RESEARCH ARTICLE

On the clique number of noisy random geometric graphs

Matthew Kahle1 | Minghao Tian2 | Yusu Wang3

1Department of Mathematics, The Ohio State University, Columbus, Ohio, USA
2Computer Science and Engineering Department, The Ohio State University, Columbus, Ohio, USA
3Halıcıoğlu Data Science Institute, University of California San Diego, San Diego, California, USA

Correspondence
Matthew Kahle, Department of Mathematics, The Ohio State University, Columbus, OH, USA.
Email: mkahle@math.osu.edu

Funding information
National Science Foundation, Grant/Award Numbers: DMS-1547357, DMS-1352386, CCF-1740761, RI-1815697

Abstract
Let $G_n$ be a random geometric graph, and then for $q, p \in [0, 1)$ we construct a $(q,p)$-perturbed noisy random geometric graph $G_n^{q,p}$ where each existing edge in $G_n$ is removed with probability $q$, while each non-existent edge in $G_n$ is inserted with probability $p$. We give asymptotically tight bounds on the clique number $\omega(G_n^{q,p})$ for several regimes of parameter.

KEYWORDS
random graphs, clique number, random perturbation

1 | INTRODUCTION AND STATEMENT OF RESULTS

The random geometric graph $G(\mathcal{X}_n; r) = G_{\mathbb{R}^d}(X_1, X_2, \ldots, X_n; r)$ has vertices $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ where the $X_i$ are $d$-dimensional variables sampled from a common probability distribution $\nu$ on $\mathbb{R}^d$, and an edge $(X_i, X_j)$ is added whenever $X_i$ and $X_j$ are within Euclidean distance $r = r(n) > 0$ to each other. Often, only mild assumptions on the underlying probability distribution on $\mathbb{R}^d$ are required, for example a bounded, measurable, density function. See Penrose’s monograph [16] for a comprehensive overview of random geometric graphs.

Random geometric graphs are useful in applications, for example, wireless networks, transportation networks [2, 14] and so forth. We are mainly interested here in adding noise to a random geometric graph, in the sense of randomly adding either long-range edges, deleting short-range edges, or both. Such graphs also arise naturally in applications, for example, in modeling biological epidemics and collective social processes [17]. See [10] for work on contagion dynamics on noisy geometric networks.

The clique number of a graph $G$, denoted $\omega(G)$, is the order of the largest complete subgraph in $G$. Clique numbers of various models of random graphs are well studied. For example, it is a...
Some definitions and notation

1.1

where they were called “ER-perturbed random geometric graphs.” More precisely, we call close in total variation distance to the Erdős–Rényi model.

In Chapter 6 in [16], studies maximum degree, clique number, and chromatic number of random geometric graphs. He establishes laws of large numbers for maximum degree and clique number. He also shows in this chapter that for random geometric graphs sampled from well-behaved probability distributions on metric spaces, for example, from a uniform distribution on the unit cube $[0,1]^d$, the clique number and the maximum degree are of the same order of magnitude. We will see that such a correspondence does not generally hold for noisy random geometric graphs.

We note that our $(q,p)$-perturbed random geometric graphs can be thought of adding random perturbations to a base “structured network,” similar to small-world network models. For example, in the popular model proposed by Watts and Strogatz [18], the underlying structured network is a ring or lattice, and instead of deleting or adding edges, a random rewiring of some edges constitutes the perturbation. As mentioned in [15], the model we study in this article can be considered as a generalization of the Watts and Strogatz model. In particular, the study of a localized clique number called the edge clique number, can potentially be used to differentiate between the “original edges” and long “short-cuts” added, and thus to detect and remove the long short-cuts. See Section 6 for further discussion.

1.1 Some definitions and notation

Before we state our main results, we first state some more precise definitions.

Definition 1.1 (Random geometric graph [12]). Given a sequence of independent random points $X_1, X_2, \ldots$ in $\mathbb{R}^d$ sampled from a common probability distribution $\nu$ with bounded density function $f$ (that is, for any Borel set $A \subseteq \mathbb{R}^d$, $\nu(A) = \int_A f(x) \, dx$), and a positive distance $r = r(n) > 0$, we construct a random geometric graph $G(\mathcal{X}_n; r)$ with vertex set $\mathcal{X}_n = \{X_1, \ldots, X_n\}$, where distinct $X_i$ and $X_j$ are adjacent when $\|X_i - X_j\| \leq r$. Here $\|\cdot\|$ may be any norm on $\mathbb{R}^d$.

Denote $\sigma$ as the essential supremum of the probability density function $f$ of $\nu$, that is

$$\sigma := \sup \left\{ t : \int_{\{y : f(y) > t\}} \, dx > 0 \right\}.$$
We call $\sigma$ the maximum density of $v$. Denote $B_s(x) := \{y \in \mathbb{R}^d : ||y - x|| \leq s\}$ as the ball centered at $x \in \mathbb{R}^d$ of radius $s$. Also set $\theta = \int_{B_s(o)} dx$, where $o$ is the origin of $\mathbb{R}^d$; that is, $\theta$ is the volume of any radius-1 ball in $\mathbb{R}^d$.

We now introduce our $(q, p)$-perturbed noisy random geometric graphs $G^{qp}(\mathcal{X}_n; r)$.

**Definition 1.2** ($(q, p)$-perturbed noisy random geometric graph). Given a random geometric graph $G(\mathcal{X}_n; r)$ as in Definition 1.1, the $(q, p)$-perturbed noisy random geometric graph $G^{qp}(\mathcal{X}_n; r)$ is obtained by deleting each existing edge in $G(\mathcal{X}_n; r)$ independently with probability $q$ as well as inserting each non-existent edge in $G(\mathcal{X}_n; r)$ independently with probability $p$.

Note that the order of applying the above two types of perturbations does not matter since they are applied to two disjoint sets respectively. This process can be applied to any graph, and we call it a $(q, p)$-perturbation. The resulting graph $G^{qp}(\mathcal{X}_n; r)$ is called a $(q, p)$-perturbation of $G(\mathcal{X}_n; r)$, or simply a noisy random geometric graph.

Throughout this article, we use the standard Bachmann–Landau notation (asymptotic notation). That is, for real valued functions $f(n)$ and $g(n)$, as $n \to \infty$, we say

1. $f(n) = O(g(n))$ if there exist constants $c > 0$ and $n_0 \in \mathbb{N}$ such that $|f(n)| \leq cg(n)$ for all $n \geq n_0$;
2. $f(n) = o(g(n))$ if for every $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $|f(n)| < cg(n)$ for all $n \geq n_0$;
3. $f(n) = \Omega(g(n))$ if there exists constants $c > 0$ and $n_0 \in \mathbb{N}$ such that $f(n) \geq cg(n)$ for all $n \geq n_0$;
4. $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.

We also use the notation $f \ll g$ to mean that $f(n)/g(n) \to 0$ as $n \to \infty$, $f \lesssim g$ to mean that there exists a constant $C > 0$ such that $f(n)/g(n) < C$ for all sufficiently large $n$, $f \gtrsim g$ to mean that there exists a constant $c > 0$ such that $f(n)/g(n) > c$ for all sufficiently large $n$, and $f \sim g$ to mean that $f \lesssim g$ and $f \gtrsim g$.

Recall that a clique in any graph $G$ is a set of vertices which are pairwise connected. In this article, we use the standard notation $\omega(G)$ in graph theory to denote the clique number of $G$, which is the largest cardinality of a clique in $G$.

Many properties of $G(\mathcal{X}_n; r)$ are qualitatively different depending on which distance $r = r(n)$ is chosen. In some sense, the distance $r$ here plays a role similar to the edge-inserting probability $p(n)$ in Erdős–Rényi random graphs $G(n, p)$. We consider the following three regimes for the quantity $nr^d$:

I. (subcritical) $nr^d \leq n^{-\alpha}$ for some fixed $\alpha > 0$;
II. ("critical or nearly critical" or "thermodynamic") $n^{-\epsilon} \ll nr^d \ll \log n$ for all $\epsilon > 0$;
III. ("supercritical") $\sigma nr^d / \log n \to t \in (0, \infty)$.

In continuum percolation it is more standard to reserve critical for the case $nr^d \to \lambda_c$ for some special constant $\lambda_c > 0$, so our terminology may be slightly nonstandard.

We often use the terminology almost surely (or a.s.): In particular, if $\xi_1, \xi_2, \ldots$ is a sequence of random variables and $k_1, k_2, \ldots$ is a sequence of positive numbers, then "a.s. $\xi_n \geq k_n$" means that $\lim_{n \to \infty} \mathbb{P}[\xi_n \geq k_n] = 1$. The other direction a.s. $\xi_n \leq k_n$ is defined similarly. Moreover, a.s. $\xi_n \leq k_n$ means that there exist $C_1 > 0$ such that $\lim_{n \to \infty} \mathbb{P}[\xi_n \leq C_1 k_n] = 1$. Similarly, we define a.s. $\xi_n \geq k_n$ and a.s. $\xi_n \sim k_n$. We also use the terminology with high probability (or w.h.p.): specifically, if $A_1, A_2, \ldots$ is a sequence of events, then "$A_n$ happens with high probability" means that $\lim_{n \to \infty} \mathbb{P}[A_n] = 1$. 


1.1.1 Assumptions and notations for the remainder of the article

In what follows, unless specified explicitly, we assume the following setting throughout, which we refer to as the standard-setting-R:

- The space we consider is the $d$-dimensional Euclidean space $\mathbb{R}^d$ with a fixed dimension $d$, equipped with some arbitrary norm $\| \cdot \|$ on $\mathbb{R}^d$.
- $\theta = \int_{B_1(\mathbf{0})} dx$ is the volume of the unit ball $B_1(\mathbf{0}) = \{ x \in \mathbb{R}^d : \| x \| \leq 1 \}$.
- $\beta$ is the so-called Besicovitch constant of $(\mathbb{R}^d, \| \cdot \|)$ (see Section 2.2).
- $\nu$ is a probability distribution with finite maximum density $\sigma$; and $X_1, X_2, \ldots$ are independent random variables sampled from $\nu$.
- $r = (r(1), r(2), \ldots)$ is a sequence of positive real numbers such that $r(n) \to 0$ as $n \to \infty$.
- $q$ and $p = p(n)$ are real numbers between 0 and 1 (for simplicity, we only consider the case when $q$ is a fixed constant).
- $G_n, G_n^{q,p}$ denote the random geometric graph $G(X_1, \ldots, X_n; r(n))$ and its $(q,p)$-perturbation, respectively.

For any graph $G$, let $V(G)$ and $E(G)$ refer to its vertex set and edge set, and let $N_G(u)$ denote the set of neighbors of $u$ in $G$ (i.e., nodes connected to $u \in V(G)$ by edges in $E(G)$). For a subset $W \subseteq \mathbb{R}^d$, we denote the number of indices $i \in \{1, \ldots, n\}$ such that $X_i \in W$ by $\mathcal{N}(W) = \mathcal{N}_n(W)$; that is, $\mathcal{N}(W)$ is the number of points from $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ contained in $W$.

1.2 Overview of main results

We now state the main results of this article, which concerns the behavior of clique number of $(q,p)$-perturbed noisy random geometric graphs. To understand the behavior of $q$-deletion and $p$-insertions (which have different effects on the clique numbers), we first separate the insertion-only case (where the perturbation only has random insertions) and the deletion-only case (where the perturbation only has random deletions), and present results for the two cases in Theorems 1.3 and 1.4, respectively.

1.2.1 Insertion only

We first consider the clique number of $G_n^{0,p}$, where no edges in $G_n$ are removed and only new edges are added. The graph generated this way can be thought of as the union of a random geometric graph and an Erdős–Rényi random graph. Indeed, in Theorem 1.3 below, we show the interplay between those two random graphs as $p = p(n)$ increases in different regimes of $r$.

Theorem 1.3. Given a $(0, p)$-perturbed noisy random geometric graph $G_n^{0,p}$ in the standard-setting-R, the following holds:

(I) Suppose that $nr^d \leq n^\alpha$ for some fixed $\alpha \in (0, 1/\beta^2]$. Then there exist constants $C_1, C_2$ such that

(I.a) if $p \leq (1/n)^{C_1}$, then a.s.

$$\omega \left( G_n^{0,p} \right) \sim 1,$$

(I.b) and if $(1/n)^{C_1} < p \leq C_2$, then a.s.

$$\omega \left( G_n^{0,p} \right) \sim \log_{1/p} n.$$
(II) Suppose that for every $\varepsilon > 0$, $n^{-\varepsilon} \ll nr^d \ll \log n$. Then there exist constants $C_1, C_2$ such that

(II.a) if $p \leq (nr^d / \log n)^{C_1}$, then a.s.

$$\omega \left( G_n^{0,p} \right) \sim \frac{\log n}{\log(\log n.nr^d)}.$$ 

(II.b) and if $(nr^d / \log n)^{C_1} < p \leq C_2$, then a.s.

$$\omega \left( G_n^{0,p} \right) \sim \log \frac{1}{p} n.$$ 

(III) Suppose that $\sigma nr^d / \log n \to t \in (0, \infty)$ as $n \to \infty$. Then there exists a constant $C_1$ such that if $p \leq C_1$, then a.s.

$$\omega \left( G_n^{0,p} \right) \sim nr^d.$$ 

For this insertion-only case, one could view the graph $G_n^{0,p}$ as the union of a random geometric graph $G_n$ and an Erdős–Rényi random graph $G(n,p)$ on the same vertex set. The theorem above suggests that intuitively either the clique number from the random geometric graph or the one from the Erdős–Rényi random graph will dominate, depending on regimes of $nr^d$ and $p$. On the surface, this may not look surprising. However, in general, the clique number of the union $G = G_1 \cup G_2$ could be significantly larger than the clique number in each individual graph $G_i$: consider for example $G_1$ is a collection of $\sqrt{n}$ disjoint cliques, each of size $\sqrt{n}$, while $G_2$ equals to the complement of $G_1$. The union $G_1 \cup G_2$ is the complete graph and the clique number is $n$. However, the clique number of $G_1$ or $G_2$ is $\sqrt{n}$. Our results suggest that due to the randomness in each of the individual graph we are considering, with high probability such a scenario will not happen and the two types of random graph do not interact strongly.

To handle the mixture of random geometric graph with inserted edges, it is not clear how to use classical tools such as scan statistic. One of the key ideas in our article is to develop and use what we call a \textit{well-separated clique-partitions family} (see Section 2.2) to help us to decouple the interaction between the two types of hidden random structures (i.e., random geometric graph, and the $(0,p)$-perturbation). We believe that this idea is interesting for its own sake.

1.2.2 Deletion only

We now present our main result for the clique number of $G_n^{0,0}$, where we only delete edges in $G_n$ with a fixed edge-deletion probability $q$. We remark that technically speaking, the deletion-only case is easier to handle than the insertion-only case.

**Theorem 1.4.** Given a $(q,0)$-perturbed noisy random geometric graph $G_n^{0,0}$ in the standard-setting $R$ with a fixed constant $0 < q < 1$, the following holds:

(I) Suppose that $nr^d \leq n^{-\alpha}$ for some fixed $\alpha > 0$. Then a.s.

$$\omega \left( G_n^{0,0} \right) \sim 1.$$ 

(II) Suppose that $n^{-\varepsilon} \ll nr^d \ll \log n$ for all $\varepsilon > 0$. Then a.s.

$$\omega \left( G_n^{0,0} \right) \sim \log \frac{\log n}{\log(\log n.nr^d)}.$$
Suppose that $\sigma_{nr^d} / \log n \to t \in (0, \infty)$. Then a.s.

$$\omega\left(G_n^{q,0}\right) \preceq \log(nr^d).$$

Furthermore, there exists a constant $T > 0$ such that if $\sigma_{nr^d} \geq T \log n$, then a.s.

$$\omega\left(G_n^{q,0}\right) \sim \log(nr^d).$$

1.2.3 Combined case

The above insertion-only and deletion-only cases in fact represent key technical challenges. When there are both random insertions and deletions, we can derive some bounds on the clique number by simply combining the above results and some technical lemmas later in the article together with the monotone property of clique number. For example, we have the following result when $nr^d$ is in the subcritical regime.

