Packing nearly optimal Ramsey $R(3, t)$ graphs

He Guo and Lutz Warnke*

December 20, 2017; revised November 26, 2018

Abstract

In 1995 Kim famously proved the Ramsey bound $R(3, t) \geq ct^2/\log t$ by constructing an n-vertex graph that is triangle-free and has independence number at most $C\sqrt{n \log n}$. We extend this celebrated result, which is best possible up to the value of the constants, by approximately decomposing the complete graph K_n into a packing of such nearly optimal Ramsey $R(3, t)$ graphs.

More precisely, for any $\epsilon > 0$ we find an edge-disjoint collection $(G_i)_i$ of n-vertex graphs $G_i \subseteq K_n$ such that (a) each G_i is triangle-free and has independence number at most $C_\epsilon\sqrt{n \log n}$, and (b) the union of all the G_i contains at least $(1 - \epsilon)\binom{n}{2}$ edges. Our algorithmic proof proceeds by sequentially choosing the graphs G_i via a semi-random (i.e., Rödl nibble type) variation of the triangle-free process.

As an application, we prove a conjecture in Ramsey theory by Fox, Grinshpun, Liebenau, Person, and Szabó (concerning a Ramsey-type parameter introduced by Burr, Erdős, Lovász in 1976). Namely, denoting by $s_r(H)$ the smallest minimum degree of r-Ramsey minimal graphs for H, we close the existing logarithmic gap for $H = K_3$ and establish that $s_r(K_3) = \Theta(r^2 \log r)$.

1 Introduction

The 1947 paper of Erdős [10] on the diagonal Ramsey number $R(t, t)$ is often considered the start of the probabilistic method, where $R(s, t)$ is defined as the smallest integer $n \in \mathbb{N}$ such that every red-blue colouring of the edges of the complete n-vertex graph K_n contains either a red K_s or a blue K_t. In general, the estimation of $R(s, t)$ and other Ramsey-type parameters is known to be notoriously difficult.

One of the celebrated results in Ramsey theory is $R(3, t) = \Theta(t^2/\log t)$, and this special case has repeatedly served as a testbed for the development of new tools and techniques in probabilistic combinatorics. Indeed, complementing the basic bound $R(3, t) = O(t^2)$ of Erdős and Szekeres [14], in 1961 Erdős [11] used a sophisticated random greedy alteration argument to prove $R(3, t) = \Omega(t^2/(\log t)^2)$. This lower bound was subsequently reproved (or only slightly improved) using the Lovász Local Lemma [31], a basic analysis of the triangle-free process[1] [13], large deviation inequalities [21], and differential equations [32]. Furthermore, in 1980 Ajtai, Komlós, and Szemerédi [1, 2] invented the influential semi-random method (nowadays also called Rödl nibble approach) to prove the upper bound $R(3, t) = O(t^2/\log t)$. But it was not until 1995, when Kim [20] famously proved the matching lower bound $R(3, t) = \Omega(t^2/\log t)$ by analyzing a semi-random variation of the triangle-free process[2] (combining several of the aforementioned ideas with martingale concentration); for this major breakthrough he also received the Fulkerson Prize in 1997. But the story does not end here: advancing the differential equation method, in 2008 Bohman [5] reproved $R(3, t) = \Omega(t^2/\log t)$ by analyzing the triangle-free process itself (and his analysis was recently further improved in [7, 15]).

In this paper we refine the powerful techniques developed for $R(3, t) = \Theta(t^2/\log t)$ to determine the order of magnitude of another Ramsey-type parameter introduced in 1976 by Burr, Erdős, and Lovász [8], proving a conjecture of Fox, Grinshpun, Liebenau, Person, and Szabó [16] (in particular, analogous to Kim's $R(3, t)$-result, we again remove the last redundant logarithmic factor from existing bounds).

*School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA. E-mail: `he.guo@gatech.edu`, `warnke@math.gatech.edu`. Research partially supported by NSF Grant DMS-1703516 and a Sloan Research Fellowship.

[1] The triangle-free process (proposed by Bollobás and Erdős) proceeds as follows: starting with an empty n-vertex graph, in each step a single edge is added, chosen uniformly at random from all non-edges which do not create a triangle.

[2] Kim's semi-random variation proceeds similar to the triangle-free process: it intuitively adds a large number of carefully chosen random-like edges in each step (instead of just a single edge); see Section 2 for more details.

1.1 Main result: packing of nearly optimal Ramsey $R(3,t)$ graphs

Kim and Bohman both proved the Ramsey bound $R(3,t) = \Omega(t^2/\log t)$ by showing the existence of a triangle-free graph $G \subseteq K_n$ on n vertices with independence number $\alpha(G) = O(\sqrt{n \log n})$, which is best possible up to the value of the implicit constants. Our first theorem naturally extends their celebrated results, by approximately decomposing the complete graph K_n into a packing of such nearly optimal Ramsey $R(3,t)$ graphs.

Theorem 1. *For any $\epsilon > 0$ there exist $n_0, C, D > 0$ such that, for all $n \geq n_0$, there is an edge-disjoint collection $(G_i)_{i\in\mathcal{I}}$ of $|\mathcal{I}| = \lceil D\sqrt{n/\log n}\rceil$ triangle-free graphs $G_i \subseteq K_n$ on n vertices with $\max_{i\in\mathcal{I}} \alpha(G_i) \leq C\sqrt{n \log n}$ and $\sum_{i\in\mathcal{I}} e(G_i) \geq (1-\epsilon)\binom{n}{2}$.*

Our algorithmic proof proceeds by sequentially choosing the $|\mathcal{I}| = \Theta(\sqrt{n/\log n})$ edge-disjoint triangle-free subgraphs $G_i \subseteq K_n \setminus \bigcup_{0\leq j<i} G_j$ with $\alpha(G_i) = O(\sqrt{n \log n})$ via a semi-random variation of the triangle-free process akin to Kim [20] (see Sections 1.3 and 2 for the details). In particular, we do not only show existence of the $(G_i)_{i\in\mathcal{I}}$, but also obtain a polynomial-time randomized algorithm which constructs these subgraphs.

Theorem 1 improves a construction of Fox et.al. [16, Lemma 4.2], who used the basic Lovász Local Lemma based $R(3,t)$-approach to sequentially choose $\Theta(\sqrt{n}/\log n)$ edge-disjoint triangle-free subgraphs with $\alpha(G_i) = O(\sqrt{n}\log n)$. It is natural to suspect that applying a more sophisticated $R(3,t)$-approach in each iteration ought to give an improved packing (with smaller independence number than the LLL approach), and here the usage of the triangle-free process was proposed by Fox et.al. [16, Section 5] as early as 2013 [22, 26]. One conceptual difficulty of this approach is to control various error terms over many iterations of the triangle-free process (so that these always stay small enough to carry out the next iteration), which in turn is the main technical reason why for Theorem 1 we instead iterate a semi-random variation.

It would be interesting to know if Theorem 1 also holds with $\epsilon = 0$, i.e., if one can completely decompose K_n into nearly optimal $R(3,t)$ graphs. Perhaps rashly, we conjecture that this is indeed possible (it might be insightful to first prove a variant of Theorem 1 where the constant C does *not* depend on ϵ).

1.2 Application in Ramsey theory: $s_r(K_3)$ has order of magnitude $r^2 \log r$

Turning to our main application, we say that a graph G is *r-Ramsey for H*, denoted by $G \to (H)_r$, if any r-colouring of the edges of G contains a monochromatic copy of H. Most fundamental questions and results in Ramsey theory can be formulated in terms of various parameters of the class

$$\mathcal{M}_r(H) := \big\{G :\ G \to (H)_r \text{ and } G' \nrightarrow (H)_r \text{ for all } G' \subsetneq G\big\}$$

of graphs which are *r-Ramsey minimal for H*. For example, Ramsey's theorem [28] states that $|\mathcal{M}_r(H)| > 0$ for all graphs H, which for cliques was strengthened to $|\mathcal{M}_r(K_t)| = \infty$ by Rödl and Siggers [29]. Furthermore, the archetypal problem of estimating various Ramsey-type parameters also corresponds to the study of certain extremal parameters of $\mathcal{M}_r(H)$, since, e.g., $R(t) = R(t,t) := \min_{G\in\mathcal{M}_2(K_t)} v(G)$ is the famous diagonal Ramsey number [14, 10, 9], $R_r(t) = R(t,\ldots,t) := \min_{G\in\mathcal{M}_r(K_t)} v(G)$ is the r-coloured Ramsey number [9], and $\hat{R}_r(H) := \min_{G\in\mathcal{M}_r(H)} e(G)$ is the widely-studied r-size-Ramsey number of H (see, e.g., [12, 4, 30, 9]).

In 1976 Burr, Erdős, and Lovász [8] initiated the systematic study of other extremal parameters of $\mathcal{M}_r(H)$, including the smallest minimum degree of all r-Ramsey minimal graphs for H, denoted by

$$s_r(H) := \min_{G\in\mathcal{M}_r(H)} \delta(G).$$

As usual, the clique-case $H = K_t$ is of particular interest, where $r(t-2) < s_r(K_t) < R_r(t)$ is easy to see (cf. [17, 33]). Perhaps surprisingly, for $r = 2$ colours Burr et.al. [8] were able to prove $s_2(K_t) = (t-1)^2$, showing that the simple exponential upper bound $R_2(t) = R(t) = 2^{\Theta(t)}$ is far from the truth. For $r \geq 2$ colours the behaviour of $s_r(K_t)$ was recently investigated in detail by Fox et.al. [16]: they proved super-quadratic bounds of form $s_r(K_t) = r^2 \cdot \text{polylog}\, r$ for fixed $t \geq 3$, and also determined $s_r(K_3)$ up to a logarithmic factor (by sharpening their general estimates). In particular, they showed $cr^2 \log r \leq s_r(K_3) \leq Cr^2(\log r)^2$, and conjectured that their lower bound gives the correct order of magnitude, see [16, Conjecture 5.4].

Our second theorem proves the aforementioned conjecture of Fox, Grinshpun, Liebenau, Person, and Szabó for $s_r(K_3)$, i.e., we close the logarithmic gap and establish $s_r(K_3) = \Theta(r^2 \log r)$.

Theorem 2. *There exists $C > 0$ such that $s_r(K_3) \le Cr^2 \log r$ for all $r \ge 2$.*

Corollary 3. *We have $s_r(K_3) = \Theta(r^2 \log r)$ for $r \ge 2$.*

Using a reformulation of $s_r(K_3)$ from [16], Theorem 2 follows easily from our main packing result. Indeed, applying Theorem 1 with $\epsilon = 1/2$, say, it is routine to see that there is a constant $A > 0$ such that the following holds for each $r \ge 2$: there exists a collection of edge-disjoint triangle-free graphs $G_1, \ldots, G_r \subseteq K_{N_r}$ on $N_r := \lfloor Ar^2 \log r \rfloor$ vertices with independence number $\alpha(G_i) < N_r/r$ (as $N_r \ge n_0$, $D\sqrt{N_r/\log N_r} \ge r$ and $C\sqrt{N_r \log N_r} < N_r/r$ all hold for $A = A(n_0, C, D)$ large enough). By Theorem 1.5 and Lemma 4.1 in [16] (with $n = N_r$ and $k = 2$) this immediately implies $s_r(K_3) \le N_r$, establishing Theorem 2.

Note that the above deduction of Theorem 2 did *not* use $\sum_{i \in \mathcal{I}} e(G_i) \ge (1-\epsilon)\binom{n}{2}$, i.e., that the nearly optimal $R(3,t)$ graphs $(G_i)_{i \in \mathcal{I}}$ approximately decompose the edge-set of K_n. It would be interesting to find applications (e.g., in Ramsey theory or extremal combinatorics) where this natural packing property is useful.

1.3 Main tool: pseudo-random triangle-free subgraphs

The $R(3,t)$-proofs of Kim and Bohman both in fact construct a triangle-free graph $G \subseteq K_n$ with pseudo-random properties (see also [32, 38, 7, 15]). Our third theorem extends their intriguing results to host graphs $H \subseteq K_n$ which are far from complete, by showing that one can again construct a triangle-free subgraph $G \subseteq H$ with pseudo-random properties. Here the crux is that Theorem 4 holds under very weak assumptions,[3] that G resembles a random subgraph of H with edge-probability $\rho = \Theta(\sqrt{(\log n)/n})$, and that the edge-estimate (1) implies $\alpha(G) = O(\sqrt{n \log n})$ for many well-behaved host graphs $H \subseteq K_n$.

Theorem 4. *There exist $\beta_0, D_0 > 0$ such that, for all $\gamma, \delta \in (0,1]$, $\beta \in (0, \beta_0)$ and $C \ge D_0/(\delta^2\sqrt{\beta}\gamma)$, the following holds for all $n \ge n_0(\gamma, \delta, \beta, C)$, with $\rho := \sqrt{\beta(\log n)/n}$. For any n-vertex graph H, there exists a triangle-free subgraph $G \subseteq H$ on the same vertex-set such that*

$$e_G(A,B) = (1 \pm \delta)\rho e_H(A,B) \tag{1}$$

for all disjoint vertex-sets $A, B \subseteq V(H)$ with $|A| = |B| = \lceil C\sqrt{n \log n}\rceil$ and $e_H(A,B) \ge \gamma |A||B|$.

Our proof uses a semi-random variant of the triangle-free process to construct $G \subseteq H$, extending and simplifying Kim's $R(3,t)$-approach for the complete case $H = K_n$ (see Sections 2–3 and Theorem 9 for the details). In particular, besides handling the difficulties arising due to incomplete host graphs $H \subseteq K_n$ (by, e.g., exploiting a 'stabilization mechanism' to keep various parameters under control), the major technical difference lies in the way we analyze the properties of all large vertex-sets (by, e.g., focusing on bipartite subgraphs, applying a concentration inequality of Warnke [37], and showing concentration in (1) instead of just $e_G(A,B) \ge 1$). Together with some streamlining of Kim's arguments (by, e.g., using fewer variables, applying convenient bounded differences inequalities, and some changes to the semi-random construction), this leads to a shorter and hopefully more accessible proof even in the complete case $H = K_n$. As a by-product, we also obtain a randomized polynomial-time algorithm which constructs $G \subseteq H$ (see Remark 10).

Theorem 4 will be the main tool for establishing our main packing result Theorem 1. Let us briefly sketch the argument (deferring the details to Section 1.5). The idea is to sequentially choose the triangle-free subgraphs $G_i \subseteq H_i := K_n \setminus \bigcup_{0 \le j < i} G_j$ via Theorem 4 with $\delta \in (0,1)$, using the pseudo-random edge-estimate (1) to inductively control the number of remaining edges (between large sets) in H_i as

$$e_{H_i}(A,B) = (1-(1\pm\delta)\rho)^i \cdot |A||B| \qquad \text{for all disjoint } A, B \subseteq V(H) \text{ of size } s := \lceil C\sqrt{n \log n}\rceil, \tag{2}$$

stopping when the right hand side of (2) drops below $\epsilon|A||B|$ after $I = \Theta(\log(1/\epsilon)/\rho) = \Theta(\sqrt{n/\log n})$ steps. A double counting argument will then show that the leftover graph H_I contains at most $\epsilon\binom{n}{2}$ edges, so that $\sum_{0 \le i < I} e(G_i) = e(K_n \setminus H_I) \ge (1-\epsilon)\binom{n}{2}$. Furthermore, $e_{G_i}(A,B) = (1\pm\delta)\rho e_{H_i}(A,B) > 0$ implies $\alpha(G_i) < 2s = O(\sqrt{n \log n})$, completing this rough proof sketch of Theorem 1 (assuming Theorem 4).

