Check for updates

Adv. Appl. Probab. 1–27 (2026)
*doi:*10.1017/apr.2026.10076

RARE EVENT PROBABILITIES IN RANDOM GEOMETRIC GRAPHS

PRABHANKA DEKA,* *Beijing International Center for Mathematical Research, Peking University*
FANGZHOU LUO,** *** *AND*
BAICHUAN WU,** **** *Peking University*

Abstract

We study rare events in spherical and Gaussian random geometric graphs in high dimensions. In these models, the vertices correspond to points sampled uniformly at random on the d-dimensional unit sphere or correspond to d-dimensional standard Gaussian vectors, and edges are added between two vertices if the inner product between their corresponding points is greater than a threshold t_p, chosen such that the probability of having an edge is equal to p. We focus on two problems: (a) the probability that the random geometric graph is a complete graph and (b) the probability of observing an atypically large number of edges. We obtain asymptotically exponential decay rates depending on n and d of the probabilities of these rare events through a combination of geometric and probabilistic arguments.

Keywords: Random geometric graph; rare event; clique; large deviation; coupling; high-dimensional probability

2020 Mathematics Subject Classification: Primary 60C05
Secondary 05C80

1. Introduction

Random graphs and networks have become fundamental tools for modeling complex systems across a wide range of disciplines, including computer science, biology, sociology, and physics. These models help capture the intricate connectivity patterns observed in real-world networks, such as the internet, social interactions, and neural connections. The study of random graphs provides insights into the structural properties, robustness, and dynamical behaviors of large-scale networks, making them indispensable in modern data science and engineering applications. The mathematical study of random graphs traces its origins to the seminal work of Erdös and Rényi, who introduced a simple yet powerful model where edges are included independently with a fixed probability [13]. Since then, many versions of this simple model have been introduced and studied, such as the inhomogeneous random graph [7], the Chung–Lu model [1], and the configuration model [6], to name a few. Further, dynamic models, such as preferential attachment networks [5], where new vertices arrive one at a time and connect

Received 15 December 2025 UTC; accepted 1 July 2026 UTC.
* Postal address: Beijing International Center for Mathematical Research, Peking University, 100871 Beijing, China. Email: deka@bicmr.pku.edu.cn
** Postal address: School of Mathematical Sciences, Peking University, 100871 Beijing, China.
*** Postal address: Email: luo_joe2005@stu.pku.edu.cn
**** Postal address: Email: 2200010917@stu.pku.edu.cn

© The Author(s), 2026. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

TABLE 1. Lower and upper bounds for $\log \mathbb{P}(\mathcal{E})$.

Regime	Lower bound	Upper bound
$n^2 \lesssim d$	$-Cn^2$	$-cn^2$
$n^2/\log n \lesssim d \lesssim n^2$	$-Cn^2$	$-cn\sqrt{d}$
$n^{4/3+o(1)} \lesssim d \lesssim n^2/\log n$	$-Cn\sqrt{d\log n}$	$-cn\sqrt{d}$
$\log n \lesssim d \lesssim n^{4/3+o(1)}$	$-Cn\sqrt{d\log n}$	$-cn\sqrt{d\log n}$
$d \lesssim \log n$	$-Cnd$	$-cnd$

to the existing network, were developed because most real-world networks grow in size over time, and explain the scale-free degree distributions observed in many real-world networks. Another important class of random graphs, random geometric graphs (RGGs), introduced in [14], incorporate spatial structure by embedding vertices in a metric space and connecting them on the basis of geometric proximity. One of the main differences between the RGG and the other models mentioned is the nature of the dependence between the edges. In the Erdős–Rényi model (and most other models mentioned above), all edges are independent. However, in the RGG, because of the latent geometry, if the edges (i, j) and (j, k) are present in the graph, it is more likely that the edge (i, k) is also present, leading to a more intricate dependence between the edges. This phenomenon, often referred to as *clustering*, is captured by the clustering coefficient of the network.

In this work, we study two related RGG models, where n vertices are sampled uniformly at random on the d-dimensional unit sphere or where the n vertices correspond to d-dimensional standard Gaussian vectors. Edges are formed between pairs of vertices whose inner product exceeds a threshold t_p, chosen such that each edge appears with probability p. Following terminology in [4], we call the two models the *spherical RGG* (SRGG) or the *Gaussian RGG* (GRGG). These models differ from the more commonly studied RGG variant of Gilbert [14], where vertices are generated via a Poisson point process of intensity λ in $\mathbb{R}^d$ and two points are connected if their Euclidean distance is below a threshold radius r. The SRGG model is particularly relevant in high-dimensional statistics and machine learning, where inner-product-based similarity measures are widely used, such as in kernel methods and neural networks.

Although the model is fairly easy to describe, the dependence between the edges introduces major challenges in understanding and analyzing even very simple questions. In this paper, we investigate rare events in this model, focusing on two key problems: (a) the probability that the RGG is a complete graph and (b) the probability of observing an atypically large number of edges. These questions are motivated by applications in outlier detection, hypothesis testing, and understanding the extremal behavior of RGGs.

Previous works [10, 15, 16, 22, 23] studied large deviation principles of certain geometric functionals in various models, where they fixed d and let n go to infinity. We focus our attention on the dense RGG in high dimensions, where the edge probability $0 < p \le 1/2$ is fixed, and we let both n and d go to infinity. For the first problem, let $\mathcal{E}$ be the event that the SRGG on n vertices is a complete graph. Table 1 summarizes the bounds from our main results, Theorems 1 and 2, for various regimes of n and d, where the lower bound and upper bound columns list the dependence of $\log \mathbb{P}(\mathcal{E})$ on n and d. The constants c and C appearing in each row are possibly different, but do not depend on n and d. In the regimes $d \ge n^2$ and $d \le n^{4/3+o(1)}$, we obtain matching lower and upper bounds. In the intermediate regime, there is a discrepancy between the lower and upper bounds, which our proof technique was unable to address.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

For the upper tail large deviation of the number of edges, we obtain matching lower and upper bounds in all regimes. We consider the event that the number of edges present in an RGG exceeds $(1+\varepsilon)$ times the expected number of edges. The probability of this event decays as $\exp(-C_1 n^2)$ when $d \gtrsim \sqrt{n}$ and as $\exp(-C_2 n\sqrt{d})$ when $d \lesssim \sqrt{n}$. Here C_1 and C_2 are constants that depend on p and ε, but are independent of n and d. In very high dimensions, the decay is of order $\exp(-\Omega(n^2))$, similar to the Erdös-Rényi model, while in low dimensions, the two models behave differently.

1.1. Related works

There has been increasing interest in studying and understanding RGGs, especially in high dimensions. One of the first works that considered RGGs in high dimensions is [11], where the clique number in the regime $d \gg \log n$ is studied. An important question in high dimensions is understanding when the geometry of the underlying space is preserved. For dense RGGs ($p = \Theta(1)$), Bubeck *et al.* [9] established a sharp phase transition: geometry is detectable when the latent dimension $d \ll n^3$, but when $d \gg n^3$, this latent geometry is lost and the RGG becomes statistically indistinguishable from the Erdös-Rényi model. For sparse RGGs ($p = O(1/n)$), Bubeck *et al.* [9] presented a polynomial time test that can distinguish between RGG and Erdös-Rényi models when $d \ll (\log n)^3$, and conjectured that geometry is lost once $d \gg (\log n)^3$, and there has been much progress toward proving this in recent years (see [2], [8] and [19]).

Chatterjee and Harel [10] studied the large deviation principles for the number of edges in the RGG model where the vertices are generated by a Poisson point process of intensity n on the torus $\mathbb{T}^d$, and edges are added between vertices whose Euclidean distance is less than a threshold r_n. The distance threshold is chosen such that the expected number of edges grows as n^α for some $\alpha \in (0, 2)$. Their results show that conditioned on the graph having an atypically large number of edges, there exists a small ball that contains a large clique, and they further obtain a large deviation principle for the upper tail of the number of edges.

Ma *et al.* [20] recently considered an analogue of the SRGG called a *random sphere graph* to obtain a new lower bound on off-diagonal Ramsey numbers, which is the first exponential improvement over the classical lower bound obtained by Erdös in 1947. The proof is based on estimating the probability of the existence of smaller cliques in the random sphere graph (see Theorem 3.1 in [20]). Hunter *et al.* [17] provided a simpler proof and strengthened the results of [20] by considering the GRGG.

1.2. Organization of the paper

The rest of the paper is organized as follows. We first introduce important definitions and notation in Subsection 1.3. We state our main results in Section 2, and present an outline of our proof along with a few necessary tools in Section 3. Finally, in Sections 4 and 5, we prove the theorems stated in Section 2.

1.3. Definitions and notation

In this section, we introduce the models that we consider in this paper along with some necessary definitions and notation. A graph $G = (V, E)$ is a pair, where V is the set of vertices and $E \subseteq V \times V$ is the set of edges. The graphs we consider are simple, undirected graphs, i.e. the collection of edges consists of unordered pairs without repetitions and self-loops. A random graph $\mathcal{G}_n$ is a probability measure on the space of all simple graphs with n vertices.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

RGGs. For some $p \in [0, 1]$, we consider the following two RGG models.

(i) *Spherical model.* Let $(X_1, \dots, X_n)$ be independent and identically distributed (i.i.d.) random variables having distribution Unif($\mathbb{S}^{d-1}$). Let t_p be such that $\mathbb{P}(\langle X_1, X_2\rangle \geq t_p) = p$. Define the SRGG $\mathcal{G}_{\mathrm{S}}(n, p)$ to be the random graph on vertices $V(\mathcal{G}_{\mathrm{S}}(n, p)) = [n]$ and edges $E(\mathcal{G}_{\mathrm{S}}(n, p))$ given by $(i, j) \in E(\mathcal{G}_{\mathrm{S}}(n, p))$ if and only if $\langle X_i, X_j\rangle \geq t_p$.

(ii) *Gaussian model.* Let $(\tilde{X}_1, \dots, \tilde{X}_n)$ be i.i.d. random variables having distribution $\mathcal{N}(0, I_d)$, the standard normal distribution on $\mathbb{R}^d$. Let $\tilde{t}_p$ be such that $\mathbb{P}(\langle \tilde{X}_1, \tilde{X}_2\rangle \geq \tilde{t}_p) = p$. Define the GRGG $\mathcal{G}_{\mathrm{G}}(n, p)$ to be the random graph on vertices $V(\mathcal{G}_{\mathrm{G}}(n, p)) = [n]$ and edges $E(\mathcal{G}_{\mathrm{G}}(n, p))$ given by $(i, j) \in E(\mathcal{G}_{\mathrm{G}}(n, p))$ if and only if $\langle \tilde{X}_i, \tilde{X}_j\rangle \geq \tilde{t}_p$.

Note that the thresholds t_p and $\tilde{t}_p$ also depend on the dimension, d, of the underlying space, but we suppress that dependence in our notation. We write $t_{p,d}$ and $\tilde{t}_{p,d}$ when the dependence on the dimension needs to be made explicit. We can also think of these models as random graphs in the corresponding space ($\mathbb{S}^{d-1}$ or $\mathbb{R}^d$, respectively) by treating vertex i as being located at X_i or $\tilde{X}_i$. Note that in both models, because of the choice of *thresholds* t_p and $\tilde{t}_p$, the probability of having an edge between two vertices i and j is given by

$$\mathbb{P}((i, j) \in E(\mathcal{G}_{\cdot}(n, p))) = p.$$

Throughout this paper, we will work with RGGs in high dimensions, i.e. $d = d(n)$ grows with n.

Remark 1. For the special case $p = 1/2$, where we have $t_{1/2} = \tilde{t}_{1/2} = 0$, the two models induce the same random graph measure because of the spherical symmetry of the d-dimensional standard normal distribution.

Measures on $\mathbb{S}^{d-1}$ and $\mathbb{R}^d$. We write μ_{S}^{d-1} for the uniform probability measure on $\mathbb{S}^{d-1}$ and μ_{G}^{d} for the standard d-dimensional Gaussian probability measure on $\mathbb{R}^d$. When it is clear which case we are working on, we drop the subscript. Similarly, when the dimension is clear from the context, we drop the superscript.

Edge-weighted graphs. An *edge-weighted* graph $H(V)$ is a graph on vertex set V of size $|V| = n$, where each unordered pair (v, w) with $v, w \in V$ is assigned an $\mathbb{R}$-valued weight w_{vw}^H. We omit the superscript H when the graph is clear from the context. For two edge-weighted graphs $H(V)$ and $H'(V)$ on the same vertex set V, we write $H \leq H'$ if $w_{vw}^H \leq w_{vw}^{H'}$ for all $v, w \in V$. We list two classes of edge-weighted graphs that we use.

(i) The *constant-weighted graph* with all edge weights equal to c, which we denote by $c\mathbf{1}_n$.

(ii) The *inner-product-weighted graph* $H^U(n)$, where for each $v \in V$, we associate a point $x_v \in U \subseteq \mathbb{R}^d$, and for each unordered pair (v, w), we assign a weight $w_{vw} = \langle x_v, x_w\rangle$.