**Corollary 1.5.** Given a $(q,p)$-perturbed noisy random geometric graph $G_n^{q,p}$ in the standard-setting-$R$ with a fixed constant $0 < q < 1$ and suppose that $nr^d \leq n^{-\alpha}$ for some fixed $\alpha \in (0, 1/\beta^2]$, then there exist constants $C_1, C_2$ such that

(a) if $p \leq (1/n)^{C_1}$, then a.s.

$$\omega\left(G_n^{q,p}\right) \sim 1,$$

(b) and if $(1/n)^{C_1} < p \leq C_2$, then a.s.

$$\omega\left(G_n^{q,p}\right) \sim \log_{1/p} n.$$

The complete list of results can be found in Theorem 5.1 of Section 5.

2 Preliminaries and well-separated clique-partitions family

In this section, we first state in Section 2.1 some existing results/tools that we will use frequently throughout the article. We will then define in Section 2.2 a new object called well-separated clique-partitions family which we will need later in the arguments.

2.1 Some standard results and tools

For the proofs in this article, we need some bounds on the binomial and Poisson distributions.

**Lemma 2.1** (Lemma 3.6 in [12]). Let $Z$ be either binomial or Poisson and $k \geq \mu := \mathbb{E}[Z]$. Then

$$\left(\frac{\mu}{ek}\right)^k \leq \mathbb{P}[Z \geq k] \leq \left(\frac{e\mu}{k}\right)^k.$$
Lemma 2.2 (Chernoff–Hoeffding theorem [16]). Suppose \( n \in \mathbb{N}, \alpha \in (0, 1) \) and \( 0 < k < n \). Let \( X \sim \text{Bin}(n, \alpha) \) be either a binomial random variable with mean \( \mu = na \) or \( X \sim \text{Poisson}(\mu) \) be a Poisson random variable with mean \( \mu > 0 \). If \( k \geq \mu \), then

\[
\mathbb{P}[X \geq k] \leq \exp\left(-\mu H\left(\frac{k}{\mu}\right)\right).
\]

where \( H : [0, \infty] \to [0, \infty) \) is a function defined by \( H(0) = 1 \) and \( H(a) = 1 - a + a \log a \).

One object we use frequently in our proofs is the following (generalized) scan statistic defined in [12]. Recall that for any set \( U \subseteq \mathbb{R}^d \), \( \mathcal{N}(U) \) is the number of points from \( \mathcal{X}_n = \{X_1, \ldots, X_n\} \) contained in \( U \).

Definition 2.3 ((Generalized) scan statistic [12]). For any set \( W \subseteq \mathbb{R}^d \), we define \( M_W \) by \( M_W := \max_{x \in \mathbb{R}^d} \mathcal{N}(x + rW) \) where \( rW = \{rw : w \in W\} \) is the scaled set of \( W \).

In other words, \( M_W \) is the maximum number of points in \( \mathcal{X}_n \) in any translate of \( rW \).

2.2 Well-separated clique-partitions family

This section discuss one main technique we will use to bound the clique number of \( G_{n,q}^\delta \) from above. The main challenge here is to disentangle two types of randomness—the location of vertices and the \((q, p)\)-perturbation. In particular, our model allows vertices even far away to each other to become connected. To solve this, we develop a novel approach using what we call a well-separated clique-partitions family (to be defined shortly) to help us to decouple the interaction between these two types of hidden random structures.

To set up the stage, we first recall the Besicovitch covering lemma which has a lot of applications in measure theory [6].

Definition 2.4 (Packings, covers, and partitions). (1) A packing is a countable collection \( B \) of pairwise disjoint closed balls in \( \mathbb{R}^d \). Such a collection \( B \) is a packing w.r.t. a set \( P \subseteq \mathbb{R}^d \) if the centers of the balls in \( B \) lie in the set \( P \), and it is a \( \delta \)-packing if all of the balls in \( B \) have radius \( \delta \). (2) A set \( \{A_1, \ldots, A_\ell\}, A_\ell \subseteq \mathbb{R}^d, \) covers \( P \) if \( P \subseteq \bigcup_j A_i \). (3) Given a set \( A \), we say that \( A \) is partitioned into \( A_1, A_2, \ldots, A_k \), if \( A = A_1 \cup \cdots \cup A_k \) and \( A_i \cap A_j = \emptyset \) for any \( i \neq j \).

Lemma 2.5 (Besicovitch covering lemma [7]). There exists a constant \( \beta = \beta(d) \in \mathbb{N} \) such that for any set \( P \subseteq \mathbb{R}^d \) and \( \delta > 0 \), there are \( \beta \) number of \( \delta \)-packings w.r.t. \( P \), denoted by \( \{B_1, \ldots, B_\beta\} \), whose union also covers \( P \).

We call the constant \( \beta \) above the Besicovitch constant. Note that \( \beta \) depends only on the dimension \( d \) and is not dependent of \( \delta \).

Definition 2.6 (Well-separated clique-partitions family). Given a geometric graph \( G^* \) in \((\mathbb{R}^d, \|\cdot\|)\) with vertex set \( V \) and edge set \( E \), a family \( P = \{P_i\}_{i \in \Lambda} \), where \( P_i \subseteq V \) and \( \Lambda \) is the index set of \( P_i \)'s, forms a well-separated clique-partitions family of \( G^* \) if:

1. \( V = \bigcup_{i \in \Lambda} P_i \).
2. \( \forall i \in \Lambda, P_i \) can be partitioned as \( P_i = C_{i,1}^{(i)} \cup C_{i,2}^{(i)} \cup \cdots \cup C_{i,m_i}^{(i)} \) where...
(2a) $\forall j \in [1, m_1], \text{there exist } v^{(i)}_j \in V \text{ such that } C_j^{(i)} \subseteq B_{r/2}(v^{(i)}_j) \cap V.$

(2b) For any $j_1, j_2 \in [1, m_1]$ with $j_1 \neq j_2$, $d_H(C_{j_1}^{(i)}, C_{j_2}^{(i)}) > r$, where $d_H$ is the Hausdorff distance between two sets in $\mathbb{R}^d$ with respect to norm $\|\cdot\|$.

We also call $C_1^{(i)} \cup C_2^{(i)} \cup \cdots \cup C_m^{(i)}$, a clique-partition of $P_i$ (w.r.t. $G^*$), and its size (cardinality) is $m_i$. The size of the well-separated clique-partitions family $P$ is its cardinality $|P| = |\Lambda|.$

In the above definition, (2a) implies that each $C_j^{(i)}$ spans a clique in the geometric graph $G^*$; thus we call $C_j^{(i)}$ as a clique in $P_i$ and $C_1^{(i)} \cup C_2^{(i)} \cup \cdots \cup C_m^{(i)}$ a clique-partition of $P_i$. (2b) means that there are no edges in $G^*$ between any two cliques of $P_i$. As a result, any edge in its corresponding $(q,p)$-perturbation $G_h^{q,p}$ between such cliques must come from $(q,p)$-perturbation ($p$-insertion). We will leverage this fact significantly later when bounding clique numbers. See Figure 1.

We have the following existence theorem of a well-separated clique-partitions family with constant size only depending on the dimension $d$.

**Theorem 2.7.** There exists a well-separated clique-partitions family $P = \{P_i\}_{i \in \Lambda}$ of any geometric graph $G^*$ with $|\Lambda| \leq \beta^2$, where $\beta = \beta(d)$ is the Besicovitch constant of $\mathbb{R}^d$.

**Proof.** To prove the theorem, first imagine we grow an $r/2$-ball around each node in $V \subseteq \mathbb{R}^d$ (the vertex set of $G^*$). By the Besicovitch covering lemma (Lemma 2.5), we have a family of $(r/2)$-packings w.r.t. $V$, $B = \{B_1, B_2, \ldots, B_{a_1}\}$, whose union covers $V$. Here, the constant $a_1$ satisfies $a_1 \leq \beta(d)$.

Each $B_i$ contains a collection of disjoint $r/2$-balls centered at a subset of nodes in $V$, and let $V_i \subseteq V$ denote the centers of these balls. For any $u, v \in V_i$, we have $\|u - v\| > r$ as otherwise, $B_{r/2}(u) \cap B_{r/2}(v) \neq \emptyset$ meaning that the $r/2$-balls in $B_i$ are not all pairwise disjoint. Now consider the collection of $r$-balls centered at all nodes in $V_i$. Applying Besicovitch covering lemma to $V_i$ again with $\delta = r$, we now obtain a family of $r$-packings w.r.t. $V_i$, denoted by $D^{(i)} = D_1^{(i)} \cup \cdots \cup D_{a_2}^{(i)}$, whose union covers $V_i$. Here, the constant $a_2^{(i)}$ satisfies $a_2^{(i)} \leq \beta(d)$ for each $i \in [1, a_1]$.

Now each $D^{(i)}$ contains a set of disjoint $r$-balls centered at a subset of nodes $V_j^{(i)} \subseteq V_i$ of $V_i$. First, we claim that $\bigcup_j V_j^{(i)} = V_i$. This is because that $B_i$ is an $r/2$-packing which implies that $\|u - v\| > r$ for any two nodes $u, v \in V_i$. In other words, the $r$-ball around any node from $V_i$ contains no other nodes.

![Figure 1](image.png) Points in the solid balls are $P_1$, and those in dashed balls are $P_2$. Each adapts a clique-partition of size $m_1 = m_2 = 4$. Assuming that all nodes in $G^*$ are shown in this figure, then $P = \{P_1, P_2\}$ forms a well-separated clique-partitions family of $G^*$. 
in $V_i$. As the union of $r$-balls $D^{(i)}_1 \cup \cdots \cup D^{(i)}_{c_2}$ covers $V_i$ by construction, it is then necessary that each node $V_i$ has to appear as the center in at least one $D^{(i)}_j$ (i.e., in $V^{(i)}_j$). Hence $\bigcup_j V^{(i)}_j = V_i$.

Now for each vertex set $V^{(i)}_j$, let $P^{(i)}_j \subseteq V$ denote all points from $V$ contained in the $r/2$-balls centered at points in $V^{(i)}_j$. As $\bigcup_j V^{(i)}_j = V_i$, we have $\bigcup_j P^{(i)}_j = \bigcup_{i \in V_i} (Br_{r/2}(v) \cap V)$. It then follows that $\bigcup_{i \in [1,a]} (\bigcup_{j \in [1,\alpha]} P_j^{(i)}) = V$ as the union of the family of $r/2$-packings $B = \{B_1, B_2, \ldots, B_{r_i}\}$ covers all points in $V$ (recall that $B_{r}$ is just the set of $r/2$-balls centered at points in $V_i$).

Clearly, each $P_j^{(i)}$ adapts a clique-partition: Indeed, for each $V^{(i)}_j$, any two nodes in $V^{(i)}_j$ are at least distance $2r$ apart (as the $r$-balls centered at nodes in $V^{(i)}_j$ are disjoint), meaning that the $r/2$-balls around them are more than distance $r$ away, in the Hausdorff metric. In other words, $P = \left\{ P_j^{(i)}, i \in [1, \alpha_1], j \in [1, \alpha_2^{(i)}] \right\}$ forms a well-separated clique-partitions family of $G^*$. Finally, since $\alpha_1, \alpha_2^{(i)} \leq \beta(d) = \beta$, the cardinality of $P$ is thus bounded by $\beta^2$.

3 | PROOF OF THEOREM 1.3

In this section, we focus on estimating the order of $\omega\left(G_n^{0,p}\right)$, the clique number of $G_n^{0,p}$. Note that for any set $W \subseteq \mathbb{R}^d$, the generalized scan statistic $M_W$ (see Definition 2.3) is the maximum number of points in the vertex set $X_n = \{X_1, X_2, \ldots, X_n\}$ in any translate of $rW$. Set $W_{1/2} := B_{1/2}(\mathbf{0})$ and $W_1 := B_1(\mathbf{0})$ where $\mathbf{0}$ is the origin. It is obvious that $\omega\left(G_n^{0,p}\right) \geq M_{W_{1/2}}$. Thus, the lower bound can be directly derived by using the results related to the generalized scan statistic in [12]. However, getting an upper bound is much more challenging, since unlike $\omega(G_n) \leq M_{W_1}$ holds in $G_n$, the vertices in a clique of $G_n^{0,p}$ can come from everywhere in the space.

3.1 | Proof of Part (I)—Subcritical regime

In this section, we discuss the order of $\omega(G_n^{0,p})$ in the regime $nr^d \leq n^{-\alpha}$ for some fixed $\alpha \in (0, 1/\beta^2]$. First, we define the long-edges in $G_n^{0,p}$.

**Definition 3.1** (long-edges). An edge $(u, v)$ in a $(0, p)$-perturbed noisy random geometric graph $G_n^{0,p}$ is a long-edge if and only if $\|u - v\| > 3r$.

Clique in $G_n^{0,p}$ can be classified into the following two types:

- Type-I clique: does not contain any long-edges.
- Type-II clique: contains at least one long-edge.

In what follows, we derive upper bounds for each type of cliques separately. Intuitively, Type-I cliques primarily depend on the underlying random geometric graph, while Type-II sees a stronger effect of the Erdős–Rényi-perturbation. We use $3r$ as the threshold in the definition of long-edges, to intuitively decouple the interaction of the local neighborhood of $u$ and $v$ within the random geometric graphs. The lower bounds are easier to derive and can be found later in Section 3.1.2.

3.1.1 | Type-I cliques

The case of Type-I cliques is rather simple to handle: note that vertices of one Type-I clique are contained within a ball of radius $3r$ centered at some vertex $X_i \in X_n$. Thus, to bound the size of such clique
from above, it suffices to estimate the number of vertices in each of the \( n \) number of \( 3r \)-ball centered at some vertex in \( \mathcal{X}_n \). Set \( W_3 := B_3(o) \). We have the following lemma, which gives a uniform upper bound of number of vertices in any \( 3r \)-ball. It is a simplified variant of Lemma 3.8 in [12]. We include its simple proof for completeness.

**Lemma 3.2.** If \( nr^d \leq n^{-\alpha} \) then \( \mathbb{P} [ M_{W_3} \leq \lfloor 4/\alpha \rfloor ] = 1 + O(n^{-3}) \).

**Proof.** For some fixed integer \( k \), we have the following inequality.

\[
\mathbb{P} [ M_{W_3} \geq k + 1 ] \leq \mathbb{P} [ \exists i : \mathcal{N}(B_{6r}(X_i)) \geq k + 1 ] \leq n \mathbb{P} [ \mathcal{N}(B_{6r}(X_1)) \geq k + 1 ] .
\]

Furthermore, note that

\[
\mathbb{P} [ \mathcal{N}(B_{6r}(X_1)) \geq k + 1 ] \leq \mathbb{P} [ \text{Bin} (n, \sigma \theta (6r)^d) \geq k ] \leq \left( \frac{e \sigma \theta^d (nr)^d}{k} \right)^k = O(n^{-ka}).
\]

Recall that \( \sigma \) is the maximum density of \( v \) and \( \theta = \int_{B_r(o)} dx \) are introduced in the standard setting-R at the end of Section 1.1. The second inequality holds due to Lemma 2.1. Now pick \( k = \lfloor 4/\alpha \rfloor \). We then have \( \mathbb{P} [ M_{W_3} \leq \lfloor 4/\alpha \rfloor ] = 1 + O(n^{-3}) \) as required.

It then follows that the size of Type-I cliques can be bounded from above by \( \lfloor 4/\alpha \rfloor \) almost surely.

### 3.1.2 Type-II cliques

Now let’s consider the Type-II cliques, which is significantly more challenging to handle. Recall that \( W_1 = B_1(o) \). We can use the same argument in Lemma 3.2 to derive the following lemma which gives a uniform upper bound of the number of points in any \( r \)-ball.

**Lemma 3.3.** If \( nr^d \leq n^{-\alpha} \) then \( \mathbb{P} [ M_{W_1} \leq \lfloor 4/\alpha \rfloor ] = 1 + O(n^{-3}) \).

We now introduce a local version of the clique number called edge clique number.

**Definition 3.4** (Edge clique number). Given a graph \( G = (V, E) \), for any edge \((u, v) \in E\), its edge clique number \( \omega_{u,v}(G) \) is defined as

\[
\omega_{u,v}(G) = \text{the largest size of any clique in } G \text{ containing } (u, v).
\]

We are now ready to bound the size of all the type-II cliques in \( G_{n,p}^\beta \). More precisely, the following theorem first bound the edge clique number for all long edge \((u, v)\). This is the key theorem in this Section, and we include its proof in the next subsection.

