AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **87(3)** (2023), Pages 365–390

On the random greedy linear uniform hypergraph packing*

Fang Tian†

Department of Applied Mathematics
Shanghai University of Finance and Economics
Shanghai, 200433, China
tianf@mail.shufe.edu.cn

Zi-Long Liu

School of Optical-Electrical and Computer Engineering
University of Shanghai for Science and Technology
Shanghai, 200093, China
liuzl@usst.edu.cn

Xiang-Feng Pan

School of Mathematical Sciences
Anhui University
Anhui, 230601, China
xfpan@ahu.edu.cn

Abstract

For any fixed integer $k \geqslant 3$, a hypergraph H is k-uniform if each edge is a set of k vertices, and is said to be linear if any two distinct edges intersect in at most one vertex. A k-clique in a graph is a complete subgraph on k vertices. The random greedy k-clique removal algorithm starts with a complete graph on vertex set $[n] = \{1, 2, \dots, n\}$, and iteratively removes the edges of a uniformly chosen k-clique. The process terminates once the remaining graph contains no k-cliques, say after M steps. Let $E(M)$ be the edge set of the graph when the process terminates. This process is equivalently viewed as creating a random linear k-uniform hypergraph, which is also called the random greedy linear hypergraph packing algorithm, starting with vertex set $[n]$ and no edges,

* This work was partially supported by NSFC (No. 12071274)
† Corresponding Author

ISSN: 2202-3518

and adding the vertex set of each chosen k-clique as a new edge in the hypergraph at each step. This algorithm generates a linear k-uniform hypergraph with M edges. A special case of a conjecture proposed by Bennett and Bohman implies that with high probability $|E(M)| = n^{2-\frac{2}{k+1}+o(1)}$. Fewer results are known for the cases $k \geqslant 4$. In this paper, we directly show that $|E(M)| \leqslant n^{2-\frac{1}{k(k-1)-2}+o(1)}$ for $k \geqslant 3$, which implies $M \geqslant \frac{n^2}{k(k-1)} - n^{2-\frac{1}{k(k-1)-2}+o(1)}$. This upper bound on $|E(M)|$ equals $n^{\frac{7}{4}+o(1)}$ when $k = 3$, coinciding with an upper bound obtained by Grable for random triangle packing. We also show the bound is a natural barrier in our proof.

1 Introduction

Hypergraphs, which are also known as set systems and block designs, are fundamental to the study of complex discrete systems. Let k and ℓ be given fixed integers such that $2 \leqslant \ell \leqslant k-1$. A hypergraph H on vertex set $[n]$ is a *k-uniform hypergraph* (*k-graph* for short) if each edge is a set of k vertices, and is said to be *linear* if every pair of distinct edges intersect in at most one vertex. A k-graph is called a partial Steiner (n, k, ℓ)-system, if every subset of size ℓ (*ℓ-set* for short) lies in at most one edge of H. In particular, partial Steiner $(n, k, 2)$-systems are linear hypergraphs. Linear hypergraphs are the subject of much study [1, 3, 14, 15].

Random greedy processes are classical mathematical models, and the power usually goes beyond the probabilistic method used in previous work [10]. A k-clique in a graph is a complete subgraph on k vertices. The *random greedy k-clique removal algorithm* starts with a complete graph on the vertex set $[n] = \{1, 2, \ldots, n\}$, denoted by $\mathrm{G}(0)$, and $\mathrm{G}(i+1)$ is the remaining graph from $\mathrm{G}(i)$ by selecting one k-clique uniformly at random out of all k-cliques in $\mathrm{G}(i)$ and deleting all its edges from the edge set $E(i)$ of $\mathrm{G}(i)$. The process terminates once the remaining graph contains no k-cliques. Let $M = \min\{i : \mathrm{G}(i)$ is k-clique free$\}$, and $E(M)$ be the set of edges left unsaturated by the produced k-cliques, which are related via

$$|E(M)| = \binom{n}{2} - \binom{k}{2} M.$$

This process is equivalently viewed as creating a random linear k-graph, which is also called the *random greedy linear hypergraph packing algorithm*, starting with vertex set $[n]$ and no edges, and adding the vertex set of each chosen k-clique as a new edge in the hypergraph at each step. This algorithm generates a linear k-graph with M edges. It is an important special case of the *random greedy hypergraph matching algorithm* [2]. As usual, we say some property holds *with high probability* (*w.h.p.* for short) if the probability that it holds tends to 1.

Bennett and Bohman [2] proposed a conjecture on the random greedy hypergraph matching algorithm, which would imply that *w.h.p.* $|E(M)| = n^{2-2/(k+1)+o(1)}$. It is exactly $|E(M)| = n^{3/2+o(1)}$ proposed by Bollobás and Erdős [7] when $k = 3$, which is also called the *random greedy triangle-removal algorithm* [6]. The next few results all assume $k = 3$. Spencer [13] and independently Rödl and Thoma [12] showed *w.h.p.* $|E(M)| = o(n^2)$. Grable [9] improved this bound to $|E(M)| \leqslant n^{7/4+o(1)}$. Bohman et al. [5] introduced the critical interval method for proving dynamic concentration. They [6] confirmed the exponent in a breakthrough by generalizing the approach in [5]. Fewer results are known on $|E(M)|$ when $k \geqslant 4$.

In this paper, we directly discuss the structure of the algorithm, using a heuristic assumption to find the trajectories of an ensemble of random variables when the process evolves. Compared with the random triangle-removal process, it is challenging to analyze the one-step change of the number of k-cliques in terms of some auxiliary variables. Finally, we obtain our main results.

Theorem 1.1. *Given a fixed integer $k \geqslant 3$, consider the random greedy k-clique removal algorithm on the vertex set $[n]$. Let $E(M)$ be the edge set of the graph when the process terminates and M be the number of edges in the generated linear k-graph. With high probability, there exists some positive constant λ such that*

$$M \geqslant \frac{n^2}{k(k-1)} - \frac{\sqrt[3]{2}}{k(k-1)} n^{2-\frac{1}{k(k-1)-2}} \log^{\lambda} n,$$

which implies that

$$|E(M)| \leqslant n^{2-\frac{1}{k(k-1)-2}+o(1)}.$$

We make no attempt to optimize the coefficient $\sqrt[3]{2}$ here. We also show our result corresponds to the inherent barrier of the process. There is a gap between our upper bound and the conjectured upper bound $|E(M)| = n^{2-2/(k+1)+o(1)}$ from [2]. The main obstacle is to find better variables to characterize the one-step change of the number of k-cliques. The analysis still works for the case of $k = 3$ to give $|E(M)| \leqslant n^{7/4+o(1)}$, which coincides with the bound in [5, 9]. We believe it is not easy to improve Theorem 1.1.

The remainder of this paper is organized as follows. In the next section, notation and some lemmas for analyzing the random linear k-graph packing algorithm are presented. In Section 3, we discuss the evolution of the process in detail and estimate the trajectories of these random variables. We formally prove the concentration of our random variables in Section 4.

2 Notation and some lemmas

All asymptotics in this paper are with respect to $n \to \infty$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an arbitrary probability space. Let $(\mathcal{F}_i)_{i\geqslant 0}$ be the filtration given by the evolution of the process. Let $\eta, N > 0$ be constants and $\{X(i)\}_{i\geqslant 0}$ be a sequence of random variables, and let $\Delta X = X(i+1) - X(i)$ denote the one-step change for the random variable $X(i)$. The pair $\{X(i), \mathcal{F}_i\}_{i\geqslant 0}$ is then called a submartingale or a supermartingale if $X(i)$ is $\mathcal{F}_i$-measurable and $\mathbb{E}[\Delta X|\mathcal{F}_i] \geqslant 0$ or $\mathbb{E}[\Delta X|\mathcal{F}_i] \leqslant 0$ for all $i \geqslant 0$, respectively.

We say that $\{X(i), \mathcal{F}_i\}_{i\geqslant 0}$ is (η, N)-bounded if $X(i)-\eta \leqslant X(i+1) \leqslant X(i)+N$ for all $i \geqslant 0$. Furthermore, for two positive-valued functions f, g on the variable n, we write $f \ll g$ to denote $\lim_{n\to\infty} f(n)/g(n) = 0$ and $f \sim g$ to denote $\lim_{n\to\infty} f(n)/g(n) = 1$. Let $a = b \pm c$ be short for $a \in [b-c, b+c]$, $\binom{S}{b} = \emptyset$ and $\binom{a}{b} = 0$ if $b > |S|$ and $b > a$. We also use the standard asymptotic notation o, O, Ω and Θ. All logarithms are natural, and the floor and ceiling signs are omitted whenever they are not crucial.

For $2 \leqslant m \leqslant k$, $u \in [n]$ and $U_m = \{u_1, \cdots, u_m\} \in \binom{[n]}{m}$, let $\mathcal{N}_u = \mathcal{N}_u(i) = \{x \in [n] : xu \in E(i)\}$, $\mathcal{N}_{U_m} = \mathcal{N}_{U_m}(i) = \cap_{j=1}^{m} \mathcal{N}_{u_j}$. Let $\mathcal{K}_m(i)$ be the set of m-cliques in $\mathrm{G}(i)$ and $\mathbf{Q}_m(i) = |\mathcal{K}_m(i)|$. Our goal is to estimate the random variable $\mathbf{Q}_k(i)$. Fix one $U_m \in \binom{[n]}{m}$, and define the random variable $\mathbf{R}_{k,U_m} = \mathbf{R}_{k,U_m}(i)$ to be

$$\mathbf{R}_{k,U_m} = \begin{cases} \left|\left\{A \in \binom{\mathcal{N}_{U_m}}{k-m} : A \text{ induces a } (k-m)\text{-clique in } \mathrm{G}(i)\right\}\right|, \ 2 \leqslant m \leqslant k-1; \\ \mathbf{1}_{U_k}, \ m = k. \end{cases} \tag{2.1}$$

Sometimes, for short, we will suppress i. For $2 \leqslant m \leqslant k-1$, $\mathbf{R}_{k,U_m}$ counts the number of $(k-m)$-cliques in $\mathrm{G}(i)$ such that every vertex in the sets counted by $\mathbf{R}_{k,U_m}$ is in $\mathcal{N}_{U_m}$; particularly $\mathbf{R}_{k,U_{k-1}} = |\mathcal{N}_{U_{k-1}}|$ is the codegree of the vertex subset U_{k-1}. Also $\mathbf{1}_{U_k}$ is the indicator random variable with $\mathbf{1}_{U_k} = 1$ if the subgraph induced by U_k in $\mathrm{G}(i)$ is complete, instead $\mathbf{1}_{U_k} = 0$ otherwise.

Suppose that U_k is the vertex set of the k-clique chosen from $\mathrm{G}(i)$ at $(i+1)$-th step, and let $U_m \in \binom{U_k}{m}$ with $2 \leqslant m \leqslant k$. Define

$$\mathbf{Q}_{k,U_m}(i) = \left|\left\{A \in \mathcal{K}_k(i) \mid A \cap U_k = U_m\right\}\right|,$$

namely, $\mathbf{Q}_{k,U_m}(i)$ denotes the number of k-cliques in $\mathrm{G}(i)$ that exactly contains the vertices U_m in U_k. In particular, $\mathbf{Q}_{k,U_k}(i) = 1$. Thus, we have

$$\mathbf{Q}_k(i) - \mathbf{Q}_k(i+1) = \sum_{m=2}^{k} \Big(\sum_{U_m \in \binom{U_k}{m}} \mathbf{Q}_{k,U_m}(i) \Big). \tag{2.2}$$

By inclusion-exclusion formula, we have

$$\begin{aligned} \mathbf{Q}_{k,U_m}(i) = \mathbf{R}_{k,U_m} &+ \sum_{T_1 \in \binom{U_k \setminus U_m}{1}} (-1)^1 \mathbf{R}_{k,U_m \cup T_1} + \dots \\ &+ \sum_{T_{k-m-1} \in \binom{U_k \setminus U_m}{k-m-1}} (-1)^{k-m-1} \mathbf{R}_{k,U_m \cup T_{k-m-1}} + (-1)^{k-m} \mathbf{R}_{k,U_k}, \end{aligned} \tag{2.3}$$

where $\mathbf{R}_{k,U_m}$ in (2.1) equals the number of extensions to one copy of k-clique from the set U_m. Note that

$$\sum_{U_m \in \binom{U_k}{m}} \Big(\sum_{T_j \in \binom{U_k \setminus U_m}{j}} \mathbf{R}_{k,U_m \cup T_j} \Big) = \binom{m+j}{m} \sum_{U_{m+j} \in \binom{U_k}{m+j}} \mathbf{R}_{k,U_{m+j}}$$

for the integer j with $0 \leqslant j \leqslant k-m$ because each element in $\mathbf{R}_{k,U_{m+j}}$ on the right side is counted $\binom{m+j}{m}$ times on the left side. Summing the above corresponding displays (2.3) for all $U_m \in \binom{U_k}{m}$ with $2 \leqslant m \leqslant k$ altogether into the equation (2.2), it follows that

$$
\begin{aligned}
&\mathbf{Q}_k(i) - \mathbf{Q}_k(i+1) \\
&= \sum_{U_2 \in \binom{U_k}{2}} \mathbf{R}_{k,U_2} + \sum_{U_3 \in \binom{U_k}{3}} \left[(-1)^1 \binom{3}{2} + (-1)^0 \binom{3}{3} \right] \mathbf{R}_{k,U_3} + \dots \\
&\quad + \sum_{U_{k-1} \in \binom{U_k}{k-1}} \left[(-1)^{k-3} \binom{k-1}{2} + \dots + (-1)^0 \binom{k-1}{k-1} \right] \mathbf{R}_{k,U_{k-1}} \\
&\quad + \left[(-1)^{k-2} \binom{k}{2} + \dots + (-1)^0 \binom{k}{k} \right] \mathbf{R}_{k,U_k}.
\end{aligned}
$$