We believe that variants of Theorems 1 and 4 also hold for many other forbidden graphs (using semi-random variants of the H-free process [25, 6, 34, 35, 27]); we hope to return to this topic in a future work.

[3] Note that Theorem 4 does *not* require the host graph H to be approximately degree or codegree regular. Furthermore, even if $G \subseteq H$ was a random subgraph with edge-probability ρ, then by standard calculations we would only expect the edge-estimate (1) to hold for vertex-sets $A, B \subseteq V(H)$ where the number of edges $e_H(A,B)$ is reasonably large (see Remark 11 for the details, which also indicates that the constant C in Theorem 4 has the correct dependence on γ, δ, β).

1.4 Organization of the paper

The remainder of this paper is organized as follows. In Section 1.5 we use Theorem 4 to state and prove some extensions of our main packing result Theorem 1. In Section 2 we introduce a semi-random variation of the triangle-free process and state our main result for this Rödl nibble type construction (that implies our main tool Theorem 4, see Section 2.4), which is then subsequently proved in Section 3.

1.5 Further results

Our methods allow us to extend Theorem 1 to $R(3,t)$-packings of graphs which are far from complete. Our fourth theorem shows that if $H \subseteq K_n$ only satisfies certain uniformity conditions on its edge distribution (that resemble a weak form of pseudo-randomness, see (3) below), then we can still approximately decompose H into a packing of nearly optimal Ramsey $R(3,t)$ graphs (again by an efficient randomized algorithm).

Theorem 5. *For all $\epsilon, \xi, C_0 > 0$ there exist $n_0, C_1, D > 0$ such the following holds for all $n \geq n_0$. If H is an n-vertex graph satisfying*

$$\min_{\substack{\text{disjoint } A,\, B \subseteq V(H):\\ |A|=|B|=\lceil C_0\sqrt{n\log n}\rceil}} \frac{e_H(A,B)}{|A||B|} \geq \xi, \tag{3}$$

then there is an edge-disjoint collection $(G_i)_{i\in\mathcal{I}}$ of $|\mathcal{I}| = \lceil D\sqrt{n/\log n}\rceil$ triangle-free subgraphs $G_i \subseteq H$ with $V(G_i) = V(H)$, $\max_{i\in\mathcal{I}} \alpha(G_i) \leq C_1\sqrt{n\log n}$ and $\sum_{i\in\mathcal{I}} e(G_i) \geq (1-\epsilon)e(H)$.

Note that the case $H = K_n$ and $\xi = C_0 = 1$ implies Theorem 1. Furthermore, the case $H = G_{n,p}$, $\xi = p/2$ and $C_0 = 1$ routinely implies the following sparse analogue of Theorem 1 for binomial random graphs $G_{n,p}$.

Corollary 6. *For any $p \in (0,1]$ and $\epsilon > 0$ there exist $C, D > 0$ such that, with probability at least $1 - o(1)$, the following event holds: there exists an edge-disjoint collection $(G_i)_{i\in\mathcal{I}}$ of $|\mathcal{I}| = \lceil D\sqrt{n/\log n}\rceil$ triangle-free graphs $G_i \subseteq G_{n,p}$ on n vertices with $\max_{i\in\mathcal{I}} \alpha(G_i) \leq C\sqrt{n\log n}$ and $\sum_{i\in\mathcal{I}} e(G_i) = (1\pm\epsilon)p\binom{n}{2}$.*

We conjecture that Corollary 6 (with $|\mathcal{I}| = \lceil Dp\sqrt{n/\log n}\rceil$ and constants $C, D > 0$ depending only on ϵ) holds for much sparser random graphs $G_{n,p}$ with edge-probabilities of form $p = p(n) \geq n^{-1/2+o(1)}$, say.[4]

We conclude the introduction with the short proof of Theorem 5, which proceeds by sequentially choosing the graphs $G_i \subseteq H \setminus \bigcup_{0\leq j<i} G_j$ via Theorem 4 (generalizing the argument sketched in Section 1.3). The reader mainly interested in the proof of Theorem 4 may perhaps wish to skip straight to Section 2.

Proof of Theorem 5 (assuming Theorem 4). We may assume $\epsilon < 1$ (as decreasing ϵ gives a stronger conclusion). For concreteness, set $\delta := 1/4$, $\gamma := \epsilon^2\xi$, $\beta := \beta_0/2$ and $C := \max\{C_0,\, D_0/(\delta^2\sqrt{\beta}\gamma)\}$, where β_0, D_0 are defined as in Theorem 4. Let $C_1 := 3C$, $s := \lceil C\sqrt{n\log n}\rceil$, $\rho := \sqrt{\beta(\log n)/n}$, and $I := \lceil \log(1/\epsilon)/(\rho(1-\delta))\rceil$.

Define $H_0 := H$. Let $\mathfrak{S}$ denote the set of all pairs (A,B) of disjoint vertex-sets $A, B \subseteq V(H)$ with $|A| = |B| = s$. Combining a 'handshaking lemma' like double counting argument with the assumed lower bound (3), writing $t := \lceil C_0\sqrt{n\log n}\rceil$ it follows that

$$\frac{e_{H_0}(A,B)}{|A||B|} = \frac{\sum_{\tilde{A}\subseteq A, \tilde{B}\subseteq B:\, |\tilde{A}|=|\tilde{B}|=t} e_H(\tilde{A},\tilde{B})}{s^2\cdot\binom{s-1}{t-1}\binom{s-1}{t-1}} \geq \frac{\binom{s}{t}\binom{s}{t}\cdot\xi t^2}{s^2\binom{s-1}{t-1}\binom{s-1}{t-1}} = \xi \qquad \text{for all } (A,B)\in\mathfrak{S}. \tag{4}$$

The plan is to sequentially choose the graphs $(G_i)_{0\leq i<I}$ with $G_i \subseteq H_i$ such that, setting $H_{i+1} := H_i \setminus G_i$ (which ensures that all the G_i are edge-disjoint), for all $0 \leq i \leq I$ we inductively have

$$\frac{e_{H_i}(A,B)}{e_{H_0}(A,B)} \in \Big[\big(1-(1+\delta)\rho\big)^i,\ \big(1-(1-\delta)\rho\big)^i\Big] \qquad \text{for all } (A,B)\in\mathfrak{S}. \tag{5}$$

Turning to the details, note that inequality (5) holds trivially for $i = 0$. Given H_i with $0 \leq i \leq I-1$ satisfying (5), by combining the definition of I with $(1+2\delta)/(1-\delta) = 2$ and (4) it follows for $n \geq n_0(\beta)$ that, say,

$$\frac{e_{H_i}(A,B)}{|A||B|} \geq e^{-(1+2\delta)\rho(I-1)}\cdot\frac{e_{H_0}(A,B)}{|A||B|} \geq \epsilon^2\cdot\xi = \gamma \qquad \text{for all } (A,B)\in\mathfrak{S}. \tag{6}$$

[4] The range of $p = p(n)$ in this conjecture is essentially best possible, since it is well-known that typically $\alpha(G_{n,p}) \gg \sqrt{n\log n}$ for $p \ll \sqrt{(\log n)/n}$. Furthermore, although we have not checked all details, it seems that our proofs can be modified to verify the conjecture for $p \geq n^{-\delta}$, where $\delta > 0$ is some small constant; so the main question is whether $p \geq n^{-1/2+o(1)}$ suffices.

Using Theorem 4, for $n \ge n_0(\epsilon, \xi, \delta, \beta, C)$ we can thus find a triangle-free subgraph $G_i \subseteq H_i$ with $e_{G_i}(A,B) = (1 \pm \delta)\rho e_{H_i}(A,B) > 0$ for all $(A,B) \in \mathfrak{S}$. Hence $\alpha(G_i) < 2s \le 3C\sqrt{n \log n}$, say. Furthermore, noting $e_{H_{i+1}}(A,B) = e_{H_i}(A,B) - e_{G_i}(A,B)$, it is immediate that $H_{i+1} = H_i \setminus G_i$ maintains (5).

Finally, for the number of edges of $\bigcup_{0 \le i < I} G_i = H_0 \setminus H_I$, by (5) and definition of I it follows that

$$e_{H_0 \setminus H_I}(A,B) \ge \bigl(1 - e^{-(1-\delta)\rho I}\bigr) \cdot e_{H_0}(A,B) \ge (1-\epsilon) e_{H_0}(A,B) \qquad \text{for all } (A,B) \in \mathfrak{S}. \tag{7}$$

Using a double counting argument similar to (4), in view of (7) and $H_0 = H$ we infer

$$e(H_0 \setminus H_I) = \frac{\sum_{(A,B)\in\mathfrak{S}} e_{H_0 \setminus H_I}(A,B)}{2\binom{n-2}{s-1}\binom{n-2-(s-1)}{s-1}} \ge (1-\epsilon) \cdot \frac{\sum_{(A,B)\in\mathfrak{S}} e_{H}(A,B)}{2\binom{n-2}{s-1}\binom{n-2-(s-1)}{s-1}} = (1-\epsilon)e(H),$$

completing the proof of $\sum_{0 \le i < I} e(G_i) = e(H_0 \setminus H_I) \ge (1-\epsilon)e(H)$. □

2 The nibble: semi-random triangle-free process

The remainder of this paper is devoted to the proof of our main tool Theorem 4. Given an n-vertex graph H with vertex-set $V = V(H)$ and edge-set $E(H)$, inspired by Kim [20] our strategy is to incrementally construct the triangle-free edge-set of $G \subseteq H$ using a semi-random variation of the triangle-free process (adding large chunks of random-like edges in each step; see also Footnotes 1–2 on page 1). One key difference to [20, 5] is that our approach only uses edges from the host graph H (and not the complete graph K_n). In particular, deferring the details to Section 2.1, the rough plan of our Rödl nibble type construction is to step-by-step build up a 'random' set of edges $E_i \subseteq E(H)$ and a triangle-free subset $T_i \subseteq E_i$; we also keep track of a set

$$O_i \subseteq \{e \in E(H) \setminus E_i : \ e \text{ does not form a triangle with any two edges of } E_i\} \tag{8}$$

of 'open' edges that can still be added. The idea of each step is to choose a small number of random edges $\Gamma_{i+1} \subseteq O_i$ so that only a few new triangles are created in $E_{i+1} = E_i \cup \Gamma_{i+1}$. This allows us to find an edge-subset $\Gamma'_{i+1} \subseteq \Gamma_{i+1}$, with $|\Gamma'_{i+1}| \approx |\Gamma_{i+1}|$, such that $T_{i+1} = T_i \cup \Gamma'_{i+1}$ remains triangle-free.[5] After

$$I := \lceil n^\beta \rceil \tag{9}$$

such alteration-method based steps, we eventually obtain a triangle-free graph $G = (V, T_I) \subseteq H$, which intuitively ought to be 'random enough' to resemble (many features of) a random subgraph of H.

2.1 Details of the nibble construction

Turning to the details of the nibble construction, consistent with (8) we start with

$$O_0 := E(H) \quad \text{ and } \quad E_0 := T_0 := \Gamma_0 := \varnothing. \tag{10}$$

In step $i+1 \ge 1$ we then set

$$E_{i+1} := E_i \cup \Gamma_{i+1}, \tag{11}$$

where each edge $e \in O_i$ is included in Γ_{i+1}, independently, with probability

$$p := \sigma/\sqrt{n}. \tag{12}$$

(The definition of the deterministic parameter $\sigma \ll 1$ is deferred to (35) in Section 2.3.) Note that $T_i \cup \Gamma_{i+1}$ is not necessarily triangle-free, since two or three edges of a triangle could enter via $\Gamma_{i+1} \subseteq O_i$ (one edge is not enough by (8) and $T_i \subseteq E_i$), i.e., via the following set of 'bad' pairs and triples of Γ_{i+1}–edges:

$$\mathcal{B}_{i+1} := \bigl\{\{wu, wv\} \subseteq \Gamma_{i+1} : \ uv \in T_i, \ |\{u,v,w\}| = 3\bigr\} \ \cup \ \bigl\{\{uv, vw, wu\} \subseteq \Gamma_{i+1} : \ |\{u,v,w\}| = 3\bigr\}, \tag{13}$$

[5] For the construction of T_{i+1} it might seem overly complicated to define O_i with respect to E_i (and not T_i). However, this slightly wasteful definition actually simplifies the analysis: e.g., for the purpose of tracking various auxiliary variables, it intuitively is easier to understand the effect of adding the random edges Γ_{i+1} (rather than some subset $\Gamma'_{i+1} \subseteq \Gamma_{i+1}$). Using an inclusion in (8) might also seem overly complicated, but it again simplifies the analysis: by removing some extra edges it actually becomes easier to prove concentration (see the 'stabilization mechanism' discussion around (21) and Lemma 19).

where we write $xy = \{x, y\}$ for brevity. To avoid triangles in T_{i+1} by alteration, we thus take $\mathcal{D}_{i+1}$ to be a maximal collection of pairwise edge-disjoint elements of $\mathcal{B}_{i+1}$ (say the first one in lexicographic order to resolve ties; any other deterministic choice also works, see Remark 7 and Section 3.5), and then set[6]

$$T_{i+1} := T_i \cup \big(\Gamma_{i+1} \setminus E(\mathcal{D}_{i+1})\big), \tag{14}$$

where we write $E(\mathcal{D}_{i+1}) := \bigcup_{\alpha \in \mathcal{D}_{i+1}} \alpha$ for the set of edges in the pairs and triples of $\mathcal{D}_{i+1}$. Note that T_{i+1} is indeed triangle-free by maximality of $\mathcal{D}_{i+1} \subseteq \mathcal{B}_{i+1}$. Defining

$$Y_{uv}(i) := \{uw \in O_i :\ vw \in E_i\} \,\cup\, \{vw \in O_i :\ uw \in E_i\}, \tag{15}$$

we now turn to the open edge-set $O_{i+1} \subseteq O_i \setminus \Gamma_{i+1}$: by (8) the set $C^{(1)}_{i+1} \cup C^{(2)}_{i+1} \subseteq O_i$ of newly 'closed' edges (that form a triangle with some two edges of E_{i+1}) is given by

$$C^{(1)}_{i+1} := \{f \in O_i :\ Y_f(i) \cap \Gamma_{i+1} \neq \varnothing\}, \tag{16}$$

$$C^{(2)}_{i+1} := \{uv \in O_i :\ \text{there is } w \text{ s.t. } uw \in \Gamma_{i+1},\ vw \in \Gamma_{i+1}\}. \tag{17}$$

Mimicking a technical idea of Alon, Kim and Spencer [3], we intuitively increase the set of closed edges (via the random set S_{i+1} below) in order to add a 'stabilization mechanism' to our construction,[7] and define

$$C_{i+1} := C^{(1)}_{i+1} \cup S_{i+1}, \tag{18}$$

$$O_{i+1} := O_i \setminus \big(\Gamma_{i+1} \cup C_{i+1} \cup C^{(2)}_{i+1}\big), \tag{19}$$

where each edge $e \in O_i$ is included in S_{i+1}, independently, with 'stabilization' probability

$$\hat{p}_{e,i} := 1 - (1-p)^{\max\{2q_i(\pi_i + \sqrt{\sigma})\sqrt{n} - |Y_e(i)|,\, 0\}}. \tag{20}$$

(The definition of the deterministic parameters q_i, π_i is deferred to (36)–(37) in Section 2.3.) Roughly put, the main point of the technical definitions of S_{i+1} and $\hat{p}_{e,i}$ will be that all the conditional probabilities

$$\mathbb{P}(e \notin C_{i+1} \mid O_i, E_i) = \mathbb{P}(e \notin C^{(1)}_{i+1} \mid O_i, E_i) \cdot (1 - \hat{p}_{e,i}) = (1-p)^{\max\{2q_i(\pi_i + \sqrt{\sigma})\sqrt{n},\, |Y_e(i)|\}} \tag{21}$$

can inductively be made equal and thus independent of the history (by only maintaining a weak upper bound on $\max_e |Y_e(i)|$; see (45), (62) and Lemma 19), which in turn helps to keep various error terms under control.