**Theorem 3.5.** Given an \((0,p)\)-perturbed noisy random geometric graph \( G_{n}^{0,p} \) in the standard-setting-R and suppose that \( nr^d \leq n^{-\alpha} \) for some fixed \( \alpha \in (0, 1/\beta^2) \), then

(a) There exist constants \( C_1, C_2 > 0 \) which depend on the Besicovitch constant \( \beta \) and \( \alpha \) such that if

\[
p \leq C_1(1/n)^{C_2}
\]

then, with high probability, for all long-edge \((u, v)\) in \( G_{n}^{0,p} \), its edge clique number \( \omega_{u,v}(G_{n}^{0,p}) \leq 1 \).
There exists a constant $\xi > 0$ which depends on the Besicovitch constant $\beta$ and $\alpha$ such that if $(1/n)^\xi \leq p < 1$, then, with high probability, for all long-edge $(u,v)$ in $G_n^{0,p}$, its edge clique number $\omega_{u,v}(G_n^{0,p}) \preceq \log_{1/p} n$.

3.1.3 Proof of Theorem 3.5

Proof of part (a) of Theorem 3.5

Given any vertex $y$, let $B_r^y(y) \subseteq X_n$ denote $B_r(y) \cap X_n$. Now consider a long-edge $(u,v)$. Set $A_{uv} = X_n \setminus \left( B_r^u(u) \cup B_r^v(v) \right)$ and $B_{uv} = B_r^u(u) \cup B_r^v(v)$. Denote $A_{uv} = A_{uv} \cup \{u\} \cup \{v\}$; easy to check that $X_n = A_{uv} \cup B_{uv}$.

Let $G|_S$ denote the subgraph of $G$ spanned by a subset $S$ of its vertices. Given any set $C$, let $C|_S = C \cap S$ be the restriction of $C$ to another set $S$. Now consider a subset of vertices $C \subseteq X_n$: obviously, $C = C|_{A_{uv}} \cup C|_{B_{uv}}$.

Set $N_{\text{max}} := \lceil * \rceil 4/\alpha$. Denote $F$ to be the event that “for every $v \in X_n$, the ball $B_r(v) \cap X_n$ contains at most $N_{\text{max}}$ points”; and $F^c$ denotes the complement event of $F$. By Lemma 3.3, we know that, $\Pr[F^c] = O(n^{-3})$.

Let $K \geq 8\beta^2$ be an integer to be determined. By applying the pigeonhole principle and the union bound, we have:

$$\Pr \left[ \omega_{u,v} \left( G_n^{0,p} \right) \geq K | F \right] \leq \Pr \left[ \omega_{u,v} \left( G_n^{0,p} |_{\tilde{A}_{uv}} \right) \geq K/2 | F \right] + \Pr \left[ \omega_{u,v} \left( G_n^{0,p} |_{B_{uv}} \right) \geq K/2 | F \right]. \quad (3.2)$$

Next, we bound the two terms on the right hand side of Equation (3.2) separately in Cases (i) and (ii) below.

Case (i): Bounding the first term in Equation (3.2)

Directly bounding the edge clique number using points from $\tilde{A}_{uv}$ is challenging, as two types of edges are involved (a “local” edge from random geometric graph, or a randomly inserted Erdős–Rényi type edge). Hence we will use the well-separated clique-partitions family introduced earlier, to consider only special types of cliques where the number of combinatorial choices these two types of edges can induce is limited. We apply Theorem 2.7 for points in $\tilde{A}_{uv}$. This gives us a well-separated clique-partitions family $\mathcal{P} = \{P_i\}_{i \in \Lambda}$ of $A_{uv}$ with $|\Lambda| \leq \beta^2$ being a constant. See Figure 2A. Augment each $P_i$ to

![Figure 2](image-url)
FIGURE 3  The red dashed lines and the edge $uv$ form a possible clique in some well-separated clique partition $P_i$. The points in the small balls are the nodes falling in $r/2$-balls (and thus they are all pairwise connected in the base random geometric graph $G_n$). All the dashed lines are the randomly inserted edges (independently with probability $p$).

\[ \hat{P}_i = P_i \cup \{u\} \cup \{v\}. \]  

Suppose there is a clique $C$ in $G_n^{0,p}|_{\hat{A}_{uv}}$, then as $\bigcup_i \hat{P}_i = \hat{A}_{uv}$, we have $C = \bigcup_{i \in \Lambda} C|_{\hat{P}_i}$, implying that $|C| \leq \sum_{i \in \Lambda} |C|_{\hat{P}_i}$. Hence again by applying pigeonhole principle and the union bound, we derive the following inequality:

\[ P \left[ \omega_{u,v} \left( G_n^{0,p}|_{\hat{A}_{uv}} \right) \geq k/2 \big| F \right] \leq \sum_{i=1}^{\Lambda} P \left[ \omega_{u,v} \left( G_n^{0,p}|_{\hat{P}_i} \right) \geq k/(2|\Lambda|) \big| F \right]. \]  

(3.3)

Now for arbitrary $i \in \Lambda$, consider $G_n^{0,p}|_{\hat{P}_i}$, the induced subgraph of $G_n^{0,p}$ spanned by vertices in $\hat{P}_i$. Note, $G_n^{0,p}|_{\hat{P}_i}$ can be viewed as generated by inserting each edge not in $G_n^{0,p}|_{\hat{A}_{uv}} \cup \{(u,v)\}$ with probability $p$. Recall from Definition 2.6 that each $P_i$ adapts a clique-partition $C_j^{(i)} \cup \cdots \cup C_m^{(i)}$, where every $C_j^{(i)}$ is contained in an $r/2$-ball, and all such balls are $r$-separated (w.r.t Hausdorff distance).

Now fix any $i \in \Lambda$. For simplicity of the argument below, set $m = m_i$, and let $N_j = |C_j^{(i)}|$ denote the number of points in the $j$th cluster $C_j^{(i)}$. Note that obviously, $m \leq |P_i| \leq |X_n| = n$ for any $i \in \Lambda$. We also know that if event $F$ has already happened, then $N_j \leq N_{\max}$.

Observe that the induced subgraph $G_n^{0,p}|_{\hat{P}_i}$ consists of a set of cliques (each clique is spanned by some $C_j^{(i)}$ with edges coming from the base random geometric graph $G_n$), $u$, $v$, edge $(u,v)$, and inserted edge between them with insertion probability $p$ (see Figure 3).

Now set $k := \lfloor \star |K/2|\Lambda \rfloor - 2$. Since $K \geq 8\beta^2$, easy to see that $k \geq 1$. For every set $S$ of $k+2$ vertices in this graph $G_n^{0,p}|_{\hat{P}_i}$, let $A_S$ be the event “$S$ is clique in $G_n^{0,p}|_{\hat{P}_i}$ containing $(u,v)$ given $F$” and $I_S$ its indicator random variable. Set

\[ I = \sum_{|S|=k+2} I_S, \]

and note that $I$ is the number of cliques of size $(k+2)$ in $G_n^{0,p}|_{\hat{P}_i}$ containing $(u,v)$ given $F$. It follows from Markov inequality that:

\[ P \left[ \omega_{u,v} \left( G_n^{0,p}|_{\hat{P}_i} \right) \geq k+2 \big| F \right] = P[I > 0] \leq \mathbb{E}[I]. \]  

(3.4)
On the other hand, using linearity of expectation, we have:

\[
\mathbb{E}[I] = \sum_{|S|=k+2} \mathbb{E}[I_S] = p^{2k} \sum_{x_1+\cdots+x_m=k \atop 0 \leq x_i \leq N_i} \binom{N_1}{x_1} \binom{N_2}{x_2} \cdots \binom{N_m}{x_m} p^{(k^2 - \sum_{i=1}^m x_i^2)/2} \\
\leq p^{2k} \sum_{x_1+\cdots+x_m=k \atop 0 \leq x_i \leq N_{\max}} \binom{N_{\max}}{x_1} \binom{N_{\max}}{x_2} \cdots \binom{N_{\max}}{x_m} p^{(k^2 - \sum_{i=1}^m x_i^2)/2}. \tag{3.5}
\]

To estimate this quantity, we have the following lemma:

**Lemma 3.6.** If \(1 \leq k \leq N_{\max}\) and \(p\) is less than or equal to

\[
\min \left\{ \frac{1}{\sqrt{e}} \left( \frac{1}{n^3 m} \right)^{\frac{1}{4}} \left( \frac{k}{N_{\max}} \right)^{\frac{1}{4}} \frac{1}{2e k^3} \left( \frac{1}{n^3 m^2} \right)^{\frac{1}{4}} \frac{1}{N_{\max}} \frac{1}{e^4} \left( \frac{1}{n^3 m^k} \right)^{\frac{4}{7}} \left( \frac{k}{N_{\max}} \right)^{\frac{4}{7}} \right\}, \tag{3.6}
\]

then we have that \(\mathbb{E}[I] = O(n^{-3})\).

The proof of this lemma is rather technical, and can be found in Appendix A.1.

Note that \(\alpha \leq 1/\beta^2\), thus \(2N_{\max} = 2|\Lambda| \leq 4/\alpha \geq 8\beta^2\). Note that if \(K \in [8\beta^2, 2N_{\max}]\), then it is easy to check that the assumption \(1 \leq k \leq N_{\max}\) in Lemma 3.6 holds.

Furthermore, \(|\Lambda| \leq \beta^2\) (which is a constant) and \(m = |P_i| \leq |\mathcal{X}_n| = n\). One can then verify that there exist constants \(c_1^*\) and \(c_2^*\) (which depend on the Besicovitch constant \(\beta\) and \(\alpha\), such that if

\[
p \leq c_1^* \cdot (1/n)^{c_2^*/K},
\]

then the conditions in Equation (3.6) will hold. Thus, combining this with Lemma 3.6 and Equation (3.4), we know that

If \(8\beta^2 \leq K \leq 2N_{\max}\) and \(p \leq c_1^* \cdot (1/n)^{c_2^*/K}\),

then \(\forall i \in \Lambda, \mathbb{P} \left[ \omega_{a,v}(G_n^{0,p}|P_i) \geq k + 2|F| \right] = O(n^{-3}). \tag{3.7}\)

On the other hand, note that

\[
\mathbb{P} \left[ \omega_{a,v}(G_n^{0,p}|P_i) \geq K/(2|\Lambda|)|F| \right] = \mathbb{P} \left[ \omega_{a,v}(G_n^{0,p}|P_i) \geq k + 2|F| \right].
\]

As \(|\Lambda|\) is a constant, by Equation (3.3), we obtain that

if \(\forall i \in \Lambda, \mathbb{P} \left[ \omega_{a,v}(G_n^{0,p}|P_i) \geq K/(2|\Lambda|)|F| \right] = O(n^{-3})\), then

\[
\mathbb{P} \left[ \omega_{a,v}(G_n^{0,p}|\hat{\mathcal{A}}_n) \geq K/2 |F| \right] = O(|\Lambda|n^{-3}) = O(n^{-3}). \tag{3.8}\)

It then follows from Equations (3.7) and (3.8) that

If \(8\beta^2 \leq K \leq 2N_{\max}\) and \(p \leq c_1^* \cdot (1/n)^{c_2^*/K}\),

then \(\mathbb{P} \left[ \omega_{a,v}(G_n^{0,p}|\hat{\mathcal{A}}_w) \geq K/2 |F| \right] = O(n^{-3}). \tag{3.9}\)
Finally, suppose \( K > K_0 = 2N_{\text{max}} \). Using Equation (3.9), we know that if \( p \leq c_1^a \cdot (1/n)^{\beta_2/K_0} \) and \( K > K_0 \) (in which case note also that \( c_1^a \cdot (1/n)^{\beta_2/K_0} \leq c_1^b \cdot (1/n)^{\beta_2/K} \)), then

\[
\Pr \left[ \omega_{u,v} \left(G_n^{0, p} \mid \tilde{\Lambda}_w \right) \geq K/2 \mid F \right] \leq \Pr \left[ \omega_{u,v} \left(G_n^{0, p} \mid \tilde{\Lambda}_w \right) \geq K_0/2 \mid F \right] = O(n^{-3}).
\]

Combining this with Equation (3.9), we thus obtain that:

If \( K \geq 8\beta^2 \) and \( p \leq \min \left\{ c_1^a \cdot (1/n)^{\beta_2/(2N_{\text{max}})}, \ c_1^b \cdot (1/n)^{\beta_2/K} \right\} \), then

\[
\Pr \left[ \omega_{u,v} \left(G_n^{0, p} \mid \tilde{\Lambda}_w \right) \geq K/2 \mid F \right] = O(n^{-3}). \tag{3.10}
\]

Case (ii): Bounding the second term in Equation (3.2)

First recall that \( B_{uv} = B_{\gamma}^{X_r}(u) \cup B_{\gamma}^{X_r}(v) \) (see Figure 2B).

On one hand, imagine we now build the following random graph \( \tilde{G}_{uv}^{\text{local}} = (\hat{V}, \hat{E}) \): The vertex set \( \hat{V} \) is simply \( B_{uv} \). To construct the edge set \( \hat{E} \), first, add edges between all pairs of distinct vertices in \( B_{\gamma}^{X_r}(u) \) and do the same thing for \( B_{\gamma}^{X_r}(v) \); that is, every two vertices in \( B_{\gamma}^{X_r}(u) \) or \( B_{\gamma}^{X_r}(v) \) are now connected by an edge. Next, add edge \((u, v)\). Finally, insert each crossing edge \((x, y)\) with \( x \in B_{\gamma}^{X_r}(u) \) and \( y \in B_{\gamma}^{X_r}(v) \) with probability \( p \).

On the other hand, consider the graph \( G_n^{0, p} \mid B_{uv} \), the induced subgraph of \( G_n^{0, p} \) spanned by vertices in \( B_{uv} \). We can imagine that the graph \( G_n^{0, p} \mid B_{uv} \) was produced by first taking the induced subgraph \( G_n \mid B_{uv} \), and then insert crossing edges \((x, y)\) each with probability \( p \). Since \((u, v)\) is a long-edge, by Definition 3.1, we know that there are no edges between nodes in \( B_{\gamma}^{X_r}(u) \) and \( B_{\gamma}^{X_r}(v) \) in \( G_n \mid B_{uv} \). Since every two vertices in \( B_{\gamma}^{X_r}(u) \) or \( B_{\gamma}^{X_r}(v) \) are not necessarily connected by an edge in \( G_n \mid B_{uv} \), we know that

\[
\Pr \left[ \omega_{u,v} \left(G_n^{0, p} \mid B_{uv} \right) \geq K/2 \mid F \right] \leq \Pr \left[ \omega_{u,v} \left(\tilde{G}_{uv}^{\text{local}} \right) \geq K/2 \mid F \right]. \tag{3.11}
\]

Using a similar argument as in case (i) (the missing details can be found in Appendix A.2), we have that there exist constants \( c_1^b, c_2^b > 0 \) which depend on the Besicovitch constant \( \beta \) and \( \alpha \) such that

If \( K \geq 8\beta^2 \) and \( p \leq c_1^b \cdot (1/n)^{\beta_2/K} \), then

\[
\Pr \left[ \omega_{u,v} \left(G_n^{0, p} \mid B_{uv} \right) \geq K/2 \mid F \right] = O(n^{-3}).
\]

Pick \( K = 2N_{\text{max}} = 2[4/\alpha] \geq 8\beta^2 \) (by condition \( \alpha \in (0, 1/\beta^2) \)). Note that \( K = O(1) \) in this case. Thus, combining the above bound with Equations (3.11), (3.10), and (3.2), there exist constants \( C_1 = \min\{c_1^a, c_1^b\} \) and \( C_2 = \max\{c_2^a, c_2^b\} \) such that if \( p \) satisfies conditions in Equation (3.1), then

\[
\Pr \left[ \omega_{u,v} \left(G_n^{0, p} \right) \geq K \mid F \right] + \Pr[\neg F] = O(n^{-3}).
\]

Finally, by applying the union bound, this means:

\[
\Pr \left[ \text{for all long-edge } (u, v), \ \omega_{u,v} \left(G_n^{0, p} \right) \geq K \right] = O(n^{-1}).
\]

Thus with high probability, we have that for all long-edge \((u, v)\), \( \omega_{u,v}(G_n^{0, p}) = O(1) \) as long as Equation (3.1) holds. This completes the proof of Part (a) of Theorem 3.5.
Proof of part (b) of Theorem 3.5
We use the same strategy in the proof of part (a). That is, we again try to bound the two terms on
the right hand side of Equation (3.2) from above respectively. The key difference here is to give an
alternative estimate of Equation (3.5) in case (i) and its counterpart in case (ii) under the new constraint
of \( p \).