Since $\sum_{j=2}^{r} (-1)^{r-j} \binom{r}{j} = (-1)^r (r-1)$ for any given integer $r \geqslant 2$ and $\mathbf{R}_{k,U_k} = 1$, we have

$$
\begin{aligned}
\mathbf{Q}_k(i) - \mathbf{Q}_k(i+1) = &\sum_{U_2 \in \binom{U_k}{2}} \mathbf{R}_{k,U_2} - 2 \sum_{U_3 \in \binom{U_k}{3}} \mathbf{R}_{k,U_3} + \dots \\
&+ (-1)^{k-1}(k-2) \sum_{U_{k-1} \in \binom{U_k}{k-1}} \mathbf{R}_{k,U_{k-1}} + (-1)^k (k-1). \qquad (2.4)
\end{aligned}
$$

Then the expectation $\mathbb{E}[\Delta \mathbf{Q}_k | \mathcal{F}_i]$ of $\Delta \mathbf{Q}_k$ is

$$
\begin{aligned}
&\mathbb{E}[\Delta \mathbf{Q}_k | \mathcal{F}_i] \\
&= -\sum_{U_k \in \mathcal{K}_k(i)} \frac{\sum_{U_2 \in \binom{U_k}{2}} \mathbf{R}_{k,U_2} + \dots + (-1)^{k-1}(k-2) \sum_{U_{k-1} \in \binom{U_k}{k-1}} \mathbf{R}_{k,U_{k-1}} + (-1)^k (k-1)}{\mathbf{Q}_k(i)} \\
&= (-1)^{k+1}(k-1) - \frac{1}{\mathbf{Q}_k(i)} \sum_{U_2 \in \mathcal{K}_2(i)} (\mathbf{R}_{k,U_2})^2 + \dots + \frac{(-1)^k (k-2)}{\mathbf{Q}_k(i)} \sum_{U_{k-1} \in \mathcal{K}_{k-1}(i)} (\mathbf{R}_{k,U_{k-1}})^2,
\end{aligned}
\qquad (2.5)
$$

where the last equality is true because

$$
\sum_{U_k \in \mathcal{K}_k(i)} \sum_{U_m \in \binom{U_k}{m}} \mathbf{R}_{k,U_m} = \sum_{U_m \in \mathcal{K}_m(i)} (\mathbf{R}_{k,U_m})^2
$$

for $2 \leqslant m \leqslant k-1$ by double counting.

We also need the following lemmas to establish dynamic concentration on variables $\mathbf{Q}_k(i)$ and $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$.

Lemma 2.1 (Lemma 2.3 in [6]). *Let $a_1, \dots, a_\ell \in \mathbb{R}$ and $a \in \mathbb{R}$. Suppose that $|a_i - a| \leqslant \varepsilon$ for all $1 \leqslant i \leqslant \ell$. Then,*

$$
\frac{(\sum_{i=1}^{\ell} a_i)^2}{\ell} \leqslant \sum_{i=1}^{\ell} a_i^2 \leqslant \frac{(\sum_{i=1}^{\ell} a_i)^2}{\ell} + 4\ell\varepsilon^2.
$$

Lemma 2.2 (Hoeffding and Azuma [11]). *Suppose a sequence of random variables $\{X(i)\}_{i\geqslant 0}$ is a supermartingale (respectively, submartingale) and $|X(i)-X(i-1)| < c_i$. Then for any positive integer ℓ and any positive real number a,*

$$\mathbb{P}\big[X(\ell)-X(0)\geqslant a\big]\leqslant \exp\left[\frac{-a^2}{2\sum_{i=1}^{\ell} c_i^2}\right];$$
$$\left(\textit{respectively, } \mathbb{P}\big[X(\ell)-X(0)\leqslant -a\big]\leqslant \exp\left[\frac{-a^2}{2\sum_{i=1}^{\ell} c_i^2}\right]\right).$$

Lemma 2.3 (Lemma 6 and Lemma 7 in [4]). *Suppose $\{X(i)\}_{i\geqslant 0}$ is an (η, N)-bounded supermartingale (respectively, submartingale) with initial value 0 and $\eta \leqslant N/10$. Then for any positive integer ℓ and any positive real number a with $a < \eta\ell$,*

$$\mathbb{P}\big[X(\ell)\geqslant a\big]\leqslant \exp\left[-\frac{a^2}{3\ell\eta N}\right]. \quad \left(\textit{respectively } \mathbb{P}\big[X(\ell)\leqslant -a\big]\leqslant \exp\left[-\frac{a^2}{3\ell\eta N}\right].\right).$$

Finally, we will use the Chernoff bound, stated below, in Section 5 to explain why we believe our bound is not best possible.

Lemma 2.4 ([8]). *For $X \sim \mathrm{Bin}(n,p)$ and any $0 < \xi \leqslant np$, $\mathbb{P}\big[|X-np| > \xi\big] < 2\exp\left[-\frac{\xi^2}{3np}\right]$.*

3 Estimates on the variables in $\mathbf{G}(i)$

A pseudo-random heuristic for divining the evolution of variables plays a central role in the understanding of graph processes, as shown in [2, 5, 6]. As well as estimating the random variable $\mathbf{Q}_k(i)$, we also estimate the likely values of the auxiliary random variables $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$ throughout the process. We assume the process produces a graph whose variables are roughly the same as they are in the binomial random graph $\mathcal{G}(n,p)$ with the same edge density. Rescale the number of steps i to be $t = t_i = \frac{i}{n^2}$ and introduce a notion of edge density as

$$p = p_i = p(t) = 1 - \frac{k(k-1)i}{n^2} = 1 - k(k-1)t. \tag{3.1}$$

Note that p can be viewed as either a continuous function of t or as a function of the discrete variable i. We pass between these interpretations without comment. With this notation, we have

$$|E(i)| = \binom{n}{2} - \binom{k}{2}i = \binom{n}{2} - \frac{1}{2}(1-p)n^2 = \frac{1}{2}(n^2p - n). \tag{3.2}$$

Hence the number of edges in $\mathrm{G}(i)$ with edge density p is approximately equal to the one in $\mathcal{G}(n,p)$ up to the negligible linear term when p lies in some proper range.

Under the assumption that $\mathrm{G}(i)$ resembles $\mathcal{G}(n,p)$, for a fixed integer $k \geqslant 3$, $2 \leqslant m \leqslant k-1$ and $U_m \in \binom{[n]}{m}$, we anticipate that the expressions of $\mathbf{Q}_k(i)$ and $\mathbf{R}_{k,U_m}$ are

$$\mathbf{Q}_k(i) \sim \frac{n^k}{k!} p^{\binom{k}{2}} \quad \text{and} \quad \mathbf{R}_{k,U_m} \sim \frac{n^{k-m}}{(k-m)!} p^{\binom{k}{2}-\binom{m}{2}},$$

where $\frac{n^k}{k!} p^{\binom{k}{2}}$ approximately counts the expected number of k-cliques in $\mathcal{G}(n,p)$; $\binom{n-m}{k-m} p^{\binom{k}{2}-\binom{m}{2}} \sim \frac{n^{k-m}}{(k-m)!} p^{\binom{k}{2}-\binom{m}{2}}$ approximately counts the expected number of $(k-m)$-cliques in which every vertex is in $\mathcal{N}_{U_m}$. Our main results below will establish these results under certain conditions. A detailed analysis and the proof of concentration will follow in the next section.

Theorem 3.1. *Given a fixed integer $k \geqslant 3$, let $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$; then there exist positive constants μ, γ_m and λ such that, with high probability,*

$$\mathbf{Q}_k(i) \leqslant \begin{cases} \frac{n^3}{6} p^3 + \frac{n^2}{3} p, \ k = 3, \\ \frac{n^k}{k!} p^{\binom{k}{2}} + \frac{n^{k-1}}{2} p^{\binom{k}{2}-4}, \ k \geqslant 4; \end{cases} \tag{3.3}$$

$$\mathbf{Q}_k(i) \geqslant \frac{n^k}{k!} p^{\binom{k}{2}} - \sigma^2 n^{\alpha} p^{-1} \log^{\mu} n; \tag{3.4}$$

$$\mathbf{R}_{k,U_m} = \frac{n^{k-m}}{(k-m)!} p^{\binom{k}{2}-\binom{m}{2}} \pm \sigma n^{\beta_m} \log^{\gamma_m} n, \quad 2 \leqslant m \leqslant k-1; \tag{3.5}$$

holding for every $i \leqslant i_0$ with $i_0 = \frac{n^2}{k(k-1)} - \frac{\sqrt[3]{2}}{k(k-1)} n^{2-\frac{1}{k(k-1)-2}} \log^{\lambda} n$, where

$$\alpha = k - \frac{\binom{k}{2}+1}{2\binom{k}{2}-2}, \tag{3.6}$$

$$\beta_m = k - m - \frac{\binom{k}{2}-\binom{m}{2}}{2\binom{k}{2}-2}, \quad 2 \leqslant m \leqslant k-1; \tag{3.7}$$

and the error function $\sigma = \sigma_i = \sigma(t)$ is defined to be

$$\sigma(t) = 1 - k(k-1)\log p(t) \tag{3.8}$$

with initial value $\sigma(0) = 1$ and it grows slowly.

Theorem 3.1 is proved in Section 4. It implies that with high probability these random variables are concentrated around the trajectories we have heuristically divined until at least step i_0. The dynamic concentration in turn show that the process produces a graph of size at most $|E(i_0)|$ with high probability. We make no attempt to optimize the constants μ, λ and γ_m in all error terms with $2 \leqslant m \leqslant k-1$. There are many choices of them that can be balanced to satisfy certain inequalities. If we choose them to satisfy, for $k \geqslant 3$, for example, the inequalities

$$\left[\binom{k}{2}+1\right]\lambda > \mu + 3; \tag{3.9}$$

$$\left[\binom{k}{2}-\binom{m}{2}\right]\lambda > \gamma_m + 2, \quad 2 \leqslant m \leqslant k-1; \tag{3.10}$$

$$\left[\binom{m}{2}-1\right]\lambda + \gamma_m > \gamma_2, \quad 3 \leqslant m \leqslant k-1; \tag{3.11}$$

$$\gamma_2 > \frac{1}{2}; \tag{3.12}$$

then we will see that these choices are sufficient for our proof of Theorem 3.1 in the next section. We do not replace them with their actual values. This is for the interest of understanding the role of these constants played in the calculations.

Proof of Theorem 1.1. We recover the number of edges when $i = i_0$ to be

$$|E(i_0)| = \binom{n}{2} - \binom{k}{2} i_0 \sim \frac{\sqrt[3]{2}}{2} n^{2-\frac{1}{k(k-1)-2}} \log^\lambda n.$$

Theorem 1.1 follows directly from Theorem 3.1 by $|E(M)| \leqslant |E(i_0)|$ and $M \geqslant i_0$ with room to spare in the power of the logarithmic factor. $\square$

Remark 3.2. We now show that in each of (3.3)–(3.5), the expression is asymptotically dominated by the first term, with the second term negligible compared to the first term. According to (3.1), define

$$p_0 = p_{i_0} = 1 - \frac{k(k-1)i_0}{n^2} = \sqrt[3]{2} n^{-\frac{1}{k(k-1)-2}} \log^\lambda n. \tag{3.13}$$

Since $i \leqslant i_0$ in Theorem 3.1, we have $p \geqslant p_0$ in (3.13). When $k \geqslant 3$, the main terms in (3.3) clearly dominate the second terms with $\lambda > 0$. We also have $\frac{n^k}{k!} p^{\binom{k}{2}} \gg \sigma^2 n^\alpha p^{-1} \log^\mu n$ in (3.4) because $\sigma = O(\log n)$ in (3.8), where α is in (3.6), λ and μ satisfy the relation as the equation shown in (3.9). Thus, it follows that $\mathbf{Q}_k(i) = (1+o(1)) n^k p^{\binom{k}{2}}/k!$ in (3.3) and (3.4). Similarly, $\mathbf{R}_{k,U_m} = (1 \pm o(1)) \frac{n^{k-m}}{(k-m)!} p^{\binom{k}{2}-\binom{m}{2}}$ in (3.5) for $2 \leqslant m \leqslant k-1$ and $k \geqslant 3$ because $\frac{n^{k-m}}{(k-m)!} p^{\binom{k}{2}-\binom{m}{2}} \gg \sigma n^{\beta_m} \log^{\gamma_m} n$, where β_m is in (3.7), λ and γ_m satisfy the relation as the equation shown in (3.10). For $k = 3$, we have $\alpha = 2$ in (3.6) and $\beta_2 = \frac{1}{2}$ in (3.7), which coincide with the errors of the number of triangles and the codegree for any two vertices in [5].