Remark 7. *Note that each step of our nibble construction requires only randomized polynomial time (since we can easily find a* maximal *edge-disjoint collection $\mathcal{D}_{i+1} \subseteq \mathcal{B}_{i+1}$ by a deterministic greedy algorithm).*

2.2 Pseudo-random intuition: trajectory equations

In this informal section we give a heuristic explanation of the differential equation that predicts the behaviour of (O_i, E_i) for $0 \le i \le I \approx n^\beta$. Inspired by [32, 20], our main non-rigorous ansatz is that the edge-sets (O_i, E_i) should resemble properties of a random subgraph of H with two types of edges, where

$$\mathbb{P}(e \in O_i) \approx q_i \quad \text{ and } \quad \mathbb{P}(e \in E_i) \approx \pi_i/\sqrt{n} \tag{22}$$

are approximately independent. We now derive properties of q_i, π_i that are consistent with this ansatz. For example, combining $E_{i+1} = E_i \cup \Gamma_{i+1}$ with the random construction of $\Gamma_{i+1} \subseteq O_i$, we expect to have

$$\mathbb{P}(e \in E_{i+1}) - \mathbb{P}(e \in E_i) = \mathbb{P}(e \in \Gamma_{i+1} \mid e \in O_i)\,\mathbb{P}(e \in O_i) \approx p \cdot q_i = \sigma q_i/\sqrt{n}, \tag{23}$$

[6]The standard alteration approach of removing one edge from each element of $\mathcal{B}_{i+1}$ seems harder to analyze: e.g., removing the edges of a maximal edge-disjoint collection $\mathcal{D}_{i+1} \subseteq \mathcal{B}_{i+1}$ greatly facilitates the technical calculations in Section 3.5.

[7]Kim uses a different stabilization mechanism in [20, Section 5.1]: instead of introducing the random sets S_j, he deterministically modifies the underlying graphs in each step (by temporarily adding some extra edges and vertices), mimicking an earlier 'regularization' idea from [19]. We find our randomized approach more elegant, and easier to implement algorithmically.

which together with (22) and $E_0 = \varnothing$ suggests that

$$\pi_{i+1} - \pi_i \approx \sigma q_i \quad \text{and} \quad \pi_0 \approx 0. \tag{24}$$

Furthermore, with lots of hand-waving, by (19) we intuitively have $O_i \setminus O_{i+1} = \Gamma_{i+1} \cup C_{i+1} \cup C^{(2)}_{i+1} \approx C_{i+1}$ (since each closed edge in $C^{(2)}_{i+1}$ requires the presence of at least two random edges from $\Gamma_{i+1} \subseteq O_i$). As (22) suggests $\mathbb{E}\,|Y_e(i)| \lesssim 2q_i\pi_i\sqrt{n}$, by the stabilization mechanism (21) and $p = \sigma/\sqrt{n}$ we thus loosely expect that

$$\mathbb{P}(e \in O_{i+1} \mid O_i, E_i) \approx \mathbb{P}(e \notin C_{i+1} \mid O_i, E_i) = (1-p)^{2q_i(\pi_i+\sqrt{\sigma})\sqrt{n}} \approx 1 - 2\sigma q_i \pi_i \quad \text{for } e \in O_i,$$

where we bluntly ignored the $\sqrt{\sigma}$-term in the exponent. Similar to (23), using (22) we thus ought to have

$$q_{i+1} - q_i \approx \mathbb{P}(e \in O_{i+1}) - \mathbb{P}(e \in O_i) \approx -2\sigma q_i \pi_i \cdot \mathbb{P}(e \in O_i) \approx -2\sigma q_i^2 \pi_i. \tag{25}$$

To extract the behaviour of π_I from (24) and (25), we further assume that $\pi_i \approx \Psi(i\sigma)$ holds for some smooth function $\Psi(x)$, where $\sigma \ll 1$ is tiny. Using Taylor series, in view of (24) and $O_0 = E(H)$ this suggests that

$$q_i \approx \Psi'(i\sigma) \quad \text{and} \quad q_0 \approx 1. \tag{26}$$

Together with (25) and the initial values from (24) and (26), this leads to the second order differential equation $\Psi''(x) = -2\Psi'(x)^2\Psi(x)$ with $\Psi'(0) = 1$ and $\Psi(0) = 0$, which in turn reduces to the simple ODE

$$\Psi'(x) = e^{-\Psi^2(x)} \quad \text{and} \quad \Psi(0) = 0. \tag{27}$$

Noting the implicit solution $x = \int_0^{\Psi(x)} e^{t^2}\,dt$, it now is easy to derive that $\Psi(x) \approx \sqrt{\log x}$ as $x \to \infty$ (see, e.g., the proof of (57) in Appendix A). Since $I \approx n^\beta$ is sufficiency large compared to σ (which will be of form $\sigma = (\log n)^{-\Theta(1)}$, see (35) in Section 2.3), this makes it plausible that

$$\pi_I \approx \Psi(I\sigma) \approx \sqrt{\log(I\sigma)} \approx \sqrt{\beta \log n}. \tag{28}$$

Finally, since by construction we expect $|E_{i+1} \setminus E_i| \approx |T_{i+1} \setminus T_i|$ to hold for all $0 \le i < I$, the edge-sets E_I and T_I ought to share many properties. Together with (22) and (28) this intuitively suggests

$$\mathbb{P}(e \in T_I) \approx \mathbb{P}(e \in E_I) \approx \sqrt{\beta(\log n)/n}, \tag{29}$$

making the pseudo-random edge-estimate (1) plausible for $G = (V, T_I)$ with $T_I \subseteq E_I \subseteq E(H)$.

2.3 Definitions and parameters

In this section we formally define several variables and parameters used in our analysis of the nibble construction. We start with two standard notions from graph theory: for any edge-subset $S \subseteq \binom{V}{2}$ we write

$$S(A,B) := \{ab \in S :\ a \in A,\ b \in B\}, \tag{30}$$
$$N_S(v) := \{w \in V :\ vw \in S\}, \tag{31}$$

where $A, B \subseteq V$ are vertex-disjoint. For all pairs of distinct vertices $u, v \in V$ we then define

$$X_{uv}(i) := N_{O_i}(u) \cap N_{O_i}(v), \tag{32}$$
$$Z_{uv}(i) := N_{E_i}(u) \cap N_{E_i}(v), \tag{33}$$

where $|X_{uv}(i)|$ and $|Z_{uv}(i)|$ intuitively correspond to an 'open codegree' and the usual codegree, respectively (note that $|Y_{uv}(i)|$ defined in (15) corresponds to a 'mixed codegree').

Guided by Section 2.2, we define $\Psi(x)$ as the unique solution to the differential equation

$$\Psi'(x) = e^{-\Psi^2(x)} \quad \text{and} \quad \Psi(0) = 0, \tag{34}$$

as suggested by (27). With the heuristics (22) in mind, we then introduce the parameters

$$\sigma := (\log n)^{-2}, \tag{35}$$

$$q_i := \Psi'(i\sigma) = e^{-\Psi^2(i\sigma)}, \tag{36}$$

$$\pi_i := \sigma + \sum_{j=0}^{i-1} \sigma q_j = \pi_{i-1} + \sigma q_{i-1}\mathbb{1}_{\{i\ge 1\}}, \tag{37}$$

making (24) and (26) rigorous (starting with $\pi_0 = \sigma > 0$ leads to cleaner formulae later on). With foresight, for $i \le I$ we also introduce the 'relative error' parameter

$$\tau_i := 1 - \frac{\delta \pi_i}{2\pi_I} = \tau_{i-1} - \frac{\delta\sigma q_{i-1}}{2\pi_I}\mathbb{1}_{\{i\ge 1\}}, \tag{38}$$

which slowly degrades from $\tau_0 = 1 - o(\delta)$ to $\tau_I = 1 - \delta/2$.

With an eye on Theorem 4, for concreteness we introduce the absolute constants[8]

$$D_0 := 108 \quad \text{and} \quad \beta_0 := 1/14, \tag{39}$$

as well as the set-sizes (with $s_0 \ll s$) and idealized edge-probability

$$s := \lceil C\sqrt{n \log n}\rceil, \qquad s_0 := \lfloor \sigma^4 q_I^2 s\rfloor, \qquad \text{and} \qquad \rho := \sqrt{\beta(\log n)/n}, \tag{40}$$

and, recalling $O_0 = E(H)$, the collection of 'relevant' pairs of large vertex-sets

$$\mathfrak{S}_{s,\gamma} := \{(A,B) :\ \text{disjoint } A,B \subseteq V \text{ with } |A| = |B| = s \text{ and } |O_0(A,B)| \ge \gamma|A||B|\}. \tag{41}$$

2.4 Main nibble result: pseudo-random properties

In this section we state our main nibble result Theorem 9, which implies our main tool Theorem 4 and establishes various pseudo-random properties of $(O_i, E_i, T_i, \Gamma_i)_{0\le i\le I}$. The following event is of core interest:

$$\mathcal{T}_I := \big\{|T_I(A,B)| = (1\pm\delta)\rho|O_0(A,B)| \text{ for all } (A,B) \in \mathfrak{S}_{s,\gamma}\big\}. \tag{42}$$

Indeed, it implies the conclusion of Theorem 4 with $G = (V, T_I)$ since the edge-set $T_I \subseteq E_I \subseteq E(H) = O_0$ is triangle-free. To get a handle on $\mathcal{T}_I$, in view of Section 2.1 it is natural that we also require some control over the other edge-sets $(E_i, O_i, \Gamma_i)_{0\le i\le I}$. To this end we introduce the 'good' events

$$\mathfrak{X}_i := \mathcal{N}_i \cap \mathcal{P}_i \cap \mathcal{Q}_i^+ \cap \mathcal{Q}_i \quad \text{and} \quad \mathfrak{X}_{\le i} := \bigcap_{0\le j\le i} \mathfrak{X}_j, \tag{43}$$

where the following auxiliary events encapsulate various pseudo-random properties:

$$\mathcal{N}_i := \big\{|N_{O_i}(v)| \le q_i n \text{ and } |N_{\Gamma_i}(v)| \le 2\sigma q_{i-1}\sqrt{n} \text{ for all } v \in V\big\}, \tag{44}$$

$$\mathcal{P}_i := \big\{|X_{uv}(i)| \le q_i^2 n,\ |Y_{uv}(i)| \le 2q_i\pi_i\sqrt{n}, \text{ and } |Z_{uv}(i)| \le i(\log n)^9 \text{ for all } u,v \in V \text{ with } u \ne v\big\}, \tag{45}$$

$$\mathcal{Q}_i^+ := \big\{|O_i(A,B)| \le q_i|A||B| \text{ for all disjoint } A,B \subseteq V \text{ with } |A|,|B| \ge s_0\big\}, \tag{46}$$

$$\mathcal{Q}_i := \big\{\tau_i q_i|O_0(A,B)| \le |O_i(A,B)| \le q_i|O_0(A,B)| \text{ for all } (A,B) \in \mathfrak{S}_{s,\gamma}\big\}. \tag{47}$$

In words, the above events give bounds for degree-like variables ($\mathcal{N}_i$), codegree-like variables ($\mathcal{P}_i$), and the number of open edges ($\mathcal{Q}_i^+$ and $\mathcal{Q}_i$). A subtle but important point is that $\mathcal{N}_i$, $\mathcal{P}_i$ and $\mathcal{Q}_i^+$ only guarantee one-sided concentration, i.e., ensure upper bounds but no matching lower bounds (which can fail badly, for example, $|Y_{uv}(i)| = 0$ holds when $uv \in E_i$). Merely $\mathcal{Q}_i$ guarantees two-sided concentration, which is harder to prove, but crucial for establishing the edge-estimate from $\mathcal{T}_I$ (see the heuristic below Theorem 9).

With $\tau_i \approx 1$ and $O_0 = E(H) \subseteq E(K_n)$ in mind, most of the bounds in (42) and (44)–(47) can easily be guessed by the pseudo-random heuristics (22) and (29) from Section 2.2 (the $|N_{\Gamma_i}(v)|$-bound is one exception: based on $\mathbb{E}\,|N_{\Gamma_i}(v)| = p \cdot \mathbb{E}\,|N_{O_{i-1}}(v)|$, it contains an extra factor of 2 to avoid additive error terms; another exception is the $|Z_{uv}(i)|$-bound: it relaxes the prediction $\mathbb{E}\,|Z_{uv}(i)| \lesssim \pi_i^2 = O(\log n)$ for technical reasons).

Inspecting (44)–(47) in the special case $i = 0$, it is not difficult to see that the good event $\mathfrak{X}_0 = \mathfrak{X}_{\le 0}$ always holds (by combining $q_0 = 1 \ge \tau_0$ and $\sigma, q_{-1}, \pi_0 \ge 0$ with $E_0 = T_0 = \Gamma_0 = \varnothing$).

[8]To make this paper easier to read, we have made no attempt to optimize the constants D_0, β_0 in (39).

Remark 8. *The event $\mathfrak{X}_0$ holds deterministically for any n-vertex host graph H.*

Our main nibble result (which is at the heart of this paper) states that, under fairly natural constraints, the pseudo-random events $\mathcal{T}_I$ and $\mathfrak{X}_{\le I}$ both hold with very high probability. Recall that $I \approx n^{\beta}$, and that all pairs $(A,B) \in \mathfrak{S}_{s,\gamma}$ of vertex-sets satisfy $|O_0(A,B)| \ge \gamma s^2$ and $|A| = |B| = s \approx C\sqrt{n \log n}$.