For our results, we are interested in the special cases of the *inner-product-weighted graph* with $U = \mathbb{S}^{d-1}$ (respectively $U = \mathbb{R}^d$), and the points associated with the vertices are chosen independently on U with distribution μ_{S}^{d-1} (respectively μ_{G}^{d}). We call these *random inner-product-weighted graphs*, whose law will be denoted by $\mathcal{H}_{\mathrm{S}}(n)$ in the spherical case and by $\mathcal{H}_{\mathrm{G}}(n)$ in the Gaussian case.

Asymptotic notation. Throughout this paper, we use Landau notation in the standard sense. Given two sequences f_n and g_n with $g_n > 0$, we say $f_n = O(g_n)$ or $f_n \lesssim g_n$ if there is a constant

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

M such that $|f_n| \le Mg_n$ for all n large enough, and say $f_n = o(g_n)$ if $\lim_{n\to\infty} f_n/g_n = 0$. We write $f_n \asymp g_n$ if $f_n = O(g_n)$ and $g_n = O(f_n)$.

Indicator functions. Throughout this paper, we use $1_A(\cdot)$ as the indicator function of a subset A, and $1\{\mathcal{B}\}$ as the indicator of an event $\mathcal{B}$.

2. Main results

In this section, we present our main results. Our first result is about the rare event that the SRGG is a complete graph. We will always assume that $p \le 1/2$ for technical reasons.

Theorem 1. *Let $G \sim \mathcal{G}_{\mathrm{S}}(n,p)$. Let $\mathcal{E}$ be the event that G is a complete graph on n vertices, i.e.*

$$\mathcal{E} = \mathcal{E}(n,d,p) := \{E(G) = \{(i,j) : 1 \le i \ne j \le n\}\} = \bigcap_{1\le i\ne j\le n} \{\langle X_i, X_j\rangle \ge t_p\}. \tag{2.1}$$

(i) ***Lower bound***. *There exists $C > 0$ such that for all n, $d \ge 2$,*

$$\mathbb{P}(\mathcal{E}) \ge \exp\left(-C\min\left(n^2, n\sqrt{d\log n}, nd\right)\right). \tag{2.2}$$

(ii) ***Upper bound***. *There exists $c > 0$ such that for all n, $d \ge 2$,*

$$\mathbb{P}(\mathcal{E}) \le \exp\left(-c\min\left(n^2, n\sqrt{d\log\left(\frac{n^{1-o(1)}}{d^{3/4}} \vee e\right)}, nd\right)\right). \tag{2.3}$$

We have a similar result for the same event for the GRGG.

Theorem 2. *Let $G \sim \mathcal{G}_{\mathrm{G}}(n,p)$. Let $\tilde{\mathcal{E}}$ be the event that G is a complete graph on n vertices, i.e.*

$$\tilde{\mathcal{E}} = \tilde{\mathcal{E}}(n,d,p) := \{E(G) = \{(i,j) : 1 \le i \ne j \le n\}\} = \bigcap_{1\le i\ne j\le n} \left\{\langle \tilde{X}_i, \tilde{X}_j\rangle \ge \tilde{t}_p\right\}. \tag{2.4}$$

(i) ***Lower bound***. *There exists $C > 0$ such that for all n, $d \ge 2$,*

$$\mathbb{P}(\tilde{\mathcal{E}}) \ge \exp\left(-C\min\left(n^2, n\sqrt{d\log n}, nd\right)\right). \tag{2.5}$$

(ii) ***Upper bound***. *There exists $c > 0$ such that for all n, $d \ge 2$,*

$$\mathbb{P}(\tilde{\mathcal{E}}) \le \exp\left(-c\min\left(n^2, n\sqrt{d\log\left(\frac{n^{1-o(1)}}{d^{3/4}} \vee e\right)}, nd\right)\right). \tag{2.6}$$

The next set of theorems presents bounds on the probability that the RGGs have an atypically large number of edges. Recall that in both models, the expected number of edges is $\binom{n}{2}p$. For any fixed $\varepsilon > 0$ such that $(1+\varepsilon)p < 1$, we obtain exponential rates of decay for the probability of the event $\left\{|E(\mathcal{G}_\cdot(n,p))| \ge (1+\varepsilon)\mathbb{E}\left[|E(\mathcal{G}_\cdot(n,p))|\right]\right\}$ as follows.

Theorem 3. *Let $G \sim \mathcal{G}_{\mathrm{S}}(n,p)$ and $|E(G)|$ be the number of edges in G. For any $\varepsilon > 0$, define the event*

$$\mathcal{I}_\varepsilon := \left\{|E(G)| \ge (1+\varepsilon)\binom{n}{2}p\right\}. \tag{2.7}$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Then there exist constants $c, C > 0$ such that

$$\exp\left(-c\min\left(n^2, n\sqrt{d}\right)\right) \le \mathbb{P}(\mathcal{I}_\varepsilon) \le \exp\left(-C\min\left(n^2, n\sqrt{d}\right)\right). \tag{2.8}$$

We also have a similar result for the GRGG.

Theorem 4. *Let $G \sim \mathcal{G}_{\mathrm{G}}(n, p)$ and let $|E(G)|$ be the number of edges in G. For any $\varepsilon > 0$, define the event*

$$\tilde{\mathcal{I}}_\varepsilon := \left\{|E(G)| \ge (1+\varepsilon)\binom{n}{2}p\right\}. \tag{2.9}$$

Then there exist constants $c, C > 0$ such that

$$\exp\left(-c\min\left(n^2, n\sqrt{d}\right)\right) \le \mathbb{P}(\tilde{\mathcal{I}}_\varepsilon) \le \exp\left(-C\min\left(n^2, n\sqrt{d}\right)\right). \tag{2.10}$$

Remark 2. The constants c and C both depend on the edge probability p and the deviation rate ε. For a more explicit characterization of the constants, see Propositions 10 and 11 for the lower bounds, and Remark 7 for the upper bounds.

Remark 3. While our results are not strong enough to describe the law of the SRGG (respectively the GRGG) conditioned on any one of the events $\mathcal{E}$ or $\mathcal{I}_\varepsilon$ (respectively $\tilde{\mathcal{E}}$ or $\tilde{\mathcal{I}}_\varepsilon$), we can still obtain some insight into the potential causes of these rare events from our proofs. When the dimension is very large, it is well known that the geometric structure vanishes from the RGG, making it indistinguishable from the Erdős–Rényi graph [9]. So for large d, the rare event probabilities behave like the Erdős–Rényi graph. In lower dimensions, these rare events can be achieved by aligning all the associated vectors weakly in some fixed direction. Conditioning on this alignment leads to our lower bounds. If these rare events were to occur, the sum of all pairwise inner products, $\sum_{i\neq j}\langle X_i, X_j\rangle$ (respectively $\sum_{i\neq j}\langle \tilde{X}_i, \tilde{X}_j\rangle$), would be larger than expected, and we obtain an upper bound by analyzing this statistic.

3. Tools and proof outline

3.1. Coupling the two models

As noted in Remark 1, the SRGG and the GRGG give the same random graph measure for the special case $p = 1/2$. While the two models differ for $p \neq 1/2$, there is a natural way to couple the two models using the same set of Gaussian vectors and the observation that when $\tilde{X} \sim \mathcal{N}(0, I_d)$, the projection of $\tilde{X}$ onto $\mathbb{S}^{d-1}$, $X = \tilde{X}/\|\tilde{X}\|$, is distributed uniformly on $\mathbb{S}^{d-1}$. Although this coupling is standard, most existing works focus on either the Gaussian model or the spherical model, choosing the one that is easier to deal with for the specific problem. For example, [11] works with the Gaussian model, where the independence of the Gaussian coordinates is more helpful, while [18] works with the spherical model, where the arguments rely on properties of the spherical harmonics, an orthonormal basis for functions on the sphere. Similarly to the proofs in [4], we use this coupling in both directions, leveraging the Gaussian model for some arguments, and turning to the spherical model for others. In this section, we derive results that relate the probabilities of observing a clique in the two models, and similar results for the probability of having an atypically large number of edges. As a consequence, it will suffice to obtain lower and upper bounds for only one of the two models,

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

combined with the results of this section, to complete the proofs of our theorems. The following lemma captures the asymptotic behavior of the thresholds t_p and $\tilde{t}_p$ as the dimension d goes to infinity.

Lemma 1. *For fixed $p \in (0, 1)$, we have*

$$\tilde{t}_p = (\bar{\Phi}^{-1}(p) + o(1))\sqrt{d} \quad \text{and} \quad t_p = (\bar{\Phi}^{-1}(p) + o(1))/\sqrt{d},$$

where $\bar{\Phi}(\cdot)$ is the tail of the standard Gaussian distribution.

Proof. Recall that $\tilde{t}_p$ is defined via $\mathbb{P}(\langle \tilde{X}_1, \tilde{X}_2 \rangle \geq \tilde{t}_p) = p$, where $\tilde{X}_1$ and $\tilde{X}_2$ are two independent $\mathcal{N}(0, I_d)$ random variables. Then we have

$$\langle \tilde{X}_1, \tilde{X}_2 \rangle = \sum_{j=1}^{d} \tilde{X}_1^{(j)} \tilde{X}_2^{(j)},$$

where $\tilde{X}_i^{(j)}$, the jth coordinate of $\tilde{X}_i$, is a standard one-dimensional Gaussian random variable. Note that the collection $\{\tilde{X}_i^{(j)} : i \in \{1, 2\}, 1 \leq j \leq d\}$ is independent. Further, $(\tilde{X}_1^{(j)} \tilde{X}_2^{(j)}, 1 \leq j \leq d)$ are independent, and $\mathbb{E}[\tilde{X}_1^{(j)} \tilde{X}_2^{(j)}] = 0$ and $\mathrm{Var}(\tilde{X}_1^{(j)} \tilde{X}_2^{(j)}) = 1$ for all $1 \leq j \leq d$. So $\mathrm{Var}(\langle \tilde{X}_1, \tilde{X}_2 \rangle) = d$. Using the central limit theorem, we get

$$\frac{\langle \tilde{X}_1, \tilde{X}_2 \rangle}{\sqrt{d}} \Rightarrow \mathcal{N}(0, 1) \tag{3.1}$$

as $d \to \infty$. Thus

$$p = \mathbb{P}\left(\langle \tilde{X}_1, \tilde{X}_2 \rangle \geq \tilde{t}_p\right) = \mathbb{P}\left(\frac{\langle \tilde{X}_1, \tilde{X}_2 \rangle}{\sqrt{d}} \geq \frac{\tilde{t}_p}{\sqrt{d}}\right) \to \mathbb{P}\left(\mathcal{N}(0, 1) \geq \frac{\tilde{t}_p}{\sqrt{d}}\right) = \bar{\Phi}\left(\frac{\tilde{t}_p}{\sqrt{d}}\right),$$

which proves the first claim. To see the second part, we use the coupling representation and observe that

$$\langle X_1, X_2 \rangle = \frac{\langle \tilde{X}_1, \tilde{X}_2 \rangle}{\|\tilde{X}_1\|_2 \|\tilde{X}_2\|_2}.$$

By the law of large numbers, $d^{-1/2}\|\tilde{X}_i\|_2 \xrightarrow{\mathrm{P}} 1$ for $i = 1, 2$. Combining with the central limit theorem in (3.1) and using Slutsky's lemma, we get

$$\sqrt{d}\,\langle X_1, X_2 \rangle \Rightarrow \mathcal{N}(0, 1) \tag{3.2}$$

as $d \to \infty$. The second claim then follows from a similar argument. □

Denote by $\mathbb{P}_{n,d}$ the law of n independent d-dimensional standard Gaussian vectors. Let $\tilde{\mathbf{X}} = (\tilde{X}_1, \ldots, \tilde{X}_n) \sim \mathbb{P}_{n,d}$, and recall that $X_i := \tilde{X}_i / \|\tilde{X}_i\|$ is distributed uniformly on $\mathbb{S}^{d-1}$. For any $1 \leq m \leq n$, define the events

$$\tilde{\mathcal{E}}(m, d, p) = \bigcap_{1 \leq i \neq j \leq m} \left\{\langle \tilde{X}_i, \tilde{X}_j \rangle \geq \tilde{t}_p\right\},$$

$$\mathcal{E}(m, d, p) = \bigcap_{1 \leq i \neq j \leq m} \left\{\langle X_i, X_j \rangle \geq t_p\right\},$$

where $\tilde{t}_p$ and t_p are as defined earlier. The following lemma establishes a relationship between the probabilities of the two events $\tilde{\mathcal{E}}$ and $\mathcal{E}$ using the coupling described above.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Lemma 2. *For fixed $p < 1/2$ and any small $\delta > 0$, there exist constants c_δ and Δ_δ, where c_δ and Δ_δ both go to 0 as δ goes to 0, and $p < q < p + \Delta_\delta$ such that the following inequalities hold:*

$$\mathbb{P}_{n,d}(\tilde{\mathcal{E}}(n, d, p)) \le \exp(-c_\delta dn) + 2^n \mathbb{P}_{n,d}(\mathcal{E}((1-\delta)n, d, q)), \tag{3.3}$$

$$\mathbb{P}_{n,d}(\mathcal{E}(n, d, p)) \le \exp(-c_\delta dn) + 2^n \mathbb{P}_{n,d}(\tilde{\mathcal{E}}((1-\delta)n, d, q)). \tag{3.4}$$