As a supplement, we show the claim below that is required to analyze Theorem 3.1 in the next section.

Claim: Assuming the estimates in (3.5) hold on $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$, we have

$$\sum_{U_m \in \mathcal{K}_m(i)} (\mathbf{R}_{k,U_m})^2 \geqslant \frac{m! \binom{k}{m}^2 \mathbf{Q}_k^2(i)}{n^m p}, \tag{3.14}$$

and

$$\sum_{U_m \in \mathcal{K}_m(i)} (\mathbf{R}_{k,U_m})^2 \leqslant \begin{cases} \frac{2!\binom{k}{2}^2 \mathbf{Q}_k^2(i)}{n^2 p} + 2\sigma^2 n^{2k-3} p \log^{2\gamma_2} n, & m = 2; \\ \frac{n^m p}{m!}\left(\frac{n^{k-m}}{(k-m)!} p^{\binom{k}{2}-\binom{m}{2}} + \sigma n^{\beta_m} \log^{\gamma_m} n\right)^2, & 3 \leqslant m \leqslant k-1. \end{cases} \tag{3.15}$$

Proof of Claim. Firstly, we prove $\mathbf{Q}_m(i) \leqslant \frac{n^m}{m!}p$ for $2 \leqslant m \leqslant k-1$. In fact, note that $\mathbf{Q}_2(i) = |E(i)| \sim \frac{n^2}{2}p$ shown in (3.2) when $p \geqslant p_0$ in (3.13), and $\mathbf{Q}_m(i) \leqslant \frac{n}{m}\mathbf{Q}_{m-1}(i)$ is clearly true because each element in $\mathcal{K}_{m-1}(i)$ has at most n possibilities to become an element in $\mathcal{K}_m(i)$, while each element in $\mathcal{K}_m(i)$ corresponds to exactly m elements in $\mathcal{K}_{m-1}(i)$. Then we recursively achieve $\mathbf{Q}_m(i) \leqslant \frac{n^m}{m!}p$ for $2 \leqslant m \leqslant k-1$.

By Lemma 2.1, for any $U_m \in \mathcal{K}_m(i)$ with $2 \leqslant m \leqslant k-1$, we have

$$\sum_{U_m \in \mathcal{K}_m(i)} (\mathbf{R}_{k,U_m})^2 \geqslant \frac{\left(\sum_{U_m \in \mathcal{K}_m(i)} \mathbf{R}_{k,U_m}\right)^2}{\mathbf{Q}_m(i)}.$$

Note that $\sum_{U_m \in \mathcal{K}_m(i)} \mathbf{R}_{k,U_m} = \binom{k}{m}\mathbf{Q}_k(i)$ because each element on the right side is counted $\binom{k}{m}$ times on the left side, thus we have the lower bound in (3.14) by $\mathbf{Q}_m(i) \leqslant \frac{n^m}{m!}p$.

For the upper bound of $\sum_{U_2 \in \mathcal{K}_2(i)} (\mathbf{R}_{k,U_2})^2$, we have $\beta_2 = k - \frac{5}{2}$ in (3.7) and $\mathbf{Q}_2(i) \sim \frac{n^2}{2}p$, then it follows that

$$\begin{aligned} \sum_{U_2 \in \mathcal{K}_2(i)} (\mathbf{R}_{k,U_2})^2 &\leqslant \frac{\left(\sum_{U_2 \in \mathcal{K}_2(i)} \mathbf{R}_{k,U_2}\right)^2}{\mathbf{Q}_2(i)} + 4\mathbf{Q}_2(i)\left(\sigma n^{\beta_2} \log^{\gamma_2} n\right)^2 \\ &\sim \frac{2!\binom{k}{2}^2 \mathbf{Q}_k^2(i)}{n^2 p} + 2\sigma^2 n^{2k-3} p \log^{2\gamma_2} n, \end{aligned}$$

when $m = 2$ by Lemma 2.1 and $\sum_{U_2 \in \mathcal{K}_2(i)} \mathbf{R}_{k,U_2} = \binom{k}{2}\mathbf{Q}_k(i)$. For $3 \leqslant m \leqslant k-1$, by the estimates in (3.5) and $\mathbf{Q}_m(i) \leqslant \frac{n^m}{m!}p$, we have the trivial upper bound of $\sum_{U_m \in \mathcal{K}_m(i)} (\mathbf{R}_{k,U_m})^2$ in (3.15). □

4 Proof of Theorem 3.1

4.1 The critical interval method

Recall the outline of the critical interval method [2, 5, 6] used to control some graph parameters as the process evolves. Let the stopping time τ be the smallest index i such that any one of the random variables violates its corresponding trajectory. In our situation, this would be the smallest index i such that one of (3.3)–(3.5) fails. Let the event $\mathcal{E}_X$ be of the form $X(i) = x(i) \pm e(i)$ for all $i \leqslant i_0$, where $X(i)$ is some random variable, $x(i)$ is the expected trajectory and $e(i)$ is the error term. We show that the event $\{\tau = i_0\}$ holds by means of $\{\tau = i_0\} = \cap_{X\in\mathcal{I}}\mathcal{E}_X$, where $|\mathcal{I}|$ is polynomial in n.

For each such random variable $X(i)$, we define a critical interval I_X for its bound (upper and lower) that has one endpoint at the bound we are trying to maintain and the other slightly closer to the expected trajectory of the random variable. Consider a fixed step $j < i_0$ such that $X(j) \in I_X$. Define the stopping time $\tau_{X,j}$ to be $\tau_{X,j} = \min\{i_0, \max\{j,\tau\}, \text{the smallest } i \geqslant j \text{ such that } X(i) \notin I_X\}$, which will make it possible to establish the martingale condition and apply the martingale inequalities in Lemma 2.2 and Lemma 2.3. We will bound the probability of the event that the designated variable crosses its critical interval in the process, and an application of the union bound over all steps j will show that the probability of the occurrence of any event in the collection is small.

4.2 Tracking $\mathbf{Q}_k(i)$

For the upper bound of $\mathbf{Q}_3(i)$ as the equation shown in (3.3), it has been shown in [5] by taking a critical interval as $\mathcal{I}^u_{\mathbf{Q}_3} = \mathcal{I}^u_{\mathbf{Q}_3}(i) = (\frac{n^3}{6}p^3 + \frac{1}{4}n^2p, \frac{n^3}{6}p^3 + \frac{1}{3}n^2p)$. For the upper bound of $\mathbf{Q}_k(i)$ as the equation shown in (3.3) when $k \geqslant 4$, we introduce a critical interval as

$$\mathcal{I}^u_{\mathbf{Q}_k} = \mathcal{I}^u_{\mathbf{Q}_k}(i) = \Big(\frac{n^k}{k!}p^{\binom{k}{2}} + Bn^{k-1}p^{\binom{k}{2}-4}, \frac{n^k}{k!}p^{\binom{k}{2}} + \frac{n^{k-1}}{2}p^{\binom{k}{2}-4}\Big), \tag{4.1}$$

where

$$B = \frac{1}{2} - \frac{1}{2\binom{k}{2}} + \frac{1}{3\binom{k}{2}(k-4)!} < \frac{1}{2}. \tag{4.2}$$

Consider a fixed step $j < i_0$. Suppose $\mathbf{Q}_k(j) \in \mathcal{I}^u_{\mathbf{K}_k}(j)$. Define

$$\tau^u_{\mathbf{Q}_k,j} = \min\Big\{i_0, \max\{j,\tau\}, \text{the smallest } i \geqslant j \text{ such that } \mathbf{Q}_k(i) \notin \mathcal{I}^u_{\mathbf{Q}_k}\Big\}.$$

Let $j \leqslant i < \tau^u_{\mathbf{Q}_k,j}$. Hence, all calculations in this subsection are conditioned on the estimates in (3.5).

By the equation in (2.5), it follows that

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] &= (-1)^{k+1}(k-1) - \frac{1}{\mathbf{Q}_k(i)}\sum_{U_2\in\mathcal{K}_2(i)}(\mathbf{R}_{k,U_2})^2 + \dots \\ &\qquad + \frac{(-1)^k(k-2)}{\mathbf{Q}_k(i)}\sum_{U_{k-1}\in\mathcal{K}_{k-1}(i)}(\mathbf{R}_{k,U_{k-1}})^2 \\ &< (-1)^{k+1}(k-1) - \frac{2\binom{k}{2}^2\mathbf{Q}_k(i)}{n^2p} \\ &\qquad + \frac{2}{\mathbf{Q}_k(i)}\frac{n^3p}{3!}\Big(\frac{n^{k-3}}{(k-3)!}p^{\binom{k}{2}-3} + \sigma n^{\beta_3}\log^{\gamma_3} n\Big)^2 + O\big(n^{k-4}p^{\binom{k}{2}-1}\big),\end{aligned}$$

where $\sum_{U_2\in\mathcal{K}_2(i)}(\mathbf{R}_{k,U_2})^2$ and $\sum_{U_3\in\mathcal{K}_3(i)}(\mathbf{R}_{k,U_3})^2$ are replaced by the equations in (3.14) and (3.15), and the last term $O(n^{k-4}p^{\binom{k}{2}-1})$ comes from $\sum_{U_4\in\mathcal{K}_4(i)}(\mathbf{R}_{k,U_4})^2$ in (3.14) that dominates all the remaining terms.

Since $\mathbf{Q}_k(i)\in\mathcal{I}^u_{\mathbf{Q}_k}$ in (4.1), we further have

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] < (-1)^{k+1}(k-1) - \frac{2\binom{k}{2}^2 n^{k-2}}{k!}p^{\binom{k}{2}-1} - 2\binom{k}{2}^2 Bn^{k-3}p^{\binom{k}{2}-5}\\ + \frac{k!n^{k-3}}{3(k-3)!^2}p^{\binom{k}{2}-5} + O\big(\sigma n^{\beta_3}p^{-2}\log^{\gamma_3} n\big),\end{aligned} \tag{4.3}$$

where the term $O(n^{k-4}p^{\binom{k}{2}-1})$ is absorbed into the term $O(\sigma n^{\beta_3}p^{-2}\log^{\gamma_3} n)$ with $\beta_3 = k-3+\frac{\binom{k}{2}-3}{2\binom{k}{2}-2}$ in (3.7).

For all i with $j\leqslant i<\tau^u_{\mathbf{Q}_k,j}$, define the sequence of random variables to be

$$\mathbf{U}(i) = \mathbf{Q}_k(i) - \frac{n^k}{k!}p^{\binom{k}{2}} - \frac{n^{k-1}}{2}p^{\binom{k}{2}-4}. \tag{4.4}$$

Claim 4.1: The sequence $\mathbf{U}(j),\mathbf{U}(j+1),\ldots,\mathbf{U}(\tau^u_{\mathbf{Q}_k,j})$ is a supermartingale and the maximum one step $\Delta\mathbf{U}$ is $O(\sigma n^{k-\frac{5}{2}}\log^{\gamma_2} n)$.