Theorem 9 (Main nibble result)**.** *For all $\gamma, \delta \in (0,1]$, $\beta \in (0, \beta_0)$ and $C \ge D_0/(\delta^2\sqrt{\beta}\gamma)$ the following holds for $n \ge n_0(\gamma,\delta,\beta,C)$: we have $\mathbb{P}(\mathcal{T}_I \cap \mathfrak{X}_{\le I}) \ge 1 - n^{-\omega(1)}$ for any n-vertex host graph H.*

Proof of Theorem 4. If the event $\mathcal{T}_I$ holds, then the triangle-free graph $G := (V, T_I)$ has the claimed properties by (42), $V = V(H)$ and $T_I \subseteq E_I \subseteq E(H) = O_0$, so Theorem 9 completes the proof. □

Remark 10. *In view of $I = O(n^{\beta_0})$ and Remark 7, the nibble thus yields a randomized polynomial time algorithm (with error probability $\le n^{-\omega(1)}$) for constructing the triangle-free $G \subseteq H$ from Theorem 4.*

Remark 11. *The heuristic edge-estimate* (29) *suggests that in Theorem 9 the dependence of the constant C on δ, β, γ is qualitatively best possible, since it would also naturally arise if $G = (V, T_I) \subseteq H$ was a random subgraph with edge-probability $\rho \approx \sqrt{\beta(\log n)/n}$. Indeed, for all $(A,B) \in \mathfrak{S}_{s,\gamma}$ the expected number of edges between A and B would then be at least $\lambda_{A,B} := \mathbb{E}\,|T_I(A,B)| = \rho|O_0(A,B)| \ge \rho\cdot\gamma s^2 \ge \sqrt{\beta}\gamma C\cdot s\log n$, and the probability that the event $\mathcal{T}_I$ from* (42) *fails would therefore be (using a union bound and standard Chernoff bounds) at most $\sum_{(A,B)\in\mathfrak{S}_{s,\gamma}} e^{-\Theta(\delta^2\lambda_{A,B})} \le n^{2s-\Omega(\delta^2\sqrt{\beta}\gamma C s)} = o(1)$ for $C = \Omega(1/(\delta^2\sqrt{\beta}\gamma))$ large enough.*

We defer the proof of Theorem 9 to Section 3, and now just outline a brief heuristic argument that illustrates how the event $\mathfrak{X}_{\le I} \subseteq \bigcap_{0\le i\le I} \mathcal{Q}_i$ is instrumental for establishing the edge-estimate from $\mathcal{T}_I$ (which seems informative). Similar to (29), in view of Section 2.1 we expect that in each step only few edges are removed due to the creation of triangles, which intuitively suggests

$$|T_{i+1}(A,B) \setminus T_i| \approx |E_{i+1}(A,B) \setminus E_i|.$$

Combining the construction of $E_{i+1} \setminus E_i = \Gamma_{i+1} \subseteq O_i$ with the event $\mathcal{Q}_i$ and $\tau_i \approx 1$, we also expect that

$$|E_{i+1}(A,B) \setminus E_i| = |\Gamma_{i+1}(A,B)| \approx p \cdot |O_i(A,B)| \approx p \cdot q_i |O_0(A,B)|.$$

Recalling $p = \sigma/\sqrt{n}$ and $\rho = \sqrt{\beta(\log n)/n}$, using the definition (37) of π_I and the approximation $\pi_I \approx \sqrt{\beta \log n}$ from (28) it now becomes plausible that

$$|T_I(A,B)| = \sum_{0\le i<I} |T_{i+1}(A,B) \setminus T_i| \approx \frac{\sum_{0\le i<I} \sigma q_i}{\sqrt{n}} \cdot |O_0(A,B)| \approx \frac{\pi_I}{\sqrt{n}} \cdot |O_0(A,B)| \approx \rho|O_0(A,B)|,$$

as suggested by $\mathcal{T}_I$ (Section 3.5 contains a rigorous version of this heuristic argument).

2.5 Tools and auxiliary estimates

In this preparatory section we gather, for later reference, some results that will be used throughout the proof of Theorem 9 (mostly probabilistic and combinatorial tools, and ending with some auxiliary estimates). On a first reading the reader may perhaps wish to skip straight to Section 3.

We start with a convenient version of the bounded differences inequality [23, 24, 36] for Bernoulli variables. Note that the upper tail estimate (48) for decreasing functions does *not* have an extra Ct term in the exponent like (49). Remarks 13–14 are well-known, see, e.g., [24, Theorem 2.3, 3.8, and 3.9] or [36, Corollary 1.4]. Inequality (48) can be deduced from the arguments in [23, Lemma 7.14], but this monotone version does not seem to be widely known; in Appendix A we thus include a simple proof for completeness.

Theorem 12. *Let $(\xi_\alpha)_{\alpha\in\mathcal{I}}$ be a finite family of independent random variables with $\xi_\alpha \in \{0,1\}$. Let $f : \{0,1\}^{|\mathcal{I}|} \to \mathbb{R}$ be a function, and assume that there exist numbers $(c_\alpha)_{\alpha\in\mathcal{I}}$ such that the following holds for all $z = (z_\alpha)_{\alpha\in\mathcal{I}} \in \{0,1\}^{|\mathcal{I}|}$ and $z' = (z'_\alpha)_{\alpha\in\mathcal{I}} \in \{0,1\}^{|\mathcal{I}|}$: $|f(z) - f(z')| \le c_\beta$ if $z_\alpha = z'_\alpha$ for all $\alpha \ne \beta$. Define $X := f\big((\xi_\alpha)_{\alpha\in\mathcal{I}}\big)$ and $\lambda := \sum_{\alpha\in\mathcal{I}} c_\alpha^2\, \mathbb{P}(\xi_\alpha = 1)$. Then, for all $t \ge 0$,*

$$\mathbb{P}(X \ge \mathbb{E}\,X + t) \le \exp\left(-\frac{t^2}{2\lambda}\right) \tag{48}$$

if the function f is decreasing (i.e., that $f(z) \le f(z')$ whenever $z_\alpha \ge z'_\alpha$ for all $\alpha \in \mathcal{I}$).

Remark 13. *Define* $C := \max_{\alpha \in \mathcal{I}} c_\alpha$*. If we drop the assumption that* f *is decreasing, then*

$$\mathbb{P}(X \le \mathbb{E}\, X - t) \le \exp\Big(-\frac{t^2}{2(\lambda + Ct)}\Big). \tag{49}$$

Remark 14. *In the special case* $X = \sum_{\alpha \in \mathcal{I}} \xi_\alpha$ *we have* $C = c_\alpha = 1$ *and* $\lambda = \mathbb{E}\, X$*. Standard Chernoff bounds (or applying (48)-(49) to the decreasing function* $-X$*) then show that in this case* $\mathbb{P}(X \le \mathbb{E}\, X - t)$ *and* $\mathbb{P}(X \ge \mathbb{E}\, X + t)$ *are at most the right hand side of (48) and (49), respectively.*

For random variables with a special combinatorial form (based on the occurrence of events with 'limited overlaps') we shall use the following Chernoff-type upper tail inequality, which is a convenient corollary of a more general result by Warnke [37, Theorem 9]. Note that the exponent of (50) scales with $1/C$.

Theorem 15. *Let* $(\xi_i)_{i \in \mathfrak{S}}$ *be a finite family of independent random variables with* $\xi_i \in \{0, 1\}$*. Let* $(Y_\alpha)_{\alpha \in \mathcal{I}}$ *be a finite family of variables* $Y_\alpha := \mathbb{1}_{\{\xi_i = 1 \text{ for all } i \in \alpha\}}$ *with* $\sum_{\alpha \in \mathcal{I}} \mathbb{E}\, Y_\alpha \le \mu$*. Define* $Z_C := \max \sum_{\alpha \in \mathcal{J}} Y_\alpha$*, where the maximum is taken over all* $\mathcal{J} \subseteq \mathcal{I}$ *with* $\max_{\beta \in \mathcal{J}} |\{\alpha \in \mathcal{J} : \alpha \cap \beta \neq \varnothing\}| \le C$*. Then, for all* $C, t > 0$*,*

$$\mathbb{P}(Z_C \ge \mu + t) \le \min\bigg\{\Big(\frac{e\mu}{\mu + t}\Big)^{(\mu + t)/C}, \; \exp\Big(-\frac{t^2}{2C(\mu + t)}\Big)\bigg\}. \tag{50}$$

The following simple combinatorial lemma formalizes the intuition that we expect $\sum_i |U_i| = O(|U|)$ whenever the subsets $U_i \subseteq U$ are nearly disjoint (i.e., have small pairwise intersections).

Lemma 16. *Suppose that* $(U_i)_{i \in \mathcal{I}}$ *is a family of subsets* $U_i \subseteq U$ *with* $|U_i| \ge z > 0$ *and* $|U_i \cap U_j| \le y$ *for all* $i \neq j$*. Then* $z \ge \sqrt{4|U|y}$ *implies* $|\mathcal{I}| \le 2|U|/z$ *and* $\sum_{i \in \mathcal{I}} |U_i| \le 2|U|$*.*

Proof. Aiming at a contradiction, suppose that $|\mathcal{I}| > 2|U|/z$. Then there is $\mathcal{J} \subseteq \mathcal{I}$ with $|\mathcal{J}| = \lfloor 2|U|/z \rfloor + 1$. Observe that, for any $i \in \mathcal{J}$,

$$\sum_{j \in \mathcal{J} : i \neq j} |U_j \cap U_i| \le (|\mathcal{J}| - 1)y \le 2|U|y/z \le z/2 \le |U_i|/2. \tag{51}$$

So we obtain a contradiction by noting that

$$|U| \ge \Big|\bigcup_{i \in \mathcal{J}} U_i\Big| \ge \sum_{i \in \mathcal{J}} \Big(|U_i| - \sum_{j \in \mathcal{J} : i \neq j} |U_j \cap U_i|\Big) \ge \sum_{i \in \mathcal{J}} |U_i|/2 \ge |\mathcal{J}|z/2 > |U|. \tag{52}$$

With $|\mathcal{I}| \le 2|U|/z$ in hand, after replacing $\mathcal{J}$ with $\mathcal{I}$, note that (51) and the first three inequalities of (52) remain valid, completing the proof of $\sum_{i \in \mathcal{I}} |U_i| \le 2|U|$. □

Our final auxiliary result contains a number of convenient estimates involving the parameters q_i, π_i, σ, I defined in Section 2.3 and (9). Roughly put, (55)–(57) state that $q_i \approx q_{i+1}$, $1 - 2\sigma q_i \pi_i \approx q_{i+1}/q_i$ and $\pi_I \approx \sqrt{\log(I\sigma)}$, as predicted by (25) and (28). The technical estimates (53)–(54) can safely be ignored on a first reading. The proof of Lemma 17 is based on elementary calculus and thus deferred to Appendix A (it also establishes $q_i \ge q_I = n^{-\beta + o(1)}$, which together with $I \approx n^\beta$ and (54) motivates our choice of $\beta_0 = 1/14$).

Lemma 17. *If* $\beta \in (0, \beta_0)$*, then there exists* $\tau, n_0 > 0$ *such that, for all* $n \ge n_0$ *and* $0 \le i \le I$*,*

$$\max\Big\{\max_{j \in \{0,1,2\}} \big\{q_i \pi_i^j\big\}, \; \sqrt{\sigma}\pi_i\Big\} \le 1, \tag{53}$$

$$\min\Big\{\min_{j \in [4]} \big\{q_i^j \sqrt{n}\big\}, \; q_i^2 \sqrt{n}/I, \; q_i^3 \sqrt[4]{n}/\sqrt{I}\Big\} \ge n^\tau, \tag{54}$$

$$0 \le q_i - q_{i+1} \le 4\sigma \cdot \min\{q_i, \; q_{i+1}, \; q_i \pi_i\}, \tag{55}$$

$$\big|(1 - 2\sigma q_i \pi_i) - q_{i+1}/q_i\big| \le 8\sigma^2 q_i, \tag{56}$$

$$\big|\pi_I - \sqrt{\log(I\sigma)}\big| \le 2. \tag{57}$$

As a simple application, for $0 \le i \le I$ we now bound the stabilization probability $\hat{p}_{e,i}$ defined in (20). Since (54) implies $q_i \sqrt{\sigma}\sqrt{n} \gg 1$, by applying $(1 - p)^r \ge 1 - pr = 1 - \sigma r/\sqrt{n}$ (valid for $r \ge 1$) we infer

$$\hat{p}_{e,i} \le 1 - (1 - p)^{2q_i(\pi_i + \sqrt{\sigma})\sqrt{n}} \le 2\sigma q_i(\pi_i + \sqrt{\sigma}) \le q_i, \tag{58}$$

where we used $\sqrt{\sigma}\pi_i \le 1$ and $\sigma \ll 1$ (see (53) and (35)) for the last inequality.

3 Analyzing the nibble

In this section we prove our main nibble result Theorem 9 (which in turn implies our main tool Theorem 4, see Section 2.4) as a corollary of the following auxiliary lemma.

Lemma 18. *Under the assumptions of Theorem 9, for $n \ge n_0(\gamma,\delta,\beta,C)$ we have*

$$\mathbb{P}(\neg\mathfrak{X}_{i+1} \mid \mathfrak{X}_{\le i}) \le n^{-\omega(1)} \qquad \text{for all } 0 \le i \le I-1, \tag{59}$$

$$\mathbb{P}(\neg\mathcal{T}_I \cap \mathfrak{X}_{\le I}) \le n^{-\omega(1)}. \tag{60}$$

Proof of Theorem 9. Recalling $I \le \lceil n^{\beta_0} \rceil = n^{O(1)}$ and $\mathfrak{X}_{\le i} = \bigcap_{0\le j\le i} \mathfrak{X}_j$, note that $\mathbb{P}(\neg\mathfrak{X}_0) = 0$ (see Remark 8) and (59) readily imply $\mathbb{P}(\neg\mathfrak{X}_{\le I}) \le n^{-\omega(1)}$, which together with (60) completes the proof. □

The remainder of this section is devoted to the proof of Lemma 18: the proof of (59) with $\neg\mathfrak{X}_{i+1} = \neg\mathcal{N}_{i+1} \cup \neg\mathcal{P}_{i+1} \cup \neg\mathcal{Q}^+_{i+1} \cup \neg\mathcal{Q}_{i+1}$ is spread across Sections 3.2–3.4, and the proof of (60) is given in Section 3.5.

3.1 Preliminaries: setup and conventions

To avoid clutter, up to (and including) Section 3.4 we shall suppress the conditioning in the notation: we will write $\mathbb{P}(\cdot)$ and $\mathbb{E}(\cdot)$ as shorthand for $\mathbb{P}(\cdot \mid \mathcal{F}_i)$ and $\mathbb{E}(\cdot \mid \mathcal{F}_i)$, where $(\mathcal{F}_i)_{0\le i\le I}$ denotes the natural filtration associated with $(O_i, E_i, T_i, \Gamma_i, S_i)_{0\le i\le I}$, as usual. We will also tacitly assume that the $\mathcal{F}_i$-measurable event $\mathfrak{X}_{\le i}$ holds. Conditional on $\mathcal{F}_i$, note that by construction of the random edge-sets Γ_{i+1} and S_{i+1}, the (conditional) probability space formally consists of the $2|O_i|$ independent Bernoulli random variables $(\mathbb{1}_{\{e\in\Gamma_{i+1}\}}, \mathbb{1}_{\{e\in S_{i+1}\}})_{e\in O_i}$, with $\mathbb{P}(e \in \Gamma_{i+1}) = p = \sigma/\sqrt{n}$ and $\mathbb{P}(e \in S_{i+1}) = \hat{p}_{e,i} \le q_i$, see (58).

Using the above setup and conventions, we shall repeatedly consider random variables of form

$$X = f\Big(\big(\mathbb{1}_{\{e\in\Gamma_{i+1}\}}, \mathbb{1}_{\{e\in S_{i+1}\}}\big)_{e\in O_i}\Big). \tag{61}$$

To get a handle on the (conditional) expectation $\mathbb{E}\, X$ we will often use $O_{i+1} \subseteq O_i \setminus C_{i+1}$ together with the following key lemma, which hinges on the stabilization mechanism to equalize all (conditional) probabilities $\mathbb{P}(e \notin C_{i+1})$, see (62) below. (The extra $\sqrt{\sigma}$ term in (20) ensures that $\mathbb{P}(e \notin C_{i+1}) < q_{i+1}/q_i$ holds with plenty of elbow room, which is convenient for avoiding ugly error terms in the upper bounds of (44)–(47).)