For case (i), instead of using Lemma 3.6, we now use the following lemma, whose proof can be
found in Appendix A.3.

**Lemma 3.7.** There exists a constant \( C_3 > 0 \) depending on the Besicovitch constant \( \beta \) and \( a \) such that if \( (1/n)^{8/(3N_{\text{max}})} \leq p < 1 \) and \( K = C_3 \lceil \frac{\log 1/p}{\log n} \rceil \), then we have that \( \mathbb{E}[I] = O(n^{-3}) \).

Now choose such \( C_3 \) as specified in Lemma 3.7. We know that the following holds.

\[
\text{If } (1/n)^{8/(3N_{\text{max}})} \leq p < 1,
\quad \text{then } \mathbb{P} \left[ \omega_{u,v} \left( G_{n}^{0,p} \big| \tilde{\Delta} \right) \geq C_3 \lceil \frac{\log 1/p}{\log n} \rceil \bigg| F \right] = O(n^{-3}). \tag{3.12}
\]

For case (ii), we know that if event \( F \) has already happened, then \( |B_{uv}| \leq 2N_{\text{max}} \), where \( |B_{uv}| \) denotes the cardinality of set \( B_{uv} \). Note that if \( (1/n)^{C_3/(4N_{\text{max}}+C_3)} \leq p < 1 \), then

\[ C_3 \lceil \frac{\log 1/p}{\log n} \rceil \geq 2N_{\text{max}} \geq |B_{uv}|. \]

Hence, we obtain that:

\[
\text{If } (1/n)^{C_3/(4N_{\text{max}}+C_3)} \leq p < 1,
\quad \text{then } \mathbb{P} \left[ \omega_{u,v} \left( G_{n}^{0,p} \big| B_{uv} \right) \geq C_3 \lceil \frac{\log 1/p}{\log n} \rceil \bigg| F \right] = 0. \tag{3.13}
\]

Set \( \xi = \min \{8/(3N_{\text{max}}), C_3/(4N_{\text{max}}+C_3)\} \), which is also a constant. Thus, combining Equations (3.13), (3.12), and (3.2), we know that if \( (1/n)^{\xi} \leq p < 1 \), then

\[ \mathbb{P} \left[ \omega_{u,v} \left( G_{n}^{0,p} \right) \geq C_3 \lceil \frac{\log 1/p}{\log n} \rceil \bigg| F \right] = O(n^{-3}). \]

Finally, by a similar argument in the proof for Part (a) using the law of total probability and union bound, we can show that with high probability, we have that for any long-edge \((u,v)\), \( \omega_{u,v} \left( G_{n}^{0,p} \right) \leq \log_{1/p} n \). This completes the proof of Theorem 3.5.

3.1.4 | Finishing the proof of Part (I) of Theorem 1.3
Based on the discussion of Type-I cliques as well as Theorem 3.5 for Type-II cliques, we have the following corollary regarding the upper bound of \( \omega(G_{n}^{0,p}) \) in the subcritical regime.

**Corollary 3.8.** Given a \((0,p)\)-perturbed noisy random geometric graph \( G_{n}^{0,p} \) in the standard-setting-\( R \) and suppose that \( nr^d \leq n^{-a} \) for some fixed \( a \in \left(0,1/\rho^2\right) \), then

(i) there exist two constants \( C_1, C_2 > 0 \) such that if \( p \leq C_1 (1/n)^{C_2} \), then a.s.

\[ \omega \left( G_{n}^{0,p} \right) \leq 1, \]
(ii) and there exists a constant $\xi > 0$ such that if $(1/n)^{\xi} \leq p < 1$, then a.s.

$$\omega \left( G_{n}^{0,p} \right) \preceq \log_{1/p} n.$$ 

Part (i) of Corollary 3.8 can be derived by combining Lemma 3.2 and part (a) of Theorem 3.5, while part (ii) of Corollary 3.8 can be derived by combining Lemma 3.2 and part (b) of Theorem 3.5.

To derive a lower bound of $\omega(G_{n}^{0,p})$, we need the following result on the clique number of Erdős–Rényi random graphs (proof can be found in Appendix A.4).

**Lemma 3.9.** For Erdős–Rényi random graph $G(n, p)$ with $(1/n)^{1/11} \leq p \leq (1/n)^{1/\sqrt{n}}$, we have a.s. $\omega(G(n, p)) > [\ast] \log_{1/p} n$.

Note that $p$ here is no longer a fixed constant as in the standard literature [1, 4], thus the well-known $\omega(G(n, p)) \sim 2 \log_{1/p} n$ statement cannot be directly applied here. Not surprisingly, the standard second moment method [1] is used here, but the calculation is different. Details of the proof can be found in Appendix A.4.

Easy to see that $\mathbb{P}[\omega(G_{n}^{0,p}) \geq K] \geq \mathbb{P}[\omega(G(n, p)) \geq K]$ for any positive integer $K$. The following corollary of Lemma 3.9 gives a lower bound of $G_{n}^{0,p}$ regardless of which regime $nr^d$ belongs to.

**Corollary 3.10.** Given a $(0, p)$-perturbed noisy random geometric graph $G_{n}^{0,p}$ in the standard-setting-R and suppose that $(1/n)^{1/11} \leq p \leq (1/n)^{1/\sqrt{n}}$, then we have a.s.

$$\omega \left( G_{n}^{0,p} \right) > [\ast] \log_{1/p} n.$$ 

Note that $(1/n)^{1/\sqrt{n}} \to 1$ as $n \to \infty$. Thus, there exists a constant $C_{3} \in (0, 1)$ (very close to 1) such that $C_{3} \leq (1/n)^{1/\sqrt{n}}$ for sufficiently large $n$. Also note that the following monotone property holds.

For any $S > 0$ and $0 \leq q_{1} < q_{2} < 1$, $\mathbb{P}\left[ \omega \left( G_{n}^{0,q_{1}} \right) \geq S \right] \leq \mathbb{P}\left[ \omega \left( G_{n}^{0,q_{2}} \right) \geq S \right].$

And by Corollary 3.8 (b), we know that $\omega \left( G_{n}^{0,n^{-\epsilon}} \right) = O \left( \log_{n^{-\epsilon}} n \right) = O(1)$ a.s. Easy to see that there exists a constant $C_{1}'$ such that $p \leq (1/n)^{C_{1}'}$ implies $p \leq C_{1}(1/n)^{C_{2}}$. Also notice that $\log_{1/p} n = \Theta(1)$ for $p \in (1/n)^{C_{1}'}$, $(1/n)^{\xi}$ (if this interval exists). Thus, the lower bound of $p$ in the condition of part (b) of Corollary 3.8 (which is $(1/n)^{\xi}$) can be extended to $(1/n)^{C_{1}'}$ and the conclusion still holds. Combining these facts with Corollaries 3.8 and 3.10 concludes the proof of part (I) of Theorem 1.3.

### 3.2 Proof of Part (II)—Subcritical regime

In this section, we discuss the order of $\omega(G_{n}^{0,p})$ in the regime $n^{-\epsilon} \ll nr^d \ll \log n$ for all $\epsilon > 0$. Again, we first derive an upper bound of $\omega(G_{n}^{0,p})$ by considering two types of cliques (Type-I and Type-II) introduced in Section 3.1. The idea of the proof in this section is similar to the one in Section 3.1, although the details vary a little.
3.2.1 Type-I cliques
Recall that $W_3 = B_3(o)$. The following lemma gives an upper bound of the number of vertices in any $3r$-ball.

**Lemma 3.11.** If $n^{-e} \ll nr^d \ll \log n$ for all $e > 0$, then

$$
\mathbb{P}\left[M_{W_3} \leq \frac{5 \log n}{\log(\log n/(\sigma^d nr^d))}\right] = 1 - O(n^{-\frac{3}{2}}).
$$

The argument is rather standard, relying only on Chernoff–Hoeffding bounds, so we omit the proof.

3.2.2 Type-II cliques
Recall that $W_1 = B_1(o)$. By using a similar argument as the one used for Lemma 3.11, we can get the following lemma which gives an upper bound for the number of points in any $r$-ball for the regime of $nr^d$ under discussion.

**Lemma 3.12.** If $n^{-e} \ll nr^d \ll \log n$ for all $e > 0$, then

$$
\mathbb{P}\left[M_{W_1} \leq \frac{5 \log n}{\log(\log n/(\sigma^2 nr^d))}\right] = 1 - O(n^{-\frac{3}{2}}).
$$

Again, the argument is standard, and we omit the proof.

**Theorem 3.13.** Given an $(0,p)$-perturbed noisy random geometric graph $G_{n,p}$ in the standard-setting-R and suppose that $n^{-e} \ll nr^d \ll \log n$ for all $e > 0$, then

(a) there exist constants $C_1, C_2 > 0$ which depend on the Besicovitch constant $\beta$ such that if

$$
p \leq C_1 \cdot \left(\frac{nr^d}{\log n}\right)^{C_2},
$$

then, with high probability, for all long-edge $(u,v)$ in $G_{n,p}$, its edge clique number

$$
\omega_{u,v}(G_{n,p}) \lesssim \frac{\log n}{\log(\log n/nr^d)},
$$

(b) and there exists a constant $\xi$ which depends on the Besicovitch constant $\beta$ such that if

$$
\left(\frac{nr^d}{\log n}\right)^{\frac{1}{\xi}} \leq p < 1,
$$

then, with high probability, for all long-edge $(u,v)$ in $G_{n,p}$, its edge clique number

$$
\omega_{u,v}(G_{n,p}) \lesssim \log(1/p n).
$$

The proof of Theorem 3.13 follows the same flow as the proof of Theorem 3.5, with some minor changes. The details can be found in Appendix A.5.

3.2.3 Putting everything together for Part (II) of Theorem 1.3

To wrap up all the above results, we have the following corollary regarding the upper bound of $\omega(G_{n,p})$ in subcritical regime.
Corollary 3.14. Given a \((0,p)\)-perturbed noisy random geometric graph \(G_{n}^{0,p}\) in the standard-setting-R and suppose that \(n^{-\epsilon} \ll nr^d \ll \log n\) for all \(\epsilon > 0\), then

(i) there exist two constants \(C_1, C_2 > 0\) such that if \(p \leq C_1 (nr^d / \log n)^{C_2}\), then a.s.

\[
\omega(G_{n}^{0,p}) \lesssim \frac{\log n}{\log(\log n/nr^d)},
\]

(ii) and there exists a constant \(\xi > 0\) such that if \((nr^d / \log n)^{\xi} \leq p < 1\), then a.s.

\[
\omega(G_{n}^{0,p}) \lesssim \log_{1/p} n.
\]

Part (i) can be derived by combining Lemma 3.11 and part (a) of Theorem 3.13, while part (ii) can be derived by combining Lemma 3.11 and part (b) of Theorem 3.13.

To derive a tight bound of \(\omega(G_{n}^{0,p})\), in addition to Corollary 3.10, we also need the following lemma, which provides a lower bound of \(\omega(G_{n}^{0,p})\).

Lemma 3.15. Given a \((0,p)\)-perturbed noisy random geometric graph \(G_{n}^{0,p}\) in the standard-setting-R and suppose that \(n^{-\epsilon} \ll nr^d \ll \log n\) for all \(\epsilon > 0\), then a.s.

\[
\omega(G_{n}^{0,p}) \gtrsim \frac{\log n}{2 \log(\log n/nr^d)}.
\]

Proof of Lemma 3.15. Note that \(\omega(G_{n}^{0,p}) \geq M_{W_{1/2}}\) and \(M_{W_{1/2}}\) can a.s. be bounded from below by \(\log n/ (2 \log(\log n/nr^d))\). (Pick \(\epsilon = 1/2\) in Lemma 3.9 of [12].)

Finally, combining Lemma 3.15 with part (i) of Corollary 3.14 concludes the proof of part (II.a) of Theorem 1.3. Note that there exists a constant \(C_1\) such that \(p \leq (nr^d / \log n)^{C_1}\) implies \(p \leq (nr^d / \log n)^{C_2}\) and if \(p \in \left( (nr^d / \log n)^{C_1}, (nr^d / \log n)^{\xi} \right)\) (if this interval exists), then

\[
\log_{1/p} n = \Theta \left( \frac{\log n}{\log(\log n/nr^d)} \right).
\]

Thus we can extend the lower bound of the condition in part (ii) of Corollary 3.14 to \((nr^d / \log n)^{C_1}\) by the same reasoning at the end of the proof for subcritical regime. Combining these facts with Corollary 3.10 and part (ii) of Corollary 3.14 concludes the proof of (II.b) of Theorem 1.3.

3.3 Proof of Part (III)—“supercritical” regime

In this section, we discuss the order of \(\omega(G_{n}^{0,p})\) in the regime \(\sigma nr^d / \log n \to t \in (0, \infty)\). Again, we first derive an upper bound of \(\omega(G_{n}^{0,p})\) by considering two types of cliques (Type-I and Type-II) introduced in Section 3.1. The idea of the proof in this section is very similar to the one in Section 3.2, thus the proofs are omitted.

Set \(\tau\) be the smallest real number such that \(\tau \geq 2\) and \(\tau(\log \tau - 1) \geq 4/(2^d \theta t)\). Since \(d, t, \) and \(\theta\) are all given constants, \(\tau\) is also a constant.
3.3.1 Type-I cliques

The following lemma gives upper bounds of the number of vertices in each $r$-ball and $3r$-ball respectively. Recall that $W_1 = B_1(o)$ and $W_3 = B_3(o)$.

**Lemma 3.16.** If $\sigma n r^d / \log n \to t \in (0, \infty)$, then

\[
\begin{align*}
\mathbb{P}[M_{W_1} \leq \tau 2^d \theta \sigma n r^d] &= 1 + O(n^{-3}), \\
\mathbb{P}[M_{W_3} \leq \tau 6^d \theta \sigma n r^d] &= 1 + O(n^{-3}).
\end{align*}
\]

3.3.2 Type-II cliques

The proof of the following technical theorem is almost the same as the proof of part (a) of Theorem 3.13 thus is omitted.

**Theorem 3.17.** Given a $(0, p)$-perturbed noisy random geometric graph $G_{n,p}^0$ in the standard-setting-R and suppose that $\sigma n r^d / \log n \to t \in (0, \infty)$, then there exists a constant $C$ which depends on the Besicovitch constant $\beta$ such that if $p \leq C$ then, with high probability, for all long-edge $(u, v)$ in $G_{n,p}^0$, its edge clique number $\omega_{u,v}(G_{n,p}^0) \lesssim n r^d$.

To derive a tight bound of $\omega(G_{n,p}^0)$, in addition to Corollary 3.10, we also need the following result on lower bound.

**Lemma 3.18.** Given a $(0, p)$-perturbed noisy random geometric graph $G_{n,p}^0$ in the standard-setting-R and suppose that $\sigma n r^d / \log n \to t \in (0, \infty)$, then a.s.

\[
\omega(G_{n,p}^0) \geq \frac{1}{2} \eta \sigma n r^d,
\]

where $\eta$ is the unique solution $x \geq \theta(1/2)^d$ to $H(x/(\theta(1/2)^d)) = 1/(\theta(1/2)^d t)$ (recall that function $H$ is defined as $H(a) = 1 - a + a \log a$).

**Proof of Lemma 3.18.** Note that $\omega(G_{n,p}^0) \geq M_{W_{1/2}}$ and $M_{W_{1/2}}$ can be bounded from below by $\eta \sigma n r^d / 2$ almost surely (directly by Theorem 1.8 of [12]).

Finally, combining Lemmas 3.18, 3.16, and Theorem 3.17 concludes the proof.