Proof of Claim 4.1. To see this, for $j\leqslant i<\tau^u_{\mathbf{Q}_k,j}$, as the equation in (4.4), we have

$$\mathbb{E}[\Delta\mathbf{U}|\mathcal{F}_i] = \mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] - \frac{n^k}{k!}\Big[p_{i+1}^{\binom{k}{2}} - p_i^{\binom{k}{2}}\Big] - \frac{n^{k-1}}{2}\Big[p_{i+1}^{\binom{k}{2}-4} - p_i^{\binom{k}{2}-4}\Big].$$

Note that by (3.1), if $p=p_i=1-k(k-1)t$, then $p_{i+1}=p-\frac{k(k-1)}{n^2}$. Hence, by Taylor's expansion, we have

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{U}|\mathcal{F}_i] &= \mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] - \frac{n^k}{k!}\Big[-\binom{k}{2}\frac{k(k-1)}{n^2}p^{\binom{k}{2}-1} + O\Big(\frac{1}{n^4}p^{\binom{k}{2}-2}\Big)\Big]\\ &\quad - \frac{n^{k-1}}{2}\Big[-\Big(\binom{k}{2}-4\Big)\frac{k(k-1)}{n^2}p^{\binom{k}{2}-5} + O\Big(\frac{1}{n^4}p^{\binom{k}{2}-6}\Big)\Big]\\ &= \mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] + \frac{2\binom{k}{2}^2 n^{k-2}}{k!}p^{\binom{k}{2}-1} + \Big[\binom{k}{2}-4\Big]\binom{k}{2}n^{k-3}p^{\binom{k}{2}-5}\\ &\quad + O\big(n^{k-4}p^{\binom{k}{2}-2}\big),\end{aligned} \tag{4.5}$$

where the term $O(n^{k-5}p^{\binom{k}{2}-6})$ is absorbed into the term $O(n^{k-4}p^{\binom{k}{2}-2})$ when $p\geqslant p_0$ in (3.13). Using (4.3) and (4.5), we further have

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{U}|\mathcal{F}_i] &< (-1)^{k+1}(k-1) - \Big[2\binom{k}{2}^2 B - \binom{k}{2}^2 + 4\binom{k}{2} - \frac{k!}{3(k-3)!^2}\Big]n^{k-3}p^{\binom{k}{2}-5}\\ &\quad + O\big(n^{k-4}p^{\binom{k}{2}-2}\big) + O\big(\sigma n^{\beta_3}p^{-2}\log^{\gamma_3} n\big)\\ &< (-1)^{k+1}(k-1) - 2\binom{k}{2}n^{k-3}p^{\binom{k}{2}-5} + O\big(\sigma n^{\beta_3}p^{-2}\log^{\gamma_3} n\big),\end{aligned}$$

where the term $O(n^{k-4}p^{\binom{k}{2}-2})$ is absorbed into the term $O(\sigma n^{\beta_3}p^{-2}\log^{\gamma_3} n)$ with β_3 in (3.7), and

$$\begin{aligned}&2\binom{k}{2}^2 B - \binom{k}{2}^2 + 4\binom{k}{2} - \frac{k!}{3(k-3)!^2}\\ &= 3\binom{k}{2} - \binom{k}{2}\frac{2}{3(k-3)!}\\ &> 2\binom{k}{2}\end{aligned}$$

with B as shown in (4.2). Note that

$$\binom{k}{2}n^{k-3}p^{\binom{k}{2}-5} > O(\sigma n^{\beta_3}p^{-2}\log^{\gamma_3} n) + (-1)^{k+1}(k-1)$$

when $p \geqslant p_0$ in (3.13) with appropriate choices of λ and γ_3 such that $[\binom{k}{2}-3]\lambda > \gamma_3+2$ as the equation shown in (3.10). Hence, we have $\mathbb{E}[\Delta\mathbf{U}|\mathcal{F}_i] < 0$ and the sequence $\mathbf{U}(j), \mathbf{U}(j+1), \ldots, \mathbf{U}(\tau^u_{\mathbf{Q}_k,j})$ is a supermartingale.

Next, we show the maximum one step $\Delta\mathbf{U}$ is $O(\sigma n^{k-\frac{5}{2}}\log^{\gamma_2} n)$. With the help of the equations in (4.4) and (4.5), we have

$$\Delta\mathbf{U} = \Delta\mathbf{Q}_k + \frac{2\binom{k}{2}^2}{k!}n^{k-2}p^{\binom{k}{2}-1} + \left[\binom{k}{2} - 4\right]\binom{k}{2}n^{k-3}p^{\binom{k}{2}-5} + O\big(n^{k-4}p^{\binom{k}{2}-2}\big).$$

Apply the equation of $\Delta\mathbf{Q}_k$ in (2.4), the equation of $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ in (3.5), and β_m in (3.7) to the above display, then we finally have

$$\begin{aligned}\Delta\mathbf{U} \leqslant& -\binom{k}{2}\Big(\frac{n^{k-2}}{(k-2)!}p^{\binom{k}{2}-1} - \sigma n^{\beta_2}\log^{\gamma_2} n\Big)\\ &+ 2\binom{k}{3}\Big(\frac{n^{k-3}}{(k-3)!}p^{\binom{k}{2}-\binom{3}{2}} + \sigma n^{\beta_3}\log^{\gamma_3} n\Big) + \ldots\\ &+ \frac{2\binom{k}{2}^2}{k!}n^{k-2}p^{\binom{k}{2}-1} + \left[\binom{k}{2} - 4\right]\binom{k}{2}n^{k-3}p^{\binom{k}{2}-5} + O\big(n^{k-4}p^{\binom{k}{2}-2}\big)\\ =& O\big(\sigma n^{k-\frac{5}{2}}\log^{\gamma_2} n\big),\end{aligned}$$

where the two terms involving $n^{k-2}p^{\binom{k}{2}-1}$ cancel exactly and the term $\sigma n^{\beta_2}\log^{\gamma_2} n$ dominates all the remaining terms. The claim follows. $\square$

The number of steps in the sequence $\mathbf{U}(j), \mathbf{U}(j+1), \ldots, \mathbf{U}(\tau^u_{\mathbf{Q}_k,j})$ is $O(n^2p)$ because $|E(i)| \sim \frac{n^2}{2}p$ shown in (3.2). Note that $\mathbf{Q}_k(j) \in \mathcal{I}^u_{\mathbf{Q}_k}(j)$, by Lemma 2.2, for all i with $j \leqslant i < \tau^u_{\mathbf{Q}_k,j}$, the probability of a large deviation for $\mathbf{Q}_k(i)$ beginning at the step j is at most

$$\begin{aligned}
&\mathbb{P}\Big[\mathbf{Q}_k(i) \geqslant \frac{n^k}{k!}p^{\binom{k}{2}} + \frac{n^{k-1}}{2}p^{\binom{k}{2}-4}\Big] \\
&= \mathbb{P}\Big[\mathbf{U}(i) \geqslant 0\Big] = \mathbb{P}\Big[\mathbf{U}(i) - \mathbf{U}(j) \geqslant -\mathbf{U}(j)\Big] \\
&\leqslant \exp\Big[-\Omega\Big(\frac{(n^{k-1}p^{\binom{k}{2}-4})^2}{(n^2p)\big(\sigma n^{k-5/2}\log^{\gamma_2} n\big)^2}\Big)\Big] \\
&= \exp\Big[-\Omega\Big(\frac{np^{2\binom{k}{2}-9}}{\sigma^2\log^{2\gamma_2} n}\Big)\Big].
\end{aligned}$$

Since there are at most n^2 possible values of j in (3.1) and $p \geqslant p_0$ in (3.13), by the union bound, we have

$$n^2\exp\Big[-\Omega\Big(\frac{np^{2\binom{k}{2}-9}}{\sigma^2\log^{2\gamma_2} n}\Big)\Big] = o(1).$$

Hence, *w.h.p.*, $\mathbf{Q}_k(i)$ never crosses its critical interval $\mathcal{I}^u_{\mathbf{Q}_k}$ in (4.1).

Remark 4.1. For the lower bound of $\mathbf{Q}_k(i)$ as the equation shown in (3.4), we work with the critical interval

$$\mathcal{I}^\ell_{\mathbf{Q}_k} = \mathcal{I}^\ell_{\mathbf{Q}_k}(i) = \Big(\frac{n^k}{k!}p^{\binom{k}{2}} - \sigma^2 n^\alpha p^{-1}\log^\mu n, \frac{n^k}{k!}p^{\binom{k}{2}} - \sigma(\sigma-1)n^\alpha p^{-1}\log^\mu n\Big),$$

where α is in (3.6). The proof is shown in the appendix for reference.

4.3 Tracking $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$

We prove the dynamic concentration of $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$ in this subsection. Fix one subset $U_{m^*} \in \binom{[n]}{m^*}$ for some m^* with $2 \leqslant m^* \leqslant k-1$. We start with the upper bound of $\mathbf{R}_{k,U_{m^*}}$. Our critical interval for the upper bound of $\mathbf{R}_{k,U_{m^*}}$ is

$$\begin{aligned}
\mathcal{I}^u_{\mathbf{R}_{k,U_{m^*}}} = \mathcal{I}^u_{\mathbf{R}_{k,U_{m^*}}}(i) = \Big(&\frac{n^{k-m^*}}{(k-m^*)!}p^{\binom{k}{2}-\binom{m^*}{2}} + (\sigma-1)n^{\beta_{m^*}}\log^{\gamma_{m^*}} n, \\
&\frac{n^{k-m^*}}{(k-m^*)!}p^{\binom{k}{2}-\binom{m^*}{2}} + \sigma n^{\beta_{m^*}}\log^{\gamma_{m^*}} n\Big),
\end{aligned} \tag{4.6}$$

where $\beta_{m^*} = k - m^* - \frac{\binom{k}{2}-\binom{m^*}{2}}{2\binom{k}{2}-2}$ in (3.7). Consider a fixed step $j < i_0$. Suppose $\mathbf{R}_{k,U_{m^*}}(j) \in \mathcal{I}^u_{\mathbf{R}_{k,U_{m^*}}}(j)$. Define

$$\tau^u_{\mathbf{R}_{k,U_{m^*}},j} = \min\Big\{i_0, \max\{j,\tau\}, \text{the smallest } i \geqslant j \;\text{ such that } \mathbf{R}_{k,U_{m^*}} \notin I^u_{\mathbf{R}_{k,U_{m^*}}}\Big\}.$$

Let $j \leqslant i < \tau^u_{\mathbf{R}_{k,U_{m^*}},j}$. Hence, all calculations are conditioned on the events that the estimates in (3.3) and (3.4) hold on $\mathbf{Q}_k(i)$, and the estimates in (3.5) hold on $\mathbf{R}_{k,U_m}$ for all $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$ and $U_m \neq U_{m^*}$.

Fix one set in $\mathcal{K}_{k-m^*} \cap \mathcal{N}_{U_{m^*}}$ in $\mathrm{G}(i)$, denoted by $U^c_{m^*}$. Let $\mathbf{Q}_{k,U_{m^*},U^c_{m^*}}(i)$ be the number of k-cliques in $\mathrm{G}(i)$ such that the removal of the edges in any one of these k-cliques results in $U^c_{m^*} \notin \mathcal{K}_{k-m^*} \cap \mathcal{N}_{U_{m^*}}$ in $\mathrm{G}(i+1)$. Then, we have

$$\mathbb{E}[\Delta \mathbf{R}_{k,U_{m^*}} | \mathcal{F}_i] = - \sum_{U^c_{m^*} \in \mathcal{K}_{k-m^*} \cap \mathcal{N}_{U_{m^*}}} \frac{\mathbf{Q}_{k,U_{m^*},U^c_{m^*}}(i)}{\mathbf{Q}_k(i)}. \tag{4.7}$$

Firstly, we consider the value of $\mathbf{Q}_{k,U_{m^*},U^c_{m^*}}(i)$. Let $H \subseteq U_{m^*} \cup U^c_{m^*}$ and $\mathbf{Q}^H_{k,U_{m^*},U^c_{m^*}}(i)$ be the number of k-cliques counted by $\mathbf{Q}_{k,U_{m^*},U^c_{m^*}}(i)$ such that each k-clique K_k satisfies $K_k \cap (U_{m^*} \cup U^c_{m^*}) = H$, namely, $\mathbf{Q}^H_{k,U_{m^*},U^c_{m^*}}(i)$ denotes the number of k-cliques counted by $\mathbf{Q}_{k,U_{m^*},U^c_{m^*}}(i)$ such that each k-clique exactly contains the vertices H in $U_{m^*} \cup U^c_{m^*}$.

