq 0, \langle V_1, V_3 \rangle \geq 0]. \end{aligned}$$

First, note that this probability is the same if  $V_1, V_2, V_3$  were isotropic Gaussian vectors instead of spherical - the sign of inner products depends only on the direction but not on the magnitude of vectors. Due to rotational invariance, we can assume that  $V_1 \in \text{span}\{e_1\}$ , so  $V_1 = z_{1,1}e_1$ , where  $z_{1,1} > 0$ , and  $V_2 \in \text{span}\{e_1, e_2\}$ . So,  $V_2 = z_{2,1}e_1 + z_{2,2}e_2$ . Finally, let  $V_3 = z_{3,1}e_1 + z_{3,2}e_2 + z_{3,3}e_3$ . Since the components of an isotropic Gaussian are independent in different directions, all of the numbers  $z_{1,1}, z_{2,1}, z_{2,2}, z_{3,1}, z_{3,2}, z_{3,3}$  are independent. Furthermore,  $z_{2,1}, z_{3,1}, z_{3,2}$  are standard Gaussians and  $z_{1,1}, z_{2,2}, z_{3,3} > 0$  as these are norms of certain projections of the Gaussian vectors. We want to show that

$$\mathbb{P}[z_{3,1}z_{2,1} + z_{3,2}z_{2,2} \geq 0 | z_{3,1} \geq 0, z_{2,1} \geq 0] \in (1/2, 1).$$

We compute

$$\begin{aligned} &\mathbb{P}[z_{3,1}z_{2,1} + z_{3,2}z_{2,2} \geq 0 | z_{3,1} \geq 0, z_{2,1} \geq 0] \\ &= \int_0^{+\infty} \mathbb{P}[z_{3,1}z_{2,1} + z_{3,2}z_{2,2} \geq 0 | z_{3,1} \geq 0, z_{2,1} \geq 0, z_{3,1}z_{2,1} = x] \times \mathbb{P}[z_{3,1}z_{2,1} = x | z_{3,1} \geq 0, z_{2,1} \geq 0] dx \\ &= \int_0^{+\infty} \mathbb{P}[z_{3,2} \geq -x/z_{2,2}] \times \mathbb{P}[z_{3,1}z_{2,1} = x | z_{3,1} \geq 0, z_{2,1} \geq 0] dx. \end{aligned}$$

Note that for each  $x, z_{2,2} \geq 0$ , it holds that  $\mathbb{P}[z_{3,2} \geq -x/z_{2,2}] \in (1/2, 1)$  as  $z_{3,2}$  is a standard Gaussian and  $-x/z_{2,2} < 0$ . The conclusion follows.  $\square$

## 8 Discussion and Open Problems

### 8.1 Robust Testing With Fewer Corruptions

One of the main results of the current paper is settling (up to log factors) the regime in which testing between  $\mathbb{G}(n, p)$  and  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  with  $\epsilon \binom{n}{2} p$  corruptions is possible when  $\epsilon > 0$  is a small constant. We show via [Theorems 1.1](#) and [1.3](#) that the threshold is  $d = \tilde{\Theta}(np)$  both computationally and information-theoretically. The analogous question without corruptions, i.e.  $\epsilon = 0$ , is still open as discussed in the introduction. It is natural to interpolate between  $\epsilon = 0$  and  $\epsilon = \Theta(1)$ .

**Problem 2.** Given are  $n, d \in \mathbb{N}$ ,  $p \in [0, 1]$  and a function  $\epsilon(n, p) \geq 0$  such that  $\epsilon(n, p) \leq K_r$  for some universal constant  $K_r$ . Consider the following hypothesis testing problem. One observes a graph  $H$  and needs to decide between  $H_0 : H = \mathcal{A}(\mathbf{G})$ ,  $\mathbf{G} \sim \mathbb{G}(n, p)$  and  $H_1 : H = \mathcal{A}(\mathbf{G})$ ,  $\mathbf{G} \sim \text{RGG}(n, d, p)$ . Here,  $\mathcal{A} : \{0, 1\}^{\binom{n}{2}} \rightarrow \{0, 1\}^{\binom{n}{2}}$  is any unknown mapping such that  $\mathcal{A}(G)$  and  $G$  differ in at most  $\epsilon(n, p) \times \binom{n}{2} p$  entries for any adjacency matrix  $G \in \{0, 1\}^{\binom{n}{2}}$ .

We expect that this question will require both novel algorithmic and lower-bound ideas. Neither the spectral refutation algorithm (when  $\epsilon = \Theta(1)$ ) nor the signed triangle count algorithm (when  $\epsilon = 0$ ) seems to generalize trivially. Regarding lower-bounds, even the case of no corruptions is still open. Particularly interesting is whether there is a statistical-computational gap for any  $\epsilon(n, p)$ . There is no evidence for such a gap when  $\epsilon(n, p) = 0$  and  $\epsilon(n, p) = \Theta(1)$ .

[Theorem 1.1](#) and [Corollary 1.1](#) show that if  $\epsilon(n, p) \geq C \frac{\max(\sqrt{np}, \sqrt{\log n})(\log n)^2}{\sqrt{d}}$  for some absolute constant  $C$ , the distinguishing problem is impossible information theoretically against the (efficient) adversary from [Theorem 1.1](#). We do not expect this to be tight for all values of  $\epsilon$  (in particular, when  $\epsilon$  approaches 0). This naturally also raises a question about the optimality of [Theorem 1.1](#) and [Corollary 1.1](#). We only phrase the question for the latter.

**Problem 3.** What is the smallest value of  $\delta(d, n, p)$  for which there exists a coupling of  $G \sim \text{RGG}(n, \mathbb{S}^{d-1}, p)$ ,  $H_- \sim \mathbb{G}(n, p(1 - \delta(d, n, p)))$  and  $H_+ \sim \mathbb{G}(n, p(1 + \delta(d, n, p)))$ , such that  $H_- \subseteq G \subseteq H_+$  with high-probability.