4 PROOF OF THEOREM 1.4

In this section, we focus on deriving the order of $\omega(G_{n,0}^0)$. Note that $\omega(G_{n,0}^0) \leq M_{W_1}$. Thus, Theorem 1.4 part (I) is obvious due to Lemma 3.3. Our proof of the remaining parts of Theorem 1.4 uses the following lemma following easily from known results in the literature. Recall that $\nu$ is the probability distribution defined in Section 1.1.

**Lemma 4.1** (Lemma 3.1 in [16]). For any fixed $\rho > 0$, recall that $W_1 = B_1(o)$ (and thus $rW_1 = B_r(o)$). There exists $N = \Omega(r^{-d})$ disjoint translates $x_1 + rW_1, \ldots, x_N + rW_1$ of $rW_1$ with $\nu(x_i + rW_1) \geq (1 - \rho)\sigma \theta r^d$ for all $i = 1, \ldots, N$. 
Our proof in this section follows an approach analogous to the proof of Theorem 1.8 in [12]. We show the proof for the subcritical regime here in this section. Since we use similar techniques in the supercritical regime, the proof for that regime (part (III)) is relegated Appendix B.1.

4.1 Proof of Part (II)—Subcritical regime

In this section, we discuss the order of $\omega(G_n^{q,0})$ in the regime $n^{-\epsilon} \ll nr^d \ll \log n$ for all $\epsilon > 0$.

4.1.1 Deriving upper bound

We first focus on the upper bound of $\omega(G_n^{q,0})$. This is obtained via considering and relating to the random geometric graphs whose nodes are generated by Poisson point process. Let $N \sim \text{Poisson} ((1 + \delta)n)$ for some $\delta > 0$ (say $\delta = 1/2$).

Note that $G_N$ (random geometric graph on $N$ nodes; recall Definition 1.1) is a geometric graph ($r$-neighborhood graph) of the Poisson point process $\mathcal{P}_{(1+\delta)n}$ with intensity $(1 + \delta)n\Pi$ [16], where $f$ is the density defined in Section 1.1. Similar to $G_n^{q,0}$, we define a $(q,0)$-perturbation of $G_N$ as $G_N^{q,0}$. Set $k_n$ be an integer to be determined. Now, we have

$$
P \left[ \omega(G_n^{q,0}) \geq k_n \right] \leq \mathbb{P} \left[ \omega(G_N^{q,0}) \geq k_n \right] + \mathbb{P} \left[ N \leq n - 1 \right]
$$

for some constant $\gamma > 0$ (depending on $\delta$) by Lemma 2.2. For $y \in \mathbb{R}^d$ let $X_y$ be the set of nodes of $G_N$ falling in $B_r(y)$. Let $M_y$ be the number of points falling in $B_r(y)$ spanning a maximum clique in $G_N^{q,0}|X_y$. Define $M := \max_{y \in \mathbb{R}^d} M_y$. Easy to see

$$
P \left[ \omega(G_N^{q,0}) \geq k_n \right] = \mathbb{P} \left[ M \geq k_n \right]. \quad (4.1)$$

Fix $y \in \mathbb{R}^d$. By the property of Poisson point process, we know $|X_y| \sim \text{Poisson}(\lambda)$ where $\lambda := (1 + \delta)n\int_{B_r(y)} f(x) \, dx$. By using Markov’s inequality, we have

$$
P \left[ M_y \geq k_n \right] = \sum_{i \geq k_n} \mathbb{P} \left[ M_y \geq k_n \left| |X_y| = i \right. \right] \mathbb{P} \left[ |X_y| = i \right] \leq \sum_{i \geq k_n} \mathbb{E} \left[ \text{number of } k_n \text{-cliques in } G_N^{q,0}\left| |X_y| = i \right. \right] \frac{e^{-\lambda} \lambda^i}{i!} \leq \sum_{i \geq k_n} \left( \frac{i}{k_n} \right) (1 - q) \binom{i}{k_n} \frac{e^{-\lambda} \lambda^i}{i!} \leq \frac{k_n}{k_n!} (1 - q) \binom{k_n}{i} \cdot e^{-\lambda} \sum_{i \geq k_n} \frac{\lambda^{i-k_n}}{(i-k_n)!} = \frac{\lambda^{k_n}}{k_n!} (1 - q) \binom{k_n}{i}.
Note that $\lambda \leq (1 + \delta)\sigma \theta nr^d$. Thus,

$$
P \left[ M_y \geq k_n \right] \leq \frac{(1 + \delta)\sigma \theta nr^d}{k_n!} (1 - q)^{\frac{k_n}{2}},
$$

which does not depend on the choice of $y$. Combining this with (4.1), we have

$$
P \left[ \omega \left( G^q N \right) \geq k_n \right] \leq \frac{(1 + \delta)\sigma \theta nr^d}{k_n!} (1 - q)^{\frac{k_n}{2}} < \frac{1}{\sqrt{2\pi}} \left( \frac{(1 + \delta)\sigma \theta nr^d (1 - q)^{(k_n - 1)/2}}{k_n} \right)^{k_n}.
$$

Finally, pick

$$
k_n = 2 \log_{1/(1-q)} \left( \frac{\log n}{\log(\log n/nr^d)} \right) + 1.
$$

Since $n^{-\epsilon} \ll nr^d \ll \log n$ for all $\epsilon > 0$, easy to see that $k_n \to \infty$. Note that

$$
\frac{(1 + \delta)\sigma \theta nr^d (1 - q)^{(k_n - 1)/2}}{k_n} = \frac{(1 + \delta)\sigma \theta}{k_n} \cdot \frac{\log(\log n/nr^d)}{\log n/nr^d} \leq \frac{C}{k_n}
$$

for some constant $C > 0$. Thus, $P \left[ \omega \left( G^q N \right) \geq k_n \right] = o(1)$. Hence, we have that a.s.

$$
\omega \left( G^q N \right) \leq \log \frac{\log n}{\log(\log n/nr^d)}.
$$

4.1.2 Deriving the lower bound

Now we consider the lower bound of $\omega \left( G^q N \right)$. We first state the following well-known result on the clique number of Erdős–Rényi random graphs, which plays an important role in proving the lower bound.

**Lemma 4.2.** Suppose $p \in (0, 1)$ is a constant. For Erdős–Rényi random graph $G(n, p)$ with $n \to \infty$, we have

$$
P \left[ \omega(G(n, p)) \leq \star \right] \log_{1/p} n \leq e^{-n}.
$$

This is a direct corollary of the standard $2 \log_{1/p} n$ statement (see P. 185 [1]), thus we omit the proof.

Now let $N \sim \text{Poisson} \left( (1 - \delta')n \right)$ for some $\delta' \in (0, 1)$ (say $\delta' = 1/2$). Note that $G_N$ is an $r$-neighborhood graph of the Poisson point process $\mathcal{P}_{(1-\delta')n}$ with intensity $(1 - \delta')nf$, where $f$ is the density defined in Section 1.1. Similarly, we define a $(q, 0)$-perturbation of $G_N$ as $G^q N$. Set $k_n$ be an integer to be determined. Now, we have

$$
P \left[ \omega \left( G^q N \right) \leq k_n \right] \leq P \left[ \omega \left( G^q N \right) \leq k_n \right] + P \left[ N \geq n + 1 \right]
\leq P \left[ \omega \left( G^q N \right) \leq k_n \right] + e^{-\gamma'n}$$

where $\gamma'$ is a constant to be determined.
for some constant \( \gamma' > 0 \) (depending on \( \delta' \)) by Lemma 2.2. Now fix some constant \( \rho \in (0, 1) \) (say \( \rho = 1/2 \)). Recall \( W_{1/2} = B_{1/2}(0) \). By Lemma 4.1, there exist points \( x_1, x_2, \ldots, x_m \) with \( m = \Omega \left( r^{-d} \right) \) such that the sets \( x_i + W_{1/2} \) are disjoint and

\[
\nu (x_i + W_{1/2}) \geq \frac{(1 - \rho) \sigma \theta}{2^d} r^d.
\]

for \( i = 1, \ldots, m \) where \( \nu \) is the probability distribution defined in Section 1.1. Let \( X_i \) be the set of points of \( G_N \) falling in \( x_i + W_{1/2} \). Then, we have

\[
P \left[ \omega \left( G_N^{q,0} | x_i \right) \leq k_n \right] \leq P \left[ \omega \left( G_N^{q,0} \right) \leq k_n, \ldots, \omega \left( G_N^{q,0} | x_m \right) \leq k_n \right] = \prod_{i=1}^{m} P \left[ \omega \left( G_N^{q,0} | x_i \right) \leq k_n \right]. \tag{4.2}
\]

Note that all the points falling in any \( r/2 \)-ball span a complete graph. Thus, for each \( i \), we know the following holds.

\[
P \left[ \omega \left( G_N^{q,0} | x_i \right) \leq k_n \right] = P \left[ \omega \left( G \left( |X_i|, 1 - q \right) \right) \leq k_n \right].
\]

Set

\[
\Phi_n := \frac{\log n}{2 \log (\log n/n \sigma^d)},
\]

which goes to infinity as \( n \) grows. Note that \( |X_i| \sim \text{Poisson} \left( \tilde{\lambda} \right) \) where

\[
\frac{(1 - \delta') \sigma \theta nr^d}{2^d} \geq \tilde{\lambda} := (1 - \delta') n \cdot \nu (x_i + W_{1/2}) \geq \frac{(1 - \rho) (1 - \delta') \sigma \theta nr^d}{2^d}. \tag{4.3}
\]

The upper bound follows from the upper bound of volume of balls.

Now pick

\[
k_n := \lceil \log_{1/(1-q)} \Phi_n \rceil = \Omega \left( \log \frac{\log n}{\log (\log n/n \sigma^d)} \right).
\]

Let \( Q \sim \text{Poisson} \left( \tilde{\lambda}/e \right) \). By the law of total probability, we have

\[
P \left[ \omega \left( G \left( |X_i|, 1 - q \right) \right) \leq k_n \right] \leq P \left[ |X_i| \leq \Phi_n \right] + \sum_{j=\lceil \log_{1/(1-q)} \Phi_n \rceil}^{\infty} P \left[ \omega \left( G \left( j, 1 - q \right) \right) \leq k_n \right] P \left[ |X_i| = j \right]
\]

\[
\leq 1 - P \left[ |X_i| \geq \Phi_n + 1 \right] + \sum_{j=\lceil \log_{1/(1-q)} \Phi_n \rceil}^{\infty} P \left[ \omega \left( G \left( j, 1 - q \right) \right) \leq \lceil \log_{1/(1-q)} \rceil \right] \frac{e^{-\tilde{\lambda}} \tilde{\lambda}^j}{j!}
\]

\[
< 1 - \left( \frac{\tilde{\lambda}}{e(\Phi_n + 1)} \right)^{\Phi_n + 1} + \sum_{j=\lceil \log_{1/(1-q)} \Phi_n \rceil}^{\infty} e^{-\tilde{\lambda}} \frac{\tilde{\lambda}^j}{j!}. \tag{4.4}
\]
\[ 1 - e^{-(\Phi_n + 1) \log(e(\Phi_n + 1)/\tilde{\lambda})} + e^{-\tilde{\lambda}/e} \sum_{j=0}^{\Phi_n} \frac{e^{j \tilde{\lambda}/e}}{j!} \] 
\[ \leq 1 - e^{-(\Phi_n + 1) \log(e(\Phi_n + 1)/\tilde{\lambda})} + e^{-\tilde{\lambda}/e} \cdot \mathbb{P}[Q \geq \Phi_n] \] 
\[ < 1 - e^{-(\Phi_n + 1) \log(e(\Phi_n + 1)/\tilde{\lambda})} + e^{-((1-1/e)\tilde{\lambda})} \cdot e^{-\Phi_n \log(\Phi_n/\tilde{\lambda})}, \]  
(4.5)

where Equations (4.4) and (4.5) hold due to Lemma 2.1 (note that \( \Phi \gg \frac{\tilde{\lambda}}{e} \)) and Lemma 4.2. Routine calculations show that for \( n \) large enough, we have

\[ (\Phi_n + 1) \log(e(\Phi_n + 1)/\tilde{\lambda}) \leq \frac{1}{2} \log n + 1 \]
\[ \Phi_n \log(\Phi_n/\tilde{\lambda}) \geq \frac{1}{2} \log n - 1, \]

and \( e^{-(1-1/e)\tilde{\lambda}} \leq 1/(2e^2) \) since \( nr^d \gg n^{-e} \) for all \( e > 0 \). Thus

\[ \mathbb{P}[\omega(G_{N(\mathcal{X})}, 1 - q)] \leq k_n] < 1 - \frac{1}{2e}n^{-1/2}. \]

Plugging this back into Equation (4.2), we have

\[ \mathbb{P}[\omega(G_{q,0} N) \leq k_n] < \left(1 - \frac{1}{2e}n^{-1/2}\right)^m \leq e^{-\frac{1}{2}n^{-1/2}m}. \]

Recall \( m = \Omega(r^{-d}) \) and \( nr^d \ll \log n \), thus \( n^{-1/2}m = \Omega\left(\sqrt{n}/\log n\right) \). This implies

\[ \mathbb{P}[\omega(G_{q,0} N) \leq k_n] = o(1). \]

Since we have that \( \mathbb{P}[\omega(G_{q,0} N) \leq k_n] = o(1) \) with

\[ k_n = \Omega\left(\log \frac{\log n}{\log(\log n/nr^d)}\right), \]

we thus obtain the lower bound in Part (II) of Theorem 1.4.

5 | COMBINED CASE

In this section, we focus on bounding the clique number \( \omega(G_{q,0} N) \) of \( G_{q,0} N \), for different regimes of \( nr^d \), \( q \) and \( p \). Analogously to the monotonicity of the clique number of Erdős–Rényi random graphs [8], we have the following two monotone properties: for any positive integer \( K \),

\[ \mathbb{P}[\omega(G_{q,0} N) \leq K] \leq \mathbb{P}[\omega(G_{q,0} N) \leq K] \leq \mathbb{P}[\omega(G_{q,0} N) \leq K], \]
\[ \mathbb{P}[\omega(G_{q,0} N) \geq K] \leq \mathbb{P}[\omega(G_{q,0} N) \geq K] \leq \mathbb{P}[\omega(G_{q,0} N) \geq K]. \]

Combining these properties with Theorems 1.3,1.4 and technical lemmas (Lemma 4.2 and Corollary 3.10), we can derive some of the results showing below (part (I) and part (III.b)). Other results can
be derived by carefully choosing $K_n$ in the proof of Theorems 3.13 and 3.17 to fit the corresponding lower bound for different regimes of $nr^d$. For example, we can set some

$$K_n = \Theta \left( \log \left( \frac{\log n}{\log(\log n/nr^d)} \right) \right),$$

(the lower bound in the subcritical regime for deletion-only case; see Theorem 1.4) in the proof of Theorem 3.13 to derive part (II.a) of Theorem 5.1. For these reasons, we omit the proof of the following theorem.

**Theorem 5.1.** Given a $(q,p)$-perturbed noisy random geometric graph $G_{n,p}^q$ in the standard-setting-R with a fixed constant $0 < q < 1$, the following holds:

(I) Suppose that $nr^d \leq n^{-\alpha}$ for some fixed $\alpha \in (0, 1/\beta^2]$. Then there exist constants $C_1, C_2$ such that

(I.a) if $p \leq (1/n)^{C_1}$, then a.s.

$$\omega \left( G_{n,p}^q \right) \sim 1,$$

(I.b) and if $(1/n)^{C_1} < p \leq C_2$, then a.s.

$$\omega \left( G_{n,p}^q \right) \sim \log 1/p \cdot n.$$

(II) Suppose that $n^{-\epsilon} \ll nr^d \ll \log n$ for all $\epsilon > 0$. Then there exist constants $C_1, C_2, C_3$ such that

(II.a) if

$$p \leq (1/n)^{C_1/\log \frac{\log n}{\log(\log n/nr^d)}},$$

then a.s.

$$\omega \left( G_{n,p}^q \right) \sim \log \frac{\log n}{\log(\log n/nr^d)},$$

(II.b) and if $(nr^d/\log n)^{C_2} < p \leq C_3$, then a.s.

$$\omega \left( G_{n,p}^q \right) \sim \log 1/p \cdot n.$$

(III) There exists a constant $T > 0$ such that if $\sigma nr^d/\log n \to t \in (T, \infty)$, then there exist constant $C_1, C_2$ such that

(III.a) if $p \leq (1/n)^{C_1/\log \log n} (\log \log n/\log n)$, then a.s.

$$\omega \left( G_{n,p}^q \right) \sim \log (nr^d),$$

(III.b) and if $0 < p \leq C_2$ and $p = \Theta(1)$, then a.s.

$$\omega \left( G_{n,p}^q \right) \sim \log 1/p \cdot n.$$
6 | CONCLUDING REMARKS

In this article, we study the behavior of the clique number of noisy random geometric graphs $G_{n,p}^q$. In particular, we give the asymptotic tight bounds for the insertion-only case $G_{n,p}^{0}$ and the deletion-only case $G_{n,0}^q$ under different assumptions on $nr^d$ (Theorems 1.3 and 1.4, respectively). To obtain these results, we deploy a range of classical and new techniques: for example, we develop a novel approach based on what we call the “well-separated clique-partitions family” to handle the insertion case. Some partial results for the general case $\omega (G_{n,p}^{0})$ are also provided (Theorem 5.1). We also note that results in our article can be extended beyond the Euclidean setting: for example, in [9], noisy random geometric graphs generated from points sampled from a well-behaved doubling measure supported on a geodesic space are considered, and behaviors of the edge clique number are investigated.