Define $|H| = h$. To ensure that the removal of the edges in any one of these k-cliques results in $U^c_{m^*} \notin \mathcal{K}_{k-m^*} \cap \mathcal{N}_{U_{m^*}}$ in $\mathrm{G}(i+1)$, we have $H \cap U^c_{m^*} \neq \emptyset$ and $2 \leqslant h \leqslant k$. Choose $H \in \cup_{\rho=0}^{h-1} \binom{U_{m^*}}{\rho} \oplus \binom{U^c_{m^*}}{h-\rho}$, where $\binom{U_{m^*}}{\rho} \oplus \binom{U^c_{m^*}}{h-\rho}$ denotes the collection of sets consisting of the union of ρ vertices in U_{m^*} and $h-\rho$ vertices in $U^c_{m^*}$. Hence, $\mathbf{Q}_{k,U_{m^*},U^c_{m^*}}(i)$ is decomposed into

$$\mathbf{Q}_{k,U_{m^*},U^c_{m^*}}(i) = \sum_{h=2}^{k} \sum_{\rho=0}^{h-1} \sum_{H \in \binom{U_{m^*}}{\rho} \oplus \binom{U^c_{m^*}}{h-\rho}} \mathbf{Q}^H_{k,U_{m^*},U^c_{m^*}}(i). \tag{4.8}$$

Following the inclusion-exclusion counting technique in (2.4), we have

$$\begin{aligned}
\mathbf{Q}^H_{k,U_{m^*},U^c_{m^*}}(i) = \mathbf{1}_H \cdot \mathbf{R}_{k,H} - &\sum_{T_1 \in \binom{(U_{m^*} \cup U^c_{m^*}) \setminus H}{1}} \mathbf{1}_{H \cup T_1} \cdot \mathbf{R}_{k,H \cup T_1} + \dots \\
&+ \sum_{T_{k-h} \in \binom{(U_{m^*} \cup U^c_{m^*}) \setminus H}{k-h}} (-1)^{k-h} \mathbf{1}_{H \cup T_{k-h}} \cdot \mathbf{R}_{k,H \cup T_{k-h}} \\
= &\sum_{j=0}^{k-h} \sum_{T_j \in \binom{(U_{m^*} \cup U^c_{m^*}) \setminus H}{j}} (-1)^j \mathbf{1}_{H \cup T_j} \cdot \mathbf{R}_{k,H \cup T_j},
\end{aligned}$$

where $\mathbf{1}_{H \cup T_j}$ with $0 \leqslant j \leqslant k-h$ is the indicator random variable depending on whether the subgraph induced by $H \cup T_j$ in $\mathrm{G}(i)$ is complete or not. Combining with the equation in (4.8), we further have

$$\mathbf{Q}_{k,U_{m^*},U^c_{m^*}}(i) = \sum_{h=2}^{k} \sum_{j=0}^{k-h} \sum_{\rho=0}^{h-1} \sum_{H \in \binom{U_{m^*}}{\rho} \oplus \binom{U^c_{m^*}}{h-\rho}} \sum_{T_j \in \binom{(U_{m^*} \cup U^c_{m^*}) \setminus H}{j}} (-1)^j \mathbf{1}_{H \cup T_j} \cdot \mathbf{R}_{k,H \cup T_j}.$$

Note that

$$\begin{aligned}
&\sum_{\rho=0}^{h-1} \sum_{H \in \binom{U_{m^*}}{\rho} \oplus \binom{U^c_{m^*}}{h-\rho}} \sum_{T_j \in \binom{(U_{m^*} \cup U^c_{m^*}) \setminus H}{j}} (-1)^j \mathbf{1}_{H \cup T_j} \cdot \mathbf{R}_{k, H \cup T_j} \\
&= \sum_{\zeta=0}^{h+j} \sum_{H_{h+j} \in \binom{U_{m^*}}{\zeta} \oplus \binom{U^c_{m^*}}{h+j-\zeta}} \left[\binom{h+j}{h} - \binom{\zeta}{h} \right] (-1)^j \mathbf{1}_{H_{h+j}} \cdot \mathbf{R}_{k, H_{h+j}},
\end{aligned}$$

because each $H_{h+j} \in \binom{U_{m^*}}{\zeta} \oplus \binom{U^c_{m^*}}{h+j-\zeta}$ with $0 \leqslant \zeta \leqslant h+j$ on the right side is counted $[\binom{h+j}{h} - \binom{\zeta}{h}]$ times to be $H \cup T_j$ on the left side. It follows that

$$\mathbf{Q}_{k, U_{m^*}, U^c_{m^*}}(i) = \sum_{h=2}^{k} \sum_{j=0}^{k-h} \sum_{\zeta=0}^{h+j} \sum_{H_{h+j} \in \binom{U_{m^*}}{\zeta} \oplus \binom{U^c_{m^*}}{h+j-\zeta}} \left[\binom{h+j}{h} - \binom{\zeta}{h} \right] (-1)^j \mathbf{1}_{H_{h+j}} \cdot \mathbf{R}_{k, H_{h+j}}. \tag{4.9}$$

Thus, $\mathbf{Q}_{k, U_{m^*}, U^c_{m^*}}(i)$ is the sum of all elements in the $k \times k$ upper triangular matrix below

$$\begin{pmatrix}
\sum_{\zeta=0}^{2} \sum_{H_2 \in \binom{U_{m^*}}{\zeta} \oplus \binom{U^c_{m^*}}{2-\zeta}} (-1)^0 [\binom{2}{2} - \binom{\zeta}{2}] \mathbf{1}_{H_2} \cdot \mathbf{R}_{k, H_2} & \cdots & \sum_{\zeta=0}^{k} \sum_{H_k \in \binom{U_{m^*}}{\zeta} \oplus \binom{U^c_{m^*}}{k-\zeta}} (-1)^{k-2} [\binom{k}{2} - \binom{\zeta}{2}] \mathbf{1}_{H_k} \cdot \mathbf{R}_{k, H_k} \\
\vdots & \ddots & \vdots \\
\sum_{\zeta=0}^{k} \sum_{H_k \in \binom{U_{m^*}}{\zeta} \oplus \binom{U^c_{m^*}}{k-\zeta}} (-1)^0 [\binom{k}{k} - \binom{\zeta}{k}] \mathbf{1}_{H_k} \cdot \mathbf{R}_{k, H_k} & \cdots & 0
\end{pmatrix},$$

where the row corresponds to the index h and the column corresponds to the index j in (4.9), respectively. Summing these elements again according to all back diagonal lines, it follows that

$$\mathbf{Q}_{k, U_{m^*}, U^c_{m^*}}(i) = \sum_{h=2}^{k} \sum_{\zeta=0}^{h} \sum_{H_h \in \binom{U_{m^*}}{\zeta} \oplus \binom{U^c_{m^*}}{h-\zeta}} \sum_{s=2}^{h} (-1)^{h-s} \left[\binom{h}{s} - \binom{\zeta}{s} \right] \mathbf{1}_{H_h} \cdot \mathbf{R}_{k, H_h}, \tag{4.10}$$

where there is no $\mathbf{R}_{k, U_{m^*}}$ on the right side of (4.10) because $\mathbf{R}_{k, U_{m^*}}$ corresponds to the case when $\zeta = h$, which has coefficient zero.

Furthermore, according to the expressions of $\mathbf{R}_{k, H_h}$ for $2 \leqslant h \leqslant k-1$ in (3.5), the term $\mathbf{R}_{k, H_2}$ dominates the sum on the right side of (4.10). For $h = 2$, the only cases to consider are $(s, \zeta) = (2, 0)$, $(2, 1)$, and as the equation shown in (4.10), we have that $\mathbf{Q}_{k, U_{m^*}, U^c_{m^*}}(i)$ satisfies

$$\begin{aligned}
&\mathbf{Q}_{k, U_{m^*}, U^c_{m^*}}(i) \\
&> \left[\binom{k-m^*}{2} + m^*(k-m^*) \right] \Big(\frac{n^{k-2}}{(k-2)!} p^{\binom{k}{2}-1} - \sigma n^{\beta_2} \log^{\gamma_2} n \Big) + O\big(n^{k-3} p^{\binom{k}{2}-3}\big),
\end{aligned} \tag{4.11}$$

where $\mathbf{1}_{H_2} = 1$ always as $\zeta = |H_2 \cap U_{m^*}| \in \{0, 1\}$, $\binom{k-m^*}{2}$ enumerates $\mathbf{R}_{k,H_2}$ when $\zeta = 0$ and $s = 2$, $m^*(k-m^*)$ enumerates $\mathbf{R}_{k,H_2}$ when $\zeta = 1$ and $s = 2$, and the last term $O(n^{k-3}p^{\binom{k}{2}-3})$ comes from those terms when $h \geqslant 3$ in (4.10). Note that $\binom{k-m^*}{2} + m^*(k-m^*) = \binom{k}{2} - \binom{m^*}{2}$ and $\beta_2 = k - \frac{5}{2}$ in (3.7), combining the equations in (4.7) and (4.11), and applying the upper bound on $\mathbf{Q}_k(i)$ from (3.3) for $k \geqslant 4$, we have

$$\mathbb{E}[\Delta\mathbf{R}_{k,U_{m^*}}|\mathcal{F}_i] < -\sum_{U^c_{m^*} \in \mathcal{K}_{k-m^*} \cap \mathcal{N}_{U_{m^*}}} \frac{\left[\binom{k}{2} - \binom{m^*}{2}\right]\left(\frac{n^{k-2}}{(k-2)!}p^{\binom{k}{2}-1} - \sigma n^{k-5/2}\log^{\gamma_2} n\right) + O\left(n^{k-3}p^{\binom{k}{2}-3}\right)}{\frac{n^k}{k!}p^{\binom{k}{2}} + \frac{n^{k-1}}{2}p^{\binom{k}{2}-4}}. \tag{4.12}$$

Note that $\frac{n^k}{k!}p^{\binom{k}{2}} + \frac{n^{k-1}}{2}p^{\binom{k}{2}-4} = \frac{n^k}{k!}p^{\binom{k}{2}}[1 + O(\frac{1}{np^4})]$ and $\frac{1}{np^4} = o(1)$ when $p \geqslant p_0$ in (3.13) and $k \geqslant 4$. It is also known that the number of ways to choose $U^c_{m^*} \in \mathcal{K}_{k-m^*} \cap \mathcal{N}_{U_{m^*}}$ is $\mathbf{R}_{k,U_{m^*}}$, and $\mathbf{R}_{k,U_{m^*}} \in \mathcal{I}^u_{\mathbf{R}_{k,U_{m^*}}}$ in (4.6). Thus, it further follows from (4.12) that

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{R}_{k,U_{m^*}}|\mathcal{F}_i] &< \\ &-\frac{\left[\binom{k}{2}-\binom{m^*}{2}\right]\left(\frac{n^{k-m^*}}{(k-m^*)!}p^{\binom{k}{2}-\binom{m^*}{2}} + (\sigma-1)n^{\beta_{m^*}}\log^{\gamma_{m^*}} n\right)\left(\frac{n^{k-2}}{(k-2)!}p^{\binom{k}{2}-1} - \sigma n^{k-\frac{5}{2}}\log^{\gamma_2} n\right)}{\frac{n^k}{k!}p^{\binom{k}{2}}} \\ &+ O\left(n^{k-m^*-3}p^{\binom{k}{2}-\binom{m^*}{2}-5}\right).\end{aligned}$$

When $k = 3$ and $m^* = 2$, the denominator in (4.12) must be replaced by $\frac{1}{6}n^3p^3 + \frac{1}{3}n^2p$ in (3.3). Then the above argument holds with the final error term replaced by $O(n^{k-m^*-3}p^{\binom{k}{2}-\binom{m^*}{2}-3}) = o(1)$. Now assume that $k \geqslant 3$. Rearranging the above equation, we get

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{R}_{k,U_{m^*}}|\mathcal{F}_i] < &-\frac{\left[\binom{k}{2} - \binom{m^*}{2}\right]k(k-1)n^{k-m^*-2}}{(k-m^*)!}p^{\binom{k}{2}-\binom{m^*}{2}-1} \\ &+ \frac{\left[\binom{k}{2} - \binom{m^*}{2}\right]k!\sigma n^{k-m^*-\frac{5}{2}}\log^{\gamma_2} n}{(k-m^*)!p^{\binom{m^*}{2}}} \\ &- \frac{\left[\binom{k}{2} - \binom{m^*}{2}\right]k(k-1)(\sigma-1)}{p}n^{\beta_{m^*}-2}\log^{\gamma_{m^*}} n \\ &+ \frac{\left[\binom{k}{2} - \binom{m^*}{2}\right]k!\sigma(\sigma-1)}{p^{\binom{k}{2}}}n^{\beta_{m^*}-\frac{5}{2}}\log^{\gamma_2+\gamma_{m^*}} n \\ &+ O\left(n^{k-m^*-3}p^{\binom{k}{2}-\binom{m^*}{2}-5}\right).\end{aligned} \tag{4.13}$$

Since $\mathbf{R}_{k,U_{m^*}}(j) \in \mathcal{I}^u_{\mathbf{R}_{k,U_{m^*}}}(j)$ for $j < i$, there exists a constant $\delta \in (0,1)$ such that

$$\mathbf{R}_{k,U_{m^*}}(j) \sim \frac{n^{k-m^*}}{(k-m^*)!}p_j^{\binom{k}{2}-\binom{m^*}{2}} - (\sigma_j - \delta)n^{\beta_{m^*}}\log^{\gamma_{m^*}} n. \tag{4.14}$$

For all i with $j \leqslant i < \tau^u_{\mathbf{R}_{k,U_{m^*}},j}$, define the sequence of random variables below as

$$\mathbf{Z}_{U_{m^*}}(i) = \mathbf{R}_{k,U_{m^*}}(i) - \frac{n^{k-m^*}}{(k-m^*)!} p_i^{\binom{k}{2}-\binom{m^*}{2}} - (\sigma_i - \delta) n^{\beta_{m^*}} \log^{\gamma_{m^*}} n. \tag{4.15}$$

Claim 4.2: Removing the edges of one k-clique in $\mathrm{G}(i)$, we have

$$\mathbf{R}_{k,U_{m^*}}(i) - \mathbf{R}_{k,U_{m^*}}(i+1) < 2k^2 n^{k-m^*-1} p^{\binom{k}{2}-\binom{m^*+1}{2}}.$$

Proof of Claim 4.2. According to the definition in (2.1), when we remove the edges of one k-clique from $\mathrm{G}(i)$, there are two cases to consider. One case is to assume that the removed k-clique contains some $u \in U_{m^*}$ and some $w \in \mathcal{N}_{U_{m^*}}$; and the other case is that the removed k-clique contains two distinct elements w_1, w_2 of $\mathcal{N}_{U_{m^*}}$. In the first case, there are $m^*(k-m^*) \leqslant k^2$ choices of (u, w), and the number of $(k-m^*-1)$-cliques such that every vertex is in $\mathcal{N}_{U_{m^*}\cup\{w\}}$ is at most $\mathbf{R}_{k,U_{m^*}\cup\{w\}}$. In the second case, there are $\binom{k-m^*}{2} \leqslant k^2$ choices of (w_1, w_2), and the number of $(k-m^*-2)$-cliques such that every vertex is in $\mathcal{N}_{U_{m^*}\cup\{w_1,w_2\}}$ is at most $\mathbf{R}_{k,U_{m^*}\cup\{w_1,w_2\}}$.