Lemma 19. *We have $\mathbb{P}(e \notin C_{i+1}) - q_{i+1}/q_i \in [-3\sigma^{3/2}q_i,\ -\sigma^{3/2}q_i]$ for all $e \in O_i$.*

Proof. For any $e \in O_i$, since $|Y_e(i)| \le 2q_i\pi_i\sqrt{n}$ by $\mathfrak{X}_{\le i} \subseteq \mathcal{P}_i$, by definition of $C_{i+1} = C^{(1)}_{i+1} \cup S_{i+1}$ we have

$$\mathbb{P}(e \notin C_{i+1}) = \mathbb{P}(e \notin C^{(1)}_{i+1}) \cdot \mathbb{P}(e \notin S_{i+1}) = (1-p)^{|Y_e(i)|} \cdot (1-\hat{p}_{e,i}) = (1-p)^{2q_i(\pi_i+\sqrt{\sigma})\sqrt{n}}. \tag{62}$$

It is well-known (and easy to check) that $1 - rx \le (1-x)^r \le 1 - rx + \binom{r}{2}x^2$ for all $x \in [0,1]$ and $r \ge 2$. Using $\sqrt{n}p = \sigma \ll 1$ and $\max\{q_i, q_i\pi_i, q_i\pi_i^2\} \le 1$ (see (53)), it follows that

$$\Big|\mathbb{P}(e \notin C_{i+1}) - \big[1 - 2\sigma q_i(\pi_i + \sqrt{\sigma})\big]\Big| \le 2\sigma^2 q_i^2(\pi_i+\sqrt{\sigma})^2 = O(\sigma^2 q_i) = o(\sigma^{3/2}q_i).$$

This completes the proof since $1 - 2\sigma q_i\pi_i = q_{i+1}/q_i + o(\sigma^{3/2}q_i)$ by (56). □

To deduce concentration properties of such random variables X we shall frequently rely on the bounded differences inequality (Theorem 12). In those cases we will bound the associated parameter λ via

$$\lambda = \sum_{e\in O_i} c_e^2\, \mathbb{P}(e \in \Gamma_{i+1}) + \sum_{e\in O_i} \hat{c}_e^2\, \mathbb{P}(e \in S_{i+1}) \le p \sum_{e\in O_i} c_e^2 + q_i \sum_{e\in O_i} \hat{c}_e^2, \tag{63}$$

where the *edge-effect* c_e is an upper bound on how much X can change if we modify the indicator $\mathbb{1}_{\{e\in\Gamma_{i+1}\}}$ (alter whether e is in Γ_{i+1} or not), and the *stabilization-effect* $\hat{c}_e$ is an upper bound on how much X can change if we modify the indicator $\mathbb{1}_{\{e\in S_{i+1}\}}$ (alter whether e is in S_{i+1} or not). Moreover, the following simple observation will later allow us to control the above sum (63) of these effects.

Lemma 20. *If $\mathfrak{X}_{\le i}$ holds, then $\sum_{e\in O_i}|Y_e(i)\cap J| \le 2q_i\pi_i\sqrt{n}\cdot|J|$ for any edge-subset $J\subseteq \binom{V}{2}$.*

Proof. For any $e\in O_i$, note that $f\in Y_e(i)$ implies $e\in Y_f(i)$. As $Y_f(i)\subseteq O_i$, we infer

$$\sum_{e\in O_i}|Y_e(i)\cap J| = \sum_{f\in J}\sum_{e\in O_i}\mathbb{1}_{\{f\in Y_e(i)\}} \le \sum_{f\in J}\sum_{e\in O_i}\mathbb{1}_{\{e\in Y_f(i)\}} = \sum_{f\in J}|Y_f(i)|.$$

This completes the proof since $\mathfrak{X}_{\le i}\subseteq \mathcal{P}_i$ implies $|Y_f(i)|\le 2q_i\pi_i\sqrt{n}$. □

3.2 Event $\mathcal{N}_{i+1}$: degree-like variables $|N_{O_{i+1}}(v)|$ and $|N_{\Gamma_{i+1}}(v)|$

Recall that the event $\mathcal{N}_{i+1}$ defined in (44) concerns degree-like variables, ensuring that $|N_{O_{i+1}}(v)|\le q_{i+1}n$ and $|N_{\Gamma_{i+1}}(v)|\le 2\sigma q_i\sqrt{n}$ for all vertices v; see (31) for the definition of $N_S(v)$.

Lemma 21. *We have $\mathbb{P}(\neg\mathcal{N}_{i+1})\le n^{-\omega(1)}$.*

Proof. We start with $|N_{O_{i+1}}(v)|$. Note that $O_{i+1}\subseteq O_i\setminus C_{i+1}$ implies

$$|N_{O_{i+1}}(v)| \le \sum_{w\in N_{O_i}(v)}\mathbb{1}_{\{vw\notin C_{i+1}\}} =: X. \tag{64}$$

Since $\mathfrak{X}_{\le i}\subseteq\mathcal{N}_i$ implies $|N_{O_i}(v)|\le q_in$, using Lemma 19 we obtain

$$\mathbb{E}\,X = \sum_{w\in N_{O_i}(v)}\mathbb{P}(vw\notin C_{i+1}) \le |N_{O_i}(v)|\cdot(q_{i+1}/q_i-\sigma^{3/2}q_i) \le q_{i+1}n-\sigma^{3/2}q_i^2n. \tag{65}$$

Gearing up to apply Theorem 12 to X, we now bound the associated parameter $\lambda\le p\sum_{e\in O_i}c_e^2+q_i\sum_{e\in O_i}\hat{c}_e^2$ from (63). Set $\mathcal{X}_v:=\{v\}\times N_{O_i}(v)\subseteq O_i$, and recall that $C_{i+1}=C_{i+1}^{(1)}\cup S_{i+1}$, where $C_{i+1}^{(1)}$ depends only on Γ_{i+1} and thus is independent of S_{i+1}. The edge-effect c_e (an upper bound on how much X changes if we alter whether $e\in\Gamma_{i+1}$ or $e\notin\Gamma_{i+1}$) is thus at most the number of changes to $C_{i+1}^{(1)}\cap\mathcal{X}_v=\{vw\in\mathcal{X}_v : Y_{vw}(i)\cap\Gamma_{i+1}\neq\varnothing\}$. Since $e\in Y_{vw}(i)$ implies $vw\in Y_e(i)$ when $vw\in\mathcal{X}_v$, we infer $c_e\le|Y_e(i)\cap\mathcal{X}_v|\le|Y_e(i)|\le 2q_i\pi_i\sqrt{n}$ by $\mathfrak{X}_{\le i}\subseteq\mathcal{P}_i$. Using Lemma 20, $|\mathcal{X}_v|=|N_{O_i}(v)|\le q_in$, and $q_i\pi_i^2\le 1$ (see (53)), it follows that

$$p\sum_{e\in O_i}c_e^2 \le p\cdot 2q_i\pi_i\sqrt{n}\cdot\sum_{e\in O_i}|Y_e(i)\cap\mathcal{X}_v| \le \sigma/\sqrt{n}\cdot(2q_i\pi_i\sqrt{n})^2\cdot|\mathcal{X}_v| \le 4\sigma q_i^3\pi_i^2n^{3/2}\le 4\sigma q_i^2n^{3/2}. \tag{66}$$

By our above discussion, the stabilization-effect $\hat{c}_e$ (an upper bound on how much X changes if we alter whether $e\in S_{i+1}$ or $e\notin S_{i+1}$) is at most the number of changes to $S_{i+1}\cap\mathcal{X}_v$. Hence $\hat{c}_e\le\mathbb{1}_{\{e\in\mathcal{X}_v\}}$, so that

$$q_i\sum_{e\in O_i}\hat{c}_e^2 \le q_i\cdot|\mathcal{X}_v| \le q_i^2n \ll \sigma q_i^2n^{3/2}.$$

Noting that X is a decreasing function of the edge-indicators $(\mathbb{1}_{\{e\in\Gamma_{i+1}\}},\mathbb{1}_{\{e\in S_{i+1}\}})_{e\in O_i}$, using Theorem 12 together with the λ–bound (63) and $q_i^2n^{1/2}\ge n^\tau$ (see (54)) it follows that

$$\mathbb{P}(|N_{O_{i+1}}(v)|\ge q_{i+1}n) \le \mathbb{P}(X\ge\mathbb{E}\,X+\sigma^{3/2}q_i^2n) \le \exp\left(-\frac{\sigma^3q_i^4n^2}{2\cdot 5\sigma q_i^2n^{3/2}}\right)\le n^{-\omega(1)}.$$

Taking a union bound over all vertices v completes the proof for the $|N_{O_{i+1}}(v)|$ variables.

Finally, note that $|N_{\Gamma_{i+1}}(v)|$ is a sum of independent Bernoulli random variables with

$$\mathbb{E}\,|N_{\Gamma_{i+1}}(v)| = |N_{O_i}(v)|\cdot p \le q_in\cdot\sigma/\sqrt{n} = \sigma q_i\sqrt{n} =: \mu,$$

where we used $\mathfrak{X}_{\le i}\subseteq\mathcal{N}_i$ to bound $|N_{O_i}(v)|\le q_in$. Applying standard Chernoff bounds (see, e.g., Remark 14), using $q_i\sqrt{n}\ge n^\tau$ (see (54)) it is routine to deduce that $\mu\gg\log n$ and

$$\mathbb{P}(|N_{\Gamma_{i+1}}(v)|\ge 2\sigma q_i\sqrt{n}) = \mathbb{P}(|N_{\Gamma_{i+1}}(v)|\ge 2\mu) \le \exp\left(-\frac{\mu^2}{2\cdot 2\mu}\right) = \exp\left(-\frac{\mu}{4}\right)\le n^{-\omega(1)}.$$

Taking a union bound over all vertices v completes the proof for the $|N_{\Gamma_{i+1}}(v)|$ variables. □

3.3 Event $\mathcal{P}_{i+1}$: codegree-like variables $|X_{uv}(i+1)|$, $|Y_{uv}(i+1)|$ and $|Z_{uv}(i+1)|$

Recall that the event $\mathcal{P}_{i+1}$ defined in (45) concerns codegree-like variables, ensuring that $|X_{uv}(i+1)| \le q_{i+1}^2 n$, $|Y_{uv}(i+1)| \le 2q_{i+1}\pi_{i+1}\sqrt{n}$, and $|Z_{uv}(i+1)| \le (i+1)(\log n)^9$ for all pairs uv of vertices.

Lemma 22. *We have* $\mathbb{P}(\neg\mathcal{P}_{i+1}) \le n^{-\omega(1)}$.

Proof. We start with $|X_{uv}(i+1)|$. Recalling $O_{i+1} \subseteq O_i \setminus C_{i+1}$, note that by construction

$$|X_{uv}(i+1)| \le \sum_{w \in X_{uv}(i)} \mathbb{1}_{\{uw \notin C_{i+1} \text{ and } vw \notin C_{i+1}\}} =: X. \tag{67}$$

To estimate $\mathbb{E}\,X$, note that the event $f \notin C_{i+1}^{(1)} = \{f \in O_i : Y_f(i) \cap \Gamma_{i+1} \neq \varnothing\}$ is determined by the values of the independent indicator variables $(\mathbb{1}_{\{e\in\Gamma_{i+1}\}})_{e \in Y_f(i)}$. In view of the reasoning (62) for the value of $\mathbb{P}(e \notin C_{i+1})$, it follows by construction of $C_{i+1} = C_{i+1}^{(1)} \cup S_{i+1}$ that

$$\mathbb{P}(uw \notin C_{i+1} \text{ and } vw \notin C_{i+1}) = \mathbb{P}(uw \notin C_{i+1})\,\mathbb{P}(vw \notin C_{i+1}) \cdot (1-p)^{-|Y_{uw}(i) \cap Y_{vw}(i)|}. \tag{68}$$

Since $\mathfrak{X}_{\le i} \subseteq \mathcal{P}_i$ implies $|Y_{uw}(i) \cap Y_{vw}(i)| \le |Z_{uv}(i)| \le I(\log n)^9$ and $|X_{uv}(i)| \le q_i^2 n$, by combining (68) with Lemma 19 it follows that

$$\mathbb{E}\,X \le |X_{uv}(i)| \cdot (q_{i+1}/q_i - \sigma^{3/2} q_i)^2 \cdot (1-p)^{-I(\log n)^9} \le q_{i+1}^2 n - \sigma^{3/2} q_i^3 n, \tag{69}$$

where for the last inequality we used $pI(\log n)^9 \ll \sigma^{3/2} q_i^3 \ll 1$ (since $q_i^3\sqrt{n}/I \ge n^\tau$ by (54)) and $\sigma^3 q_i^4 \ll \sigma^{3/2} q_i^3 \sim \sigma^{3/2} q_{i+1} q_i^2$ (see (53)–(55)). With an eye on Theorem 12, we now bound the parameter $\lambda \le p\sum_{e \in O_i} c_e^2 + q_i \sum_{e\in O_i} \hat{c}_e^2$ from (63). Set $\mathcal{X}_{uv} := \{u,v\} \times X_{uv}(i) \subseteq O_i$. Analogous to the proof of Lemma 21 for $|N_{O_{i+1}}(v)|$, here we have edge-effect $c_e \le |Y_e(i) \cap \mathcal{X}_{uv}| \le |Y_e(i)| \le 2q_i\pi_i\sqrt{n}$ and stabilization-effect $\hat{c}_e \le \mathbb{1}_{\{e \in \mathcal{X}_{uv}\}}$. Similar to (66), using Lemma 20, $|\mathcal{X}_{uv}| = 2\cdot|X_{uv}(i)| \le 2q_i^2 n$ and $q_i\pi_i^2 \le 1$ it follows that

$$p\sum_{e \in O_i} c_e^2 \le \sigma/\sqrt{n} \cdot (2q_i\pi_i\sqrt{n})^2 \cdot |\mathcal{X}_{uv}| \le 8\sigma q_i^4 \pi_i^2 n^{3/2} \le 8\sigma q_i^3 n^{3/2}. \tag{70}$$

Furthermore, $q_i \sum \hat{c}_e^2 \le q_i|\mathcal{X}_{uv}| \le 2q_i^3 n \ll \sigma q_i^3 n^{3/2}$. Noting that X is a decreasing function of the edge-indicators $(\mathbb{1}_{\{e\in\Gamma_{i+1}\}}, \mathbb{1}_{\{e \in S_{i+1}\}})_{e\in O_i}$, using Theorem 12 and $q_i^3 n^{1/2} \ge n^\tau$ (see (54)) it follows that

$$\mathbb{P}(|X_{uv}(i+1)| \ge q_{i+1}^2 n) \le \mathbb{P}(X \ge \mathbb{E}\,X + \sigma^{3/2} q_i^3 n) \le \exp\left(-\frac{\sigma^3 q_i^6 n^2}{2\cdot 9\sigma q_i^3 n^{3/2}}\right) \le n^{-\omega(1)}.$$

Taking a union bound over all pairs of vertices u, v completes the proof for the $|X_{uv}(i+1)|$ variables.