Proof. The proofs of (3.3) and (3.4) follow from similar arguments, and we present the details for only (3.3). Note that we can write

$$\begin{aligned}
\tilde{\mathcal{E}}(n, d, p) &= \bigcap_{1 \le i \ne j \le n} \left\{ \langle \tilde{X}_i, \tilde{X}_j \rangle \ge \tilde{t}_p \right\} \\
&= \bigcap_{1 \le i \ne j \le n} \left\{ \left\langle \frac{\tilde{X}_i}{\|\tilde{X}_i\|}, \frac{\tilde{X}_j}{\|\tilde{X}_j\|} \right\rangle \ge \frac{\tilde{t}_p}{\|\tilde{X}_i\| \|\tilde{X}_j\|} \right\} \\
&= \bigcap_{1 \le i \ne j \le n} \left\{ \langle X_i, X_j \rangle \ge t_p \cdot \frac{\tilde{t}_p}{t_p} \cdot \frac{1}{\|\tilde{X}_i\| \|\tilde{X}_j\|} \right\}.
\end{aligned}$$

By Lemma 1, $\tilde{t}_p / t_p \sim d$. Now, suppose for some subset $V \subseteq [n]$ that we have $\|\tilde{X}_i\| \ge (1+\delta)\sqrt{d}$ for all $i \in V$. Assuming without loss of generality that $V = [m]$, where $m = |V|$, if the event

$$\bigcap_{1 \le i \ne j \le n} \left\{ \langle \tilde{X}_i, \tilde{X}_j \rangle \ge \tilde{t}_p \right\} \cap \bigcap_{1 \le i \le m} \left\{ \|\tilde{X}_i\| \ge (1+\delta)\sqrt{d} \right\}$$

holds, then we also have

$$\bigcap_{1 \le i \ne j \le m} \left\{ \langle X_i, X_j \rangle \ge t_q \right\}$$

holds for a new threshold $t_q = (1+\delta)^{-2} t_p (1 + o(1))$, where $p < q < p + \Delta_\delta$, which implies that

$$\mathbb{P}_{n,d}\left(\tilde{\mathcal{E}}(n, d, p) \cap \bigcap_{1 \le i \le m} \left\{ \|\tilde{X}_i\| \ge (1+\delta)\sqrt{d} \right\} \right) \le \mathbb{P}_{n,d}\left(\mathcal{E}(m, d, q) \right).$$

Thus, to obtain (3.3), it suffices to show that we can take the subset V above to be of size δn. To see that, first note that for a Gaussian vector $\tilde{X}$, its length $\|\tilde{X}\|$ satisfies

$$\| \|\tilde{X}\| - \sqrt{d} \|_{\psi_2} \le C,$$

where $\|\cdot\|_{\psi_2}$ is the sub-Gaussian norm (see Theorem 3.1.1 in [25]). Thus, for a single Gaussian vector $\tilde{X}$ and any $\delta > 0$, we have $\mathbb{P}(\|\tilde{X}\| > (1+\delta)\sqrt{d}) \le \exp(-Cd)$ for some constant C that depends only on δ. A simple binomial concentration gives that the probability of the event $\mathcal{B}_{n,\delta} := \{ \#\{ i : \|\tilde{X}_i\| \ge (1+\delta)\sqrt{d} \} \ge \delta n \}$ is upper-bounded by

$$\mathbb{P}(\mathcal{B}_{n,\delta}) \le \exp(-Cd\delta n) = \exp(-cdn),$$

where $c > 0$ is again a constant that depends only on δ. Let $\mathcal{J}$ denote all the subsets of $[n]$ with size $\lfloor (1-\delta)n \rfloor$. A union bound gives

$$\begin{aligned}
\mathbb{P}_{n,d}(\tilde{\mathcal{E}}(n, d, p)) &\le \mathbb{P}_{n,d}(\tilde{\mathcal{E}}(n, d, p) \cap \mathcal{B}_{n,\delta}) + \mathbb{P}_{n,d}\left(\tilde{\mathcal{E}}(n, d, p) \cap \mathcal{B}_{n,\delta}^c\right) \\
&\le \exp(-cdn) + \sum_{I \in \mathcal{J}} \mathbb{P}_{n,d}\left(\tilde{\mathcal{E}}(n, d, p) \cap \bigcap_{i \in I} \left\{ \|\tilde{X}_i\| \ge (1+\delta)\sqrt{d} \right\} \right)
\end{aligned}$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

$$\leq \exp(-cdn) + 2^n \cdot \mathbb{P}\left(\tilde{\mathcal{E}}(n, d, p) \cap \bigcap_{i \leq \lfloor (1-\delta)n \rfloor} \{\|\tilde{X}_i\| \geq (1+\delta)\sqrt{d}\}\right)$$

$$\leq \exp(-cdn) + 2^n \cdot \mathbb{P}_{n,d}(\mathcal{E}((1-\delta)n, d, q))$$

as required. □

Similarly, we obtain the following lemma, which uses the coupling to obtain a relationship between the probabilities of the edge-deviation events in the two models. For any $1 \leq m \leq n$, define the events

$$\tilde{\mathcal{I}}_\varepsilon(m, d, p) = \left\{\langle \tilde{X}_i, \tilde{X}_j \rangle \geq \tilde{t}_p, \text{ for at least } (1+\varepsilon)p\binom{m}{2} \text{ pairs of } 1 \leq i < j \leq m\right\},$$

$$\mathcal{I}_\varepsilon(m, d, p) = \left\{\langle X_i, X_j \rangle \geq t_p, \text{ for at least } (1+\varepsilon)p\binom{m}{2} \text{ pairs of } 1 \leq i < j \leq m\right\}.$$

Lemma 3. *For fixed $p < 1/2$, $\varepsilon > 0$, and any small $\delta > 0$, there exist constants c_δ and Δ_δ, where c_δ and Δ_δ both go to 0 as δ goes to 0, $p < q < p + \Delta_\delta$ and $0 < \varepsilon' < \varepsilon$ such that the following inequalities hold:*

$$\mathbb{P}(\tilde{\mathcal{I}}_\varepsilon(n, d, p)) < \exp(-c_\delta dn) + 2^n \mathbb{P}(\mathcal{I}_{\varepsilon'}((1-\delta)n, d, q)), \tag{3.5}$$

$$\mathbb{P}(\mathcal{I}_\varepsilon(n, d, p)) < \exp(-c_\delta dn) + 2^n \mathbb{P}(\tilde{\mathcal{I}}_{\varepsilon'}((1-\delta)n, d, q)). \tag{3.6}$$

The proof of this lemma is identical to that for Lemma 2, and we omit the details. Although Lemmas 2 and 3 are not tight, they allow us to estimate $\mathbb{P}(\mathcal{E})$ and $\mathbb{P}(\tilde{\mathcal{E}})$ (and similarly $\mathbb{P}(\mathcal{I}_\varepsilon)$ and $\mathbb{P}(\tilde{\mathcal{I}}_\varepsilon)$) together, since the additive error term $\exp(-cdn)$ and the multiplicative factor 2^n are not comparable to the leading-order terms for the clique probabilities.

3.2. Geometric tools

In this section, we introduce *symmetric rearrangement*, the core geometric tool used to establish the upper bounds in our theorems. While the definitions are borrowed from [12] (see also [3] for a comprehensive treatment), we restate them here for the sake of completeness. The main idea is to use measure-preserving transformations to transform arbitrary, measurable subsets of the sphere into spherical caps, defined below.

Spherical caps. For a point $v \in \mathbb{S}^{d-1}$ and $a \in [0, 1]$, a *spherical cap* centered at v with measure a, denoted by $\mathcal{A}_v^a$, is the subset of $\mathbb{S}^{d-1}$ given by

$$\mathcal{A}_v^a = \{w \in \mathbb{S}^{d-1} : \langle w, v \rangle \geq h(a)\},$$

where the *altitude* of the spherical cap, $h(a) \in [-1, 1]$, is chosen such that $\mu(\mathcal{A}_v^a) = a$ (see Figure 1) Recall that $\mu = \mu_{\mathrm{S}}^{d-1}$ is the uniform probability measure on $\mathbb{S}^{d-1}$. By definition, the cap centered at v with measure p is the neighborhood of a vertex located at v in the SRGG $\mathcal{G}_{\mathrm{S}}(n, p)$. With this notation, there is an edge between two vertices i and j if and only if X_i lies in $\mathcal{A}_{X_j}^p$ and vice versa.

Let $f_d(x)$ be the density function of a one-dimensional marginal of μ_{S}^{d-1}, given by

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

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

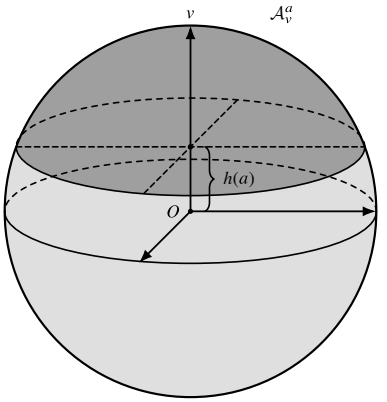


FIGURE 1. A spherical cap centered at v with measure a and height $h(a)$.

and define $\Psi_d(u) := \int_u^1 f_d(x)\,\mathrm{d}x$ to be the corresponding tail distribution function. Then $h(a) = \Psi_d^{-1}(a)$ and $t_{p,d} = \Psi_d^{-1}(p)$. If the dimension is clear from the context, we drop d from the notation and write $\Psi(p)$ and $\Psi^{-1}(p)$.

Symmetric rearrangement of a subset. We associate with $\mathbb{S}^{d-1}$ a canonical center $\mathbf{e} = (1, 0, \ldots, 0)$. Given a measurable subset $R \subset \mathbb{S}^{d-1}$, the *symmetric rearrangement* of R, denoted by $R^\bullet$, is a *spherical cap* centered at $\mathbf{e}$ that has the same measure as R does, i.e. $R^\bullet = \mathcal{A}_{\mathbf{e}}^{\mu(R)}$.

Symmetric rearrangement of a function. Let $f : \mathbb{S}^{d-1} \to \mathbb{R}_+$ be a non-negative real-valued measurable function. The *symmetric rearrangement* (also called *symmetric decreasing rearrangement* in the references in this section) of f is defined to be a function $f^\bullet : \mathbb{S}^{d-1} \to \mathbb{R}_+$ satisfying

$$\mu\big(\{x \in \mathbb{S}^{d-1} : f^\bullet(x) \geq t\}\big) = \mu\big(\{x \in \mathbb{S}^{d-1} : f(x) \geq t\}\big)$$

for each $t > 0$. Further, it is constant on each slice $\{x \in \mathbb{S}^{d-1} : \langle x, \mathbf{e}\rangle = h\}$, $-1 \leq h \leq 1$, and is non-decreasing in h. For a rigorous definition of $f^\bullet$, see Subsections 1.6 and 7.1 in [3]. It is easy to see that for any measurable subset $R \subseteq \mathbb{S}^{d-1}$, the symmetric rearrangement of its indicator function is equivalent to the indicator of $R^\bullet$.

In this section, we prove two propositions (Propositions 1 and 2) that are immediate corollaries of Theorem 2.2 in [12], restated below. Intuitively, it says that certain functionals of functions are maximized by their symmetric rearrangements. To state it formally, let $AL_2(\mathbb{R}_+^2)$ be the set of continuous functions $\psi : \mathbb{R}_+^2 \to \mathbb{R}_+$ satisfying that for every $x_1 < x_2$ and $y_1 < y_2$,

$$\psi(x_1, y_2) + \psi(x_2, y_1) \leq \psi(x_1, y_1) + \psi(x_2, y_2), \tag{3.7}$$

and ψ vanishes on the boundary, i.e. $\psi|_{x_i=0} = 0$ for $i = 1, 2$. A continuous function $\psi : \mathbb{R}_+^n \to \mathbb{R}_+$ is said to be in $AL_2(\mathbb{R}_+^n)$ if after any $n-2$ coordinates are fixed, ψ is in $AL_2(\mathbb{R}_+^2)$ in the remaining two coordinates. Note that for any $\psi \in C^2(\mathbb{R}_+^n)$, we have that ψ satisfies (3.7) if and only if $\partial_i \partial_j \psi \geq 0$ for every $i \neq j$.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Theorem 5. ([12], Theorem 2.2.) *Let $M = \mathbb{S}^{d-1}$, $f_i : M \to \mathbb{R}_+$ be n non-negative functions, ψ be a continuous function in $AL_2(\mathbb{R}_+^n)$, and $K_{i,j} : \mathbb{R}_+ \to \mathbb{R}_+$, $1 \le i < j \le n$, be increasing functions. Let*

$$I[f_1, \dots, f_n] := \int_{M^n} \psi(f_1(X_1), \dots, f_n(X_n)) \prod_{i<j} K_{i,j}(\langle X_i, X_j \rangle)\, \mathrm{d}\mu(X_1) \dots \mathrm{d}\mu(X_n),$$

where $\mathrm{d}\mu$ is the uniform measure on M. Then we have

$$I[f_1, \dots, f_n] \le I[f_1^\bullet, \dots, f_n^\bullet].$$

Remark 4. In the original statement, $K_{i,j}(\langle X_i, X_j \rangle)$ is replaced by $K_{ij}(d(X_i, X_j))$, where K_{ij} are decreasing functions and $d(\cdot, \cdot)$ is the geodesic distance on $\mathbb{S}^{d-1}$. Since $\langle X_i, X_j \rangle = \cos(d(X_i, X_j))$ is a decreasing function of $d(X_i, X_j)$, the version presented here is equivalent to the one in [12].