As the equation shown in (3.5), we have

$$\begin{aligned} \mathbf{R}_{k,U_{m^*}\cup\{w_1,w_2\}} &\sim \frac{n^{k-m^*-2}}{(k-m^*-2)!} p^{\binom{k}{2}-\binom{m^*+2}{2}}, \\ \mathbf{R}_{k,U_{m^*}\cup\{w\}} &\sim \frac{n^{k-m^*-1}}{(k-m^*-1)!} p^{\binom{k}{2}-\binom{m^*+1}{2}} \end{aligned}$$

because the estimates in (3.5) hold on $\mathbf{R}_{k,U_m}$ for all $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$ and $U_m \neq U_{m^*}$. Thus, $\mathbf{R}_{k,U_{m^*}\cup\{w_1,w_2\}} = o(\mathbf{R}_{k,U_{m^*}\cup\{w\}})$ when $p \geqslant p_0$ in (3.13) and $k \geqslant 3$. Summing the values in these two cases, we complete the proof of Claim 4.2. $\square$

Claim 4.3: $\mathbf{Z}_{U_{m^*}}(j) = 0$ and the sequence $-\mathbf{Z}_{U_{m^*}}(j), -\mathbf{Z}_{U_{m^*}}(j+1), \ldots, -\mathbf{Z}_{U_{m^*}}(\tau^u_{\mathbf{R}_{k,U_{m^*}},j})$ is an (η, N)-bounded submartingale, where

$$\begin{aligned} \eta &= 2\left[\binom{k}{2} - \binom{m^*}{2}\right] \frac{k(k-1)n^{k-m^*-2}}{(k-m^*)!} p^{\binom{k}{2}-\binom{m^*}{2}-1}, \\ N &= 3k^2 n^{k-m^*-1} p^{\binom{k}{2}-\binom{m^*+1}{2}} \end{aligned}$$

with $\eta = o(N)$ for $2 \leqslant m^* \leqslant k-1$.

Proof of Claim 4.3. As the equations shown in (4.14) and (4.15), we have $\mathbf{Z}_{U_{m^*}}(j) = 0$. For all i with $j \leqslant i < \tau^u_{\mathbf{R}_{k,U_{m^*}},j}$, we have

$$\begin{aligned} \mathbb{E}[\Delta \mathbf{Z}_{U_{m^*}} | \mathcal{F}_i] = \mathbb{E}[\Delta \mathbf{R}_{k,U_{m^*}} | \mathcal{F}_i] &- \frac{n^{k-m^*}}{(k-m^*)!}\left[p_{i+1}^{\binom{k}{2}-\binom{m^*}{2}} - p_i^{\binom{k}{2}-\binom{m^*}{2}}\right] \\ &- n^{\beta_{m^*}} \log^{\gamma_{m^*}} n \big[\sigma_{i+1} - \sigma_i\big] \end{aligned}$$

by the equation in (4.15). Note that $p = p_i = 1 - k(k-1)t$, $p_{i+1} = p - \frac{k(k-1)}{n^2}$ in (3.1), then $\sigma = \sigma_i = 1 - k(k-1)\log p$ and $\sigma_{i+1} = 1 - k(k-1)\log(p - \frac{k(k-1)}{n^2})$ in (3.8). Differentiating σ with respect to t gives

$$\sigma' = \frac{k^2(k-1)^2}{p}, \quad \sigma'' = \frac{k^3(k-1)^3}{p^2} = O(p^{-2}). \tag{4.16}$$

Expanding $\sigma_{i+1} = \sigma(t + \frac{1}{n^2})$ around t gives $\sigma_{i+1} = \sigma + \frac{\sigma'}{n^2} + O(\frac{\sigma''}{n^4})$. Thus, it follows that

$$\begin{aligned}
\mathbb{E}[\Delta \mathbf{Z}_{U_{m^*}}|\mathcal{F}_i] &= \mathbb{E}[\Delta \mathbf{R}_{k,U_{m^*}}|\mathcal{F}_i] - \frac{n^{k-m^*}}{(k-m^*)!}\Big[-\Big(\binom{k}{2} - \binom{m^*}{2}\Big)\frac{k(k-1)}{n^2}p^{\binom{k}{2}-\binom{m^*}{2}-1} \\
&\quad + O\Big(\frac{1}{n^4}p^{\binom{k}{2}-\binom{m^*}{2}-2}\Big)\Big] - n^{\beta_{m^*}}\log^{\gamma_{m^*}} n\Big[\frac{\sigma'}{n^2} + O\Big(\frac{\sigma''}{n^4}\Big)\Big] \\
&= \mathbb{E}[\Delta \mathbf{R}_{k,U_{m^*}}|\mathcal{F}_i] + \Big[\binom{k}{2} - \binom{m^*}{2}\Big]\frac{k(k-1)n^{k-m^*-2}}{(k-m^*)!}p^{\binom{k}{2}-\binom{m^*}{2}-1} \\
&\quad - \sigma' n^{\beta_{m^*}-2}\log^{\gamma_{m^*}} n + O\big(n^{k-m^*-4}p^{\binom{k}{2}-\binom{m^*}{2}-2}\big),
\end{aligned} \tag{4.17}$$

where the term $O(\sigma'' n^{\beta_{m^*}-4}\log^{\gamma_{m^*}} n)$ is absorbed into the term $O(n^{k-m^*-4}p^{\binom{k}{2}-\binom{m^*}{2}-2})$ because $\sigma'' = O(p^{-2})$ in (4.16), β_{m^*} in (3.7), $p \geqslant p_0$ in (3.13) with λ and γ_{m^*} satisfies $[\binom{k}{2} - \binom{m^*}{2}]\lambda > \gamma_{m^*} + 2$ shown in (3.10). Combining the equations in (4.13) and (4.17), we further have

$$\begin{aligned}
\mathbb{E}[\Delta \mathbf{Z}_{U_{m^*}}|\mathcal{F}_i] &< \frac{\big[\binom{k}{2} - \binom{m^*}{2}\big]k!\sigma}{(k-m^*)!p^{\binom{m^*}{2}}}n^{k-m^*-\frac{5}{2}}\log^{\gamma_2} n \\
&\quad - \frac{\big[\binom{k}{2} - \binom{m^*}{2}\big]k(k-1)\sigma}{p}n^{\beta_{m^*}-2}\log^{\gamma_{m^*}} n \\
&\quad + \frac{\big[\binom{k}{2} - \binom{m^*}{2}\big]k(k-1)}{p}n^{\beta_{m^*}-2}\log^{\gamma_{m^*}} n \\
&\quad + \frac{\big[\binom{k}{2} - \binom{m^*}{2}\big]k!\sigma(\sigma-1)}{p^{\binom{k}{2}}}n^{\beta_{m^*}-\frac{5}{2}}\log^{\gamma_2+\gamma_{m^*}} n \\
&\quad - \sigma' n^{\beta_{m^*}-2}\log^{\gamma_{m^*}} n + O\big(n^{k-m^*-3}p^{\binom{k}{2}-\binom{m^*}{2}-5}\big),
\end{aligned} \tag{4.18}$$

where the term $O(n^{k-m^*-4}p^{\binom{k}{2}-\binom{m^*}{2}-2})$ in (4.17) is absorbed into the term $O(n^{k-m^*-3}p^{\binom{k}{2}-\binom{m^*}{2}-5})$ in (4.13). At last, by (4.18), we have $\mathbb{E}[\Delta \mathbf{Z}_{U_{m^*}}|\mathcal{F}_i] < 0$

because the following inequalities

$$\frac{k!\sigma}{(k-m^*)!p^{\binom{m^*}{2}}} n^{k-m^*-\frac{5}{2}} \log^{\gamma_2} n \leqslant \frac{k(k-1)\sigma}{p} n^{\beta_{m^*}-2} \log^{\gamma_{m^*}} n, \tag{4.19}$$

$$\frac{\left[\binom{k}{2}-\binom{m^*}{2}\right]k(k-1)}{p} n^{\beta_{m^*}-2} \log^{\gamma_{m^*}} n < \frac{\sigma'}{2} n^{\beta_{m^*}-2} \log^{\gamma_{m^*}} n, \tag{4.20}$$

$$\frac{\left[\binom{k}{2}-\binom{m^*}{2}\right]k!\sigma(\sigma-1)}{p^{\binom{k}{2}}} n^{\beta_{m^*}-\frac{5}{2}} \log^{\gamma_2+\gamma_{m^*}} n < \frac{\sigma'}{4} n^{\beta_{m^*}-2} \log^{\gamma_{m^*}} n, \tag{4.21}$$

$$O\big(n^{k-m^*-3} p^{\binom{k}{2}-\binom{m^*}{2}-5}\big) < \frac{\sigma'}{4} n^{\beta_{m^*}-2} \log^{\gamma_{m^*}} n \tag{4.22}$$

are true for β_{m^*} in (3.7), $p \geqslant p_0$ in (3.13), $k \geqslant 3$ and $2 \leqslant m^* \leqslant k-1$. The first one (4.19) is clearly true when $m^* = 2$, and is also true if we choose $[\binom{m^*}{2}-1]\lambda+\gamma_{m^*} > \gamma_2$ when $m^* \geqslant 3$, which is assumed in (3.11); the second one (4.20) is always true because $\sigma' = \frac{k^2(k-1)^2}{p}$ in (4.16); the third one (4.21) is satisfied because $\sigma = O(\log n)$ in (3.8) and $[\binom{k}{2}-1]\lambda > \gamma_2 + 2$ assumed in (3.10) for $m = 2$; and the last one (4.22) is always true for $k \geqslant 3$. We have proved that the sequence $-\mathbf{Z}_{U_{m^*}}(j), -\mathbf{Z}_{U_{m^*}}(j+1), \ldots, -\mathbf{Z}_{U_{m^*}}(\tau^u_{\mathbf{R}_{k,U_{m^*}},j})$ is a submartingale for any $2 \leqslant m^* \leqslant k-1$.

In the following, we show the sequence is (η, N)-bounded. By the equation in (4.15) and the calculation in (4.17), we have

$$\begin{aligned}
&-\mathbf{Z}_{U_{m^*}}(i+1) + \mathbf{Z}_{U_{m^*}}(i) \\
&= \mathbf{R}_{k,U_{m^*}}(i) - \mathbf{R}_{k,U_{m^*}}(i+1) + \frac{n^{k-m^*}}{(k-m^*)!}\Big[p^{\binom{k}{2}-\binom{m^*}{2}}(i+1) - p^{\binom{k}{2}-\binom{m^*}{2}}(i)\Big] \\
&\quad + n^{\beta_{m^*}} \log^{\gamma_{m^*}} n\big[\sigma(i+1)-\sigma(i)\big] \\
&= \mathbf{R}_{k,U_{m^*}}(i) - \mathbf{R}_{k,U_{m^*}}(i+1) - \left[\binom{k}{2}-\binom{m^*}{2}\right]\frac{k(k-1)n^{k-m^*-2}}{(k-m^*)!} p^{\binom{k}{2}-\binom{m^*}{2}-1} \\
&\quad + \sigma' n^{\beta_{m^*}-2} \log^{\gamma_{m^*}} n + O\big(n^{k-m^*-4} p^{\binom{k}{2}-\binom{m^*}{2}-2}\big).
\end{aligned}$$

Then, we have

$$\begin{aligned}
&-\mathbf{Z}_{U_{m^*}}(i+1) + \mathbf{Z}_{U_{m^*}}(i) \\
&\geqslant -\left[\binom{k}{2}-\binom{m^*}{2}\right]\frac{k(k-1)n^{k-m^*-2}}{(k-m^*)!} p^{\binom{k}{2}-\binom{m^*}{2}-1} + O\big(n^{k-m^*-4} p^{\binom{k}{2}-\binom{m^*}{2}-2}\big) \\
&> -2\left[\binom{k}{2}-\binom{m^*}{2}\right]\frac{k(k-1)n^{k-m^*-2}}{(k-m^*)!} p^{\binom{k}{2}-\binom{m^*}{2}-1},
\end{aligned}$$

and we choose

$$\eta = 2\left[\binom{k}{2}-\binom{m^*}{2}\right]\frac{k(k-1)n^{k-m^*-2}}{(k-m^*)!} p^{\binom{k}{2}-\binom{m^*}{2}-1}.$$