Turning to the more involved $|Y_{uv}(i+1)|$ variables, note that by construction

$$|Y_{uv}(i+1)| \le \sum_{w \in X_{uv}(i)} \mathbb{1}_{\{uw \in \Gamma_{i+1} \text{ or } vw \in \Gamma_{i+1}\}} + \sum_{f \in Y_{uv}(i)} \mathbb{1}_{\{f \notin C_{i+1}\}} =: Y_{uv}^+ + Y_{uv}^*. \tag{71}$$

(To clarify: Y_{uv}^+ and Y_{uv}^* are defined by the first and second sum in (71), respectively.) Using Lemma 19 together with $\sigma q_i^2 = \sigma q_i q_{i+1} + o(\sigma^{3/2} q_i^2 \pi_i)$ (see (55)) and $\pi_i q_{i+1} = q_{i+1}\pi_{i+1} - \sigma q_i q_{i+1}$ (as $\pi_{i+1} = \pi_i + \sigma q_i$ by (37)), it follows that

$$\begin{aligned}\mathbb{E}(Y_{uv}^+ + Y_{uv}^*) &\le |X_{uv}(i)| \cdot 2p \;+\; |Y_{uv}(i)| \cdot (q_{i+1}/q_i - \sigma^{3/2} q_i)\\ &\le 2\sigma q_i^2\sqrt{n} + 2\pi_i\sqrt{n}(q_{i+1} - \sigma^{3/2} q_i^2) \le 2q_{i+1}\pi_{i+1}\sqrt{n} - \sigma^{3/2} q_i^2 \pi_i \sqrt{n}.\end{aligned} \tag{72}$$

We now estimate Y_{uv}^+ and Y_{uv}^* separately. Noting $\mathbb{E}\,Y_{uv}^+ \le 2\sigma q_i^2\sqrt{n}$ and $\sigma^2 q_i^2 \pi_i \sqrt{n} = o(\sigma q_i^2 \sqrt{n})$ (see (53)), using standard Chernoff bounds together with $\pi_i^2 \ge \pi_0^2 = \sigma^2$ and $q_i^2\sqrt{n} \ge n^\tau$ (see (54)) it follows that

$$\mathbb{P}(Y_{uv}^+ \ge \mathbb{E}\,Y_{uv}^+ + \sigma^2 q_i^2 \pi_i \sqrt{n}) \le \exp\left(-\frac{\left(\sigma^2 q_i^2 \pi_i \sqrt{n}\right)^2}{4\cdot 2\sigma q_i^2\sqrt{n}}\right) \le \exp\left(-\frac{\sigma^5 q_i^2\sqrt{n}}{8}\right) \le n^{-\omega(1)}. \tag{73}$$

For Y^*_{uv} we shall apply Theorem 12, and we thus now bound $\lambda \le p\sum_{e\in O_i} c_e^2 + q_i \sum_{e\in O_i} \hat{c}_e^2$ from (63). As usual, we have edge-effect $c_e \le |Y_e(i) \cap Y_{uv}(i)| \le |Y_e(i)| \le 2q_i\pi_i\sqrt{n}$ and stabilization-effect $\hat{c}_e \le \mathbb{1}_{\{e\in Y_{uv}(i)\}}$. Here we can significantly improve the simple worst case estimate $c_e \le |Y_e(i)|$ when $e \neq uv$. Indeed, if $e = w_1w_2$ does not intersect uv, then $c_e \le 4$ since $Y_e(i) \cap Y_{uv}(i) \subseteq \{u,v\} \times \{w_1,w_2\}$, say. Furthermore, if $e = w_1w_2$ intersects uv in one vertex, say $u = w_1$, then $c_e \le \max_f |Z_f(i)| \le I(\log n)^9$ since $Y_e(i) \cap Y_{uv}(i) \subseteq \{u\} \times [N_{E_i}(w_2) \cap N_{E_i}(v)]$. To sum up, for $e \neq uv$ we have $c_e \le \max\{4, I(\log n)^9\} \le \sigma^{-5}I$, say. Similar to (66) and (70), using Lemma 20 and $|Y_{uv}(i)| \le 2q_i\pi_i\sqrt{n}$ it follows that

$$p\sum_{e\in O_i} c_e^2 \le \sigma/\sqrt{n} \cdot \Big((2q_i\pi_i\sqrt{n})^2 + \sigma^{-5}I \cdot 2q_i\pi_i\sqrt{n} \cdot |Y_{uv}(i)|\Big) \le 8\sigma^{-4}q_i^2\pi_i^2 I\sqrt{n}.$$

Furthermore, using $\pi_i \ge \sigma$ and $I \ge 1$ we obtain $q_i\sum_{e\in O_i} \hat{c}_e^2 \le q_i|Y_{uv}(i)| \le 2q_i^2\pi_i\sqrt{n} \ll \sigma^{-4}q_i^2\pi_i^2 I\sqrt{n}$. Noting that Y^*_{uv} is decreasing, using Theorem 12 and $q_i^2\sqrt{n}/I \ge n^{\tau}$ (see (54)) it follows that

$$\mathbb{P}(Y^*_{uv} \ge \mathbb{E}\, Y^*_{uv} + \sigma^2 q_i^2\pi_i\sqrt{n}) \le \exp\left(-\frac{\sigma^4 q_i^4\pi_i^2 n}{2\cdot 9\sigma^{-4}q_i^2\pi_i^2 I\sqrt{n}}\right) \le n^{-\omega(1)}. \tag{74}$$

Combining the probability estimates (73) and (74) with inequalities (71)–(72) and $\sigma^2 \ll \sigma^{3/2}$, now a union bound argument (to account for all pairs of vertices u,v) completes the proof for the $|Y_{uv}(i+1)|$ variables.

Finally, for $|Z_{uv}(i+1)|$ note that the one-step difference

$$\Delta Z := |Z_{uv}(i+1)| - |Z_{uv}(i)| = \sum_{w\in X_{uv}(i)} \mathbb{1}_{\{uw\in\Gamma_{i+1} \text{ and } vw\in\Gamma_{i+1}\}} + \sum_{f\in Y_{uv}(i)} \mathbb{1}_{\{f\in\Gamma_{i+1}\}} \tag{75}$$

is a sum of independent Bernoulli random variables with

$$\mathbb{E}(\Delta Z) = |X_{uv}(i)| \cdot p^2 + |Y_{uv}(i)| \cdot p \le \sigma^2 q_i^2 + 2\sigma q_i\pi_i \le 3\sigma \ll 1, \tag{76}$$

where we used $|X_{uv}(i)| \le q_i^2 n$ and $|Y_{uv}(i)| \le 2q_i\pi_i\sqrt{n}$ for the first inequality, and $\max\{q_i^2, q_i\pi_i\} \le 1$ (see (53)) and $\sigma \ll 1$ for the last two inequalities. Inspecting (75), note that $\mathfrak{X}_{\le i} \subseteq \mathcal{P}_i$ implies $|Z_{uv}(i+1)| \le \Delta Z + i(\log n)^9$. Applying standard Chernoff bounds, using $\mathbb{E}(\Delta Z) \ll 1$ it readily follows that, say,

$$\mathbb{P}\big(|Z_{uv}(i+1)| \ge (i+1)(\log n)^9\big) \le \mathbb{P}\big(\Delta Z \ge (\log n)^9\big) \le n^{-\omega(1)}.$$

Taking a union bound over all pairs of vertices u,v completes the proof for the $|Z_{uv}(i+1)|$ variables. $\square$

Remark 23. *If desired, it would not be difficult to establish the better upper bound $|Z_{uv}(i)| \le (\log n)^2$, say (using the stochastic domination arguments leading to (95) in Section 3.5; in view (75)–(76) the main point is that, for $0 \le i \le I$, the event $\mathfrak{X}_{\le i}$ implies $\sum_{0\le j\le i}(|X_{uv}(j)|p^2 + |Y_{uv}(j)|p) = O(\log n)$). This in turn could, e.g., be used to increase the constant β_0 slightly (as we could then remove $I = \lceil n^{\beta}\rceil$ from constraint (54)).*

3.4 Event $\mathcal{Q}^+_{i+1} \cap \mathcal{Q}_{i+1}$: number $|O_{i+1}(A,B)|$ of open edges between large sets

Recall that the events $\mathcal{Q}^+_{i+1}, \mathcal{Q}_{i+1}$ defined in (46)–(47) concern the open edge-set $O_{i+1} \subseteq E(H) = O_0$, ensuring that $|O_{i+1}(A,B)| \le q_{i+1}|A||B|$ for all disjoint $A,B \subseteq V$ with $|A|,|B| \ge s_0$, and $\tau_{i+1}q_{i+1}|O_0(A,B)| \le |O_{i+1}(A,B)| \le q_{i+1}|O_0(A,B)|$ for all $(A,B) \in \mathfrak{S}_{s,\gamma}$; see (40)–(41) for the definition of s_0 and $\mathfrak{S}_{s,\gamma}$.

Turning to $|O_{i+1}(A,B)|$, note that one edge $e \in \Gamma_{i+1}$ can add up to $|Y_e(i) \cap O_i(A,B)| \le \sum_{w\in e} |N_{E_i}(w) \cap (A \cup B)|$ edges to $C^{(1)}_{i+1}(A,B) \subseteq O_i(A,B) \setminus O_{i+1}(A,B)$, which can potentially lead to large edge-effects c_e. To sidestep such technical difficulties, we now introduce the following auxiliary variables for vertex-sets $A,B \subseteq V$ with $|A| = |B|$ (to avoid clutter we suppress the dependence on A,B,i in parts of our notation):

$$\begin{aligned}
z &:= \sigma^4 q_i^2 |A|,\\
W_1 &:= \{w \in V : |N_{E_i}(w) \cap (A\cup B)| \ge z\},\\
W_2 &:= \{w \in V : |N_{\Gamma_{i+1}}(w) \cap (A\cup B)| \ge z\},\\
\hat{C}^{(1)}_{i+1} &:= \{uv \in O_i : \text{ there is } w \notin W_1 \text{ s.t. } |\{uw,vw\} \cap \Gamma_{i+1}| = |\{uw,vw\} \cap E_i| = 1\},\\
\hat{C}^{(2)}_{i+1} &:= \{uv \in O_i : \text{ there is } w \notin W_2 \text{ s.t. } uw \in \Gamma_{i+1},\ vw \in \Gamma_{i+1}\},\\
\hat{C}_{i+1} &:= \hat{C}^{(1)}_{i+1} \cup S_{i+1}.
\end{aligned}$$

Note that $\hat{C}^{(j)}_{i+1} \subseteq C^{(j)}_{i+1}$ for $j \in \{1,2\}$, and that $\hat{C}_{i+1} \subseteq C_{i+1}$. Furthermore, recalling $q_i \ge q_I$ (see (55)), using inequality (54) it is routine to check that $s_0 \gg 1$ holds, that $|A| \ge s_0$ implies $z \gg 1$, and moreover that

$$\min_{|A|\ge s_0} z/\sqrt{|A|I} \ge \sigma^4 q_i^2 \sqrt{s_0}/\sqrt{I} \gg \sigma^6 q_I^3 \sqrt[4]{n}/\sqrt{I} \gg n^{\tau/2}. \tag{77}$$

Lemma 24. *We have $\mathbb{P}(\neg \mathcal{Q}^+_{i+1}) \le n^{-\omega(1)}$.*

Proof. Mimicking the double counting argument from (4), it follows that the special case $|A| = |B|$ of $\mathcal{Q}^+_{i+1}$ implies the event $\mathcal{Q}^+_{i+1}$ in full. Hence $\neg\mathcal{Q}^+_{i+1}$ implies that $|O_{i+1}(A,B)| \le q_{i+1}|A||B|$ fails for some disjoint vertex-sets $A, B \subseteq V$ with $|A| = |B| \ge s_0$, and we shall below estimate the probability of this special case.

Recalling $\hat{C}_{i+1} \subseteq C_{i+1}$, noting $O_{i+1} \subseteq O_i \setminus C_{i+1} \subseteq O_i \setminus \hat{C}_{i+1}$ we obtain

$$|O_{i+1}(A,B)| \le |O_i(A,B) \setminus \hat{C}_{i+1}| = \sum_{f \in O_i(A,B)} \mathbb{1}_{\{f \notin \hat{C}_{i+1}\}} =: X. \tag{78}$$

To estimate $\mathbb{E}\, X$, recall that $C^{(1)}_{i+1} = \{f \in O_i : Y_f(i) \cap \Gamma_{i+1} \neq \varnothing\}$. Note that if the event $\mathcal{Q}_f := \{(f \times W_1) \cap \Gamma_{i+1} = \varnothing\}$ holds, then $f \notin \hat{C}^{(1)}_{i+1}$ implies $f \notin C^{(1)}_{i+1}$, so that $f \notin \hat{C}_{i+1}$ implies $f \notin C_{i+1} = C^{(1)}_{i+1} \cup S_{i+1}$. Since $f \notin C^{(1)}_{i+1}$ and $\mathcal{Q}_f$ are both monotone decreasing functions of the edge-indicators $(\mathbb{1}_{\{e\in\Gamma_{i+1}\}}, \mathbb{1}_{\{e \in S_{i+1}\}})_{e \in O_i}$, using Harris's inequality [18] and $\mathbb{P}(\mathcal{Q}_f) \ge (1-p)^{2|W_1|}$ it follows that

$$\mathbb{P}(f \notin C_{i+1}) \ge \mathbb{P}(f \notin \hat{C}_{i+1} \text{ and } \mathcal{Q}_f) \ge \mathbb{P}(f \notin \hat{C}_{i+1})\, \mathbb{P}(\mathcal{Q}_f) \ge \mathbb{P}(f \notin \hat{C}_{i+1}) \cdot (1-p)^{2|W_1|}.$$

Note that $\mathfrak{X}_{\le i}$ and $i < I$ imply $|N_{E_i}(u) \cap N_{E_i}(v)| = |Z_{uv}(i)| \le I(\log n)^9 =: y$ when $u \neq v$, and that (77) implies $z \gg \sqrt{|A \cup B| y}$. Using the definition of W_1 and Lemma 16 (with $\mathcal{I} = W_1$, $U = A \cup B$ and $U_w = N_{E_i}(w) \cap U$), we infer $|W_1| \le 2|A \cup B|/z = 4/(\sigma^4 q_i^2) \le q_i \sigma \sqrt{n}$ by (54), say. Similar to (69), using Lemma 19, $|O_i(A,B)| \le q_i|A||B|$, $p|W_1| \le q_i\sigma^2 \ll 1$ and $q_i q_{i+1} \sim q_i^2$ (see (55)) it is routine to deduce that

$$\mathbb{E}\, X \le |O_i(A,B)| \cdot (q_{i+1}/q_i - \sigma^{3/2} q_i) \cdot (1-p)^{-2|W_1|} \le |A||B| \cdot (q_{i+1} - \sigma^{3/2} q_i^2/2). \tag{79}$$

Gearing up to apply Theorem 12, we now bound $\lambda \le p \sum_{e \in O_i} c_e^2 + q_i \sum_{e \in O_i} \hat{c}_e^2$. Noting $\hat{C}_{i+1} \subseteq C_{i+1}$, as usual we have edge-effect $c_e \le |Y_e(i) \cap O_i(A,B)|$ and stabilization-effect $\hat{c}_e \le \mathbb{1}_{\{e \in O_i(A,B)\}}$. Here the definition of $\hat{C}_{i+1}$ allows us to improve the simple worst case estimate $c_e \le |Y_e(i)|$. Indeed, inspecting the corresponding argument for $|N_{O_{i+1}}(v)|$ from Lemma 21, we see that the edge-effect c_e (an upper bound on how much X changes if we alter whether $e \in \Gamma_{i+1}$ or $e \notin \Gamma_{i+1}$) is at most the number of changes to

$$\begin{aligned}\hat{C}^{(1)}_{i+1} \cap O_i(A,B) = \bigl\{uv \in O_i(A,B) :\ & \text{there is } w \notin W_1 \text{ s.t. either } uw \in \Gamma_{i+1},\ vw \in E_i \\ & \text{or } vw \in \Gamma_{i+1},\ uw \in E_i\bigr\}.\end{aligned} \tag{80}$$

Since any $w \notin W_1$ has at most z neighbours in $A \cup B$ via E_i–edges, we infer that $c_e \le 2z$ (the factor of two takes into account that each vertex of e could potentially play the role of w in (80) above). Similar to (66) and (70), using Lemma 20, $\sigma\pi_i \le \sqrt{\sigma} \ll 1$ (see (53)), and $|O_i(A,B)| \le q_i|A||B|$ it follows that

$$p \sum_{e \in O_i} c_e^2 \le \sigma/\sqrt{n} \cdot 2z \cdot 2q_i\pi_i\sqrt{n} \cdot |O_i(A,B)| \ll z q_i |O_i(A,B)| \le z q_i^2 |A||B|.$$

Furthermore, using $z \ge 1$ we obtain $q_i \sum \hat{c}_e^2 \le q_i |O_i(A,B)| \le z q_i |O_i(A,B)| \le z q_i^2 |A||B|$. Noting that X is decreasing, using Theorem 12 and the λ–bound (63) it follows that

$$\begin{aligned}\mathbb{P}(|O_{i+1}(A,B)| \ge q_{i+1}|A||B|) &\le \mathbb{P}(X \ge \mathbb{E}\, X + \sigma^{3/2} q_i^2 |A||B|/2) \\ &\le \exp\left(-\frac{\bigl(\sigma^{3/2} q_i^2 |A||B|/2\bigr)^2}{2 \cdot 2 z q_i^2 |A||B|}\right) = \exp\left(-\frac{\sigma^3 q_i^2 |A||B|}{16 z}\right) \le n^{-\omega(|B|)},\end{aligned} \tag{81}$$

where for the last inequality we used $z = \sigma^4 q_i^2 |A|$ and $\sigma^{-1} \gg \log n$. Finally, taking a union bound over all disjoint vertex-sets $A, B \subseteq V$ with $|A| = |B| \ge s_0$ completes the proof (as discussed). $\square$

For the 'relative error' τ_i used in the event $\mathcal{Q}_i$, see (38), we now record the following convenient bounds:

$$1 \geq \tau_i \geq \tau_I = 1 - \delta/2 \geq 1/2 \qquad \text{for all } 0 \leq i \leq I. \tag{82}$$

Lemma 25. *We have* $\mathbb{P}(\neg \mathcal{Q}_{i+1} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}) \leq n^{-\omega(1)}$.