This work represents a first step toward characterizing properties of the noisy random geometric graphs (which intuitively are generated based on two types of random processes). There are many interesting open problems. For example, the combined case is not yet completely resolved (there are still gaps in the regimes). Also currently we only provide asymptotic tight bounds, and it would be interesting to identify the exact constant for the high order terms too. It will also be interesting to study other quantities beyond the clique number.

Finally, we note that the random deletions/insertions can be viewed as “noise” on top of a base graph (which is a random geometric graph in our work). It will be interesting to see whether studies of clique numbers of other quantities can be used to “denoise” the input graph in practical applications (e.g., as in [15]). Indeed, the work of [15] showed that the shortest path metric of the random geometric graph can be recovered (with approximation guarantees) from its ER-perturbed version, if the insertion probability $p$ is small (compared to expected degree). In particular, in the high level, the work of [15] uses a quantity, called the Jaccard index, to identify what they refer to as “long-edges,” removes such edges and use the shortest distances in the denoised graph to approximate those of the underlying random geometric graphs with approximation guarantees. We believe that the edge clique number that we study in this article can be used as a more powerful way to identify such “long-edges” that can tolerate a much larger range of insertion probability $p$ than the previous work in Jaccard index. We will leave this as an interesting direction to explore in the future.

FUNDING INFORMATION

This work is partially supported by National Science Foundation Grants DMS-1547357, DMS-1352386, CCF-1740761, and RI-1815697.

REFERENCES

1. N. Alon and J. H. Spencer, The probabilistic method, Vol 4, Wiley Publishing, New York, 2016.
2. P. Blanchard and D. Volchenkov, Mathematical analysis of urban spatial networks, Springer Science & Business Media, New York, 2008.
3. B. Bollobás and P. Erdős, Cliques in random graphs, Math. Proc. Cambridge Philos. Soc. 80 (1976), no. 3, 419–427. https://doi.org/10.1017/S0305004100053056 MR 498256.
4. B. Bollobás, Random graphs, Cambridge University Press, Cambridge, 2001, 73.
5. L. Devroye, A. György, G. Lugosi, and F. Udina, High-dimensional random geometric graphs and their clique number, Electron. J. Probab. 16 (2011), 2481–2508.
6. H. Federer, Geometric measure theory, Springer, New York, 2014.
7. J. Heinonen, Lectures on analysis on metric spaces, Springer Science & Business Media, New York, 2012.
8. S. Janson, T. Łuczak, and A. Ruciński, Random graphs, John Wiley & Sons, New York, 2011, 45.
9. M. Kahle, M. Tian, Y. Wang, Local cliques in er-perturbed random geometric graphs, arXiv preprint arXiv:1810.08383, 2018.
APPENDIX A: THE MISSING PROOFS IN SECTION 3

A.1  The proof of Lemma 3.6

Since $1 \leq k \leq N_{\text{max}}$, we know that $p \leq 1$ thus is well-defined as a probability. To estimate the summation on the right hand side of Equation (3.5), we consider the quantity $x_{\text{max}} := \max_{i} \{ x_i \}$. We first enumerate all the possible cases of $(x_1, x_2, \ldots, x_m)$ when $x_{\text{max}}$ is fixed, and then vary the value of $x_{\text{max}}$.

Set $h(y) = \max_{x_{\text{max}}} \{ \sum_{i=1}^{m} x_i^2 \}$ for $y \geq \lceil *k/m \rceil$. It is the maximum value of $\sum_{i=1}^{m} x_i^2$ under the constraint $x_{\text{max}} = y$. Without loss of generality, we assume $x_1 = y$ and $y \geq x_2 \geq x_3 \geq \cdots \geq x_m \geq 0$.

We argue that $\arg \max_{x_{\text{max}}} \{ \sum_{i=1}^{m} x_i^2 \} = \{ y, y, \ldots, y, k - ry, 0, \ldots, 0 \}$, that is $x_1 = x_2 = \cdots = x_r = y, x_{r+1} = k - ry$ where $r = \lceil *k/y \rceil$.

To show this, we first consider $x_2$: if $x_2 = y$, then consider $x_3$; otherwise, $x_2 < y$, then we search for the largest index $j$ such that $x_j > 0$. Note the fact that if $x \geq y > 0$, then $(x + 1)^2 + (y - 1)^2 = x^2 + y^2 + 2(x - y) + 2 > x^2 + y^2$. So if we increase $x_2$ by 1 and decrease $x_j$ by 1, we will enlarge $\sum_{i=1}^{m} x_i^2$. After we update $x_2 = x_2 + 1, x_j = x_j - 1$, we still get a decreasing sequence $x_1 \geq x_2 \geq \cdots \geq x_m \geq 0$. If we still have $x_2 < y$, then we repeat the same procedure above (by increasing $x_2$ and decreasing $x_j$ where $j$ is the largest index such that $x_j > 0$). We repeat this process until $x_2 = y$ or $x_1 + x_2 = k$. If it is the former case (i.e., $x_2 = y$), then we consider $x_3$ and so on. Finally, we will get the sequence $x_1 = x_2 = \cdots = x_r = y, x_{r+1} = k - ry$ where $r = \lceil *k/y \rceil$ as claimed, and this setting maximizes $\sum_{i=1}^{m} x_i^2$.

Next we claim that $h(y + 1) > h(y)$. The reason is similar to the above. We update the sequence $x_1 = x_2 = \cdots = x_r = y, x_{r+1} = k - ry$ (which corresponding to $h(y)$) from $x_1$: we increase $x_1$ by 1; search the largest index $s$ such that $x_s > 0$ and decrease $x_1$ by 1. And then consider $x_2$ and so on and so forth. This process won’t stop until $x_1 = x_2 = \cdots = x_s = y + 1$ and $x_{s+1} = k - s(y + 1)$ with $s = \lceil *k/y \rceil + 1$. Thus $h(y + 1) > h(y)$.

By enumerating all the possible values of $x_{\text{max}}$, we split Equation (3.5) into three parts as follows (corresponding to the cases when $x_{\text{max}} = k, x_{\text{max}} \in \lceil *k + 1)/2, k - 1 \rceil$ and $x_{\text{max}} \in \lceil *k + 1)/2, k - 1 \rceil$).
The first term of Equation (A1) comes from the fact that if $x_{\text{max}} = k$, then all the possible cases for $(x_1, x_2, \ldots, x_m)$ are $(k, 0, 0, \ldots, 0), \ldots, (0, \ldots, 0, k)$, and there are $m$ cases all together. For each case, the value of each term in the summation is $\binom{n_{\text{max}}}{k}$, giving rise to the first term in Equation (A1).

The third term on the right hand side of Equation (A1) can be derived as follows. First, observe that

$$\sum_{\substack{x_1 + x_2 + \cdots + x_m = \lfloor k + 1 \rfloor \atop x_i \geq 0}} \binom{n_{\text{max}}}{x_1} \binom{n_{\text{max}}}{x_2} \cdots \binom{n_{\text{max}}}{x_m} = \binom{m n_{\text{max}}}{k}.$$ 

On the other hand, as $x_{\text{max}} \leq \lfloor k \rfloor (k + 1)/2 - 1 = \lfloor k \rfloor (k - 1)/2$, we have:

$$k^2 - \sum_{i=1}^{\lfloor k \rfloor (k-1)/2} x_i^2 \geq \frac{k^2 - h(x_{\text{max}})}{2} \geq \frac{k^2 - h\left(\lfloor k \rfloor \frac{k-1}{2}\right)}{2} \geq \frac{(k - 1)^2}{4},$$

where the second inequality uses the fact that $h(y)$ is an increasing function, and the last inequality comes from that $h\left(\lfloor k \rfloor \frac{k-1}{2}\right) \leq \left(\lfloor k \rfloor \frac{k-1}{2}\right)^2 + \left(\lfloor k \rfloor \frac{k-1}{2}\right)^2 + 1 \leq k^2/4 + k^2/4 + 1 = k^2/2 + 1$.

In what remains, it suffices to estimate all the three terms on the right hand side of Equation (A1). We will repeatedly use the well-known combinatorial inequality $\binom{n}{k} < (en/k)^k$. 

A.1.1 The first term of Equation (A1)

According to the assumptions in Equation (3.6), we know

$$p \leq \frac{1}{\sqrt{e}} \left( \frac{1}{n^3 m} \right)^{\frac{1}{2}} \left( \frac{k}{n_{\text{max}}} \right)^{\frac{1}{2}}.$$ 

Thus, for the first term of Equation (A1), we have:

$$p^{2k} \binom{n_{\text{max}}}{k} m < \left( \frac{1}{e^k} \left( \frac{1}{n^3 m} \right) \left( \frac{k}{n_{\text{max}}} \right)^k \right) \left( \frac{e n_{\text{max}}}{k} \right)^k m = \frac{1}{n^3}. \tag{A2}$$
A.1.2 The second term of Equation (A1)

For the second term of Equation (A1), we relax the constraint $x_{\text{max}} \geq y_i \geq 0$ to $y_i \geq 0$. Thus, we have:

$$
\sum_{y_1 + \cdots + y_{m-1} \leq k \cdot x_{\text{max}}} \left( \binom{N_{\text{max}}}{y_1} \cdots \binom{N_{\text{max}}}{y_{m-1}} \right) \leq \sum_{y_1 + \cdots + y_{m-1} \leq k \cdot x_{\text{max}}} \left( \binom{N_{\text{max}}}{y_1} \cdots \binom{N_{\text{max}}}{y_{m-1}} \right)
$$

$$= (m - 1)N_{\text{max}} \leq \left( \frac{e(m - 1)N_{\text{max}}}{k - x_{\text{max}}} \right)^{k - x_{\text{max}}} \quad \text{(A3)}
$$

Now apply (A3) to the second term of (A1), we have (starting from the second line, we replace $x_{\text{max}}$ to be $j$ for simplicity):

$$p^{2k} \sum_{x_{\text{max}} = \lceil \frac{s}{k} \rceil}^{k-1} \left( \binom{m}{1} \binom{N_{\text{max}}}{x_{\text{max}}} \sum_{y_1 + \cdots + y_{m-1} \leq (k - j) \cdot x_{\text{max}}} \left( \binom{N_{\text{max}}}{y_1} \cdots \binom{N_{\text{max}}}{y_{m-1}} \right) p^{x_{\text{max}}(k - x_{\text{max}})} \right)
$$

$$< \sum_{j = \lceil \frac{s}{k} \rceil}^{k-1} \left( \binom{eN_{\text{max}}}{j} \right) p^{2k+j(k-j)} \left( \frac{e(m - 1)N_{\text{max}}}{k - x_{\text{max}}} \right)^{k - x_{\text{max}}}
$$

$$= \sum_{j = \lceil \frac{s}{k} \rceil}^{k-1} \left( m^{k-j+1}N_{\text{max}}^k e^k \left( \frac{1}{j} \right)^j \left( \frac{1}{k-j} \right)^{k-j} p^{2k+j(k-j)} \right)
$$

$$< \sum_{j = \lceil \frac{s}{k} \rceil}^{k-1} \left( m^{k-j+1}N_{\text{max}}^k e^k \left( \frac{2}{k} \right) k^{k-j} p^{2k+j(k-j)} \right) \quad \text{(A4)}
$$

where the last inequality holds due to the inequality of arithmetic and geometric means.

Note that by tedious elementary calculation, we know $[2k + j(k - j)]/k \geq (k - j + 1)/2 \geq 1$ when $\lceil \frac{s}{k} \rceil (k + 1)/2 \leq j \leq k - 1$. Since

$$p \leq \frac{1}{2ek i} \left( \frac{1}{n^3 m^2} \right)^{1/2} \frac{k}{N_{\text{max}}} < 1
$$

by Equation (3.6), for each $j$ satisfying $\lceil \frac{s}{k} \rceil (k + 1)/2 \leq j \leq k - 1$, we have:

$$m^{k-j+1}N_{\text{max}}^k e^k \left( \frac{2}{k} \right) k^{k-j} p^{2k+j(k-j)}
$$

$$\leq m^{k-j+1}N_{\text{max}}^k e^k \left( \frac{2}{k} \right) k^{k-j} \left( \frac{1}{(2e)^k} \right)^{1/k} \left( \frac{1}{k} \right)^{k-j+1}
$$

$$\leq N_{\text{max}}^k e^k \left( \frac{2}{k} \right) k^{k-j} \left( \frac{1}{(2e)^k} \right)^{1/k} \left( \frac{1}{k} \right)^{k-j+1}
$$

$$\leq \frac{1}{kn^3} \quad \text{(A5)}
$$

where the inequality on the fourth line holds as $k \leq N_{\text{max}}$ and $(k - j + 1)/2 \geq 1$. 
A.1.3 The third term of Equation (A1)

For the third term of (A1), note that \((k-1)^2/4 + 2k > k^2/4\) and by plugging in the condition

\[
p \leq \frac{1}{e^z} \left( \frac{1}{n^3} \right)^{\frac{4}{z^2}} \left( \frac{k}{N_{\text{max}}} \right)^{\frac{4}{z}} < 1,
\]

we have

\[
\left( \frac{m N_{\text{max}}}{k} \right) \left( \frac{1}{e^z} \right) \left( \frac{1}{n^3} \right)^{\frac{4}{z^2}} + \frac{1}{n^3} \frac{1}{k} \frac{k^2}{N_{\text{max}}} = \frac{1}{n^3}.
\]

(A6)

Finally, combining (A2), (A5), and (A6), we have:

\[
E[I] \leq \frac{1}{n^3} + \frac{k}{2} \cdot \frac{1}{k n^3} + \frac{1}{n^3} = \frac{5}{2n^3}.
\]

This concludes Lemma 3.6.

A.2 The missing details in case (ii) of part (a) of Theorem 3.5

Set \(N_u := |B_{\pm}^Y(u)|\) and \(N_v := |B_{\pm}^Y(v)|\). Let \(\tilde{k} := \lfloor K/2 - 2 \rfloor\). Easy to see \(\tilde{k} \geq 1\) since \(K \geq 8\beta^2 \geq 8\).

For every set \(S\) of \((\tilde{k} + 2)\) vertices in \(G_{uv}^{\text{local}}\), let \(A_S\) be the event “\(S\) is a clique in \(G_{uv}^{\text{local}}\) containing edge \((u, v)\) given \(F\)’” and \(Y_S\) its indicator random variable.

Set

\[
Y = \sum_{|S|=\tilde{k}+2} Y_S.
\]

Then \(Y\) is the number of cliques of size \((\tilde{k} + 2)\) in \(G_{uv}^{\text{local}}\) containing edge \((u, v)\) given \(F\).

Linearity of expectation gives:

\[
E[Y] = \sum_{|S|=\tilde{k}+2} E[Y_S] = \sum_{x_1+x_2=\tilde{k}+2} \left( \begin{array}{c} N_u - 1 \\ x_1 \end{array} \right) \left( \begin{array}{c} N_v - 1 \\ x_2 \end{array} \right) p^{(x_1+1)(x_2+1)-1}.
\]

(A7)

To estimate this quantity, we first prove the following result:

Lemma A.1. If \(\tilde{k} \geq 1\) and

\[
p \leq \frac{1}{e^z} \left( \frac{1}{n^3} \right)^{\frac{4}{z^2}} \frac{\tilde{k}}{N_u + N_v},
\]

hold, then \(E[Y] = O\left(n^{-3}\right)\).