### 8.2 Embedding Random Geometric Graphs

**New Algorithms For Embedding.** [Theorem 1.1](#) can also be viewed as an approximate embedding of a sample  $H$  from  $\mathbb{G}(n, p)$  as a geometric graph. This naturally raises the question whether one can use [Algorithm 1](#) to embed a sample  $H$  from  $\text{RGG}(n, \mathbb{S}^{d-1}, p)$  (with the true vectors remaining unknown). Our martingale-based analysis fails in this scenario as the independence of variables  $H_{ji}$  is crucial. Nevertheless, the fact that the marginal distribution of vectors  $V_1, V_2, \dots, V_n$  is uniform and independent makes such an algorithm extremely appealing as it gives a certain sampling flavor of the resulting embedding: It produces a uniform sample from a certain set of  $n$ -tuples of latent vectors which nearly match the input adjacency matrix.

**Robust Embedding.** A different algorithmic question is whether there exist embedding algorithms in the setting of adversarial corruptions. One way to phrase this question (using a Frobenius norm error) is:

**Problem 4.** Let  $V = (V_1, V_2, \dots, V_n) \in \mathbb{R}^{d \times n}$  where  $V_1, V_2, \dots, V_n \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$  and  $G$  be the corresponding geometric graph with  $\epsilon \binom{n}{2} p$  edge corruptions (or  $\epsilon n$  node corruptions). On input  $G$ , find a matrix  $\widehat{W}$  such that  $\|\widehat{W} - V^T V\|_F = \|V^T V\|_F \times (f(\epsilon) + o(1))$ , where  $f(\epsilon)$  is some function such that  $\lim_{\epsilon \rightarrow 0} f(\epsilon) = 0$ .

The case of no corruptions (i.e.  $\epsilon = 0$ ) is treated in [\[LS23\]](#) via a spectral algorithm which succeeds whenever  $d = o(np)$ .

### 8.3 A Reverse Coupling?

[Algorithm 1](#) efficiently couples with high probability a sample  $H$  from  $\mathsf{G}(n, p)$  and a sample  $G$  from  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  on input  $H$  such that they differ on  $o(n^2p)$  edges. One may naturally ask the question of whether such an efficient coupling algorithm exists on input the random geometric graph.

**Problem 5.** Design an efficient algorithm which on input  $G \sim \mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  where  $d \geq np \times \log^C n$  for an appropriate constant  $C$  produces a graph  $H(G)$  such that:

1.  $d_{\text{TV}}(\mathcal{L}(H(G)), \mathsf{G}(n, p)) = o(1)$ .
2. With high probability,  $H(G)$  and  $G$  differ on  $o(n^2p)$  edges.

The question is only about efficiency as one can simply reverse the coupling in [Theorem 1.1](#).

### 8.4 Other Random Geometric Graph Distributions

We showed that the central result of the current paper [Theorem 1.1](#) has a wide range of applications, including to robust statistics, enumeration, and sharp thresholds. The proof technique, however, is specialized to spherical random geometric graphs. Naturally, one may want a more systematic approach which also works for other high-dimensional random geometric graphs such as over the torus [BB24a] or Boolean hypercube [BB23] or non-uniform distributions such as the anisotropic random geometric graphs in [EM20, BBH24]. We believe that extending our results and techniques to other geometries is an exciting open direction. Our only progress in that direction is for Gaussian random geometric graphs, which is nearly identical to the spherical case.

**Remark 8.** [Corollary 1.1](#) and [Theorem 1.1](#) hold with the exact same parameters (up to multiplicative constants) for Gaussian RGG. A sample from  $\mathsf{RGG}(n, \mathcal{N}(0, \frac{1}{d}I_d), p)$  is an  $n$ -vertex graph  $G$  defined by sampling  $Z_1, Z_2, \dots, Z_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \frac{1}{d}I_d)$  and setting  $G_{ji} = \mathbf{1}[\langle Z_i, Z_j \rangle \geq \xi_d^p]$  where  $\xi_d^p$  is such that the expected density is  $p$ . One coupling algorithm is [Algorithm 1](#) except that at the end one defines  $Z_i^{(i-1)} = \chi_i \times V_i^{(i-1)}$ , where the variables  $(\chi_i)_{i=1}^n$  are independent of everything else and are distributed as  $\|\mathcal{N}(0, \frac{1}{d}I_d)\|_2$ . Now, use that  $|\chi_i - 1| \leq \frac{\log n}{\sqrt{d}}$  with high probability (e.g. [BB24b]) and the fact that  $|\tau_d^p - \xi_d^p| = \tilde{O}(1/d)$  [BB24b]. One immediately concludes that when  $d \gg np$ , the robust testing [Problem 1](#) is information theoretically impossible. This is perhaps surprising in the sparse case due to the difference between spherical and Gaussian RGG [BB24b]. There is an efficient test that distinguishes  $\mathsf{RGG}(n, \mathcal{N}(0, \frac{1}{d}I_d), p)$  and  $\mathsf{G}(n, p)$  when  $d \gg n^{3/2}p, p = O(n^{-3/4})$  [BB24b] in the case of no corruptions, a regime in which  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  and  $\mathsf{G}(n, p)$  are conjectured to be information-theoretically indistinguishable. Yet, when testing with adversarial corruptions no such difference between  $\mathsf{RGG}(n, \mathcal{N}(0, \frac{1}{d}I_d), p)$  and  $\mathsf{RGG}(n, \mathbb{S}^{d-1}, p)$  seems to emerge.

## Acknowledgements

We are grateful to Sidhanth Mohanty for many helpful discussions on robust statics and random geometric graphs. We also thank Chenghao Guo for raising the question about efficient adversaries for the robust testing problem. We are indebted to Dheeraj Nagaraj for many conversations on random geometric graphs over the years and also specifically for discussions about the flip orientation map.

## References

- [ABH16] Emmanuel Abbe, Afonso S. Bandeira, and Georgina Hall. Exact recovery in the stochastic block model. *IEEE Trans. Inf. Theor.*, 62(1):471–487, jan 2016.
- [ADK21] Johannes Alt, Raphaël Ducatez, and Antti Knowles. Extremal eigenvalues of critical Erdős-Rényi graphs. *The Annals of Probability*, 49(3):1347 – 1401, 2021.