The proof strategy is to estimate the different contributions to $O_{i+1} = O_i \setminus (\Gamma_{i+1} \cup C_{i+1} \cup C^{(2)}_{i+1})$ separately (here $\mathcal{Q}_i^+$ will be crucial for bounding some of the large edge-effects ignored in Lemma 24).

Claim 26. *Let $\mathcal{Q}_{A,B}$ be the event that the following bounds hold:*

$$\begin{aligned}
X_1 &:= \big|O_i(A,B) \setminus \hat{C}_{i+1}\big| \in \big[|O_i(A,B)| \cdot (q_{i+1}/q_i - 4\sigma^{3/2}q_i),\ |O_i(A,B)| \cdot q_{i+1}/q_i\big],\\
X_2 &:= \big|O_i(A,B) \cap \hat{C}^{(2)}_{i+1}\big| \leq |O_i(A,B)| \cdot 2\sigma^2 q_i,\\
X_3 &:= |O_i(A,B) \cap \Gamma_{i+1}| \leq |O_i(A,B)| \cdot 2\sigma^2 q_i,\\
X_4 &:= \big|O_i(A,B) \cap (C_{i+1} \cup C^{(2)}_{i+1}) \setminus (\hat{C}_{i+1} \cup \hat{C}^{(2)}_{i+1})\big| \leq 36\sigma q_i^2 \sqrt{n}|A|.
\end{aligned}$$

Then $\mathbb{P}(\neg \mathcal{Q}_{A,B} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}) \leq n^{-\omega(s)}$ *for all vertex-sets* $(A,B) \in \mathfrak{S}_{s,\gamma}$.

Before giving the proof, we first show that Claim 26 implies Lemma 25. Using a union bound argument (to account for the $|\mathfrak{S}_{s,\gamma}| \leq n^{2s}$ vertex-sets $(A,B) \in \mathfrak{S}_{s,\gamma}$), it is enough to show that $\mathcal{Q}_{A,B} \cap \mathfrak{X}_{\leq i}$ implies $\tau_{i+1}q_{i+1}|O_0(A,B)| \leq |O_{i+1}(A,B)| \leq q_{i+1}|O_0(A,B)|$. By definition of $O_{i+1}(A,B)$ we have

$$X_1 - X_2 - X_3 - X_4 \leq |O_{i+1}(A,B)| \leq X_1.$$

Combining $\mathcal{Q}_{A,B}$ with the fact that $|O_i(A,B)| \leq q_i|O_0(A,B)|$ by $\mathfrak{X}_{\leq i} \subseteq \mathcal{Q}_i$, we readily infer the upper bound $|O_{i+1}(A,B)| \leq q_{i+1}|O_0(A,B)|$. Turning to the lower bound, using $\mathcal{Q}_{A,B}$ it follows that

$$\begin{aligned}
X_1 - X_2 - X_3 - X_4 &\geq |O_i(A,B)| \cdot \big(q_{i+1}/q_i - 8\sigma^{3/2}q_i\big) - 36\sigma q_i^2\sqrt{n}|A|\\
&\geq \left(\tau_i q_i\big(q_{i+1}/q_i - 8\sigma^{3/2}q_i\big) - \frac{36\sigma q_i^2}{\gamma C\sqrt{\log n}}\right) \cdot |O_0(A,B)|\\
&\geq \left(\tau_i - \frac{45\sigma q_i}{\gamma C\sqrt{\log n}}\right) \cdot q_{i+1}|O_0(A,B)| \geq \tau_{i+1} \cdot q_{i+1}|O_0(A,B)|,
\end{aligned}$$

where for the second inequality we used $|O_i(A,B)| \geq \tau_i q_i|O_0(A,B)|$ (by $\mathfrak{X}_{\leq i} \subseteq \mathcal{Q}_i$) and $|O_0(A,B)| \geq \gamma|A||B| \geq \gamma C\sqrt{\log n} \cdot \sqrt{n}|A|$, for the third inequality we used $\tau_i \leq 1$ (see (82)), $\sigma^{1/2} \ll 1/\sqrt{\log n}$, and $q_i \sim q_{i+1}$ (see (55)), and for the last inequality we used $\sqrt{\log n} \sim \sqrt{\log(I\sigma)/\beta} \sim \pi_I/\sqrt{\beta}$ (see (57)), $\gamma C/\sqrt{\beta} \geq D_0/\delta^2 \geq 91/\delta$ (by assumption and (39)) and $\tau_i - \delta\sigma q_i/\pi_I = \tau_{i+1}$ (see (38)). This completes the proof of Lemma 25 (assuming Claim 26).

Proof of Claim 26. We start with $X_1 = |O_i(A,B) \setminus \hat{C}_{i+1}|$. Since $s \geq s_0$, the upper tail argument for $X = X_1$ defined in (78) carries over from Lemma 24, with $\mathbb{E} X_1 \leq |O_i(A,B)|(q_{i+1}/q_i - \sigma^{3/2}q_i/2)$ and $\lambda \leq 2zq_i|O_i(A,B)|$, say. In particular, noting that here $|O_i(A,B)| \geq \tau_i q_i|O_0(A,B)| \geq \gamma\tau_i q_i|A||B|$, an application of Theorem 12 along the lines of (81) gives

$$\mathbb{P}(X_1 \geq |O_i(A,B)|q_{i+1}/q_i) \leq \exp\left(-\frac{\big(\sigma^{3/2}q_i|O_i(A,B)|/2\big)^2}{2 \cdot 2zq_i|O_i(A,B)|}\right) \leq \exp\left(-\frac{\gamma\tau_i\sigma^3q_i^2|A||B|}{16z}\right) \leq n^{-\omega(s)}, \tag{83}$$

where for the last inequality we used $z = \sigma^4q_i^2|A|$, $\tau_i \geq 1/2$ (see (82)), $\gamma\sigma^{-1} \gg \log n$ and $|B| = s$. For the lower tail of X_1 we proceed similarly. Since $\hat{C}_{i+1} \subseteq C_{i+1}$, using Lemma 19 we obtain

$$\mathbb{E} X_1 = \sum_{e \in O_i(A,B)} \mathbb{P}(e \notin \hat{C}_{i+1}) \geq \sum_{e \in O_i(A,B)} \mathbb{P}(e \notin C_{i+1}) \geq |O_i(A,B)| \cdot (q_{i+1}/q_i - 3\sigma^{3/2}q_i).$$

Furthermore, the edge-effect and stabilization-effect estimates from the proof of Lemma 24 again carry over, giving $\lambda \le 2zq_i|O_i(A,B)|$ and $\max_{e\in O_i}\max\{c_e,\hat{c}_e\} \le 2z$, say. Applying inequality (49) of Remark 13 (with $C=2z$), it follows similarly to (83) that

$$\begin{aligned}\mathbb{P}\big(X_1 \le |O_i(A,B)|(q_{i+1}/q_i - 4\sigma^{3/2}q_i)\big) &\le \mathbb{P}\big(X_1 \le \mathbb{E}\,X_1 - \sigma^{3/2}q_i|O_i(A,B)|\big)\\ &\le \exp\left(-\frac{\big(\sigma^{3/2}q_i|O_i(A,B)|\big)^2}{2\big(2zq_i|O_i(A,B)| + 2z\cdot\sigma^{3/2}q_i|O_i(A,B)|\big)}\right)\\ &\le \exp\left(-\frac{\gamma\tau_i\sigma^3q_i^2|A||B|}{8z}\right) \le n^{-\omega(s)}.\end{aligned} \tag{84}$$

Turning to $X_2 = |O_i(A,B)\cap\hat{C}^{(2)}_{i+1}|$, note that by construction of $\hat{C}^{(2)}_{i+1}$ we have

$$X_2 = \sum_{e\in O_i(A,B)} \mathbb{1}_{\{e\in\hat{C}^{(2)}_{i+1}\}} \le \sum_{ab\in O_i(A,B)}\sum_{w\in V\setminus W_2}\mathbb{1}_{\{\{wa,wb\}\subseteq\Gamma_{i+1}\}} =: X_2^+. \tag{85}$$

Gearing up to apply Theorem 15 to X_2^+, in view of $\Gamma_{i+1}\subseteq O_i$ we define

$$\begin{aligned}\mathcal{I} &:= \big\{\{wa,wb\}\subseteq O_i :\ ab\in O_i(A,B),\ w\in V,\ |\{a,b,w\}|=3\big\},\\ \mathcal{K} &:= \{\{wa,wb\}\in\mathcal{I} :\ w\notin W_2,\ \{wa,wb\}\subseteq\Gamma_{i+1}\}.\end{aligned}$$

Since $p^2\cdot|X_{ab}(i)| \le \sigma^2q_i^2 \le \sigma^2q_i$ by $\mathfrak{X}_{\le i}\subseteq\mathcal{P}_i$ and $q_i\le 1$ (see (53)), we obtain

$$\sum_{\alpha\in\mathcal{I}}\mathbb{E}\,\mathbb{1}_{\{\alpha\subseteq\Gamma_{i+1}\}} = p^2\sum_{ab\in O_i(A,B)}\sum_{v\in V}\mathbb{1}_{\{\{va,vb\}\subseteq O_i\}} = p^2\sum_{ab\in O_i(A,B)}|X_{ab}(i)| \le \sigma^2q_i\cdot|O_i(A,B)| =: \mu.$$

Furthermore, since $\mathcal{K}$ only contains edge-pairs $\{wa,wb\}$ with $\{a,b\}\subseteq N_{\Gamma_{i+1}}(w)\cap(A\cup B)$ where the 'central vertex' w satisfies $w\notin W_2$ and thus $|N_{\Gamma_{i+1}}(w)\cap(A\cup B)|\le z$, for all $\beta\in\mathcal{K}$ we see that

$$|\{\alpha\in\mathcal{K} :\ \alpha\cap\beta\neq\varnothing\}| \le \sum_{f\in\beta}|\{\alpha\in\mathcal{K} :\ f\in\alpha\}| \le \sum_{f\in\beta}\sum_{v\in f\setminus W_2}|N_{\Gamma_{i+1}}(v)\cap(A\cup B)| \le 2\cdot 2\cdot z.$$

It follows that $X_2^+ = \sum_{\alpha\in\mathcal{K}}\mathbb{1}_{\{\alpha\subseteq\Gamma_{i+1}\}} \le Z_{4z}$, where Z_{4z} is defined as in Theorem 15. Applying first (85) and then inequality (50) with $C=4z$, using $|O_i(A,B)|\ge\gamma\tau_iq_i|A||B|$ it follows similarly to (83) that

$$\mathbb{P}(X_2 \ge 2\sigma^2q_i|O_i(A,B)|) \le \mathbb{P}(Z_{4z}\ge 2\mu) \le \exp\Big(-\frac{\mu^2}{2\cdot 4z\cdot 2\mu}\Big) \le \exp\Big(-\frac{\gamma\tau_i\sigma^2q_i^2|A||B|}{16z}\Big) \le n^{-\omega(s)}. \tag{86}$$

We next turn to $X_3 = |O_i(A,B)\cap\Gamma_{i+1}|$, which is a sum of independent Bernoulli random variables with $\mathbb{E}\,X_3 = |O_i(A,B)|\cdot p \ll \sigma^2q_i|O_i(A,B)| =: t$, as $q_i\sqrt{n}\ge n^{\tau}$ by (54). Applying standard Chernoff bounds, using $|O_i(A,B)|\ge\gamma\tau_iq_i|A||B|$ and $z\ge 1$ it follows by comparison with the last inequality of (86) that

$$\mathbb{P}(X_3 \ge 2\sigma^2q_i|O_i(A,B)|) \le \mathbb{P}(X_3\ge\mathbb{E}\,X_3+t) \le \exp\Big(-\frac{t^2}{2\cdot 2t}\Big) \le \exp\Big(-\frac{\gamma\tau_i\sigma^2q_i^2|A||B|}{4}\Big) \le n^{-\omega(s)}. \tag{87}$$

Finally, X_4 is a more difficult variable: assuming that $\mathcal{N}_{i+1}\cap\mathcal{P}_{i+1}\cap\mathfrak{X}_{\le i}$ holds, we shall bound X_4 by *deterministic* counting arguments (here the edge-effects can potentially be fairly large, so concentration inequalities seem less effective). Noting $C_{i+1}\setminus\hat{C}_{i+1} = C^{(1)}_{i+1}\setminus\hat{C}^{(1)}_{i+1}$, similarly to (85) we obtain

$$\begin{aligned}X_4 &\le \sum_{e\in O_i(A,B)}\mathbb{1}_{\{e\in C^{(1)}_{i+1}\setminus\hat{C}^{(1)}_{i+1}\}} + \sum_{e\in O_i(A,B)}\mathbb{1}_{\{e\in C^{(2)}_{i+1}\setminus\hat{C}^{(2)}_{i+1}\}}\\ &\le \sum_{w\in W_1}\Big(|O_i(N_{\Gamma_{i+1}}(w)\cap A,\ N_{E_i}(w)\cap B)| + |O_i(N_{\Gamma_{i+1}}(w)\cap B,\ N_{E_i}(w)\cap A)|\Big)\\ &\qquad + \sum_{w\in W_2}|O_i(N_{\Gamma_{i+1}}(w)\cap A,\ N_{\Gamma_{i+1}}(w)\cap B)|.\end{aligned} \tag{88}$$

Using the upper bound estimate from $\mathfrak{X}_{\le i} \subseteq \mathcal{Q}_i^+$ when $\min\{|N_{\Gamma_{i+1}}(v) \cap A|, |N_{E_i}(v) \cap B|\} \ge z$ holds (note that $z = \sigma^4 q_i^2 s \ge s_0$), and a trivial estimate otherwise, it follows that

$$\begin{aligned} &|O_i(N_{\Gamma_{i+1}}(w) \cap A,\ N_{E_i}(w) \cap B)| \\ &\quad \le q_i |N_{\Gamma_{i+1}}(w) \cap A||N_{E_i}(w) \cap B| + z \max\{|N_{\Gamma_{i+1}}(w) \cap A|, |N_{E_i}(w) \cap B|\} \\ &\quad \le \big(q_i |N_{\Gamma_{i+1}}(w)| + z\big) \cdot |N_{E_i \cup \Gamma_{i+1}}(w) \cap (A \cup B)|. \end{aligned} \tag{89}$$