Using this theorem, we establish two results for SRGGs. Recall that for two edge-weighted graphs H_1 and H_2 on the same node set, we write $H_1 \le H_2$ if and only if $w_{ij}^{H_1} \le w_{ij}^{H_2}$ for all $i \ne j$. The following proposition states that for any fixed edge-weighted graph H_0 and an edge-weighted RGG H, the probability of the event $\{H \ge H_0\}$ under constraints $X_i \in A_i$ is optimized when the A_i's are *spherical caps* centered at the same point.

Proposition 1. *Let $A_i \subset \mathbb{S}^{d-1}$, $i \in [n]$, be measurable, and let $A_i^\bullet$ be their symmetric rearrangements. Let $H \sim \mathcal{H}_S(n)$. Then for any edge-weighted graph H_0 on the same node set $[n]$,*

$$\mathbb{P}\left(\{H \ge H_0\} \cap \bigcap_{i=1}^n \{X_i \in A_i\} \right) \le \mathbb{P}\left(\{H \ge H_0\} \cap \bigcap_{i=1}^n \{X_i \in A_i^\bullet\} \right). \tag{3.8}$$

Proof. Take $f_i(X) = 1_{A_i}(X)$, $i \in [n]$, and set $\psi(y_1, \dots, y_n) = y_1 y_2 \cdots y_n$. Further, set $K_{i,j}(x) = 1\{x \ge w_{ij}^{H_0}\}$. Then the left-hand side of (3.8) is equal to $I[f_1, \dots, f_n]$ and the right-hand side is equal to $I[f_1^\bullet, \dots, f_n^\bullet]$, and the conclusion immediately follows from Theorem 5. □

Proposition 2. *Let f be a non-negative function on $\mathbb{S}^{d-1}$, and let $X_1, \dots, X_n$ be i.i.d. uniformly on $\mathbb{S}^{d-1}$. Then for any $0 \le p \le 1$ and $\theta \ge 0$, we have*

$$\mathbb{E} \exp\left(\theta \left(\sum_{i=1}^n f(X_i) + \sum_{i<j} 1\{\langle X_i, X_j \rangle \ge t_p\} \right) \right) \le \mathbb{E} \exp\left(\theta \left(\sum_{i=1}^n f^\bullet(X_i) + \sum_{i<j} 1\{\langle X_i, X_j \rangle \ge t_p\} \right) \right).$$

Proof. Take $f_i(X) = \exp(\theta f(X))$ for all $i \in [n]$, and $\psi(y_1, \dots, y_n) = y_1 y_2 \cdots y_n$. Also set $K_{i,j}(x) = \exp(\theta 1\{x \ge t_p\})$. By the definition of symmetric rearrangements and using that $\exp(\cdot)$ is an increasing function, we have $f_i^\bullet(X) = \exp(\theta f^\bullet(X))$. It is a simple to verify that the conditions of Theorem 5 are satisfied, and the result follows immediately. □

3.3. Proof outline

The estimation of the probability of the RGGs being a complete graph and of the edge deviation events follows a similar approach. Instead of individually proving Theorems 1 and 2 (respectively Theorems 3 and 4), we use the coupling lemmas established in Subsection 3.1 to reduce the analysis to only one of the two models. For some bounds, the spherical model turns out to be simpler to handle, whereas for others, we work with the Gaussian model.

The lower bounds for the probability of observing a complete graph are proved in Subsection 4.1. We first obtain an Erdős–Rényi-type lower bound, $p^{\binom{n}{2}}$, for the spherical model

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

using a purely geometric argument. Note that this lower bound does not depend on the dimension d. The remaining bounds in (2.2) and (2.5) are proved by considering the Gaussian model, which is easier to handle with our method. To obtain these bounds, we consider the *'bias'* event that the first coordinates of all the vectors $\tilde{X}_1, \ldots, \tilde{X}_n$ are larger than $c\sqrt[4]{d \log n}$ for some constant $c > 0$. Conditioned on this event, we show that the GRGG is a clique with a probability bounded away from 0. We apply a similar idea to establish a lower bound on the edge deviation probability (see Subsection 5.1).

The upper bounds are proved in Subsection 4.2 and are more involved. We observe that conditioned on the GRGG being a complete graph, the statistic $S = \sum_{i \neq j} \langle \tilde{X}_i, \tilde{X}_j \rangle$ is atypically large. This statistic is easier to handle since it is the sum of products of independent standard Gaussian random variables, and we use a Bayesian argument to get one upper bound. We obtain another upper bound by considering the SRGG on n vertices to be obtained by a graph process, where the vertices are added one at a time. This gives us a corresponding subset-valued random process $(A_k)_{k \geq 0}$ starting from $A_0 = \mathbb{S}^{d-1}$, and conditioned on $X_1, \ldots, X_k$, A_k is the subset of $\mathbb{S}^{d-1}$ where X_{k+1} can be in so that the SRGG on the first $k+1$ vertices is still a complete graph. By construction, $A_{k+1} = A_k \cap \mathcal{A}^p_{X_{k+1}}$. We use symmetric rearrangements of subsets to transform the process $(A_k)_{k \geq 0}$ to a spherical-cap-valued random process. This reduced process is more tractable, and we prove one of the bounds in Theorems 1 and 2 by analyzing how the caps shrink with the arrival of each new vertex.

For the edge deviation event $\mathcal{I}_\varepsilon$, we first obtain a bound on the moment-generating function (MGF) of the number of edges in the RGG. Defining

$$f_k(\cdot) = \sum_{i=1}^{k} 1_{\mathcal{A}^p_{X_i}}(\cdot),$$

we can write the number of edges in the SRGG as $|E(G)| = \sum_{i=1}^{n} f_{i-1}(X_i)$. To get an upper bound on the MGF of $|E(G)|$, we use symmetric rearrangement of functions to transform f_k into a more tractable process, and then bound the MGF of $|E(G)|$ by the MGF of this reduced process. Use of the exponential Markov's inequality then completes the proofs.

4. Estimating clique probability

This section contains the proofs of Theorems 1 and 2, along with the statements and proofs of various lemmas and propositions that we use.

4.1. Lower bounds

In this section, we obtain the lower bounds in Theorems 1 and 2. The three different bounds in the exponent are derived via different techniques. First, we prove a lower bound that does not depend on the dimension d. As a consequence, we see that the clique probability for the SRGG is at least as large as that for the Erdős–Rényi graph, which matches the intuition that the RGGs are more clustered than the Erdős–Rényi graph.

Proposition 3 *For all n, $d \geq 2$ and $0 \leq p \leq \frac{1}{2}$,*

$$\mathbb{P}(\mathcal{E}) \geq p^{\binom{n}{2}} \geq \exp(-cn^2).$$

We prove the above proposition by using the following result for the spherical model. We drop the subscript and superscript from μ_{S}^{d-1} for simplicity.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Lemma 4. *Let μ denote the uniform probability measure on $\mathbb{S}^{d-1}$. Suppose $A \subset \mathbb{S}^{d-1}$ is convex, and $x \in A$. Then*

$$\mu\big(A \cap \mathcal{A}_x^p\big) \geq p \cdot \mu(A).$$

Proof. Recall that $\Psi(x) = \mathbb{P}(X_1^{(1)} > x)$ is the tail distribution function of a one-dimensional marginal of μ. We construct a contraction mapping $\varphi = \varphi_{p,x} : \mathbb{S}^{d-1} \setminus \{-x\} \to \mathcal{A}_x^p$ as follows. For every $y \in \mathbb{S}^{d-1} \setminus \{-x\}$, set $\varphi_{p,x}(y)$ to be the unique point z on the geodesic segment connecting x and y such that $\langle x, z\rangle = \Psi^{-1}(p \cdot \Psi(\langle x, y\rangle))$. Then $\mu \circ \varphi = p \cdot \mu$, and since A is convex with $x \in A$, $\varphi(A) \subset A$. So we immediately conclude that

$$\mu\big(A \cap \mathcal{A}_x^p\big) \geq \mu(\varphi(A)) = p \cdot \mu(A). \qquad \square$$

Proof of Proposition 3. Define the sets $A_k = \bigcap_{i=1}^k \mathcal{A}_{X_i}^p$, and consider the events

$$\mathcal{E}_k = \bigcap_{1 \leq i \neq j \leq k} \{\langle X_i, X_j\rangle \geq t_p\}, \quad 1 \leq k \leq n.$$

Note that $\mathcal{E}_n = \mathcal{E}$, and

$$\mathbb{P}(\mathcal{E}) = \prod_{i=1}^{n-1} \mathbb{P}(\mathcal{E}_{i+1} | \mathcal{E}_i).$$

Further, observe that $\mathbb{P}(\mathcal{E}_{k+1}|\mathcal{E}_k) = \mathbb{E}[X_{k+1} \in A_k | \mathcal{E}_k] = \mathbb{E}[\mu(A_k)|\mathcal{E}_k]$. Using Lemma 4 recursively, we obtain $\mathbb{E}[\mu(A_k)|\mathcal{E}_k] \geq p^k$, which we plug into the previous equation to conclude that

$$\mathbb{P}(\mathcal{E}) \geq p^{\binom{n}{2}} \geq p^{n^2}. \qquad \square$$

Note that the result above does not depend on d. But for k large enough, a typical A_k is too 'small', in the sense that A_{k+1} has nearly the same measure as A_k instead of $p\mu(A_k)$. We obtain a dimension-dependent bound in this regime by introducing a biasing argument, where we condition on the event that all the points have a slightly large first coordinate. The technique can be used in both the spherical case and the Gaussian case. We choose the Gaussian case for simplicity.

Proposition 4. *There exists a constant $C > 0$ such that for all $n, d \geq 2$ and $\log n < d$,*

$$\mathbb{P}(\tilde{\mathcal{E}}) \geq \frac{1}{2} \exp\left(-Cn\sqrt{d \log n}\right).$$

Proof. Suppose $\tilde{X}_i = (\tilde{X}_i^{(1)}, \ldots, \tilde{X}_i^{(d)})$, where the $\tilde{X}_i^{(k)}$'s are i.i.d. standard Gaussian random variables. For any constant $C > 0$, define the biased event

$$\mathcal{B}_C := \bigcap_{i=1}^n \left\{\tilde{X}_i^{(1)} \geq (\tilde{t}_p + C\sqrt{d \log n})^{1/2}\right\}.$$

Recall that the tail of the standard Gaussian distribution satisfies $\bar{\Phi}(x) \geq (1/10x)\mathrm{e}^{-x^2/2}$ for all $x \geq 2$ (see Proposition 2.1.2 in [25]), the $\tilde{X}_i^{(1)}$'s are independent, and $\tilde{t}_p = (\bar{\Phi}^{-1}(p) + o(1))\sqrt{d}$, so we have

$$\mathbb{P}(\mathcal{B}_C) = \bar{\Phi}\left(\left(\tilde{t}_p + C\sqrt{d \log n}\right)^{1/2}\right)^n \geq \exp\left(-Cn\sqrt{d \log n}\right).$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Conditionally on $\mathcal{B}_C$ happening, $\tilde{\mathcal{E}}$ fails only if

$$\sum_{\ell=2}^{d} \tilde{X}_i^{(\ell)} \tilde{X}_j^{(\ell)} < \tilde{t}_p - \tilde{X}_i^{(1)} \tilde{X}_j^{(1)} \leq -C\sqrt{d \log n}$$

for some $i, j \in [n]$. Since $(\tilde{X}_i^{(\ell)} \tilde{X}_j^{(\ell)})_{2 \leq \ell \leq d}$ are independent, mean zero, subexponential random variables, we can apply Bernstein's inequality (see Theorem 2.8.1 in [25]) to get

$$\mathbb{P}\left(\left| \sum_{\ell=2}^{d} \tilde{X}_i^{(\ell)} \tilde{X}_j^{(\ell)} \right| \geq C\sqrt{d \log n} \right) \leq 2 \exp\left[-c \min\left(C^2 \log n, C\sqrt{d \log n} \right) \right].$$

For $\log n < d$ and appropriately chosen constant C, the right-hand side above is bounded by $2n^{-3}$. Then, by a union bound on all possible pairs (i, j), we have

$$\mathbb{P}(\tilde{\mathcal{E}} | \mathcal{B}_C) \geq 1 - \binom{n}{2} 2n^{-3} \geq \frac{1}{2}.$$

Combining the above two inequalities, we get

$$\mathbb{P}(\tilde{\mathcal{E}}) = \mathbb{P}(\tilde{\mathcal{E}} | \mathcal{B}_C) \mathbb{P}(\mathcal{B}_C) \geq \frac{1}{2} \exp\left(-Cn\sqrt{d \log n} \right),$$

which completes the proof. □

If n is exponentially large in d, then the union bound used in the proof above will be redundant, and we can establish a better bound by adding a small bias to all coordinates.