Proof. Easy to see that if \(\tilde{k} > N_u + N_v - 2\), then the summation on the right hand side of Equation (A7) is 0. Now we move on to the case when \(\tilde{k} \leq (N_u - 1) + (N_v - 1) < N_u + N_v\). Easy to see \(p < 1\) in this case. Thus, the right hand side of (A7) can be bounded from above by:

\[
\sum_{x_1+x_2=\tilde{k}+2} \left( \begin{array}{c} N_u - 1 \\ x_1 \end{array} \right) \left( \begin{array}{c} N_v - 1 \\ x_2 \end{array} \right) p^{(x_1+1)(x_2+1)-1}
\]
The proof of Lemma 3.7

where the last inequality holds due to condition (A8).

Easy to see that there exist two constants $c_1$ and $c_2$ which depend on the Besicovitch constant $\beta$ and $\alpha$, such that

if $K \geq 8\beta^2$ and $p \leq c_1 \cdot (1/n)^{c_2/K}$, then the conditions in Equation (A8) will hold.

On the other hand, we have

$$\mathbb{P} \left[ \alpha_{n,v} \left( \mathcal{G}_{uv}^{local} \right) \geq K/2 | F \right] = \mathbb{P}[Y > 0] \leq \mathbb{E}[Y].$$

Thus, by Lemma A.1, we know that

If $K \geq 8\beta^2$ and $p \leq c_1 \cdot (1/n)^{c_2/K}$, then $\mathbb{P} \left[ \alpha_{n,v} \left( \mathcal{G}_{uv}^{local} \right) \geq K/2 | F \right] = O(n^{-3}).$ (A9)

A.3 The proof of Lemma 3.7

Proof. By using a similar argument as in Appendix A.1, it is easy to see that the maximum value of $\sum_{i=1}^{m} x_i^2$, under the constraints $\sum_{i=1}^{m} x_i = k$ and $x_i \in [0, N_{\max}]$ for each $i \in [1, m]$, is $rN_{\max}^2 + (k - rN_{\max})^2$ where $r = \lfloor \frac{k}{N_{\max}} \rfloor$. The maximum can be achieved when $(x_1, x_2, \ldots, x_m) = (N_{\max}, \ldots, N_{\max}, k - rN_{\max}, \ldots).$ Thus, we have

$$\mathbb{E}[I] \leq p^{2k} \sum_{x_1 + x_2 + \ldots + x_m = k \atop x_i \geq 0} \left( N_{\max} \begin{array}{c} \frac{N_{\max}}{x_1} \\ \frac{N_{\max}}{x_2} \\ \vdots \\ \frac{N_{\max}}{x_m} \end{array} \right) \cdot p^{(k^2 - \sum_{i=1}^{m} x_i^2)/2}$$

$$\leq \sum_{x_1 + x_2 + \ldots + x_m = k \atop x_i \geq 0} \left( N_{\max} \begin{array}{c} \frac{N_{\max}}{x_1} \\ \frac{N_{\max}}{x_2} \\ \vdots \\ \frac{N_{\max}}{x_m} \end{array} \right) \cdot p^{(k^2 - \sum_{i=1}^{m} x_i^2)/2} + k^{2k}$$

$$= \left( \frac{mN_{\max}}{k} \right)^k p^{(k^2 - \sum_{i=1}^{m} x_i^2)/2 + k}$$

$$\leq \left( \frac{emN_{\max}}{k} \right)^k p^{(kN_{\max} + 1) - \frac{k+1}{2}N_{\max}}$$

$$\leq \left( \frac{emN_{\max}}{k} \right)^k p^{k \left( \frac{N_{\max} + 1}{1} \right) - \frac{k+1}{2}N_{\max} + \frac{k}{2}N_{\max} + 1}$$

$$= \left( \frac{emN_{\max}}{k} \right)^k p^{\frac{1}{2}k - \frac{3}{2}N_{\max} + 1}.$$ (A10)

where Equation (A10) holds since $k/N_{\max} \geq r > k/N_{\max} - 1$. 

(A11)
Pick a constant $C_3$, which only depends on the Besicovitch constant $\beta$ and $\alpha$, such that
\[
\frac{C_3[\star] \log_{1/p} n}{2|\Lambda|} - 3 \geq 16 \log_{1/p} n \geq 6N_{\text{max}}.
\]
This can be done since $N_{\text{max}}$ is a constant and $(1/n)^{8/(3N_{\text{max}})} \leq p < 1$ implies $\log_{1/p} n \geq 3N_{\text{max}}/8$. Set $K = C_3[\star] \log_{1/p} n$. Recall that $k = [\star] K/(2|\Lambda|) - 2$, thus $k \geq 16 \log_{1/p} n \geq 6N_{\text{max}}$. Also note that $m \leq n$. Hence, we have the following inequality.
\[
\left(\frac{em_{\text{max}}}{k} p^{\frac{k-1}{2m_{\text{max}}+1}}\right)^k < \left(\frac{en_{\text{max}}}{k} p^{\frac{1}{k}}\right)^k \leq \left(\frac{en_{\text{max}}}{k} n^{-3}\right)^k = O(n^{-3}). \tag{A12}
\]
Finally, combining (A11) and (A12), we have $\mathbb{E}[I] = O(n^{-3})$. □

A.4 Proof of Lemma 3.9

We will use the standard second moment method to prove this lemma. For completeness, we first state the second moment method. For those who are familiar with this method, our main technical step is to estimate the summation on the right hand side of Equation (A13).

Definition A.2 (Symmetric random variables). We say random variables $Z_1, \ldots, Z_m$, where $Z_i$ is the indicator random variable for event $U_i$, are symmetric if for every $i \neq j$ there is a measure preserving mapping of the underlying probability space that permutes the $m$ events and sends event $U_i$ to event $U_j$.

Let $Z$ be a nonnegative integral-valued random variable, and suppose we have a decomposition $Z = Z_1 + \cdots + Z_m$, where $Z_i$ is the indicator random variable for event $U_i$ and $Z_1, \ldots, Z_m$ are symmetric. For indices $i, j$, write $i \sim j$ if $i \neq j$ and the events $U_i, U_j$ are not independent. For any fixed index $i$, we set
\[
\Delta^* := \sum_{j \sim i} \mathbb{P}[U_j | U_i],
\]
and note that by the symmetry of $Z_i$, $\Delta^*$ is independent of the index $i$ (thus we are not denoting it by $\Delta^*_i$).

Theorem A.3 (The second moment method [1]). If $\mathbb{E}[Z] \to \infty$ and $\Delta^* = o(\mathbb{E}[Z])$ as $m \to \infty$, then $\mathbb{P}[Z = 0] \to 0$.

Now we are ready to prove Lemma 3.9.

Proof of Lemma 3.9. Set $k = [\star] \log_{1/p} n$. Now consider all the $k$-set $S_i$ of vertices in $G(n, p)$. Let $U_i$ be the event “$S_i$ is a clique” and $Z_i$ its indicator random variable. (All $k$-sets “look the same” so that the $Z_i$’s are symmetric.) $I$ is a finite index set enumerating all the $k$-sets in $G(n, p)$. Set
\[
Z = \sum_{i \in I} Z_i,
\]
so that $Z$ is the number of $k$-cliques in $G(n, p)$. Linearity of expectation gives:
\[
\mathbb{E}[Z] = \sum_{i \in I} \mathbb{E}[Z_i] = \binom{n}{k} p^{\binom{k}{2}}.
\]
Easy to see that

$$
\Delta^* = \sum_{j=1}^{k-1} \mathbb{P}[U_j | U_1] = \sum_{\ell'=2}^{k-1} \binom{k}{\ell'} \binom{n-k}{\ell'} \frac{1-\binom{n-k}{\ell'}}{k-\ell'}.
$$

Since \( k = \lceil \log_1/n \rceil \) and \( p \leq \left(\frac{1}{n}\right)\lceil \sqrt{n} \rceil \), we know that \( p^{k-1} > p^k > 1/n, k \leq n^{1/4} \) and \( \log k / \log n \leq 1/4 \). Also note that \( p \geq \left(\frac{1}{n}\right)^{1/11} \). Easy to see that

$$
k + 1 > \log_{1/p} n \geq \frac{1}{\sqrt{k}} = 11,
$$

which further implies \( k > 10 \).

Note that for sufficiently large \( n \), we have \( n - k > n/2 \). Thus, using \( p^{k-1} > 1/n \) as derived earlier, we have:

$$
\mathbb{E}[Z] = \left(\binom{n}{k}\right) \binom{k}{\ell} \binom{n-k}{\ell} \frac{(n-k)^{\ell}}{k^\ell} > \left(\binom{n}{2k}\right) \frac{k^{\ell}}{n^{\ell}} = n^{\frac{\ell}{2}} \mathcal{Z} > n^{\frac{\ell}{2}} \to \infty.
$$

To apply Theorem A.3, it suffices to estimate the term \( \Delta^*/\mathbb{E}[Z] \).

$$
\frac{\Delta^*}{\mathbb{E}[Z]} = \sum_{\ell'=2}^{k-1} \binom{k}{\ell'} \binom{n-k}{\ell'} \frac{1-\binom{n-k}{\ell'}}{k-\ell'} \frac{(n-k)^{\ell}}{k^\ell}. \tag{A13}
$$

We estimate the summation on the right hand side term by term. Let

$$
g(\ell) := \binom{k}{\ell} \binom{n-k}{k-\ell} \frac{1-\binom{n-k}{\ell}}{n^{\ell}}.
$$

Note that for \( \ell \in [2, k-1] \), we have

$$
g(\ell) = \binom{k}{\ell} \binom{n-k}{k-\ell} \frac{1-\binom{n-k}{\ell}}{n^{\ell}} \leq \binom{k}{\ell} \binom{n-k}{k-\ell} \frac{1}{(n-k)^{\ell}} \leq \binom{k}{\ell} \frac{1}{(n-k)^{\ell}} \frac{\ell}{n^{\ell/2}} \leq \binom{k}{\ell} \frac{\ell}{n^{\ell/2}} \frac{\ell}{2}.
$$

Now set

$$
h(\ell) = -\ell \log \frac{\ell}{ek^2} + \frac{1}{2} \ell (\ell - 1),
$$

and thus by the above inequality we have \( g(\ell) \leq n^{h(\ell)} \). We claim that \( \forall \ell \in [2, k-1], h(\ell) \leq \max \{ h(2), h(k-1) \} \). We then further use \( h(2) \) and \( h(k-1) \) to derive an upper bound on \( g(l) \).
Indeed, by the following direct calculation, we can easily prove this:

Note that its derivative with respect to $\ell$ is

$$h'(\ell) = -\frac{\log \ell + \log(n-k) - 2 \log k}{\log n} + \frac{2\ell - 1}{2k}. $$

Further calculate its second derivative:

$$h''(\ell) = -\frac{1}{\log n} \frac{1}{\ell} + \frac{1}{k}. $$

Note that $\ell_0 = k/\log n$ is the only solution of $h''(\ell) = 0$. Easy to check that $\ell_0 \leq k - 1$. Therefore, we have the following two cases:

**Case (i)** If $\ell_0 < 2$, then $h'(\ell)$ is strictly increasing on $\ell \in [2, k - 1]$;

**Case (ii)** If $\ell_0 \in [2, k - 1]$, then $h'(\ell)$ is strictly decreasing on $[2, \ell_0]$ and strictly increasing on $[\ell_0, k - 1]$.

Note that $h'(2) < -\frac{\log 2 + \log(n/2) - 2 \log k}{\log n} + \frac{3}{2k} = -1 + \frac{2 \log k}{\log n} + \frac{3}{2k} > -1 + \frac{1}{2} + \frac{3}{20} < 0$.

Thus in either case $h'(\ell)$ can become 0 at most once within $\ell \in [2, k - 1]$, and we have $\max_{\ell \in [2,k-1]} h(\ell) = \max \{ h(2), h(k-1) \}$.

Routine calculation shows that (using that $n - k > n/2$), for $n$ large enough:

$$h(2) < -\frac{2 \left( \log(2(n/2)) - 2 \log k \right)}{\log n} + \frac{1}{k} = -2 + \frac{1}{k} + \frac{2 \log k}{\log n} + \frac{4 \log k}{\log n} < -\frac{1}{2}. $$

$$h(k-1) < -\frac{(k-1) \left( \log(n/2) - 1 - \log k - \log(k/(k-1)) \right)}{\log n} + \frac{k^2 - 3k + 2}{2k}$$

$$< \left[ \frac{k^2 - 3k + 2}{2k} - (k-1) \right] + \frac{k(1 + \log 2)}{\log n} + \frac{k \log k}{\log n} + \frac{k \log(k/(k-1))}{\log n}$$

$$< -\frac{1}{2} + \frac{1}{10} - \frac{k}{6} < -\frac{1}{2}. $$

Thus, $\forall \ell \in [2, k - 1]$, we have $g(\ell) < n^{-1/2}$. It then follows that

$$\sum_{\ell=2}^{k-1} g(\ell) < k \cdot n^{-1/2} \leq n^{1/2} \cdot n^{-1} = n^{-1}. $$

Hence by Equation (A13), we have $\Delta^*/\mathbb{E}[Z] < n^{-1/4} \rightarrow 0$, and therefore $\mathbb{P}[Z = 0] \rightarrow 0$ by Theorem A.3.

**A.5 | Proof of Theorem 3.13**

**A.5.1 | Proof of part (a)**

We use the same notation $\tilde{A}_{uv}$ and $B_{uv}$ as in the proof of Theorem 3.5. Now we set

$$N_{\max} := \frac{5 \log n}{\log(\log n/((\sigma \theta 2^n n r^d))). $$
Again, denote \( F \) to be the event that “for every \( v \in \mathcal{X}_n \), the ball \( B_r(v) \cap \mathcal{X}_n \) contains at most \( N_{\text{max}} \) points”; and \( F^c \) denotes the complement event of \( F \). By Lemma 3.12, we know that, \( \mathbb{P}[F^c] = O(n^{-3}) \).

Let \( K_n \) be a positive number to be determined such that \( K_n \to \infty \) as \( n \to \infty \). By applying the pigeonhole principle and the union bound, we have:

\[
\mathbb{P} \left[ \omega_{u,v} \left( G_n^{0,p} | \tilde{A}_n \right) \geq K_n | F \right] \\
\leq \mathbb{P} \left[ \omega_{u,v} \left( G_n^{0,p} | \tilde{A}_n \right) \geq K_n/2 | F \right] + \mathbb{P} \left[ \omega_{u,v} \left( G_n^{0,p} | \tilde{B}_n \right) \geq K_n/2 | F \right].
\]  
(A14)

**A.5.2 Case (i): Bounding the first term in Equation (A14)**

Applying Theorem 2.7 for points in \( A_{uv} \) gives a well-separated clique-partitions family \( P = \{ P_i \}_{i \in \Lambda} \) of \( A_{uv} \) with \( |\Lambda| \leq \beta^2 \) being a constant. Augment each \( P_i \) to \( \tilde{P}_i = P_i \cup \{ u \} \cup \{ v \} \). Check Figure 2A.

Again, by applying pigeonhole principle and the union bound, we have:

\[
\mathbb{P} \left[ \omega_{u,v} \left( G_n^{0,p} | \tilde{A}_n \right) \geq K_n/2 | F \right] \leq \sum_{i=1}^{|\Lambda|} \mathbb{P} \left[ \omega_{u,v} \left( G_n^{0,p} | \tilde{P}_i \right) \geq K_n/(2|\Lambda|) | F \right].
\]  
(A15)

Now set \( k_n := [3|K_n/(2|\Lambda|)] - 2 \). Easy to see that \( k_n \to \infty \) as \( n \to \infty \). Same as in the proof for Theorem 3.5, we have

\[
\mathbb{P} \left[ \omega_{u,v} \left( G_n^{0,p} | \tilde{P}_i \right) \geq k_n + 2 | F \right] \\
\leq p^{2k_n} \sum_{x_1 + x_2 + \ldots + x_m = k_n} \binom{N_{\text{max}}}{x_1} \binom{N_{\text{max}}}{x_2} \ldots \binom{N_{\text{max}}}{x_m} \left( \frac{n^d}{\log n} \right)^{m} = O(n^{-3}).
\]  
(A16)

where \( m \leq n \) is the number of \( C_i^{(j)} \) in the clique-partition \( \tilde{P}_i \).