- [AKS98] Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '98, page 594–598, USA, 1998. Society for Industrial and Applied Mathematics.
- [BB23] Kiril Bangachev and Guy Bresler. Random Algebraic Graphs and Their Convergence to Erdos-Renyi, 2023.
- [BB24a] Kiril Bangachev and Guy Bresler. Detection of  $L_\infty$  Geometry in Random Geometric Graphs: Suboptimality of Triangles and Cluster Expansion. In Shipra Agrawal and Aaron Roth, editors, *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 427–497. PMLR, 30 Jun–03 Jul 2024.
- [BB24b] Kiril Bangachev and Guy Bresler. On the fourier coefficients of high-dimensional random geometric graphs. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 549–560, New York, NY, USA, 2024. Association for Computing Machinery.
- [BDDD<sup>+</sup>21] Marian Boguna, Ivan Bonamassa, Manlio De Domenico, Shlomo Havlin, Dmitri Krioukov, and M Ángeles Serrano. Network geometry. *Nature Reviews Physics*, 3(2):114–135, 2021.
- [BBH24] Matthew Brennan, Guy Bresler, and Brice Huang. Threshold for detecting high dimensional geometry in anisotropic random geometric graphs. *Random Structures & Algorithms*, 64(1):125–137, 2024.
- [BBN20] Matthew Brennan, Guy Bresler, and Dheeraj Nagaraj. Phase transitions for detecting latent geometry in random graphs. *Probability Theory and Related Fields*, 178:1215–1289, 12 2020.
- [BDER14] Sébastien Bubeck, Jian Ding, Ronen Eldan, and Miklós Rácz. Testing for high-dimensional geometry in random graphs. *Random Structures & Algorithms*, 49, 11 2014.
- [BGBK19] Florent Benaych-Georges, Charles Bordenave, and Antti Knowles. Largest eigenvalues of sparse inhomogeneous erdős-rÉnyi graphs. *The Annals of Probability*, 47(3):pp. 1653–1676, 2019.
- [BGBK20] Florent Benaych-Georges, Charles Bordenave, and Antti Knowles. Spectral radii of sparse random matrices. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 56(3):2141 – 2161, 2020.
- [BHK<sup>+</sup>19] Boaz Barak, Samuel Hopkins, Jonathan Kelner, Pravesh K. Kothari, Ankur Moitra, and Aaron Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM Journal on Computing*, 48(2):687–735, 2019.
- [BKK<sup>+</sup>92] Jean Bourgain, Jeff Kahn, Gil Kalai, Yitzhak Katznelson, and Nathan Linial. The influence of variables in product spaces. *Israel Journal of Mathematics*, 77(1-2):55–64, February 1992.
- [BMR21] Jess Banks, Sidhanth Mohanty, and Prasad Raghavendra. *Local Statistics, Semidefinite Programming, and Community Detection*, pages 1298–1316. Society for Industrial and Applied Mathematics, 2021.
- [BP13] Milan Bradonjic and Will Perkins. On sharp thresholds in random geometric graphs. In *International Workshop on Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, 2013.
- [DdC22] Quentin Duchemin and Yohann de Castro. Random geometric graph: Some recent developments and perspectives, 2022.
- [DdNS22] Jingqiu Ding, Tommaso d’Orsi, Rajai Nasser, and David Steurer. Robust recovery for stochastic block models. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 387–394, Los Alamitos, CA, USA, feb 2022. IEEE Computer Society.
- [DGLU11] Luc Devroye, András György, Gábor Lugosi, and Frederic Udina. High-Dimensional Random Geometric Graphs and their Clique Number. *Electronic Journal of Probability*, 16:2481 – 2508, 2011.

- [EAM21] Ahmed El Alaoui and Andrea Montanari. On the computational tractability of statistical estimation on amenable graphs. *Probability theory and related fields*, 181(4):815–864, 2021.
- [EM20] Ronen Eldan and Dan Mikulincer. *Information and Dimensionality of Anisotropic Random Geometric Graphs*, pages 273–324. Springer International Publishing, Cham, 2020.
- [ES16] Ernesto Estrada and Matthew Sheerin. Consensus dynamics on random rectangular graphs. *Physica D: Nonlinear Phenomena*, 323-324:20–26, 2016. Nonlinear Dynamics on Interconnected Networks.
- [FabJB99] Ehud Friedgut and appendix by Jean Bourgain. Sharp thresholds of graph properties, and the  $\text{-sat}$  problem. *Journal of the American Mathematical Society*, 12:1017–1054, 1999.
- [FK96] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proceedings of the American Mathematical Society*, 124(10):2993–3002, 1996.
- [Fre75] David A. Freedman. On Tail Probabilities for Martingales. *The Annals of Probability*, 3(1):100 – 118, 1975.
- [GG06] B. T. Graham and G. R. Grimmett. Influence and sharp-threshold theorems for monotonic measures. *The Annals of Probability*, 34(5):1726 – 1745, 2006.
- [GG11] Benjamin Graham and Geoffrey Grimmett. Sharp thresholds for the random-cluster and Ising models. *The Annals of Applied Probability*, 21(1):240 – 265, 2011.
- [Gil61] E. N. Gilbert. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):533–543, 1961.
- [Gri06] Geoffrey. Grimmett. *The random-cluster model*. Grundlehren der mathematischen Wissenschaften, 333. Springer, Berlin, 2006.
- [GRK05] Ashish Goel, Sanatan Rai, and Bhaskar Krishnamachari. Monotone properties of random geometric graphs have sharp thresholds. *The Annals of Applied Probability*, 15(4):2535–2552, 2005.
- [GV16] Olivier Guédon and Roman Vershynin. Community detection in sparse networks via grothendieck’s inequality. *Probability theory and related fields*, 165(3-4):1025–1049, 2016.
- [Hae12] Martin Haenggi. *Random geometric graphs and continuum percolation*, page 221–245. Cambridge University Press, 2012.
- [HRH02] Peter D Hoff, Adrian E Raftery, and Mark S Handcock. Latent space approaches to social network analysis. *Journal of the American Statistical association*, 97(460):1090–1098, 2002.
- [HRP08] Desmond J. Higham, Marija Rašajski, and Nataša Pržulj. Fitting a geometric graph to a protein–protein interaction network. *Bioinformatics*, 24(8):1093–1099, 03 2008.
- [Jer92] Mark Jerrum. Large cliques elude the metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992.
- [Kuc95] Ludek Kucera. Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2):193–212, 1995. Combinatorial optimization 1992.
- [LMSY22] Siqi Liu, Sidhanth Mohanty, Tselil Schramm, and Elizabeth Yang. Testing thresholds for high-dimensional sparse random geometric graphs. In *Proceedings of the 54th Annual ACM Symposium on Theory of Computing*, STOC 2022, New York, NY, USA, 2022. Association for Computing Machinery.
- [LMSY23] Siqi Liu, Sidhanth Mohanty, Tselil Schramm, and Elizabeth Yang. Local and global expansion in random geometric graphs. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 817–825, New York, NY, USA, 2023. Association for Computing Machinery.