Proposition 5. *There exists a constant $C > 0$ such that for all $n, d \geq 2$,*

$$\mathbb{P}(\tilde{\mathcal{E}}) \geq \exp(-Cnd).$$

Proof. Consider a different event

$$\mathcal{E}_0' := \bigcap_{\substack{1 \leq i \leq n \\ 1 \leq \ell \leq d}} \left\{ \tilde{X}_i^{(\ell)} \geq \left(\tilde{t}_p / d \right)^{1/2} \right\},$$

which immediately implies $\tilde{\mathcal{E}}$. Recall that $\tilde{t}_p / d = O(d^{-1/2})$, so we have

$$\mathbb{P}(\tilde{\mathcal{E}}) \geq \bar{\Phi}\left(\left(\tilde{t}_p / d \right)^{1/2} \right)^{nd} \geq 2^{-(1+o(1))nd} \geq \exp(-Cnd).$$ □

Combining the results of Propositions 3, 4, and 5 via the coupling result (Lemma 2) gives us the desired lower bounds in Theorems 1 and 2.

4.2. Upper bounds

We can infer from the proof of Proposition 4 that in some regimes (e.g. $d \gtrsim \sqrt{n}$) the event $\tilde{\mathcal{E}}$ happens if there is a small bias for a majority of the Gaussian vectors. To capture a similar phenomenon from a macroscopic perspective, we introduce a statistic $S := \sum_{1 \leq i \neq j \leq n} \langle \tilde{X}_i, \tilde{X}_j \rangle$. The following lemmas cover some properties of this statistic.

Lemma 5. *There exists a constant $c_1 = c_p > 0$ such that for all $n, d \geq 2$,*

$$\mathbb{E}[S | \tilde{\mathcal{E}}] \geq c_1 n^2 \sqrt{d}.$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Proof. If $p < 1/2$, we can deduce from Lemma 1 that

$$\mathbb{E}[S|\tilde{\mathcal{E}}] \geq (n^2 - n)\tilde{t}_p \geq c_1 n^2 \sqrt{d}.$$

For the case $p = 1/2$, $\tilde{t}_p = 0$, and we use a geometric argument to obtain the result. Note that in this case, the two RGG models coincide. So we first consider the spherical model and prove that $\mathbb{P}(\mathcal{E}|\,\langle X_1, X_2\rangle = x)$ is monotonically increasing in x. For $0 \leq x < x' \leq 1$, fix X_1, X_2 so that $\langle X_1, X_2\rangle = x$, and choose X_2' to lie on the geodesic connecting X_1 and X_2 such that $\langle X_1, X_2'\rangle = x'$. By construction, X_2' is a convex combination of X_1 and X_2, so if $X_3, \ldots, X_n$ are sampled uniformly at random on $\mathbb{S}^{d-1}$, independently of X_1, X_2, and X_2', the event that $(X_1, X_2, X_3, \ldots, X_n)$ forms a clique implies that $(X_1, X_2', X_3, \ldots, X_n)$ also forms a clique. Then we get

$$\mathbb{P}(\mathcal{E}|\,\langle X_1, X_2\rangle = x') \geq \mathbb{P}(\mathcal{E}|\,\langle X_1, X_2\rangle = x).$$

Recalling that we denote by f_d the density of $\langle X_1, X_2\rangle$, we find that Bayes's formula gives

$$\begin{aligned}
\mathbb{E}[\langle X_1, X_2\rangle|\mathcal{E}] &= \mathbb{P}(\mathcal{E})^{-1} \int_0^1 \mathbb{P}(\mathcal{E}|\,\langle X_1, X_2\rangle = x) x f_d(x)\,\mathrm{d}x \\
&\geq \mathbb{P}(\mathcal{E})^{-1} \left(\int_0^1 \mathbb{P}(\mathcal{E}|\,\langle X_1, X_2\rangle = x) f_d(x)\,\mathrm{d}x \right) \left(\int_0^1 x f_d(x)\,\mathrm{d}x \right) \left(\int_0^1 f_d(x)\,\mathrm{d}x \right)^{-1} \\
&= \mathbb{E}[\langle X_1, X_2\rangle|\langle X_1, X_2\rangle \geq 0].
\end{aligned}$$

The inequality in the second line follows from Chebyshev's integral inequality (see Chapter 9 in [21]). Further, the first two terms in the second line cancel out, and the remaining two terms are equal to the conditional expectation in the final line. Finally, recall that when $p = 1/2$, $t_p = \tilde{t}_p = 0$. Hence, under the coupling $X_i = \tilde{X}_i/\|\tilde{X}_i\|$, the event $\tilde{\mathcal{E}}$ coincides exactly with $\mathcal{E}$. Moreover, since $\|\tilde{X}_1\|$ and $\|\tilde{X}_2\|$ are independent of $X_1, \ldots, X_n$, we obtain

$$\begin{aligned}
\mathbb{E}[\langle \tilde{X}_1, \tilde{X}_2\rangle|\tilde{\mathcal{E}}] &= \mathbb{E}[\|\tilde{X}_1\| \cdot \|\tilde{X}_2\| \cdot \langle X_1, X_2\rangle\,|\mathcal{E}] = \mathbb{E}\|\tilde{X}_1\| \cdot \mathbb{E}\|\tilde{X}_2\| \cdot \mathbb{E}[\langle X_1, X_2\rangle|\mathcal{E}] \\
&\geq \mathbb{E}\|\tilde{X}_1\| \cdot \mathbb{E}\|\tilde{X}_2\| \cdot \mathbb{E}[\langle X_1, X_2\rangle|\langle X_1, X_2\rangle \geq 0] = \mathbb{E}[\langle \tilde{X}_1, \tilde{X}_2\rangle|\langle \tilde{X}_1, \tilde{X}_2\rangle \geq 0].
\end{aligned}$$

Since $\langle \tilde{X}_1, \tilde{X}_2\rangle$ has asymptotic distribution $\mathcal{N}(0, d)$, the term on the right is bounded below by $c_1\sqrt{d}$ for some $c_1 > 0$. Replacing $(\tilde{X}_1, \tilde{X}_2)$ by any pair $(\tilde{X}_i, \tilde{X}_j)$, $i \neq j$, we obtain the same bound. Summing up over all i, j gives $\mathbb{E}[S|\tilde{\mathcal{E}}] \geq c_1 n^2 \sqrt{d}$. □

Since S is the sum of i.i.d. subexponential random variables, it concentrates around its mean, and we have the following estimate using Bernstein's inequality.

Lemma 6. *There exists a constant $c_2 > 0$ such that for all n, $d \geq 2$ and $t > 0$,*

$$\mathbb{P}(|S| \geq t) \leq 2 \exp\left(-c_2 \min\left(\frac{t^2}{n^2 d}, \frac{t}{n} \right) \right).$$

Proof. Note that we can decompose the sum as

$$S = \sum_{k=1}^{d} \left(\sum_{1 \leq i \neq j \leq n} \tilde{X}_i^{(k)} \tilde{X}_j^{(k)} \right) =: \sum_{k=1}^{d} Y^{(k)},$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

where $Y^{(k)}$ are i.i.d. mean zero random variables with the same distribution as $Y := (\sum_{i=1}^n Y_i)^2 - \sum_{i=1}^n Y_i^2$, in which the Y_i's are i.i.d. standard normal variables. For every positive integer q, the following moment estimate on Y holds:

$$\|Y\|_{L_q} \le \left\| \left(\sum_{i=1}^n Y_i \right)^2 \right\|_{L_q} + \sum_{i=1}^n \|Y_i^2\|_{L_q} \le 6nq.$$

This coincides with property (b) in Proposition 2.7.1 in [25]. It follows that Y is a subexponential variable with $\|Y\|_{\psi_1} \lesssim n$, where $\|\cdot\|_{\psi_1}$ is the subexponential norm (see Subsection 2.7 in [25]). Then Lemma 6 follows from Bernstein's inequality. □

Lemmas 5 and 6 together give the following upper bound.

Proposition 6. *There exists a constant $c > 0$ such that for all n, $d \ge 2$,*

$$\mathbb{P}(\tilde{\mathcal{E}}) \le \exp\left(-c \min\left(n^2, n\sqrt{d}\right)\right).$$

Proof. We use the following extension of Markov's inequality for the tail of positive random variables. Suppose A is measurable and ξ is a real-valued random variable such that $a = \mathbb{E}[\xi|A]$ is finite. Then for any $b < a$,

$$(a-b)\mathbb{P}(A) = \mathbb{E}[(\xi - b)1_A] \le \mathbb{E}(\xi - b)^+ \Rightarrow \mathbb{P}(A) \le \frac{1}{a-b}\mathbb{E}(\xi - b)^+, \tag{4.1}$$

where $(x)^+ = \max(x, 0)$. Take $A = \tilde{\mathcal{E}}$, $\xi = S$, and $b = a/2$. By Lemma 5, since $a = \mathbb{E}[S|\tilde{\mathcal{E}}]$, $b \ge c_1 n^2\sqrt{d}/2$. With use of Lemma 6 in the first line below, the right-hand side of (4.1) can be bounded by

$$\begin{aligned}
\mathbb{E}(S-b)_+ &\le \int_b^{+\infty} \left(2\exp(-c_2 t/n) + 2\exp\left(-c_2 t^2/n^2 d\right)\right) \mathrm{d}t \\
&\le \frac{2n}{c_2}\exp(-c_2 b/n) + \frac{n^2 d}{cb}\exp\left(-c_2 b^2/n^2 d\right) \\
&\le C(n+\sqrt{d})\exp\left(-c\min\left(n^2, n\sqrt{d}\right)\right)
\end{aligned}$$

for some $c > 0$. Therefore

$$\mathbb{P}(\tilde{\mathcal{E}}) \le \frac{1}{a-b}\mathbb{E}[(S-b)_+] \le \exp\left(-(c+o(1))\min\left(n^2, n\sqrt{d}\right)\right).$$ □

Remark 5. In principle, we can consider the same argument in the spherical model and use the statistic $\sum_{i\ne j}\langle X_i, X_j\rangle$ to derive a similar upper bound. But in the spherical model, $\sum_{i\ne j}\langle X_i, X_j\rangle$ is not the sum of i.i.d. random variables, which presents a minor obstacle. So we chose to obtain the bounds for the Gaussian model and rely on Lemma 2 to get bounds of the same order for the spherical model.

Another way of dealing with the upper bound is to analyze the convex set-valued random process A_k, defined in the proof of Proposition 3. If we define H_k to be the induced subgraph of $\mathcal{G}_S(n, p)$ on the last k vertices $n-k+1, \ldots, n$, then we have the inequality

$$\mathbb{P}(\mathcal{E}) \le \mathbb{P}\left(\{H_{n-k} \ge t_{p,d}\mathbf{1}_{n-k}\} \cap \bigcap_{i=k+1}^{n} \{X_i \in A_k\}\right).$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

This reduces analyzing $\mathcal{E}$ to the study of the random process A_k, which is still quite complicated. The following proposition allows us to compare A_k with its symmetric rearrangement $A_k^\bullet$, which is easier to analyze. Define $\mathcal{C}_0 = \mathbb{S}^{d-1}$, and iteratively define $\mathcal{C}_{k+1} := (\mathcal{C}_k \cap \mathcal{A}^p_{X_{k+1}})^\bullet$. It is a sequence of spherical caps all centered at $\mathbf{e}$.

Proposition 7. *For any $0 \le k \le n$, we have*

$$\mathbb{P}(\mathcal{E}) \le \mathbb{P}\left(\{H_{n-k} \ge t_{p,d}\mathbf{1}_{n-k}\} \cap \bigcap_{i=k+1}^{n} \{X_i \in \mathcal{C}_k\} \right). \tag{4.2}$$

Proof. When $k = 0$, $H_n \ge t_{p,d}\mathbf{1}_n$ is exactly the event $\mathcal{E}$, and $\mathcal{C}_0 = \mathbb{S}^{d-1}$, so the two sides of (4.2) are equal. To prove the statement, we show that the right-hand side of (4.2) is non-decreasing in k. Note that for any $0 \le k \le n-1$,

$$\begin{aligned}
&\mathbb{P}\left(\{H_{n-k} \ge t_{p,d}\mathbf{1}_{n-k}\} \cap \bigcap_{i=k+1}^{n} \{X_i \in \mathcal{C}_k\} \right) \\
&= \mathbb{P}\left(\{H_{n-k-1} \ge t_{p,d}\mathbf{1}_{n-k-1}\} \cap \bigcap_{i=k+1}^{n} \left\{X_i \in \left(\mathcal{C}_k \cap \mathcal{A}^p_{X_{k+1}}\right)\right\} \right) \\
&\le \mathbb{P}\left(\{H_{n-k-1} \ge t_{p,d}\mathbf{1}_{n-k-1}\} \cap \bigcap_{i=k+2}^{n} \left\{X_i \in \left(\mathcal{C}_k \cap \mathcal{A}^p_{X_{k+1}}\right)\right\} \right) \\
&\le \mathbb{P}\left(\{H_{n-k-1} \ge t_{p,d}\mathbf{1}_{n-k-1}\} \cap \bigcap_{i=k+2}^{n} \{X_i \in \mathcal{C}_{k+1}\} \right).
\end{aligned}$$

Here, in the first inequality we drop $Bigg\{X_{k+1} \in (\mathcal{C}_k \cap \mathcal{A}^p_{X_{k+1}})\}$, and in the second inequality, we apply Proposition 1, which completes the proof. □

Defining $\eta_k := \mu(\mathcal{C}_k)$, and dropping the event $\{H_{n-k} \ge t_{p,d}\mathbf{1}_{n-k}\}$ on the right-hand side of (4.2), we get

$$\mathbb{P}(\mathcal{E}) \le \mathbb{E}\left(\eta_k^{n-k}\right).$$

$(\eta_k)_{k\ge 0}$ is a random, decreasing process, and the following lemma quantifies the difference $\eta_k - \eta_{k+1}$. Recall that $h(\eta)$ is the *altitude* of a spherical cap with measure η.