On the other hand, since

$$\begin{aligned}
&-\mathbf{Z}_{U_{m^*}}(i+1) + \mathbf{Z}_{U_{m^*}}(i) \\
&\leqslant \mathbf{R}_{k,U_{m^*}}(i) - \mathbf{R}_{k,U_{m^*}}(i+1) + \sigma' n^{\beta_{m^*}-2} \log^{\gamma_{m^*}} n + O\big(n^{k-m^*-4} p^{\binom{k}{2}-\binom{m^*}{2}-2}\big),
\end{aligned}$$

applying Claim 4.2, note that

$$\sigma' n^{\beta_{m^*}-2} \log^{\gamma_{m^*}} n + O(n^{k-m^*-4} p^{\binom{k}{2}-\binom{m^*}{2}-2}) = o(n^{k-m^*-1} p^{\binom{k}{2}-\binom{m^*+1}{2}});$$

thus we choose

$$N = 3k^2 n^{k-m^*-1} p^{\binom{k}{2}-\binom{m^*+1}{2}}.$$

This completes the proof of Claim 4.3. $\square$

Take $a = n^{\beta_{m^*}} \log^{\gamma_{m^*}} n$ and $\ell = O(n^2 p)$, then $a = o(\eta\ell)$ when $p \geqslant p_0$ and taking λ, γ_{m^*} to satisfy $[\binom{k}{2} - \binom{m^*}{2}]\lambda > \gamma_{m^*} + 2$ assumed in (3.10). Combining with Claim 4.3, the assumptions of Lemma 2.3 hold. Applying Lemma 2.3 to $-\mathbf{Z}_{U_{m^*}}(j), -\mathbf{Z}_{U_{m^*}}(j+1), \ldots, -\mathbf{Z}_{U_{m^*}}(\tau^u_{\mathbf{R}_{k,U_{m^*}},j})$ yields that

$$\begin{aligned}
&\mathbb{P}\Big[\mathbf{R}_{k,U_{m^*}} \geqslant \frac{n^{k-m^*}}{(k-m^*)!} p^{\binom{k}{2}-\binom{m^*}{2}} + \sigma n^{\beta_{m^*}} \log^{\gamma_{m^*}} n\Big] \\
&= \mathbb{P}\Big[-\mathbf{Z}_{U_{m^*}}(i) \leqslant -n^{\beta_{m^*}} \log^{\gamma_{m^*}} n\Big] \\
&\leqslant \exp\Big[-\Omega\Big(\frac{n^{2\beta_{m^*}} \log^{2\gamma_{m^*}} n}{n^{k-m^*} \cdot n^{k-m^*-1}}\Big)\Big] \\
&= \exp\Big[-\Omega\Big(n^{\frac{\binom{m^*}{2}-1}{\binom{k}{2}-1}} \log^{2\gamma_{m^*}} n\Big)\Big].
\end{aligned}$$

For each $m^* = 2, \ldots, k-1$, note that the number of ways to choose j and $U_{m^*} \in \binom{[n]}{m^*}$ is at most n^{m^*+2}, then we also have

$$n^{m^*+2} \exp\Big[-\Omega\Big(n^{\frac{\binom{m^*}{2}-1}{\binom{k}{2}-1}} \log^{2\gamma_{m^*}} n\Big)\Big] = o(1),$$

which is clearly true when $3 \leqslant m^* \leqslant k-1$, or $\gamma_2 > \frac{1}{2}$ shown in (3.12) when $m^* = 2$. Finally taking the union bound over m^* when $2 \leqslant m^* \leqslant k-1$ completes the proof. In conclusion, *w.h.p.*, none $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{n}{m}$ with $2 \leqslant m \leqslant k-1$ crosses its critical interval $\mathcal{I}^u_{\mathbf{R}_{k,U_m}}$ defined in (4.6).

Remark 4.2. The argument for the lower bound of $\mathbf{R}_{k,U_m}$ in (3.5) for any $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$ is the symmetric analogue of the above analysis.

5 $|E(M)| \leqslant n^{2-1/(k(k-1)-2)+o(1)}$ is a natural barrier

The bound in Theorem 1.1 for $k=3$ is $|E(M)| \leqslant n^{7/4+o(1)}$, which is same with the one in [5, 9]. There is a gap between the upper bound $|E(M)| \leqslant n^{2-1/(k(k-1)-2)+o(1)}$ and the conjectured upper bound $|E(M)| \leqslant n^{2k/(k+1)+o(1)}$ from [2]. We will show that the expression $n^{2-1/(k(k-1)-2)}$ corresponds to a natural barrier of the process in our proof.

To illustrate this, as stated in Theorem 3.1, the process $\mathrm{G}(i)$ resembles $\mathcal{G}(n,p)$ when $i \leqslant i_0$ and $p = 1 - \frac{k(k-1)i}{n^2}$; the standard variation $\sigma n^{\beta_m} \log^{\gamma_m} n$ of $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$ would be as large as their main trajectory $\frac{n^{k-m}}{(k-m)!} p^{\binom{k}{2}-\binom{m}{2}}$ when p is around p_0 in (3.13) (up to logarithmic factors). At that time i_0, the power of $n^{k-m} {p_0}^{\binom{k}{2}-\binom{m}{2}}$ is exactly equal to β_m in (3.7), which means that the control over $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ is lost up to logarithmic factors. The main obstacle to improve our main results is that the behaviors of the parameters $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$ do not allow us to further analyze the process, while the following means that it is definitely possible to further improve these results.

In fact, we apply Lemma 2.4 at the critical point i_0 to $\mathbf{R}_{k,U_m}(i_0)$ with $\xi = \sigma n^{\beta_m} \log^{\gamma_m} n$, where $\beta_m = k - m - \frac{\binom{k}{2}-\binom{m}{2}}{k(k-1)-2}$ in (3.7). It follows that

$$
\begin{aligned}
&\sum_{U_m \in \binom{[n]}{m}, 2 \leqslant m \leqslant k-1} \mathbb{P}\left[\left|\mathbf{R}_{k,U_m}(i_0) - \frac{n^{k-m}}{(k-m)!} {p_0}^{\binom{k}{2}-\binom{m}{2}}\right| > \sigma n^{\beta_m} \log^{\gamma_m} n\right] \\
&< 2 \sum_{m=2}^{k-1} \binom{n}{m} \exp\left[-\frac{(k-m)!(\sigma n^{\beta_m} \log^{\gamma_m} n)^2}{3n^{k-m} {p_0}^{\binom{k}{2}-\binom{m}{2}}}\right] \\
&= \sum_{m=2}^{k-1} \binom{n}{m} \exp\left[-O\left(n^{k-m-\frac{\binom{k}{2}-\binom{m}{2}}{k(k-1)-2}+o(1)}\right)\right] \\
&= O\left(n^{k-1} \exp\left[-n^{1-\frac{\binom{k}{2}-\binom{k-1}{2}}{k(k-1)-2}+o(1)}\right]\right) \\
&= o(1),
\end{aligned}
$$

where $f(m) = k - m - \frac{k(k-1)-m(m-1)}{2k(k-1)-4}$ is decreasing in m and the series is dominated by the term $m = k-1$. Thus, the probability of the event that there exists one $U_m \in \binom{[n]}{m}$ with $2 \leqslant m \leqslant k-1$ such that the control over $\mathbf{R}_{k,U_m}$ loses when $p = p_0$ is very low. In order to obtain better results on $|E(M)|$, as demonstrated in [6], it is necessary to design new random variables such that their variations decrease as the process evolves. This is not easy and we will consider this problem in future work.

Appendix: Lower bound of $\mathbf{Q}_k(i)$ (for Remark 4.1)

For the lower bound of $\mathbf{Q}_k(i)$, we work with the critical interval

$$
\mathcal{I}^{\ell}_{\mathbf{Q}_k} = \mathcal{I}^{\ell}_{\mathbf{Q}_k}(i) = \left(\frac{n^k}{k!} p^{\binom{k}{2}} - \sigma^2 n^{\alpha} p^{-1} \log^{\mu} n, \frac{n^k}{k!} p^{\binom{k}{2}} - \sigma(\sigma-1) n^{\alpha} p^{-1} \log^{\mu} n\right), \quad (1)
$$

where α is shown in (3.6). Consider a fixed step $j < i_0$. Similarly, suppose $\mathbf{Q}_k(j) \in \mathcal{I}^{\ell}_{\mathbf{Q}_k}(j)$ and define

$$\tau^{\ell}_{\mathbf{Q}_k,j} = \min\big\{i_0, \max\{j,\tau\}, \text{the smallest } i \geqslant j \text{ such that } \mathbf{Q}_k(i) \notin \mathcal{I}^{\ell}_{\mathbf{Q}_k}\big\}.$$

Let $j \leqslant i < \tau^{\ell}_{\mathbf{Q}_k,j}$. All calculations in this subsection are conditioned on the estimates in (3.5).

By the equations shown in (2.5), we get the estimate on $\mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i]$ in reverse direction,

$$\begin{aligned}
\mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] &= (-1)^{k+1}(k-1) \\
&\quad - \frac{1}{\mathbf{Q}_k(i)}\sum_{U_2\in\mathcal{K}_2(i)}(\mathbf{R}_{k,U_2})^2 + \cdots + \frac{(-1)^k(k-2)}{\mathbf{Q}_k(i)}\sum_{U_{k-1}\in\mathcal{K}_{k-1}(i)}(\mathbf{R}_{k,U_{k-1}})^2 \\
&> (-1)^{k+1}(k-1) - \frac{1}{\mathbf{Q}_k(i)}\Big(\frac{2!\binom{k}{2}^2\mathbf{Q}_k^2(i)}{n^2p} + 2\sigma^2 n^{2k-3}p\log^{2\gamma_2} n\Big) \\
&\quad + \frac{12\binom{k}{3}^2\mathbf{Q}_k(i)}{n^3p} + O(n^{k-4}p^{\binom{k}{2}-11}) \\
&= (-1)^{k+1}(k-1) - \frac{2\binom{k}{2}^2\mathbf{Q}_k(i)}{n^2p} + \frac{12\binom{k}{3}^2\mathbf{Q}_k(i)}{n^3p} + O(\sigma^2 n^{k-3}p^{-\binom{k}{2}+1}\log^{2\gamma_2} n),
\end{aligned}$$

where $\sum_{U_2\in\mathcal{K}_2(i)}(\mathbf{R}_{k,U_2})^2$ and $\sum_{U_3\in\mathcal{K}_3(i)}(\mathbf{R}_{k,U_3})^2$ are replaced by the equations in (3.14) and (3.15), the term $O(n^{k-4}p^{\binom{k}{2}-11})$ comes from $\sum_{U_4\in\mathcal{K}_4(i)}(\mathbf{R}_{k,U_4})^2$ in (3.15) that dominates all the remaining terms.

Since $\mathbf{Q}_k(i) \in \mathcal{I}^{\ell}_{\mathbf{Q}_k}$ as the equation shown in (1), we further have

$$\begin{aligned}
\mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] &> (-1)^{k+1}(k-1) - \frac{2\binom{k}{2}^2\big(\frac{n^k}{k!}p^{\binom{k}{2}} - \sigma(\sigma-1)n^{\alpha}p^{-1}\log^{\mu} n\big)}{n^2p} \\
&\quad + \frac{12\binom{k}{3}^2\big(\frac{n^k}{k!}p^{\binom{k}{2}} - \sigma^2 n^{\alpha}p^{-1}\log^{\mu} n\big)}{n^3p} + O\big(n^{k-3}\sigma^2 p^{-\binom{k}{2}+1}\log^{2\gamma_2} n\big) \\
&= (-1)^{k+1}(k-1) - \frac{2\binom{k}{2}^2 n^{k-2}}{k!}p^{\binom{k}{2}-1} + \frac{2\binom{k}{2}^2\sigma(\sigma-1)n^{\alpha-2}\log^{\mu} n}{p^2} \\
&\quad + \frac{12\binom{k}{3}^2 n^{k-3}}{k!}p^{\binom{k}{2}-1} + O\big(\sigma^2 n^{k-3}p^{-\binom{k}{2}+1}\log^{2\gamma_2} n\big),
\end{aligned} \tag{2}$$

where α is in (3.6), and the term $O(\sigma^2 n^{\alpha-3}p^{-2}\log^{\mu} n)$ is absorbed into the term $O(\sigma^2 n^{k-3}p^{-\binom{k}{2}+1}\log^{2\gamma_2} n)$.