- [LR23] Suqi Liu and Miklós Z. Rácz. A probabilistic view of latent space graphs and phase transitions. *Bernoulli*, 29(3):2417 – 2441, 2023.
- [LS23] Shuangping Li and Tselil Schramm. Spectral clustering in the gaussian mixture block model, 2023.
- [Mar74] G. A. Margulis. Probabilistic characteristics of graphs with large connectivity. *Probl. Peredachi Inf.*, 1974.
- [MB73] David D McFarland and Daniel J Brown. Social distance as a metric: a systematic introduction to smallest space analysis. *Bonds of pluralism: The form and substance of urban social networks*, 6:2, 1973.
- [MM11] Colin McDiarmid and Tobias Müller. Counting disk graphs. *Electronic Notes in Discrete Mathematics*, 38:627–632, 2011. The Sixth European Conference on Combinatorics, Graph Theory and Applications, EuroComb 2011.
- [MNS15] Elchanan Mossel, Joe Neeman, and Allan Sly. Consistency thresholds for the planted bisection model. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC ’15, page 69–75, New York, NY, USA, 2015. Association for Computing Machinery.
- [MRW24] Sidhanth Mohanty, Prasad Raghavendra, and David X. Wu. Robust recovery for stochastic block models, simplified and generalized. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 367–374, New York, NY, USA, 2024. Association for Computing Machinery.
- [MS16] Andrea Montanari and Subhabrata Sen. Semidefinite programs on sparse random graphs and their application to community detection. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC ’16, page 814–827, New York, NY, USA, 2016. Association for Computing Machinery.
- [MSLC01] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444, 2001.
- [MWZ23] Cheng Mao, Alexander S. Wein, and Shenduo Zhang. Detection-recovery gap for planted dense cycles. In Gergely Neu and Lorenzo Rosasco, editors, *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pages 2440–2481. PMLR, 12–15 Jul 2023.
- [MWZ24] Cheng Mao, Alexander S. Wein, and Shenduo Zhang. Information-theoretic thresholds for planted dense cycles, 2024.
- [MZ24] Cheng Mao and Shenduo Zhang. Impossibility of latent inner product recovery via rate distortion, 2024.
- [O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, West Nyack, 1 edition, 2014.
- [Pen03] Mathew Penrose. *Random Geometric Graphs*. Oxford University Press, 05 2003.
- [Per24] Will Perkins. Searching for (sharp) thresholds in random structures: where are we now?, 2024.
- [PJ09] Victor M. Preciado and Ali Jadbabaie. Spectral analysis of virus spreading in random geometric networks. *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 4802–4807, 2009.
- [Rus81] Lucio Russo. On the critical percolation probabilities. *Probability Theory and Related Fields*, 56:229–237, 06 1981.
- [Rus82] Lucio Russo. An approximate zero-one law. *Zeitschrift f r Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 61(1):129–139, 1982.

- [RV23] Mirabel E. Reid and Santosh S. Vempala. The  $k$ -cap process on geometric random graphs. In Gergely Neu and Lorenzo Rosasco, editors, *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pages 3469–3509. PMLR, 12–15 Jul 2023.
- [Sau21] Lisa Sauermann. On the speed of algebraically defined graph classes. *Advances in Mathematics*, 380:107593, 2021.
- [SGN<sup>+</sup>21] Damian Szkłarczyk, Annika L Gable, Katerina C Nastou, David Lyon, Rebecca Kirsch, Sampo Pyysalo, Nadezhda T Doncheva, Marc Legeay, Tao Fang, and Peer Bork. The string database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, 49(D1):D605–D612, 2021.
- [SKB08] M Ángeles Serrano, Dmitri Krioukov, and Marián Boguná. Self-similarity of complex networks and hidden metric spaces. *Physical review letters*, 100(7):078701, 2008.
- [SW22] Tselil Schramm and Alexander S. Wein. Computational barriers to estimation from low-degree polynomials. *The Annals of Statistics*, 50(3), 6 2022.
- [Tal94] Michel Talagrand. On Russo’s Approximate Zero-One Law. *The Annals of Probability*, 22(3):1576 – 1587, 1994.
- [War68] Hugh E. Warren. Lower bounds for approximation by nonlinear manifolds. *Transactions of the American Mathematical Society*, 133(1):167–178, 1968.

## A Proof of Lemma 2.1

We will use the fact that the density of  $\mathcal{D}_d$  is  $f_d(x) := \frac{\Gamma(d/2)}{\Gamma(d/2-1/2)\sqrt{\pi}}(1-x^2)^{\frac{d-3}{2}}$  and, for some absolute constant  $C_\Gamma$ ,  $\frac{\sqrt{d}}{C_\Gamma} \leq \frac{\Gamma(d/2)}{\Gamma(d/2-1/2)\sqrt{\pi}} \leq C_\Gamma\sqrt{d}$  [BDER14]. Part A) appears explicitly in [BDER14].