Lemma 7. *For any $\eta \ge 0$, let $\eta' = \mu_{\mathrm{S}}^{d-1}(\mathcal{A}^\eta_{\mathbf{e}} \cap \mathcal{A}^p_Y)$, where Y is uniformly distributed on $\mathbb{S}^{d-1}$. Then there exist $c, c_0 > 0$ such that for any $\delta \in (0, c)$ and $\eta \in [\exp(-c_0 d\delta^2/2), 1/2]$,*

$$\mathbb{P}\left(\eta - \eta' \ge c\delta \exp\left(-\frac{d}{2}(1+3\delta^2)h(\eta)^2 \right) \right) \ge 1 - \exp\left(-d\delta^2/2\right).$$

Recall that $\Psi(\cdot)$ is the tail distribution function of a one-dimensional marginal of μ_{S}^{d-1}, i.e. $\Psi(t) = \mathbb{P}(X_1^{(1)} > t)$, where $X_1^{(1)}$ is the first coordinate of X_1. To prove Lemma 7 we use the following distributional approximation result obtained by Sodin [24].

Lemma 8. *There exist constants $C, C_1, C_2 > 0$ and a sequence $\varepsilon_d \searrow 0$ such that for all $t \in [0, C]$,*

$$(1-\varepsilon_d)\bar{\Phi}(\sqrt{d}t)\exp\left(-C_1 dt^4\right) \le \Psi(t) \le (1+\varepsilon_d)\bar{\Phi}(\sqrt{d}t)\exp\left(-C_2 dt^4\right).$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Proof of Lemma 7. Note that η' depends only on $Y^{(1)}$, the first coordinate of Y. Recall the density of $Y^{(1)}$ is given by

$$f_d(x) = \frac{\Gamma(d/2)}{\Gamma((d-1)/2)\sqrt{\pi}}(1-x^2)^{(d-3)/2}.$$

From Lemma 8, we know that $\mathbb{P}(Y^{(1)} > \delta) \le \exp(-d\delta^2/2)$. Next we show that if $Y^{(1)} \le \delta$ then $\eta - \eta' \ge c\delta \exp(-(d/2)(1+3\delta^2)h(\eta)^2)$. Since η' is monotonically increasing in $Y^{(1)}$, it suffices to consider $Y = (\delta, -\sqrt{1-\delta^2}, 0, \dots, 0)$. Note that $\{(x, y, \dots) \in \mathbb{S}^{d-1} : x > h(\eta), y > \delta x/\sqrt{1-\delta^2}\} \subseteq \mathcal{A}_{\mathbf{e}}^{\eta} \setminus \mathcal{A}_Y^p$. Further, the two-dimensional marginal of μ_{S}^{d-1} on the first two coordinates has density function

$$f_d(x, y) = \frac{d-2}{2\pi}(1-x^2-y^2)^{(d-4)/2}.$$

By direct computation, when $Y^{(1)} \le \delta$ we have

$$\begin{aligned}
\mu\big(\mathcal{A}_{\mathbf{e}}^{\eta} \setminus \mathcal{A}_{Y^{(1)}}^p\big) &\ge \int_{x^2+y^2\le 1} 1_{\{x>h(\eta), y>\delta/\sqrt{1-\delta^2}x\}} f_d(x, y)\,\mathrm{d}x\,\mathrm{d}y \\
&\ge cd \int_{h(\eta)}^{1/2} \left(\int_{\delta x/\sqrt{(1-\delta^2)}}^{\sqrt{1-x^2}} (1-x^2-y^2)^{(d-4)/2} dy \right) \mathrm{d}x \\
&= cd \int_{h(\eta)}^{1/2} (1-x^2)^{(d-3)/2} \left(\int_{\delta x/\sqrt{(1-\delta^2)(1-x^2)}}^{1} (1-t^2)^{(d-4)/2}\,\mathrm{d}t \right) \mathrm{d}x \\
&\ge cd \int_{h(\eta)}^{1/2} (1-x^2)^{d/2} \delta x (1-2\delta^2 x^2)^{d/2}\,\mathrm{d}x \\
&\ge c\delta(1-(1+2\delta^2)h(\eta)^2)^{d/2} \\
&\ge c\delta \exp\left(-\frac{d}{2}(1+3\delta^2)h(\eta)^2\right).
\end{aligned}$$

For the last two inequalities to hold, we need $h(\eta) < c'\delta$ for some constant c' small enough. This can be achieved by taking c_0 to be sufficiently small in the assumption, which completes the proof. □

Using Lemma 7, we obtain the following two rates of decay of the sequence $(\eta_n)_{n\ge 1}$ as n goes to infinity, for different regimes of d.

Proposition 8. *For a sufficiently small constant $c > 0$ and $\log n \lesssim d \lesssim n^{4/3}$,*

$$\mathbb{P}\left(\eta_n \ge \exp\left(-c\sqrt{d}\cdot\frac{\log n - \log d^{3/4}}{\sqrt{\log n}}\right)\right) \le \exp(-cn\sqrt{d\log n}).$$

Proof. We first select a constant $0 < c \ll 1$ and $\delta = c\sqrt[4]{\log n/d}$. Let $I \subset [n]$ contain all indices i such that the first coordinate of X_i is smaller than δ. Since I is a Bernoulli process with success parameter $p_0 \ge 1 - \exp(-d\delta^2/2) \ge 1 - \exp(-c\sqrt{d\log n})$,

$$\mathbb{P}(|I| \le n/2) \le \sum_{i=0}^{n/2} p_0^i (1-p_0)^{n-i} \binom{n}{i} \le 2^n (1-p_0)^{n/2} \le \exp(-cn\sqrt{d\log n}).$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Let $N = (\log n - \log d^{3/4})/\sqrt{\log n}$. On the event $|I| \geq n/2$ we prove $\eta_n < \exp(-c\sqrt{d}N)$. Since $\eta_1 = p \leq 1/2$, we may assume $\eta_i \in [\exp(-c\sqrt{d\log n}), 1/2]$ for any $1 \leq i \leq n$. From the proof of Lemma 7, we deduce that for $i \in I$, $\eta_i - \eta_{i+1} > c\delta \exp(-(d/2)(1+3\delta^2)h(\eta_i)^2)$. Denote $I_k = I \cap \{i : \eta_i \in (e^{-k-1}, e^{-k}]\}$ for any integer $1 \leq k \leq \sqrt{d\log n}$. Using Lemma 8, we get that for any $i \in I_k$,

$$e^{-k-1} \leq C \exp\left(-\frac{d}{2}h(\eta_i)^2\right). \tag{4.3}$$

Then

$$\eta_i - \eta_{i+1} \geq c\delta e^{-(1+3\delta^2)k},$$

and therefore

$$|I_k| \leq \left\lceil \frac{e^{-k}}{c\delta e^{-(1+3\delta^2)k}} \right\rceil \leq 1 + \frac{1}{c\delta} e^{3\delta^2 k}. \tag{4.4}$$

Summing the above inequality for $0 \leq k \leq c\sqrt{d}N$ gives

$$\sum_{k=0}^{c\sqrt{d}N} |I_k| \leq c\sqrt{d}N + 1 + \frac{e^{3c\delta^2\sqrt{d}N}}{c\delta(e^{3\delta^2}-1)} \leq \frac{n}{2}.$$

The last inequality holds for sufficiently small $c > 0$. Therefore $\eta_n < \exp(-c\sqrt{d}N)$ when $|I| \geq n/2$. □

Proposition 9. *For sufficiently small constant $c > 0$, and $d \lesssim \log n$,*

$$\mathbb{P}(\eta_n \geq \exp(-cd)) \leq \exp(-cnd).$$

Proof. The proof is similar to that of Proposition 8. We select a small constant $\delta \ll 1$, and define I and I_k for $1 \leq k \leq d$ as before. In this regime, I is again a Bernoulli process, but with parameter $p_0 \geq 1 - \exp(-cd)$. Therefore $\mathbb{P}(|I| \leq n/2) \leq \exp(-cnd)$. The inequalities (4.3) and (4.4) still hold, and we can sum over k to get that

$$\sum_{k=0}^{cd} |I_k| \leq 1 + cd + \frac{e^{3c\delta^2 d}}{c\delta(e^{3\delta^2}-1)} \leq \frac{n}{2}$$

holds for a small enough constant c. Thus $\eta_n \leq e^{-cd}$ if $|I| \geq n/2$. □

Finally, we complete the proof of the upper bounds in Theorems 1 and 2. Substituting Proposition 8 into (4.2), when $\log n \lesssim d \lesssim n^{4/3}$, we obtain

$$\mathbb{P}(\mathcal{E}) \leq \mathbb{E}\left(\eta_{n/2}^{n/2}\right) \leq \mathbb{P}\left(\eta_n \geq \exp\left(-c\sqrt{d} \cdot \frac{\log n - \log d^{3/4}}{\sqrt{\log n}}\right)\right)$$
$$+ \exp\left(-c\sqrt{d} \cdot \frac{\log n - \log d^{3/4}}{\sqrt{\log n}} \cdot \frac{n}{2}\right) \leq \exp\left(-cn\sqrt{d \log \frac{n^{1-o(1)}}{d^{3/4}}}\right).$$

Similarly, in the regime $d \lesssim \log n$, we use Proposition 9 to get

$$\mathbb{P}(\mathcal{E}) \leq \mathbb{E}\left(\eta_{n/2}^{n/2}\right) \leq \mathbb{P}(\eta_n \geq \exp(-cd)) + \exp\left(-cd \cdot \frac{n}{2}\right) \leq \exp(-cnd).$$

The results above, along with Proposition 6 and Lemma 2, complete the proof of the upper bounds in Theorems 1 and 2.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Remark 6. Since we lose some information during the symmetric rearrangement transformations, the upper bound we finally get is not tight when $n^{4/3} \lesssim d \lesssim n^2$. The Bayesian argument gives a matching upper bound when $n^2 \lesssim d$. Meanwhile, in the intermediate regime, it is not immediately clear what the correct order of decay should be, and is left as an open problem for the future.

5. Estimating edge deviation probability

In this section we present the proofs of Theorems 3 and 4, our main results on the large deviation for edges of the SRGG and GRGG models.

5.1. Lower bounds

The first lower bound is a corollary of Proposition 3, the Erdős–Rényi-type lower bound for the clique probability. It follows from noting that conditioned on the existence of a large enough clique, the event $\mathcal{I}_\varepsilon$ happens with a probability bounded away from zero.

Proposition 10. *For all n and $d \geq 2$,*

$$\mathbb{P}(\mathcal{I}_\varepsilon) \geq \exp\left[-\left(\frac{p\log\frac{1}{p}}{2(1-p)}\varepsilon + o(1)\right)n^2\right].$$

Proof. Let $\alpha = \sqrt{p\varepsilon/(1-p)}$, and let $\mathcal{E}_{\lceil \alpha n+1\rceil}$ be as defined in the proof of Proposition 4.1. We can infer from the proof that

$$\mathbb{P}(\mathcal{E}_{\lceil \alpha n+1\rceil}) \geq p^{\binom{\lceil \alpha n+1\rceil}{2}} \geq \exp\left(-\frac{p\log\frac{1}{p}}{2(1-p)}\varepsilon n^2 - n\right). \tag{5.1}$$

Note that conditioned on $\mathcal{E}_{\lceil \alpha n+1\rceil}$ happening, the connection probability of edges not included in this small clique remains equal to p. That implies if $G \sim \mathcal{G}_{\mathrm{S}}(n,p)$,

$$\mathbb{E}[|E(G)||\mathcal{E}_{\lceil \alpha n+1\rceil}] \geq \binom{\lceil \alpha n+1\rceil}{2} + p\left[\binom{n}{2} - \binom{\lceil \alpha n+1\rceil}{2}\right] \geq p(1+\varepsilon)\binom{n}{2} + 1.$$

Hence

$$\mathbb{P}\left(|E(G)| \geq p(1+\varepsilon)\binom{n}{2}\middle|\mathcal{E}_{\lceil \alpha n+1\rceil}\right) \geq \frac{1}{\binom{n}{2}}\left[\mathbb{E}[|E(G)||\mathcal{E}_{\lceil \alpha n+1\rceil}] - p(1+\varepsilon)\binom{n}{2}\right] \geq \frac{1}{n^2}.$$

Combining with (5.1) completes the proof. □

For the second lower bound, we apply the biasing argument as we did in the proof of Proposition 4. Again, we prove the result for the Gaussian model and use Lemma 3 to obtain the result for the spherical model.