If \( K_n \leq 2N_{\text{max}} \), then \( k_n \in [1, N_{\text{max}}] \). By applying Lemma 3.6, we have that if \( 1 \leq k_n \leq N_{\text{max}} \), then there exist constants \( c_1^d \) and \( c_2^d \) (which depend on the Besicovitch constant \( \beta \)), such that if

\[
p \leq c_1^d \cdot \left( \frac{1}{n} \right)^{c_2^d/K_n} \frac{K_n}{N_{\text{max}}},
\]

then the right hand side of Equation (A16) is \( O(n^{-3}) \).

Thus, following the same argument in the proof of part (a) of Theorem 3.5, we have

If \( K_n \leq 2N_{\text{max}} \) and \( p \leq c_1^d \cdot (1/n)^{c_2^d/K_n} (K_n/N_{\text{max}}) \),

then \( \mathbb{P} \left[ \omega_{u,v} \left( G_n^{0,p} | \tilde{A}_n \right) \geq K_n/2 | F \right] = O(n^{-3}) \).  
(A17)

Finally, suppose \( K_n > K_0 = 2N_{\text{max}} \). Using Equation (A17), we know that if

\[
p \leq c_1^d \cdot \left( \frac{1}{n} \right)^{c_2^d/K_0} \frac{K_0}{N_{\text{max}}} = 2c_1^d \left( \frac{\sigma \theta^d (m^d)}{\log n} \right)^{c_2^d/K_0} = \left( 2c_1^d \cdot (\sigma \theta^d)^{c_2^d/N_{\text{max}}} \right) \left( \frac{m^d}{\log n} \right)^{c_2^d/N_{\text{max}}},
\]

\( (\sigma \theta^d)^{c_2^d/N_{\text{max}}} \),
and \( K_n > K_0 \), then

\[
\mathbb{P} \left[ \omega_{u,v} \left( G^0_{n,\rho} | \Lambda_n \right) \geq K_n / 2 \right] \leq \mathbb{P} \left[ \omega_{u,v} \left( G^0_{n,\rho} | \Lambda_n \right) \geq K_0 / 2 \right] = O(n^{-3}).
\]

Set \( C_1^a := 2c_1^a \cdot (\sigma \theta^2 d)^{c_2^a/10} \) and \( C_2^a := c_2^a / 10 \) be two constants. Combining this with Equation (A17), we thus obtain that:

If \( K_n \to \infty \) and \( p \leq \min \left\{ C_1^a \cdot (nr^d / \log n)^{C_2^a}, c_1^b \cdot (1/n)^{c_2^b/K_0 (K_n/N_{\max})} \right\} \),

then

\[
\mathbb{P} \left[ \omega_{u,v} \left( G^0_{n,\rho} | \Lambda_n \right) \geq K_n / 2 \right] = O(n^{-3}).
\] (A18)

**A.5.3 Case (ii): Bounding the second term in Equation (A14)**

Recall that \( B_{uv} = B_{r}^{X_k}(u) \cup B_{r}^{X_k}(v) \) (see Figure 2B). We again use the notation \( \tilde{\sigma}^{local}_{uv} \) defined in the proof of part (a) of Theorem 3.5. Set \( N_u := |B_{r}^{X_k}(u)| \) and \( N_v := |B_{r}^{X_k}(v)| \). Let \( \tilde{k}_n:= \lfloor * \rfloor K_n/2 - 2 \). Easy to see \( \tilde{k}_n \geq 1 \). Using the same argument as in Case (ii) in the proof of Theorem 3.5, we have

\[
\mathbb{P} \left[ \omega_{u,v} \left( \tilde{\sigma}^{local}_{uv} \right) \geq K_n / 2 \right] \leq \sum_{x_1 + x_2 = \tilde{k}_n} \binom{N_u-1}{x_1} \binom{N_v-1}{x_2} p^{(x_1+1)(x_2+1)-1}.
\] (A19)

By applying Lemma A.1, we know that there exist constants \( c_1^b \) and \( c_2^b \) (which depend on the Besicovitch constant \( \beta \)), such that if \( K_n \to \infty \) and

\[
p \leq c_1^b \cdot \left( \frac{1}{n} \right)^{c_2^b/K_0} \frac{K_n}{N_{\max}},
\]

then the right hand side of Equation (A19) is \( O(n^{-3}) \). That is,

if \( K_n \to \infty \) and \( p \leq c_1^b \cdot (1/n)^{c_2^b/K_0 (K_n/N_{\max})} \),

then

\[
\mathbb{P} \left[ \omega_{u,v} \left( \tilde{\sigma}^{local}_{uv} \right) \geq K_n / 2 \right] = O(n^{-3}).
\] (A20)

Pick

\[
K_n = 4N_{\max} = \frac{20 \log n}{\log(\log n / (\sigma \theta^2 dnr^d))} = \frac{20 \log n}{\log(\log n / nr^d)} + \text{const}.
\]

Note that this makes the first term of the constraint on \( p \) in Equation (A18) dominate. Thus, combining Equations (A20), (A18), and (A14), there exist constants \( C_1^c = \min \{ C_1^a, c_1^b \} \) and \( C_2^c = \max \{ C_2^a, c_2^b / 10 \} \) such that if \( p \) satisfies conditions in Equation (3.14), then

\[
\mathbb{P} \left[ \omega_{u,v} \left( G^0_{n,\rho} \right) \geq K_n \right] \leq \mathbb{P} \left[ \omega_{u,v} \left( G^0_{n,\rho} \right) \geq K_n \right] + \mathbb{P}[F^c] = O(n^{-3}).
\]

Finally, by applying the union bound, we derive that with high probability, for each of the \( O(n^2) \) long-edge \((u,v)\), its edge clique number

\[
\omega_{u,v}(G^0_{n,\rho}) \leq \frac{\log n}{\log(\log n / nr^d)}
\]

as long as Equation (3.14) holds. This completes the proof of Part (a) if Theorem 3.13.
We again try to bound the two terms on the right hand side of Equation (A14) from above separately. For case (i), our result relies on the following lemma.

**Lemma A.4.** There exists a constant \( C_3 > 0 \) depending on the Besicovitch constant \( \beta \) such that if \((nr^d / \log n)^{4/15} \leq p < 1 \) and \( K_n = C_3 [\ast] \log_{1/p} n \), then we have

\[
P \left[ \omega_{u'v'} \left( G_n \theta | \tilde{\theta}_{u'v'} \right) \geq K_n/2 \middle| F \right] = O(n^{-3}).
\]

**Proof.** By a similar argument in Appendix A.3, we know that

\[
P \left[ \omega_{u'v'} \left( G_n \theta | \tilde{\theta}_{u'v'} \right) \geq K_n/2 \middle| F \right] \leq \frac{1}{\sqrt{2\pi}} \left( \frac{em_{\max}}{k_n} p^{-\frac{1}{2}} k_n^{-\frac{3}{2}} n_{\max}^{-1} \right) k_n,
\]

where \( k_n = [\ast] K_n/(2|\Lambda|) - 2 \). Pick a constant \( C_3 \), which only depends on the Besicovitch constant \( \beta \), such that

\[
\frac{C_3 [\ast] \log_{1/p} n}{2|\Lambda|} - 3 \geq 16 \log_{1/p} n \geq 6n_{\max}.
\]

This can be done since we have \( \log n/n^d \to \infty \) and \((nr^d / \log n)^{4/15} \leq p < 1 \) and thus

\[
\log_{1/p} n \geq \frac{15 \log n}{4 \log(\log n/n^d)} = \frac{3}{8} \frac{5 \log n}{\log(\log n/n^d)} > \frac{3}{8} \frac{5 \log n}{\log n/(\sigma \theta^2 n^d)} = \frac{3}{8} n_{\max}.
\]

Set \( K_n = C_3 [\ast] \log_{1/p} n \). Recall that \( k_n = [\ast] K_n/(2|\Lambda|) - 2 \), thus \( k_n \geq 16 \log_{1/p} n \geq 6n_{\max} \). Finally, we use Equation (A22) (with \( k \) being replaced by \( k_n \)) to complete the proof.

Now pick such \( C_3 \) in Lemma A.4. We know that the following statement holds.

\[
\text{If } (nr^d / \log n)^{4/15} \leq p < 1,
\]

\[
\text{then } P \left[ \omega_{u'v'} \left( G_n \theta | \tilde{\theta}_{u'v'} \right) \geq C_3 [\ast] \log_{1/p} n/2 \middle| F \right] = O(n^{-3}).
\]

For case (ii), we know that if event \( F \) has already happened, then \(|B_{uv}| \leq 2n_{\max} \). Note that if \((nr^d / \log n)^{C_3/42} \leq p < 1 \), then we have

\[
\frac{C_3 [\ast] \log_{1/p} n}{2} > 2 \frac{5 \log n}{(1/2) \log(\log n/n^d)} + \left( \frac{\log n}{\log(\log n/n^d)} - \frac{C_3}{2} \right)
\]

\[
> 2 \frac{5 \log n}{\log n/(\sigma \theta^2 n^d)} = 2n_{\max} \geq |B_{uv}|.
\]

Set \( \xi = \min \{4/15, C_3/42 \} \). Hence, we obtain that:

\[
\text{If } (nr^d / \log n)^{\xi} \leq p < 1, \text{ then } P \left[ \omega_{u'v'} \left( G_n \theta | \tilde{\theta}_{u'v'} \right) \geq C_3 [\ast] \log_{1/p} n/2 \middle| F \right] = 0.
\]
Thus, combining Equations (A23), (A22), and (A14), we know that if \((nr^d/\log n)^x \leq p < 1\), then with high probability, for every long-edge \((u,v)\), its edge clique number \(\omega_{u,v}(G_n^{0,p}) \lesssim \log_{1/p} n\). This completes the proof of Theorem 3.13.

APPENDIX B: THE MISSING PROOFS IN SECTION 4

B.1 Proof of Part (III)—Supercritical regime

In this section, we discuss the order of \(\omega(G_n^{0})\) in the regime \(\sigma nr^d/\log n \to t \in (0, \infty)\).

B.1.1 Proof of upper bound

We first focus on the upper bound of \(\omega(G_n^{0})\). Let \(N\) be a random variable sampled from \(Poisson((1+\delta)n)\) for some \(\delta > 0\) (say \(\delta = 1/2\)). Note that \(G_N\) is an \(r\)-neighborhood graph of the Poisson point process \(P_{(1+\delta)n}\) with intensity \((1+\delta)n\). Completely analogously to the proof of upper bound in Section 4.1, we have

\[
\mathbb{P}\left[\omega\left(G_N^{0}\right) \geq k_n\right] < \frac{1}{\sqrt{2\pi}} \left(\frac{1}{k_n}\right)^{k_n}.
\]

Finally, pick \(k_n = 3 \log_{1/(1-q)} nr^d\). Since in the supercritical regime \(\sigma nr^d/\log n \to t \in (0, \infty)\), routine calculations show that \(\mathbb{P}\left[\omega\left(G_N^{0}\right) \geq k_n\right] = o(1)\). Hence, a.s.

\[
\omega\left(G_N^{0}\right) \lesssim \log(nr^d).
\]

B.1.2 Proof of lower bound

Now let us move on to the lower bound of \(\omega(G_n^{0})\). For this regime, we need slightly stronger condition on the range of \(t\). That is, \(\sigma nr^d \geq T \log n\) for some constant \(T > 0\) to be determined.

Completely analogously to the proof of lower bound in Section 4.1, let \(N\) be a random variable sampled from \(Poisson(1-\delta')n\) for some \(\delta' \in (0, 1)\) (say \(\delta' = 1/2\)). Set \(k_n\) to be an integer to be determined. Now, we have

\[
\mathbb{P}\left[\omega\left(G_N^{0}\right) \leq k_n\right] \leq \mathbb{P}\left[\omega\left(G_N^{0}\right) \leq k_n\right] + e^{-\gamma'n},
\]

for some constant \(\gamma' > 0\) (depending on \(\delta'\)) by Lemma 2.2. Now fix some constant \(\rho \in (0, 1)\) (say \(\rho = 1/2\)). Recall \(W_{1/2} = B_{1/2}(0)\). By Lemma 4.1, there exist points \(x_1, x_2, \ldots, x_m\) with \(m = \Omega(r^{-d}) \geq 1\) such that the sets \(x_i + W_{1/2}\) are disjoint and

\[
\nu\left(x_i + W_{1/2}\right) \geq (1-\rho)\sigma r^d.
\]

for \(i = 1, \ldots, m\). Let \(X_i\) be the set of points of \(G_N\) falling in \(x_i + W_{1/2}\). Then, we have

\[
\mathbb{P}\left[\omega\left(G_N^{0}\right) \leq k_n\right] \leq \prod_{i=1}^{m} \mathbb{P}\left[\omega\left(G_N^{0} | x_i\right) \leq k_n, \ldots, \omega\left(G_N^{0} | x_m\right) \leq k_n\right] = \prod_{i=1}^{m} \mathbb{P}\left[\omega\left(G_N^{0} | x_i\right) \leq k_n\right].
\]
Easy to see that all the points falling in any \( \frac{r}{2} \)-ball span a clique in \( G_N \). Thus, for each \( i \), we have
\[
P \left[ \omega \left( G_N^{q,0} \mid X_i \right) \leq k_n \right] = P \left[ \omega \left( G \left( |X_i|, 1 - q \right) \right) \leq k_n \right].
\]

Set
\[
\Phi_n := \frac{\left(1 - \rho\right)(1 - \delta')(\sigma \theta nr^d)}{2^{d+1}},
\]
which goes to infinity as \( n \) grows. Note that \( |X_i| \sim \text{Poisson}(\bar{\lambda}) \) where \( \bar{\lambda} := (1 - \delta')n \cdot \nu(x_i + W_{1/2}) \geq 2 \Phi_n \) (see Equation (4.3)). Now pick \( k_n := \lceil * \rceil \log_{1/(1-q)} \Phi_n = \Omega \left( \log(nr^d) \right) \). By the law of total probability, we have
\[
P \left[ \omega \left( G\left( |X_i|, 1 - q \right) \right) \leq k_n \right] \leq \mathbb{P} \left[ |X_i| \leq \Phi_n \right] + \sum_{j=\lceil * \rceil}^{\infty} \mathbb{P} \left[ \omega \left( G(j, 1 - q) \right) \leq \lceil * \rceil \log_{1/(1-q)} j \right] \frac{e^{-\bar{\lambda}j}}{j!} < e^{-\frac{1}{10} \bar{\lambda}} + \sum_{j=\lceil * \rceil}^{\infty} e^{-\bar{\lambda}j} \frac{e^{-\bar{\lambda}j}}{j!} \leq e^{-\frac{1}{10} \bar{\lambda}} + e^{-\bar{\lambda}} \sum_{j=0}^{\infty} \left( \bar{\lambda} / e \right)^j \frac{j!}{j!} = e^{-\frac{1}{10} \bar{\lambda}} + e^{-\bar{\lambda}} \cdot e^{\bar{\lambda}/e} < 2e^{-\frac{1}{10} \bar{\lambda}}.
\]

Inequality (B1) holds due to Lemmas 2.2 and 4.2. Now set
\[
T := \frac{10 \cdot 2^d}{(1 - \rho)(1 - \delta') \theta}.
\]

Note that \( \sigma nr^d \geq T \log n \). Then
\[
e^{-\frac{1}{10} \bar{\lambda}} \leq e^{-\frac{\left(1 - \rho\right)(1 - \delta') \theta}{10 \cdot 2^d} (\sigma nr^d)} \leq e^{-\log n} = n^{-1}.
\]

It follows that \( \mathbb{P} \left[ \omega \left( G_N^{q,0} \right) \leq k_n \right] = o(1) \) with \( k_n = \Omega \left( \log(nr^d) \right) \), which concludes the proof of part (III) of Theorem 1.4.