### Part B)

Note that  $\mathbb{E}_{X \sim \mathcal{D}_d}[X \mid X \geq \tau_d^p] \geq \tau_d^p$ . Hence, if  $p \leq 1/4$  there is nothing to prove by Part A). If  $p \in (1/4, 1/2]$ , we compute:

$$\begin{aligned} & \mathbb{E}_{X \sim \mathcal{D}_d}[X \mid X \geq \tau_d^p] \\ & \geq \mathbb{E}_{X \sim \mathcal{D}_d}[X \mid X \geq 0] \\ & = 2 \int_0^1 \frac{\Gamma(d/2)}{\Gamma(d/2-1/2)\sqrt{\pi}} x(1-x^2)^{\frac{d-3}{2}} dx \\ & \geq 2 \frac{\Gamma(d/2)}{\Gamma(d/2-1/2)\sqrt{\pi}} \int_{1/\sqrt{d}}^{2/\sqrt{d}} x(1-x^2)^{\frac{d-3}{2}} dx \\ & \geq \frac{2\sqrt{d}}{C_\Gamma} \int_{1/\sqrt{d}}^{2/\sqrt{d}} \frac{1}{\sqrt{d}}(1-\frac{4}{d})^{\frac{d-3}{2}} dx \geq \frac{2\sqrt{d}}{C_\Gamma} \times \frac{1}{\sqrt{d}} \frac{1}{\sqrt{d}} e^{-4} \geq \frac{C' \sqrt{\log 2}}{\sqrt{d}} \geq \frac{\sqrt{\log(1/p)}}{C_B \sqrt{d}} \end{aligned}$$

for some absolute constants  $C', C_B$ .

## Part C)

We will prove that  $\mathbb{E}_{X \sim \mathcal{D}_d}[X^k \mid X \geq \tau_d^p] \leq c_k(\tau_d^p + c_k)^k$  for some absolute constant  $c_k$  depending only on  $k$ . Clearly, this is enough. Again, we begin with the explicit formula for the expectation:

$$\mathbb{E}_{X \sim \mathcal{D}_d}[X^k \mid X \geq \tau_d^p] = \frac{\int_{\tau_d^p}^1 x^k f_d(x) dx}{\int_{\tau_d^p}^1 f_d(x) dx}.$$

We now analyse the numerator in several steps. Let  $\beta_k > 0$  be a constant, depending only on  $k$ , that we will choose later.

**Small values of  $x$ .** We compute

$$\begin{aligned} & \int_{\tau_d^p}^{\tau_d^p + \frac{\beta_k}{\sqrt{d}}} x^k f_d(x) dx \\ & \leq \int_{\tau_d^p}^{\tau_d^p + \frac{\beta_k}{\sqrt{d}}} (\tau_d^p + \frac{\beta_k}{\sqrt{d}})^k f_d(x) dx \\ & \leq \int_{\tau_d^p}^1 (\tau_d^p + \frac{\beta_k}{\sqrt{d}})^k f_d(x) dx \\ & \leq (\tau_d^p + \frac{\beta_k}{\sqrt{d}})^k \int_{\tau_d^p}^1 f_d(x) dx, \end{aligned}$$

so this part of the numerator is of the desired order.

**Large Values of  $x$ .** The key argument when  $\frac{1}{2} \geq x \geq \tau_d^p + \frac{\beta_k}{\sqrt{d}}$  is that (for an appropriately chosen  $\beta_k$ ), one has

$$(\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k f_d(x - \frac{\beta_k}{\sqrt{d}}) \geq x^k f_d(x). \quad (12)$$

First, note that this is enough as it will imply that

$$\int_{\tau_d^p + \frac{\beta_k}{\sqrt{d}}}^{1/2} x^k f_d(x) dx \leq \int_{\tau_d^p + \frac{\beta_k}{\sqrt{d}}}^{1/2} (\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k f_d(x - \frac{\beta_k}{\sqrt{d}}) dx \leq (\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k \int_{\tau_d^p}^1 f_d(x) dx.$$

Going back to Eq. (12). As  $f_d(x)$  is decreasing, the inequality is trivial for  $x \leq \tau_d^p + \frac{2\beta_k}{\sqrt{d}}$ . When  $x > \tau_d^p + \frac{2\beta_k}{\sqrt{d}}$ , set  $y = x - \frac{\beta_k}{\sqrt{d}} \geq \frac{\beta_k}{\sqrt{d}}$ . Using the explicit formula for  $f_d(x)$ , we need to show that

$$\begin{aligned} & (\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k \geq (y + \frac{\beta_k}{\sqrt{d}})^k \left( \frac{1 - (y + \frac{\beta_k}{\sqrt{d}})^2}{1 - y^2} \right)^{(d-3)/2} \iff \\ & (\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k \geq (y + \frac{\beta_k}{\sqrt{d}})^k \left( 1 - \frac{2y\frac{\beta_k}{\sqrt{d}} + y^2}{1 - x^2} \right)^{(d-3)/2} \iff \\ & (\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k \geq (y + \frac{\beta_k}{\sqrt{d}})^k \left( 1 - \frac{y\frac{\beta_k}{\sqrt{d}} + y^2}{3/4} \right)^{(d-3)/2} \iff \\ & (\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k \geq (y + \frac{\beta_k}{\sqrt{d}})^k \exp(-\alpha y(y + \frac{\beta_k}{\sqrt{d}})d) \end{aligned}$$

for some absolute constant  $\alpha$  since  $y\frac{\beta_k}{\sqrt{d}} + y^2 \leq \frac{1}{2}\frac{\beta_k}{\sqrt{d}} + 1/4 \leq 2/3 < 3/4$ . Now, the function  $u(z) = z^k \exp(-\alpha dz(z + \frac{\beta_k}{\sqrt{d}}))$  is decreasing on  $[\frac{\beta_k}{\sqrt{d}}, 1]$  for a sufficiently large  $\beta_k$ . Indeed, note that its derivative is

$$\exp(-\alpha dz(z + \frac{\beta_k}{\sqrt{d}}))(kz^{k-1} - z^k \alpha d(2z + \frac{\beta_k}{\sqrt{d}})).$$

For large enough  $\beta_k$ , we have that  $kz^{k-1} \leq z^k \alpha \beta_k \sqrt{d}$  for all  $z \geq \beta_k/\sqrt{d}$ , from which the fact that the derivative is negative follows. This is enough as

$$(y + \frac{\beta_k}{\sqrt{d}})^k \exp(-\alpha y(y + \frac{\beta_k}{\sqrt{d}})d) \leq (\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k \exp(-\alpha(\tau_d^p + \frac{\beta_k}{\sqrt{d}})(\tau_d^p + \frac{2\beta_k}{\sqrt{d}})d) \leq (\tau_d^p + \frac{2\beta_k}{\sqrt{d}})^k$$

in that case.