Proposition 11. *For all n and $d \geq 2$,*

$$\mathbb{P}(\tilde{\mathcal{I}}_\varepsilon) \geq \exp\left[-\left(\frac{\bar{\Phi}^{-1}(p) - \bar{\Phi}^{-1}(p(1+\varepsilon))}{2} + o(1)\right)n\sqrt{d}\right].$$

Proof. Suppose $\tilde{X}_i = (\tilde{X}_i^{(1)}, \ldots, \tilde{X}_i^{(d)})$. Let $C := C_{n,d,p,\varepsilon} = \sqrt{\tilde{t}_{p,d} - \tilde{t}_{p(1+\varepsilon)+1/n,d-1}}$, and define the biased event

$$\mathcal{B}_C := \bigcap_{1\leq i\leq n}\{\tilde{X}_i^{(1)} > C\}.$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Then for all $1 \le i < j \le n$ we have

$$\mathbb{P}(\langle \tilde{X}_i, \tilde{X}_j \rangle \ge \tilde{t}_{p,d} | \mathcal{B}_C) \ge \mathbb{P}\left(\sum_{k=2}^{d} \tilde{X}_i^{(k)} \tilde{X}_j^{(k)} \ge \tilde{t}_{p,d} - C^2 \right) = (p(1+\varepsilon) + 1/n) \wedge 1.$$

Let $G \sim \mathcal{G}_{\mathrm{G}}(n,p)$. By summing over all i, j, we obtain

$$\mathbb{E}[|E(G)||\mathcal{B}_C] \ge p(1+\varepsilon)\binom{n}{2} + 1.$$

Similarly to in the proof of Proposition 10, we get

$$\mathbb{P}\left(|E(G)| \ge p(1+\varepsilon)\binom{n}{2} \Big| \mathcal{B}_C \right) \ge \frac{1}{n^2}. \tag{5.2}$$

Since $1/n \to 0$ as $n \to \infty$, we can deduce from Lemma 1 that

$$C^2 = \tilde{t}_{p,d} - \tilde{t}_{p(1+\varepsilon)+1/n,d-1} = \left[\bar{\Phi}^{-1}(p) - \bar{\Phi}^{-1}\left(p(1+\varepsilon)\right) + o(1) \right] \sqrt{d}. \tag{5.3}$$

Using the standard estimation for Gaussian tails, we obtain

$$\mathbb{P}(\mathcal{B}_C) = \bar{\Phi}(C)^n \ge \exp\left[-(1+o(1))C^2 n/2 \right]. \tag{5.4}$$

Combining (5.2), (5.3), and (5.4) concludes the proof. □

Combining Propositions 10 and 11 together with Lemma 3 gives the lower bounds in Theorems 3 and 4.

5.2. Upper bounds

We consider the number of edges $|E(G)|$ as the final term of the following stochastic process. Denote by $|E_k(G)|$ the number of edges between $X_1, \dots, X_k$, and define $f_k(x) = \sum_{1 \le i \le k} 1_{\mathcal{A}^p_{X_i}}(x)$. Then $f_{k-1}(X_k)$ counts the number of edges from X_k to $X_1, \dots, X_{k-1}$, and

$$|E_k(G)| - |E_{k-1}(G)| = f_{k-1}(X_k).$$

Similarly to the idea in Subsection 4.2, we will use Proposition 2 here to obtain a bound on the MGF of $|E(G)| = \sum_{k=1}^{n} f_{k-1}(X_k)$ by a sequence of symmetric rearrangements. We start by defining another sequence of functions $(g_k)_{k \ge 0}$ by setting $g_0 \equiv 0$ on $\mathbb{S}^{d-1}$, and iteratively defining

$$g_k := \left(g_{k-1} + 1_{\mathcal{A}^p_{X_k}} \right)^{\bullet}, \quad 1 \le k \le n-1.$$

Using these functions, we define an interpolating process

$$m_k := \sum_{i=1}^{k} g_{i-1}(X_i) + \sum_{i=k+1}^{n} g_k(X_i) + \sum_{k+1 \le i < j \le n} 1\{\langle X_i, X_j \rangle \ge t_p\}, \quad 1 \le k \le n. \tag{5.5}$$

By definition, we have $m_0 = |E(G)|$ and $m_n = \sum_{k=1}^{n} g_{k-1}(X_k)$. For each k, the first term in (5.5) corresponds to the symmetric rearrangement of the edges between $X_1, \dots, X_k$, the second term corresponds to the symmetric rearrangement of the edges between $X_1, \dots, X_k$ and $X_{k+1}, \dots, X_n$, and the final term counts the true edges between $X_{k+1}, \dots, X_n$ in G.

Proposition 12. *The MGFs of m_k are non-decreasing in k, i.e. for any $\theta \ge 0$, we have*

$$\mathbb{E} \exp\left(\theta |E(G)|\right) = \mathbb{E}\mathrm{e}^{\theta m_0} \le \mathbb{E}\mathrm{e}^{\theta m_1} \le \cdots \le \mathbb{E}\mathrm{e}^{\theta m_n}.$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Proof. It suffices to show that $\mathbb{E}e^{\theta m_{k-1}} \le \mathbb{E}e^{\theta m_k}$ for any $1 \le k \le n$. Fix k. Then by first conditioning on $X_1, \dots, X_k$, we get

$$\mathbb{E}[e^{\theta m_{k-1}} | X_1, \dots, X_k]$$
$$= \exp\left(\theta \sum_{i=1}^{k} g_{i-1}(X_i)\right) \mathbb{E}_{X_{k+1},\dots,X_n} \exp\left(\theta\left(\sum_{i=k+1}^{n} g_{k-1}(X_i) + \sum_{k\le i<j\le n} 1\left\{\langle X_i, X_j\rangle \ge t_p\right\}\right)\right).$$

Here $\mathbb{E}_{X_{k+1},\dots,X_n}$ means taking expectation over randomness of $X_{k+1}, \dots, X_n$, and we drop the subscript and simply write $\mathbb{E}$ to simplify the notation. Let $f(X) = g_{k-1}(X) + 1_{\mathcal{A}^p_{X_k}}(X)$, and note that $f^\bullet = g_k$. By applying Proposition 2, we obtain

$$\begin{aligned}
&\mathbb{E}\exp\left(\theta\left(\sum_{i=k+1}^{n} g_{k-1}(X_i) + \sum_{k\le i<j\le n} 1\{\langle X_i, X_j\rangle \ge t_p\}\right)\right)\\
&= \mathbb{E}\exp\left(\theta\left(\sum_{i=k+1}^{n} f(X_i) + \sum_{k+1\le i<j\le n} 1\{\langle X_i, X_j\rangle \ge t_p\}\right)\right)\\
&\le \mathbb{E}\exp\left(\theta\left(\sum_{i=k+1}^{n} f^\bullet(X_i) + \sum_{k+1\le i<j\le n} 1\{\langle X_i, X_j\rangle \ge t_p\}\right)\right)\\
&= \mathbb{E}\exp\left(\theta\left(\sum_{i=k+1}^{n} g_k(X_i) + \sum_{k+1\le i<j\le n} 1\{\langle X_i, X_j\rangle \ge t_p\}\right)\right).
\end{aligned}$$

Thus

$$\mathbb{E}[e^{\theta m_{k-1}} | X_1, \dots, X_k] \le \mathbb{E}[e^{\theta m_k} | X_1, \dots, X_k].$$

Taking the expectation over $X_1, \dots, X_k$ completes the proof. □

We complete the proof of Theorems 3 and 4 using the following upper bound on m_n.

Proposition 13. *Let m_n be defined as in (5.5). Then*

$$m_n \le (p + p\varepsilon/2)\binom{n}{2} \tag{5.6}$$

holds with probability at least

$$1 - \exp(-\min(c_1(p,\varepsilon)n^2, c_2(p,\varepsilon)n\sqrt{d})).$$

Here $c_1(p,\varepsilon)$ and $c_2(p,\varepsilon)$ are positive constants that depend only on p and ε.

We first prove Theorem 3 assuming Proposition 13, using the following general lemma on stochastic order. The proof of Proposition 13 is delayed to the end of this section.

Lemma 9. *Suppose Y and Y' are random variables satisfying $Y, Y' \le 1$, and for all $\theta \ge 0$,*

$$\mathbb{E}e^{\theta Y} \le \mathbb{E}e^{\theta Y'}. \tag{5.7}$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Then for all $\delta \geq 0$,

$$\log \mathbb{P}(Y \geq \delta) \leq \frac{\delta}{2} \log \mathbb{P}\left(Y' \geq \frac{\delta}{2}\right) + \log 2. \tag{5.8}$$

Proof. Note that for all $\theta \geq 0$,

$$\mathbb{E}e^{\theta Y'} = \mathbb{E}\left[e^{\theta Y'} 1\left\{Y' \geq \frac{\delta}{2}\right\}\right] + \mathbb{E}\left[e^{\theta Y'} 1\left\{Y' < \frac{\delta}{2}\right\}\right] \leq e^{\theta}\mathbb{P}\left(Y' \geq \frac{\delta}{2}\right) + e^{\theta(\delta/2)}.$$

Applying Markov's inequality, we obtain

$$\log \mathbb{P}(Y \geq \delta) \leq \log\left(e^{-\theta\delta}\mathbb{E}e^{\theta Y}\right) \leq -\theta\delta + \log \mathbb{E}e^{\theta Y'} \leq -\frac{\theta\delta}{2} + \log\left(1 + \mathbb{P}\left(Y' \geq \frac{\delta}{2}\right) e^{\theta - \frac{\theta\delta}{2}}\right).$$

We get (5.8) by taking $\theta = -\log \mathbb{P}(Y' \geq \delta/2) \geq 0$, since

$$\log\left(1 + \mathbb{P}\left(Y' \geq \frac{\delta}{2}\right) e^{\theta - (\theta\delta/2)}\right) = \log\left(1 + e^{-(\theta\delta/2)}\right) \leq \log 2. \qquad \square$$

Proof of Theorem 3. Let

$$Y = \frac{|E(G)|}{\binom{n}{2}} - p,\ Y' = \frac{m_n}{\binom{n}{2}} - p.$$

It is clear that $Y, Y' \leq 1$, and they satisfy (5.7) due to Proposition 12. Combining with Proposition 13, we get

$$\log \mathbb{P}(Y \geq \varepsilon) \leq \frac{\varepsilon}{2} \log \mathbb{P}\left(Y' \geq \frac{\varepsilon}{2}\right) + \log 2 \leq -\frac{\varepsilon}{2} \min\left(c_1(p, \varepsilon)n^2, c_2(p, \varepsilon)n\sqrt{d}\right) + \log 2.$$

Note that event $\mathcal{I}_\varepsilon$ is exactly $Y \geq \varepsilon$, so the proof is complete. $\square$

For the proof of Proposition 13, it is sufficient to bound each $g_{k-1}(X_k)$ individually, and then sum over k. We are going to bound an analogue of the MGF and then apply Markov's inequality. Given a measurable function $g : \mathbb{S}^{d-1} \to \mathbb{R}$, a subset $V \subset \mathbb{S}^{d-1}$, and $\lambda \in \mathbb{R}$, define

$$\mathcal{M}[g; V](\lambda) := \int_V \exp\left(\lambda g(X)\right) \mathrm{d}\mu(X).$$

Let $q = p + p\varepsilon/32$, and select a constant $a > 0$ that satisfies

$$t_{p,d} - \frac{a^2}{\sqrt{d}} > \max\left(t_{q,d-1}, \left(1 - \frac{a^2}{\sqrt{d}}\right) t_{q,d-1}\right) \tag{5.9}$$

for all d sufficiently large. Such an a exists since Lemma 1 indicates a gap of order $\Omega(d^{-1/2})$ between $t_{p,d}$ and $t_{q,d-1}$. Consider the strip L_t, and the complement of a cap V_t, defined as

$$L_t := \left\{w \in \mathbb{S}^{d-1} : |\langle w, e\rangle| \leq \frac{t}{\sqrt[4]{d}}\right\},$$

$$V_t := \left\{w \in \mathbb{S}^{d-1} : \langle w, e\rangle \leq \frac{t}{\sqrt[4]{d}}\right\}.$$

The key observation is the following lemma.

Lemma 10. *Suppose $\lambda \geq 0$. Then for any function g such that $g = g^\bullet$ almost everywhere, and vector $Y \in L_a$,*

$$\mathcal{M}\left[g + 1_{\mathcal{A}_Y^p}; V_a\right](\lambda) \leq \left(1 - q + qe^{\lambda} + 2e^{-a^2\sqrt{d}/2 + O(1) + \lambda}\right)\mathcal{M}[g; V_a](\lambda).$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Proof. Let $S_t = \{X \in \mathbb{S}^{d-1} : \langle X, e\rangle = t\}$ be a $(d-2)$-dimensional slice at t and let μ_t denote the uniform measure on S_t. We first prove that for all $|t| \le a/\sqrt[4]{d}$,

$$\int_{S_t} \exp\left[\lambda \cdot 1_{\mathcal{A}_Y^p}(X)\right] \mathrm{d}\mu_t(X) \le 1 - q + q\mathrm{e}^{\lambda}. \tag{5.10}$$

This is equivalent to showing $\mu_t(S_t \cap \mathcal{A}_Y^p) \le q$, which can be computed directly. Suppose $Y = (y, \sqrt{1-y^2}, 0, \dots, 0)$. Then

$$S_t \cap \mathcal{A}_Y^p = \left\{X = (t, u, \dots) : ty + u\sqrt{1-y^2} \ge t_{p,d}\right\}.$$

After scaling by $1/\sqrt{1-t^2}$, $S_t \cap \mathcal{A}_Y^p$ is a spherical cap on $\mathbb{S}^{d-2}$ with altitude $(t_{p,d} - ty)/\sqrt{(1-y^2)(1-t^2)}$, and so

$$\mu_t\big(S_t \cap \mathcal{A}_Y^p\big) = \Psi_{d-1}\left(\frac{t_{p,d} - ty}{\sqrt{(1-y^2)(1-t^2)}}\right).$$

Recall that $|t|, |y| \le a/\sqrt[4]{d}$, and a satisfies (5.9). This is upper-bounded by $\Psi_{d-1}(t_{q,d-1}) = q$, which proves (5.10).