For all i with $j \leqslant i < \tau^{\ell}_{\mathbf{Q}_k,j}$, define the sequence of random variables to be

$$\mathbf{L}(i) = \mathbf{Q}_k(i) - \frac{n^k}{k!}p^{\binom{k}{2}} + \sigma^2 n^{\alpha}p^{-1}\log^{\mu} n. \tag{3}$$

Claim: The sequence $\mathbf{L}(j), \mathbf{L}(j+1), \ldots, \mathbf{L}(\tau^{\ell}_{\mathbf{Q}_k,j})$ is a submartingale and the maximum one step $\Delta\mathbf{L}$ is $O(\sigma n^{k-\frac{5}{2}}\log^{\gamma_2} n)$.

Proof of Claim. Similarly, for all i with $j \leqslant i < \tau^{\ell}_{\mathbf{Q}_k,j}$, as the equation shown in (3), we have

$$\mathbb{E}[\Delta\mathbf{L}|\mathcal{F}_i] = \mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] - \frac{n^k}{k!}\Big[p_{i+1}^{\binom{k}{2}} - p_i^{\binom{k}{2}}\Big] + n^{\alpha}\log^{\mu} n\Big[\frac{\sigma_{i+1}^2}{p_{i+1}} - \frac{\sigma_i^2}{p_i}\Big].$$

Since $p = p_i = 1-k(k-1)t$, $p_{i+1} = p-\frac{k(k-1)}{n^2}$ in (3.1), then $\sigma = \sigma_i = 1-k(k-1)\log p$, $\sigma_{i+1} = 1-k(k-1)\log(p-\frac{k(k-1)}{n^2})$ in (3.8), by Taylor's expansion, we have

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{L}|\mathcal{F}_i] &= \mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] - \frac{n^k}{k!}\Big[-\binom{k}{2}\frac{k(k-1)}{n^2}p^{\binom{k}{2}-1} + O\Big(\frac{1}{n^4}p^{\binom{k}{2}-2}\Big)\Big]\\ &\quad + n^{\alpha}\log^{\mu} n\Big[\frac{2\sigma\sigma' p - \sigma^2 p'}{n^2p^2} + O\Big(\frac{\sigma^2}{n^4p^3}\Big)\Big]\\ &= \mathbb{E}[\Delta\mathbf{Q}_k|\mathcal{F}_i] + \frac{2\binom{k}{2}^2 n^{k-2}}{k!}p^{\binom{k}{2}-1} + \frac{2\sigma\sigma' n^{\alpha-2}\log^{\mu} n}{p}\\ &\quad + \frac{k(k-1)\sigma^2n^{\alpha-2}\log^{\mu} n}{p^2} + O\big(n^{k-4}p^{\binom{k}{2}-2}\big),\end{aligned} \tag{4}$$

where the term $O(n^{\alpha-4}\sigma^2p^{-3}\log^{\mu} n)$ is absorbed into the term $O(n^{k-4}p^{\binom{k}{2}-2})$ because α is shown in (3.6) and $p \geqslant p_0$ is shown in (3.13) by taking λ and μ to satisfy $[\binom{k}{2}+1]\lambda > \mu+3$ in (3.9). Combining the equations in (2) and (4), we have

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{L}|\mathcal{F}_i] &> (-1)^{k+1}(k-1) + \frac{\big[2\binom{k}{2}^2 + k(k-1)\big]\sigma^2n^{\alpha-2}\log^{\mu} n}{p^2} - \frac{2\binom{k}{2}^2\sigma n^{\alpha-2}\log^{\mu} n}{p^2}\\ &\quad + \frac{12\binom{k}{3}^2 n^{k-3}}{k!}p^{\binom{k}{2}-1} + \frac{2\sigma\sigma' n^{\alpha-2}\log^{\mu} n}{p} + O\big(\sigma^2 n^{k-3}p^{-\binom{k}{2}+1}\log^{2\gamma_2} n\big),\end{aligned}$$

where the term $O(n^{k-4}p^{\binom{k}{2}-2})$ in (4) is absorbed into the term $O(\sigma^2n^{k-3}p^{-\binom{k}{2}+1}\log^{2\gamma_2} n)$ in (2). We have

$$2\sigma\sigma' n^{\alpha-2}\log^{\mu} np^{-1} = 8\binom{k}{2}^2\sigma n^{\alpha-2}p^{-2}\log^{\mu} n$$

by $\sigma' = k^2(k-1)^2p^{-1}$ in (3.8). It follows that

$$\begin{aligned}\mathbb{E}[\Delta\mathbf{L}|\mathcal{F}_i] &> (-1)^{k+1}(k-1) + \frac{\big[2\binom{k}{2}^2 + k(k-1)\big]\sigma^2n^{\alpha-2}\log^{\mu} n}{p^2}\\ &\quad + 6\binom{k}{2}^2\sigma n^{\alpha-2}p^{-2}\log^{\mu} n\\ &\quad + \frac{12\binom{k}{3}^2 n^{k-3}}{k!}p^{\binom{k}{2}-1} + O\big(\sigma^2 n^{k-3}p^{-\binom{k}{2}+1}\log^{2\gamma_2} n\big).\end{aligned} \tag{5}$$

Note that

$$\Big[2\binom{k}{2}^2 + k(k-1)\Big]\sigma^2n^{\alpha-2}p^{-2}\log^{\mu} n > O(\sigma^2n^{k-3}p^{-\binom{k}{2}+1}\log^{2\gamma_2} n) + (-1)^{k+1}(k-1)$$

when α is shown in (3.6), $p \geqslant p_0$ in (3.13), λ, μ and γ_2 are chosen such that $[\binom{k}{2}-3]\lambda > 2\gamma_2 - \mu$. We have $\mathbb{E}[\Delta\mathbf{L}|\mathcal{F}_i] > 0$ in (5). The sequence $\mathbf{L}(j), \mathbf{L}(j+1), \ldots, \mathbf{L}(\tau^{\ell}_{\mathbf{Q}_k,j})$ is a submartingale.

Next, we show the maximum one step $\Delta\mathbf{L}$ is $O(\sigma n^{k-\frac{5}{2}}\log^{\gamma_2} n)$. As the equation shown in (3) and the calculations in (4), we have

$$\begin{aligned}\Delta\mathbf{L} =& \Delta\mathbf{Q}_k + \frac{2\binom{k}{2}^2 n^{k-2}}{k!}p^{\binom{k}{2}-1} + \frac{2\sigma\sigma' n^{\alpha-2}\log^{\mu} n}{p} \\ &+ \frac{k(k-1)\sigma^2 n^{\alpha-2}\log^{\mu} n}{p^2} + O(n^{k-4}p^{\binom{k}{2}-2}).\end{aligned}$$

Apply the equation of $\Delta\mathbf{Q}_k$ in (2.4) and the estimates on $\mathbf{R}_{k,U_m}$ for any $U_m \in \binom{[n]}{m}$ when $2 \leqslant m \leqslant k-1$ in (3.5) to the above display,

$$\begin{aligned}\Delta\mathbf{L} \leqslant& -\binom{k}{2}\Big(\frac{n^{k-2}}{(k-2)!}p^{\binom{k}{2}-1} - \sigma n^{\beta_2}\log^{\gamma_2} n\Big) \\ &+ \binom{k}{3}\Big(\frac{n^{k-3}}{(k-3)!}p^{\binom{k}{2}-\binom{3}{2}} + \sigma n^{\beta_3}\log^{\gamma_3} n\Big) + \cdots + \frac{2\binom{k}{2}^2 n^{k-2}}{k!}p^{\binom{k}{2}-1} \\ &+ \frac{2\sigma\sigma' n^{\alpha-2}\log^{\mu} n}{p} + \frac{k(k-1)\sigma^2 n^{\alpha-2}\log^{\mu} n}{p^2} + O\big(n^{k-4}p^{\binom{k}{2}-2}\big) \\ =& O(\sigma n^{k-\frac{5}{2}}\log^{\gamma_2} n),\end{aligned}$$

where the two terms involving $n^{k-2}p^{\binom{k}{2}-1}$ cancel exactly and the term $\sigma n^{\beta_2}\log^{\gamma_2} n$ dominates all the remaining terms because the terms $(2\sigma\sigma' n^{\alpha-2}\log^{\mu} n)/p$ and $(k(k-1)\sigma^2 n^{\alpha-2}\log^{\mu} n)/p^2$ are absorbed into the term $O(\sigma n^{k-\frac{5}{2}}\log^{\gamma_2} n)$ by choosing λ, μ and γ_2 such that $2\lambda + \gamma_2 > \mu + 1$ when $p \geqslant p_0$ in (3.13).

The number of steps in this sequence is also $O(n^2 p)$. Note that $\mathbf{Q}_k(j) \in \mathcal{I}^{\ell}_{\mathbf{K}_k}(j)$, for all i with $j \leqslant i < \tau^{\ell}_{\mathbf{Q}_k,j}$, Lemma 2.3 yields that the probability of such a large deviation beginning at the step j is at most

$$\begin{aligned}&\mathbb{P}\Big[\mathbf{Q}_k(i) \leqslant \frac{n^k}{k!}p^{\binom{k}{2}} - \sigma^2 n^{\alpha}p^{-1}\log^{\mu} n\Big] \\ &= \mathbb{P}\Big[\mathbf{L}(i) \leqslant 0\Big] = \mathbb{P}\Big[\mathbf{L}(i) - \mathbf{L}(j) \leqslant -\mathbf{L}(j)\Big] \\ &\leqslant \exp\Bigg[-\Omega\Bigg(\frac{\big(n^{\alpha}p^{-1}\log^{\mu} n\big)^2}{(n^2 p)\big(\sigma n^{k-5/2}\log^{\gamma_2} n\big)^2}\Bigg)\Bigg] \\ &= \exp\Bigg[-\Omega\Bigg(\frac{n^{2\alpha-2k+3}\log^{2\mu} n}{\sigma^2 p^3 \log^{2\gamma_2} n}\Bigg)\Bigg].\end{aligned}$$

There are at most n^2 possible values of j shown in (3.1), by the union bound, then we have

$$n^2 \exp\Bigg[-\Omega\Bigg(\frac{n^{2-\frac{2}{\binom{k}{2}-1}}\log^{2\mu} n}{\sigma^2 p^3 \log^{2\gamma_2} n}\Bigg)\Bigg] = o(1)$$

with α is shown in (3.6) and $p \geqslant p_0$ for $k \geqslant 3$. Hence, *w.h.p.*, $\mathbf{Q}_k(i)$ never crosses its critical interval $\mathcal{I}^{\ell}_{\mathbf{Q}_k}$ in (1), and so the lower bound on $\mathbf{Q}_k(i)$ in (3.4) is true. $\square$

Acknowledgements

The authors are immensely grateful to one of the anonymous reviewers for their detailed and helpful suggestions. Without their help, it was impossible for us to write this manuscript in its current form. Fang Tian was supported by the National Natural Science Foundation of China (Grant No. 12071274). X.-F. Pan was supported by University Natural Science Research Project of Anhui Province under Grant No. KJ2020A0001.

References

[1] J. Balogh and L. Li, On the number of linear hypergraphs of large girth, *J. Graph Theory* **93**(1) (2020), 113–141.

[2] P. Bennett and T. Bohman, A natural barrier in random greedy hypergraph matching, *Combin. Probab. Computing* **28** (2019), 816–825.

[3] V. Blinovsky and C. Greenhill, Asymptotic enumeration of sparse uniform linear hypergraphs with given degrees, *Electron. J. Comb.* **23**(3) (2016), #P3.17.

[4] T. Bohman, The triangle-free process, *Adv. Math.* **221** (2009), 1653–1677.

[5] T. Bohman, A. Frieze and E. Lubetzky, A note on the random greedy triangle-packing algorithm, *J. Comb.* **1** (2010), 477–488.

[6] T. Bohman, A. Frieze and E. Lubetzky, Random triangle removal, *Adv. Math.* **280** (2015), 379–438.

[7] B. Bollobás, To prove and conjecture: Paul Erdős and his mathematics, *Amer. J. Math.* **105**(3) (1998), 209–237.

[8] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis bases on the sum of observations, *Ann. Math. Stat.* **23** (1952), 493–507.

[9] D. Grable, On random greedy triangle packing, *Electron. J. Combin.* **4** (1997), #R11.

[10] H. Guo and L. Warnke, On the power of random greedy algorithms, *Europ. J. Combin.* **105** (2022), 103551.

[11] W. Hoeffding, Probability inequalities for sums of bounded variables, *J. Amer. Statist. Assoc.* **58** (1963), 13–30.

[12] V. Rödl and L. Thoma, Asymptotic packing and the random greedy algorithm, *Random Struct. Algor.* **8** (1996), 161–177.

[13] J. H. Spencer, Asymptotic packing via a branching process, *Random Struct. Algor.* **7** (1995), 167–172.

[14] F. Tian, On the number of linear multiparitite hypergraphs with given size, *Graphs Combin.* **37** (2021), 2487–2496.

[15] F. Tian, Connectivity of the linear hypergraph process, *Discrete Math.* **345** (2022), 113056.

(Received 22 Nov 2022; revised 2 July 2023, 31 Aug 2023)