**Extremely large values of  $x$ .** Note that as  $d = \omega(\log(1/p))$ , when  $x \in [1/2, 1]$ ,  $f_d(x) \leq (3/4)^{(d-3)/2} = o(\sqrt{\log(1/p)/d})$ .

## Part D)

The main observation is the following.

**Observation A.1.** If  $u \leq C(\log(1/p)^{-1/2}/\sqrt{d})$ , for some absolute constant  $C$ , then

$$\frac{1}{2} \leq f_d(\tau_d^p + u)/f_d(\tau_d^p) \leq 2.$$

*Proof.* We write

$$f_d(\tau_d^p + u)/f_d(\tau_d^p) = \left( \frac{1 - (\tau_d^p + u)^2}{1 - (\tau_d^p)^2} \right)^{(d-3)/2} = \left( 1 - \frac{2\tau_d^p u + u^2}{1 - (\tau_d^p)^2} \right)^{(d-3)/2}.$$

Whenever  $d = \omega(\log 1/p)$ , we have  $1 - (\tau_d^p)^2 \geq 1/2$  by part A). Whenever  $u \leq C(\log(1/p)^{-1/2}/\sqrt{d})$ , clearly  $2\tau_d^p u + u^2 \leq 2C/d$ . Hence, the expression above is in  $(1 \pm 2C/d)^{(d-3)/2} \in [1/2, 2]$ .  $\square$

Now, in particular, this means that  $f_d(\tau_d^p) \leq C' p \sqrt{d} \sqrt{\log 1/p}$  for some absolute constant  $C'$ . Indeed, notice that

$$p = \int_{\tau_d^p}^1 f_d(x) dx \geq \int_{\tau_d^p}^{\tau_d^p + C(\log(1/p)^{-1/2}/\sqrt{d})} f(x) dx \geq \int_{\tau_d^p}^{\tau_d^p + C(\log(1/p)^{-1/2}/\sqrt{d})} f(\tau_d^p)/2 dx.$$

Hence,  $f_d(x) \leq 2C' p \sqrt{d} \sqrt{\log 1/p}$  for all  $x \in [\tau_d^p - \Delta, \tau_d^p + \Delta]$  for sufficiently small  $\Delta \leq C(\log(1/p)^{-1/2}/\sqrt{d})$  and the claim follows.

## B Simple Lemmas on Graph Counting

*Proof of Observation 6.2.* This lemma appears in [BB23], but we give the proof for completeness. Consider  $H \sim \mathbb{G}(n, p)$ . First, as edges are independent, by Chernyshev's inequality,

$$\mathbb{P}\left[|E(H)| \leq \frac{1}{2} \binom{n}{2} p\right] = \mathbb{P}\left[|E(H)| \leq \frac{1}{2} \mathbb{E}[|E(H)|]\right] = o(1).$$

On the other hand, for any fixed graph  $K$  on at least  $\frac{1}{2} \mathbb{E}|E(H)| = \frac{1}{2} \binom{n}{2} p$  vertices, we have

$$\mathbb{P}[H = K] = p^{|E(K)|} (1-p)^{\binom{n}{2} - |E(K)|} \leq p^{|E(K)|} \leq \exp\left(\log p \frac{1}{2} \binom{n}{2} p\right) \leq \exp\left(-\frac{1}{5} n^2 p \log \frac{1}{p}\right).$$

Hence, a  $1/2 - o(1)$  mass of the set  $A$  must be supported on graphs with mass  $\exp\left(-\frac{1}{5}n^2p \log \frac{1}{p}\right)$  each with respect to  $G(n, p)$ . The statement follows.  $\square$

*Proof of Observation 6.3.* The number of such subgraphs of  $H$  is simply  $\sum_{r=0}^{\epsilon n^2 p} \binom{|E(H)|}{r}$ . We can trivially bound

$$\begin{aligned} \sum_{r=0}^{\epsilon n^2 p} \binom{|E(H)|}{r} &\leq \sum_{r=0}^{\epsilon n^2 p} \binom{n^2 p}{r} \leq \epsilon n^2 p \binom{n^2 p}{\epsilon n^2 p} \\ &\leq \epsilon n^2 p \left(\frac{\epsilon n^2 p}{\epsilon n^2 p}\right)^{\epsilon n^2 p} = \epsilon n^2 p (e/\epsilon)^{\epsilon n^2 p} \leq \exp(2n^2 p \epsilon \log(1/\epsilon)). \end{aligned} \quad \square$$

## C Random Geometric Graphs as Recursively Planted Erdős-Rényi

Here, we discuss in more detail [Theorem 3.1](#).

**More General Flip-Orientation Maps** In analogy to [Section 4.1](#), we can define a more general flip-orientation map. Let  $\phi(\cdot|[u, v]) : [u, v] \rightarrow [u, v]$  for  $u < \tau_d^p < v$  be the unique monotone involution such that  $\phi(u|[u, v]) = v, \phi(v|[u, v]) = u, \phi(\tau_d^p|[u, v]) = \tau_d^p$  and  $\phi(\cdot|[u, v]) : [u, \tau_d^p] \rightarrow [\tau_d^p, v]$  and  $\phi(\cdot|[u, v]) : [\tau_d^p, v] \rightarrow [u, \tau_d^p]$  are measure-preserving with respect to  $\mathcal{D}_d|_{[u, \tau_d^p]}, \mathcal{D}_d|_{[\tau_d^p, v]}$ . In particular,  $\phi(\cdot) = \phi(\cdot|[-1, 1])$  in the notation of [Section 4.1](#).