Next we show that the integral on $V_a - L_a$ is small. For those $t < -a/\sqrt[4]{d}$, the integral on $V_a - L_a$ can be bounded by that on L_a since g is non-decreasing in the direction $\mathbf{e}$ as

$$\begin{aligned}\mathcal{M}\big[g + 1_{\mathcal{A}_Y^p}; V_a - L_a\big](\lambda) &\le \mathrm{e}^{\lambda}\mathcal{M}[g; V_a - L_a](\lambda)\\ \le \mathrm{e}^{\lambda}\frac{\mu(V_a - L_a)}{\mu(V_a)}\mathcal{M}[g; V_a](\lambda) &\le 2e^{-a^2\sqrt{d}/2 + O(1) + \lambda}\mathcal{M}[g; V_a](\lambda).\end{aligned}$$

Here in the last inequality we apply Lemma 8 on $\mu(V_a - L_a)$. Multiplying (5.10) by the value of g on S_t and integrating over $|t| \le a/\sqrt[4]{d}$ gives

$$\begin{aligned}\mathcal{M}\big[g + 1_{\mathcal{A}_Y^p}; V_a\big](\lambda) &= \mathcal{M}\big[g + 1_{\mathcal{A}_Y^p}; L_a\big](\lambda) + \mathcal{M}\big[g + 1_{\mathcal{A}_Y^p}; V_a - L_a\big](\lambda)\\ &\le \big(1 - q + q\mathrm{e}^{\lambda} + 2\mathrm{e}^{-a^2\sqrt{d}/2 + O(1) + \lambda}\big)\mathcal{M}[g; V_a](\lambda).\end{aligned}$$

□

Finally, we derive an upper bound on $g_{k-1}(X_k)$ from Lemma 10, and complete the proof of Proposition 13.

Proof of Proposition 13. Let λ and c_0 be constants to be determined later, and define

$$h := \min\left(\frac{c_0\sqrt{n}}{\sqrt[4]{d}}, \frac{a}{2}\right).$$

Define the index set

$$I := \{1 \le k \le n : X_k \notin L_h\}.$$

These vectors $\{X_k\}_{k \in I}$ are atypical and have large contributions to $\mathcal{M}[g_k; V_a](\lambda)$. Consider the bad event $\mathcal{B} := \{|I| \ge p\varepsilon n/16\}$ of having too many atypical vectors. Its probability can be controlled as follows. By Lemma 8, for a single k,

$$\mathbb{P}(X_k \notin L_h) = 2\Psi(h/\sqrt[4]{d}) \le \exp\left(-\sqrt{d}h^2/2 + O(1)\right).$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

A union bound gives

$$\mathbb{P}(\mathcal{B}) \le 2^n \cdot \mathbb{P}\left(X_k \notin L_h \text{ for all } 1 \le k \le \frac{p\varepsilon n}{16}\right) \le 2^n \exp\left(\frac{p\varepsilon n}{16}\left(-\sqrt{d}h^2/2 + O(1)\right)\right)$$
$$\le \exp\left(-p\varepsilon n\sqrt{d}h^2/64\right).$$

Since $g_k = (g_{k-1} + 1_{\mathcal{A}^p_{X_k}})^\bullet$, we can use Lemma 10 to bound $\mathcal{M}[g_k; V_a](\lambda)$ as

$$\mathcal{M}[g_k; V_a](\lambda) \le \mathcal{M}\left[g_{k-1} + 1_{\mathcal{A}^p_{X_k}}; V_a\right](\lambda) \le \begin{cases} (1 - q + qe^\lambda + o(1))\mathcal{M}[g_{k-1}; V_a](\lambda), & k \notin I, \\ e^\lambda \mathcal{M}[g_{k-1}; V_a](\lambda), & k \in I. \end{cases}$$

Note that $\mathcal{M}[g_0; V_a](\lambda) = \mu(V_a) \le 1$. We can deduce from above that for all $1 \le k \le n$,

$$\mathcal{M}[g_k; V_a](\lambda) \le (1 - q + qe^\lambda + o(1))^k \exp\left(\lambda |I \cap \{1, \dots, k\}|\right).$$

In particular, on $\mathcal{B}^c$, $|I \cap \{1, \dots, k\}| \le p\varepsilon n/16$, and thus for all $1 \le k \le n$,

$$\log \mathcal{M}[g_k; V_a](\lambda) \le k \log(1 - q + qe^\lambda + o(1)) + \lambda p\varepsilon n/16.$$

From here we derive a bound on $g_{k-1}(X_k)$ by using Markov's inequality. Since $h \le a/2 = O(1)$, we have

$$\mu(V_a - V_h) \asymp \mu\left(V_h^c\right) = \exp\left(-\sqrt{d}h^2/2 + O(1)\right).$$

Then for all $k \notin I$, $X_k \in L_h$ and

$$\mathcal{M}[g_{k-1}; V_a](\lambda) \ge \mathcal{M}[g_{k-1}; V_a - V_h](\lambda) \ge \mu(V_a - V_h)e^{\lambda g_{k-1}(X_k)},$$

which implies

$$g_{k-1}(X_k) \le \frac{1}{\lambda}\log \mathcal{M}[g_{k-1}; V_a](\lambda) - \frac{1}{\lambda}\log \mu(V_a - V_h)$$
$$\le \frac{\log(1 - q + qe^\lambda + o(1))}{\lambda}(k-1) + p\varepsilon n/16 + \frac{\sqrt{d}h^2}{2\lambda} + O(1).$$

For $k \in I$, $X_k \notin L_h$, and $g_{k-1}(X_k) \le n$ is trivially true. Finally, we upper-bound m_n. Summing over k, we get that on $\mathcal{B}^c$,

$$m_n = \sum_{k \notin I} g_{k-1}(X_k) + \sum_{k \in I} g_{k-1}(X_k)$$
$$\le \frac{\log(1 - q + qe^\lambda + o(1))}{\lambda}\binom{n}{2} + p\varepsilon n^2/16 + \frac{n\sqrt{d}h^2}{2\lambda} + O(n) + p\varepsilon n^2/16.$$

Choose λ such that $(1/\lambda)\log(1 - q + qe^\lambda) \le p + p\varepsilon/16$, and $c_0 \le \sqrt{p\varepsilon\lambda}/10$. Such a λ exists since

$$\lim_{\lambda \to 0^+} \frac{\log(1 - q + qe^\lambda)}{\lambda} = q < p + p\varepsilon/16.$$

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

Then m_n is bounded by

$$(p+p\varepsilon/16)\binom{n}{2}+p\varepsilon n^2/16+p\varepsilon n^2/16+O\left(\frac{n}{\lambda}\right)+p\varepsilon n^2/16\le(p+p\varepsilon/2)\binom{n}{2}.$$

So we can conclude that (5.6) holds on $\mathcal{B}^c$. The proof of Proposition 13 is completed by noting that

$$\mathbb{P}(\mathcal{B})\le\exp\left(-p\varepsilon n\sqrt{d}h^2/64\right)\le\exp\left(-\min\left(c_1(p,\varepsilon)n^2,c_2(p,\varepsilon)n\sqrt{d}\right)\right),$$

where $c_1(p,\varepsilon)=p\varepsilon c_0^2/64$ and $c_2(p,\varepsilon)=p\varepsilon a^2/64$ are constants. □

Remark 7. We can determine the dependence of the constants in Theorem 4 (and also Theorem 3) on p and ε by estimating the constants used in the proof of Proposition 13 Take $\lambda=\varepsilon/100$ and $c_0=\sqrt{p}\varepsilon/100$. Note that $\bar{\Phi}^{-1}(p)\sim\sqrt{\log 1/p}$, so we may take

$$a^2=\frac{\bar{\Phi}^{-1}(p)-\bar{\Phi}^{-1}(p+p\varepsilon/32)}{2}=\Omega\left(\frac{\varepsilon}{\sqrt{\log(1/p)}}\right).$$

Then the corresponding constants satisfy $c_1(p,\varepsilon)=\Omega(p^2\varepsilon^3)$ and $c_2(p,\varepsilon)=\Omega(p\varepsilon^2/\sqrt{\log 1/p})$. This gives

$$\mathbb{P}(\mathcal{I}_\varepsilon)\le\exp\left(-\min\left(\Omega(p^2\varepsilon^4)n^2,\Omega\left(\frac{p\varepsilon^3}{\sqrt{\log 1/p}}\right)n\sqrt{d}\right)\right).$$

Acknowledgements

We thank Jian Ding for introducing the problem, and for valuable discussions and suggestions during the early stages of this project.

Funding information

This work was supported by the National Key R&D Program of China (grant no. 2023YFA1010103).

Competing interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

References

[1] AIELLO, W., CHUNG, F. AND LU, L. (2000). A random graph model for massive graphs. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, pp. 171–180.
[2] AVRACHENKOV, K. E. AND BOBU, A. V. (2020). Cliques in high-dimensional random geometric graphs. *Appl. Network Sci.* **5**, 92.
[3] BAERNSTEIN, A. I. (2019). *Symmetrization in Analysis*, New Mathematical Monographs, vol. 36. Cambridge University Press, Cambridge, UK.
[4] BANGACHEV, K. AND BRESLER, G. (2024). On the Fourier coefficients of high-dimensional random geometric graphs. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pp. 549–560.
[5] BARABÁSI, A.-L. AND ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.
[6] BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Comb.* **1**, 311–316.
[7] BOLLOBÁS, B., JANSON, S. AND RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Struct. Algorithms* **31**, 3–122.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press

[8] BRENNAN, M., BRESLER, G. AND NAGARAJ, D. (2020). Phase transitions for detecting latent geometry in random graphs. *Probab. Theory Related Fields* **178**, 1215–1289.
[9] BUBECK, S., DING, J., ELDAN, R. AND RÁCZ, M. Z. (2016). Testing for high-dimensional geometry in random graphs. *Random Struct. Algorithms* **49**, 503–532.
[10] CHATTERJEE, S. AND HAREL, M. (2020). Localization in random geometric graphs with too many edges. *Ann. Probab.* **48**, 574–621.
[11] DEVROYE, L., GYÖRGY, A., LUGOSI, G. AND UDINA, F. (2011). High-dimensional random geometric graphs and their clique number. *Electron. J. Probab.* **16**, 2481–2508.
[12] DRAGHICI, C. (2005). Rearrangement inequalities with application to ratios of heat kernels. *Potential Anal.* **22**, 351–374.
[13] ERDÖS, P. AND RÉNYI, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hungaricum Acad. Sci.* **5**, 17–60.
[14] GILBERT, E. N. (1961). Random plane networks. *J. Soc. Ind. Appl. Math.* **9**, 533–543.
[15] HIRSCH, C., JAHNEL, B. AND TÓBIÁS, A. (2020). Lower large deviations for geometric functionals. *Electron. Commun. Probab.* 25, Paper No. **41**, 12.
[16] HIRSCH, C. AND OWADA, T. (2023). Large deviation principle for geometric and topological functionals and associated point processes. *Ann. Appl. Probab.* **33**, 4008–4043.
[17] HUNTER, Z., MILOJEVIĆ, A. AND SUDAKOV, B. (2026). Gaussian random graphs and Ramsey numbers, arXiv preprint arXiv:2512.17718
[18] LI, S. AND SCHRAMM, T. (2026). Spectral clustering in the Gaussian mixture block model, arXiv preprint arXiv:2305.00979
[19] LIU, S., MOHANTY, S., SCHRAMM, T. AND YANG, E. (2022). Testing thresholds for high-dimensional sparse random geometric graphs. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 672–677.
[20] MA, J., SHEN, W. AND XIE, S. (2026). An exponential improvement for Ramsey lower bounds. *Invent. math.* https://doi.org/10.1007/s00222-026-01421-9
[21] MITRINOVIĆ, D. S., PEČARIĆ, J. E. AND FINK, A. M. (1993). *Classical and New Inequalities in Analysis*, Mathematics and Its Applications (East European *Series*), vol. 61. Kluwer Academic Publishers Group, Dordrecht.
[22] SCHREIBER, T. AND YUKICH, J. E. (2005). Large deviations for functionals of spatial point processes with applications to random packing and spatial graphs. *Stochastic Process. Appl.* **115**, 1332–1356.
[23] SEPPÄLÄINEN, T. AND YUKICH, J. E. (2001). Large deviation principles for Euclidean functionals and other nearly additive processes. *Probab. Theory Relat. Fields* **120**, 309–345.
[24] SODIN, S. (2007). Tail-sensitive gaussian asymptotics for marginals of concentrated measures in high dimension. In *Geometric Aspects of Functional Analysis: Israel Seminar 2004–2005*. Springer, pp. 271–295.
[25] VERSHYNIN, R. (2018). *High-Dimensional Probability*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 47. Cambridge University Press, Cambridge.

https://doi.org/10.1017/apr.2026.10076 Published online by Cambridge University Press