Define also

$$\rho_a^b(z|[u, v]) := \begin{cases} z & \text{if } \langle z, a \rangle \notin [u, v], \\ z & \text{if } \mathbb{1}[\langle z, a \rangle \geq \tau_d^p] = b, \langle z, a \rangle \in [u, v], \\ (z - \langle z, a \rangle a) \sqrt{\frac{1 - \phi^2(\langle z, a \rangle|[u, v])}{1 - \langle z, a \rangle^2}} + \phi(\langle a, z \rangle|[u, v])a & \text{if } \mathbb{1}[\langle z, a \rangle \geq \tau_d^p] \neq b, \langle z, a \rangle \in [u, v]. \end{cases} \quad (13)$$

Define  $s^b(\langle a, z \rangle|[u, v]) = s^b(\langle a, z \rangle) \times \mathbb{1}[\langle a, z \rangle \in [u, v]]$ . Also, define analogously  $\psi(\cdot|[u, v]), \kappa(\cdot|[u, v])$ . Let  $q_{[u, v]} = \frac{\mathbb{P}_{X \sim \mathcal{D}_d}[X \in [\tau_d^p, v]]}{\mathbb{P}_{X \sim \mathcal{D}_d}[X \in [u, v]]}$ . If  $b \sim \text{Bern}(q_{[u, v]})$ , then the equivalent statements of [Observation 4.1](#) and [Observation 4.2](#) hold with same argument.

**Observation C.1.** If  $b \sim \text{Bern}(q_{[u, v]}), z \sim \text{Unif}(\mathbb{S}^{d-1})$  and  $a, b, z$  are independent, then  $\rho_a^b(z|[u, v]) \sim \text{Unif}(\mathbb{S}^{d-1})$ .

**Observation C.2.** If  $b \sim \text{Bern}(q_{[u, v]}), z \sim \text{Unif}(\mathbb{S}^{d-1})$  and  $a, b, z$  are independent, then  $\mathbb{E}[s^b(\langle z, a \rangle|[u, v])] = q_{[u, v]} \times \mathbb{P}_{X \sim \mathcal{D}_d}[X \in [u, v]]$ .

Finally, in addition to [Proposition 4.1](#), one also has the following trivial bounds.

**Proposition C.1.** Suppose that  $|u - \tau_d^p|, |v - \tau_d^p| = o(1)$ . Then, for some absolute constant  $C$ :

1.  $|\phi(\langle a, z \rangle|[u, v]) - \langle a, z \rangle| \times \mathbb{1}[\langle a, z \rangle \in [u, v]] \leq (v - u),$
2.  $|\psi(\langle a, z \rangle|[u, v])| \times \mathbb{1}[\langle a, z \rangle \in [u, v]] \leq C(v - u)^2,$
3.  $|\kappa(\langle u, v \rangle)| \times \mathbb{1}[\langle u, v \rangle \in [u, v]] \leq C(v - u).$

**Defect Edges After Algorithm 1** Observe that one can choose constants  $c_1, c_2 > c$  such that

$$\frac{1}{2} = \mathbb{P}\left[\langle \tilde{V}_i, \tilde{V}_j \rangle \leq \tau_d^p \mid \tau_d^p - \frac{c_1 \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d} \leq \langle \tilde{V}_i, \tilde{V}_j \rangle \leq \tau_d^p + \frac{c_2 \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}\right].$$

Hence, one can attempt to fix the “defect edges” by using fresh Bernoulli samples as follows.

## Multiple Rounds of Edge Flipping.

1. Draw independently  $H^0 \sim \mathsf{G}(n, p)$ ,  $H^1 \sim \mathsf{G}(n, 1/2)$ ,  $H^2 \sim \mathsf{G}(n, 1/2), \dots, H^T \sim \mathsf{G}(n, 1/2)$  (where  $T = O(\frac{\log d}{\log \kappa} + \frac{\log n}{\log d})$  to be specified later).
2. Draw  $V_1^{(0)}(0), V_2^{(0)}(0), \dots, V_n^{(0)}(0) \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})$ .
3. On input  $H^0, V_1^{(0)}(0), V_2^{(0)}(0), \dots, V_n^{(0)}(0)$  apply [Algorithm 1](#) to produce  $V_1^{(0)}(1), V_2^{(0)}(1), \dots, V_n^{(0)}(1)$  (where  $V_i^{(0)}(1) = V_i^{(i-1)}(0) = V_i^{(i-1)}$  in the notation of [Algorithm 1](#)). Set  $U_1 = \frac{c_1 \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}$ ,  $L_1 = \frac{c_2 \max(\sqrt{np}, \sqrt{\log n})(\log n)^{3/2}}{d}$  for some  $c_1, c_2 > c$  so that  $\mathbb{P}_{X \sim \mathcal{D}_d}[X \leq \tau_d^p | \tau_d^p - L_1 \leq X \leq \tau_d^p + U_1] = \frac{1}{2}$ .
4. For  $t = 1, 2, \dots, T$  :
  - (a) Initialize  $F_t = \emptyset$ .
  - (b) Go in lexicographic order of the pairs  $(j, i)$  :
    - i. If  $\langle V_j^{(i-1)}(t), V_i^{(i-1)}(t) \rangle \in [\tau_d^p - L_t, \tau_d^p + U_t]$ , add  $(ji)$  to  $F_t$ .
    - ii. Update  $V_j^{(i)}(t) = \rho_{V_i^{(i-1)}(t)}^{H_{ji}^t}(V_j^{(i-1)}(t) | [\tau_d^p - L_t, \tau_d^p + U_t])$ .
  - (c) Set  $V_i^{(0)}(t+1) = V_i^{(i-1)}(t)$  for all  $i$ . For some large absolute constant  $C$ , set

$$U_{t+1} = C(L_t + U_t) \max\left(\frac{\log^{3/4} n \sqrt{np(L_t + U_t)}}{d^{1/4}}, \frac{\log n}{\sqrt{d}}\right) \text{ and } L_{t+1} \quad (14)$$

such that  $\mathbb{P}_{X \sim \mathcal{D}_d}[X \leq \tau_d^p | \tau_d^p - L_{t+1} \leq X \leq \tau_d^p + U_{t+1}] = 1/2$ .

5. Output  $V_1^{(0)}(T), V_2^{(0)}(T), \dots, V_n^{(0)}(T)$  and the associated random geometric graph  $G^T$  where  $(G^T)_{ji} = \mathbf{1}[\langle V_j^{(0)}(T), V_i^{(0)}(T) \rangle \geq \tau_d^p]$ .

One can reason similarly as in [Section 4](#) using [Observation C.2](#) and [Observation C.1](#) to show at each moment of time, the collection of vectors  $V_1^{(0)}(t), V_2^{(1)}(t), \dots, V_s^{(i_s)}(t)$  is independent uniform over  $\mathbb{S}^{d-1}$ . Furthermore, with high probability, for all  $i, j, t$ , it holds that

$$|\langle V_j^{(i-1)}(t), V_i^{(i-1)}(t) \rangle - \langle V_j^{(i-1)}(t), V_i^{(i-1)}(t) \rangle| \leq C(L_t + U_t) \max\left(\frac{\log^{3/4} n \sqrt{np(L_t + U_t)}}{d^{1/4}}, \frac{\log n}{\sqrt{d}}\right)$$

in analogy to [Lemma 3.2](#). To prove this, one reasons in the exact same way as in [Lemma 4.4](#) to show that each martingale differences sum has at most  $nC \max((L_t + U_t) \times p \times \sqrt{d} \times \sqrt{\log 1/p}, \log n)$  non-zero terms for some absolute constant  $C$ . Furthermore, one can use the bounds from [Proposition C.1](#) in place of [Proposition 4.1](#) to argue that each of the non-zero terms is bounded by  $C(L_t + U_t)/d$  and the martingale-violation terms are bounded by  $C(L_t + U_t)^2/d$  and  $C(L_t + U_t)/d^{3/2}$  respectively.

One can similarly prove that

$$\begin{aligned} & |\langle V_j^{(0)}(t+1), V_i^{(0)}(t+1) \rangle - \langle V_j^{(0)}(t), V_i^{(0)}(t) \rangle| \\ & \leq (L_t + U_t) + C(L_t + U_t) \max\left(\frac{\log^{3/4} n \sqrt{np(L_t + U_t)}}{d^{1/4}}, \frac{\log n}{\sqrt{d}}\right) \end{aligned}$$

when  $t \geq 1$ .

One concludes that with high probability, if  $\langle V_j^{(0)}(t'), V_i^{(0)}(t') \rangle \notin [\tau_d^p - L_{t'}, \tau_d^p + U_{t'}]$  for some  $t' \geq 1$  and  $t'$  is the smallest index with this property, then

$$H_{ji}^{t'-1} = \mathbf{1}[\langle V_j^{(0)}(t'), V_i^{(0)}(t') \rangle \geq \tau_d^p] = \mathbf{1}[\langle V_j^{(0)}(t), V_i^{(0)}(t) \rangle \geq \tau_d^p] \text{ for all } t > t' \text{ w.h.p.}$$

Finally, Eq. (14) implies that there will be no more defect edges after  $T = O(\frac{\log d}{\log \kappa} + \frac{\log n}{\log d})$  rounds. Indeed, note that if  $I_t := L_t + U_t$  is the length of the  $t$ -th interval,<sup>7</sup> then

$$I_1 \leq \log^3 n \times \frac{\sqrt{np}}{d} \quad \text{and} \quad I_{t+1} \leq \log^2 n \times I_t \times \max\left(\frac{\sqrt{np}I_t}{d^{1/4}}, \frac{1}{\sqrt{d}}\right).$$

Now, suppose that  $d = (\log n)^6 np \times \kappa$  for some  $\kappa > 1$ . By induction, trivially follow the following statements.

**Proposition C.2.** Suppose that  $I_t \geq \frac{1}{d\sqrt{np}}$ , so that  $\frac{\sqrt{np}I_t}{d^{1/4}} \geq \frac{1}{\sqrt{d}}$ . Then,  $I_t \leq \frac{1}{\sqrt{d}} \left( \frac{np \log^6 n}{d} \right)^{\frac{3^{t-1}}{2^t}}$ .

In particular, after  $T_1 = O(\log \kappa d) = O(\frac{\log d}{\log \kappa})$  rounds, it will be the case that  $I_{T_1} \leq \frac{1}{d\sqrt{np}} = \frac{1}{d^2}$ .

**Proposition C.3.** Suppose that  $I_t \leq \frac{1}{d\sqrt{np}}$ , so that  $\frac{\sqrt{np}I_t}{d^{1/4}} \leq \frac{1}{\sqrt{d}}$ . Then,  $I_{t+s} \leq I_t \times \left( \frac{\log^2 n}{d} \right)^s$ .

The bounds on  $F_t$  stated in Theorem 3.1 follow immediately from the above propositions combined with Lemma 2.1. In particular, after  $T_2 = \log_d n = O(\frac{\log n}{\log d})$  rounds, it will be the case that  $I_{T_1+T_2} \leq \frac{1}{n^3 \sqrt{d}}$ . By Lemma 2.1,

$$\mathbb{P}_{U, V \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(\mathbb{S}^{d-1})} \left[ X \in \left[ \tau_d^p - \frac{1}{n^3 \sqrt{d}}, \tau_d^p + \frac{1}{n^3 \sqrt{d}} \right] \right] \leq C_D \frac{p}{n^3} \sqrt{\log 1/p} = o(1/n^3).$$

Hence, after  $T = O(\frac{\log d}{\log \kappa} + \frac{\log n}{\log d})$ , there would be no more defect edges which finishes the proof of Theorem 3.1.

---

<sup>7</sup>We are intentionally loose with the log-factors here for simplicity of exposition.