With an eye on (88), we note that an analogous estimate also holds when we reverse the role of A and B in (89). Furthermore, $q_i|N_{\Gamma_{i+1}}(w)| \le 2\sigma q_i^2 \sqrt{n}$ by $\mathcal{N}_{i+1}$, and $z = \sigma^4 q_i^2 s = O(\sigma^3 q_i^2 \sqrt{n}) \ll \sigma q_i^2 \sqrt{n}$. Recalling $E_i \cup \Gamma_{i+1} = E_{i+1}$, observe that $\mathcal{P}_{i+1}$ and $i+1 \le I$ imply $|N_{E_i \cup \Gamma_{i+1}}(u) \cap N_{E_i \cup \Gamma_{i+1}}(v)| = |Z_{uv}(i+1)| \le I(\log n)^9 =: y$ when $u \neq v$, and that (77) implies $z \gg \sqrt{|A \cup B| y}$ (as $|A| = s \ge s_0$). Using the definition of W_1 and Lemma 16 (with $\mathcal{I} = W_1$, $U = A \cup B$ and $U_w = N_{E_i \cup \Gamma_{i+1}}(w) \cap U$), it follows that

$$\begin{aligned} &\sum_{w \in W_1} \Big(|O_i(N_{\Gamma_{i+1}}(w) \cap A,\ N_{E_i}(w) \cap B)| + |O_i(N_{\Gamma_{i+1}}(w) \cap B,\ N_{E_i}(w) \cap A)| \Big) \\ &\quad \le 2 \cdot 3\sigma q_i^2 \sqrt{n} \cdot \sum_{w \in W_1} |N_{E_i \cup \Gamma_{i+1}}(w) \cap (A \cup B)| \le 2 \cdot 3\sigma q_i^2 \sqrt{n} \cdot 2|A \cup B| \le 24\sigma q_i^2 \sqrt{n}|A|. \end{aligned} \tag{90}$$

Proceeding analogously to (89)–(90), using the definition of W_2 and Lemma 16 we similarly obtain

$$\begin{aligned} &\sum_{w \in W_2} |O_i(N_{\Gamma_{i+1}}(w) \cap A,\ N_{\Gamma_{i+1}}(w) \cap B)| \\ &\quad \le 3\sigma q_i^2 \sqrt{n} \cdot \sum_{w \in W_2} |N_{\Gamma_{i+1}}(w) \cap (A \cup B)| \le 3\sigma q_i^2 \sqrt{n} \cdot 2|A \cup B| \le 12\sigma q_i^2 \sqrt{n}|A|. \end{aligned} \tag{91}$$

To sum up, inserting the bounds (90)–(91) into (88), we showed that $\mathcal{N}_{i+1} \cap \mathcal{P}_{i+1} \cap \mathfrak{X}_{\le i}$ implies $X_4 \le 36\sigma q_i^2 \sqrt{n}|A|$. This completes the proof together with the probability estimates (83), (84), (86), and (87). $\square$

Remark 27. *If desired, it would not be difficult to extend the event $\mathcal{Q}_i$ to larger vertex-sets $(A,B) \in \mathfrak{S}_{\ge s,\gamma} := \bigcup_{s \le r \le n} \mathfrak{S}_{r,\gamma}$ (the above arguments all carry over, except for the modified bound $X_4 \le 3 \cdot \max_w(q_i|N_{\Gamma_{i+1}}(w)| + z) \cdot 2|A \cup B| \le 36\sigma q_i^2 \max\{\sqrt{n}, \sigma^3|B|\}|A|$, which is still strong enough to deduce Lemma 25). This in turn could, e.g., be used to also extend the event $\mathcal{T}_I$ to $(A,B) \in \mathfrak{S}_{\ge s,\gamma}$ (the proofs in Section 3.5 then carry over).*

Remark 28. *Under a mild extra assumption such as $|O_0| \ge \sigma n$, say, it would not be difficult to add two-sided bounds for the total number of open edges $|O_i|$ and edges $|T_I|$ to the events $\mathcal{Q}_i$ and $\mathcal{T}_I$. For example, much simpler variants of the above arguments then imply $\tau_i q_i |O_0| \le |O_i| \le q_i |O_0|$ (by directly estimating $|O_i \setminus C_{i+1}| - |\Gamma_{i+1}| - |C_{i+1}^{(2)}| \le |O_{i+1}| \le |O_i \setminus C_{i+1}|$, without using $\hat{C}_{i+1}$ or $\hat{C}_{i+1}^{(2)}$, nor a union bound over all vertex-sets), which in turn gives $|T_I| = (1 \pm \delta)\rho|O_0|$ by straightforward variants of the proofs in Section 3.5.*

3.5 Event $\mathcal{T}_I$: number $|T_I(A,B)|$ of edges between large sets

Recall that the event $\mathcal{T}_I$ defined in (42) concerns the triangle-free edge-set $T_I \subseteq E(H) = O_0$, ensuring that $|T_I(A,B)| = (1 \pm \delta)\rho|O_0(A,B)|$ for all $(A,B) \in \mathfrak{S}_{s,\gamma}$; see (41) for the definition of $\mathfrak{S}_{s,\gamma}$.

For $|T_I(A,B)|$ it is convenient to think of the entire nibble construction as one evolving random process. Thus, in contrast to previous sections, in Lemma 29 and Claim 30 below we shall *not* tacitly condition on $\mathcal{F}_i$.

Lemma 29. *We have $\mathbb{P}(\neg\mathcal{T}_I \cap \mathfrak{X}_{\le I}) \le n^{-\omega(1)}$.*

Since $T_I = \bigcup_{0 \le i < I}(T_{i+1} \setminus T_i)$ forms a partition, the proof strategy is to estimate the two contributions to $T_{i+1} \setminus T_i = \Gamma_{i+1} \setminus E(\mathcal{D}_{i+1})$ separately (here the deleted edges $E(\mathcal{D}_{i+1})$ will have negligible impact).

Claim 30. *Let $\mathcal{T}_{A,B}$ be the event that the following bounds hold:*

$$X := \sum_{0 \le i < I} |O_i(A,B) \cap \Gamma_{i+1}| \in \big[(1-\delta/2)\mu^-,\ (1+\delta/2)\mu^+\big],$$

$$Y := \sum_{0 \le i < I} \big|O_i(A,B) \cap E(\mathcal{D}_{i+1})\big| \le \delta^2\mu^-/9,$$

where $\mu^+ := \sum_{0\le i<I} \lfloor q_i|O_0(A,B)|\rfloor p$ *and* $\mu^- := \sum_{0\le i<I}\lceil \tau_i q_i|O_0(A,B)|\rceil p$. *Then* $\mathbb{P}(\neg\mathcal{T}_{A,B}\cap \mathfrak{X}_{\le I}) \le 3n^{-3s}$ *for all vertex-sets* $(A,B)\in\mathfrak{S}_{s,\gamma}$.

Before giving the proof, we first show that Claim 30 implies Lemma 29. Using a union bound argument (to account for the $|\mathfrak{S}_{s,\gamma}| \le n^{2s}$ vertex-sets $(A,B)\in\mathfrak{S}_{s,\gamma}$), it is enough to show that $\mathcal{T}_{A,B}$ implies $|T_I(A,B)| = (1\pm\delta)\rho|O_0(A,B)|$. Since all the $(\Gamma_{i+1})_{0\le i<I}$ are edge-disjoint, by the recursive definition (14) of T_I we have

$$X - Y \le |T_I(A,B)| \le X. \tag{92}$$

Noting $\mu^- \ge \tau_I\mu^+ = (1-\delta/2)\mu^+$ (see (82)), it follows that $\mathcal{T}_{A,B}$ implies $X \le (1+\delta/2)\mu^+$ and

$$X - Y \ge \bigl(1-\delta/2-\delta^2/9\bigr)\cdot\mu^- \ge (1-\delta+\delta^2/8)\mu^+.$$

It thus suffices to show that $\mu^+ \sim \rho|O_0(A,B)|$, where $\rho = \sqrt{\beta(\log n)/n}$. But this is routine: indeed, since $q_i|O_0(A,B)| \ge q_i\cdot\gamma s^2 \gg q_i n \gg \sqrt{n}$ by (54), and $\pi_I \sim \sqrt{\log(I\sigma)} \sim \sqrt{\beta\log n}$ by (57), using the definition (37) of π_I we readily infer

$$\begin{aligned}\mu^+ &= \sum_{0\le i<I}(q_i|O_0(A,B)|\pm 1)p \sim \sum_{0\le i<I}\sigma q_i/\sqrt{n}\cdot|O_0(A,B)|\\ &= (\pi_I-\sigma)/\sqrt{n}\cdot|O_0(A,B)| \sim \rho|O_0(A,B)|,\end{aligned}\tag{93}$$

completing the proof of Lemma 29 (assuming Claim 30).

Proof of Claim 30. We start with $X = \sum_{0\le i<I}|O_i(A,B)\cap\Gamma_{i+1}|$. Define

$$X^+_{i+1} := \mathbb{1}_{\{\mathfrak{X}_i\}}\sum_{e\in O_i(A,B)}\mathbb{1}_{\{e\in\Gamma_{i+1}\}} \qquad\text{and}\qquad X^+ := \sum_{0\le i<I}X^+_{i+1}.$$

Note that $X = X^+$ when $\mathfrak{X}_{\le I} = \bigcap_{0\le i\le I}\mathfrak{X}_i$ holds. Let $Z^+_{i+1} \overset{\mathrm{d}}{=} \mathrm{Bin}(\lfloor q_i|O_0(A,B)|\rfloor, p)$ be independent random variables (where $\overset{\mathrm{d}}{=}$ means equality in distribution, as usual). Since the $\mathcal{F}_i$-measurable event $\mathfrak{X}_i \subseteq \mathcal{Q}_i$ implies $|O_i(A,B)| \le q_i|O_0(A,B)|$, it is easy to see that $\mathbb{P}(X^+_{i+1}\ge t \mid \mathcal{F}_i) \le \mathbb{P}(Z^+_{i+1}\ge t)$ for $t\in\mathbb{R}$. Setting

$$Z^+ := \sum_{0\le i<I}Z^+_{i+1} \overset{\mathrm{d}}{=} \mathrm{Bin}\Bigl(\sum_{0\le i<I}\lfloor q_i|O_0(A,B)|\rfloor,\ p\Bigr),\tag{94}$$

a standard stochastic domination argument then shows $\mathbb{P}(X^+\ge t) \le \mathbb{P}(Z^+\ge t)$ for $t\in\mathbb{R}$, so that

$$\mathbb{P}(X\ge t\text{ and }\mathfrak{X}_{\le I}) \le \mathbb{P}(X^+\ge t) \le \mathbb{P}(Z^+\ge t).\tag{95}$$

Since $\mathfrak{X}_i$ also implies $|O_i(A,B)| \ge \tau_i q_i|O_0(A,B)|$, an analogous argument gives

$$\mathbb{P}(X\le t\text{ and }\mathfrak{X}_{\le I}) \le \mathbb{P}(Z^-\le t) \qquad\text{with}\qquad Z^- \overset{\mathrm{d}}{=} \mathrm{Bin}\Bigl(\sum_{0\le i<I}\lceil\tau_i q_i|O_0(A,B)|\rceil,\ p\Bigr).\tag{96}$$

Combining $\mu^- \ge \tau_I\mu^+ \ge \mu^+/2$ (see (82)) and (93) with $|O_0(A,B)| \ge \gamma s^2$, using $\delta^2\sqrt{\beta}\gamma\cdot C \ge D_0 = 108$ (by assumption and (39)) we have

$$\delta^2\min\{\mu^-,\mu^+\} \ge \tfrac{\delta^2}{2}\mu^+ \ge \tfrac{\delta^2}{3}\rho|O_0(A,B)| \ge \tfrac{\delta^2}{3}\sqrt{\beta(\log n)/n}\cdot\gamma C\sqrt{n\log n}\cdot s \ge 36 s\log n.\tag{97}$$

Using (94)–(96) and $\mathbb{E}\,Z^\pm = \mu^\pm$, by standard Chernoff bounds (see, e.g., Remark 14) we obtain, say,

$$\begin{aligned}\mathbb{P}\bigl(X\notin\bigl[(1-\delta/2)\mu^-,\ (1+\delta/2)\mu^+\bigr]\text{ and }\mathfrak{X}_{\le I}\bigr) &\le \mathbb{P}\bigl(Z^-\le(1-\delta/2)\mu^-\bigr) + \mathbb{P}\bigl(Z^+\ge(1+\delta/2)\mu^+\bigr)\\ &\le \exp\bigl(-\delta^2\mu^-/8\bigr) + \exp\bigl(-\delta^2\mu^+/12\bigr) \le 2n^{-3s}.\end{aligned}\tag{98}$$

Finally, turning to $Y = \sum_{0\le i<I}|O_i(A,B)\cap E(\mathcal{D}_{i+1})|$, for brevity we define

$$Y_{i+1} := |O_i(A,B)\cap E(\mathcal{D}_{i+1})| \qquad\text{and}\qquad y := \delta^2\mu^-/9.$$

Note that $Y = \sum_{0 \le i < I} Y_{i+1}$ and $Y_{i+1} \in \mathbb{N}$. Since $\mathfrak{X}_{\le i} = \bigcap_{0 \le j \le i} \mathfrak{X}_j$, a union bound argument gives

$$\begin{aligned}
\mathbb{P}\big(Y \ge \delta^2\mu^-/9 \text{ and } \mathfrak{X}_{\le I}\big) &\le \sum_{\substack{(y_1,\dots,y_I)\in\mathbb{N}^I \\ \sum_{1\le i\le I} y_i = \lceil y \rceil}} \mathbb{P}\Big(\bigcap_{0\le i<I} \big(Y_{i+1} \ge y_{i+1} \text{ and } \mathfrak{X}_{\le i+1}\big)\Big) \\
&\le \sum_{\substack{(y_1,\dots,y_I)\in\mathbb{N}^I \\ \sum_{0\le i< I} y_{i+1} = \lceil y \rceil}} \prod_{0\le i<I} \mathbb{P}\Big(Y_{i+1} \ge y_{i+1} \;\Big|\; \bigcap_{0\le j<i} \big(Y_{j+1} \ge y_{j+1} \text{ and } \mathfrak{X}_{\le j+1}\big)\Big).
\end{aligned} \tag{99}$$

Gearing up to apply Theorem 15 to Y_{i+1}, with an eye on $\mathcal{D}_{i+1} \subseteq \mathcal{B}_{i+1}$ and $T_i \subseteq E_i$ (see Section 2.1) we define

$$\begin{aligned}
\mathcal{I} := \big\{\{wu, wv\} \subseteq O_i :\ & uv \in E_i,\ |\{u,v,w\}| = 3,\ \{wu,wv\} \cap O_i(A,B) \neq \varnothing\big\} \\
&\cup\ \big\{\{uv, vw, wu\} \subseteq O_i :\ |\{u,v,w\}| = 3,\ \{uv,vw,wu\} \cap O_i(A,B) \neq \varnothing\big\}.
\end{aligned}$$

Since each edge-set $\alpha \in \mathcal{I}$ contains at least one edge from $O_i(A,B)$, when the $\mathcal{F}_i$-measurable event $\mathfrak{X}_{\le i}$ holds we infer by the usual reasoning (using, e.g., $\mathcal{P}_i \cap \mathcal{Q}_i$ and $\max\{\pi_i q_i, q_i^2\} \le 1$) that

$$\begin{aligned}
\sum_{\alpha\in\mathcal{I}} \mathbb{E}(\mathbb{1}_{\{\alpha\subseteq\Gamma_{i+1}\}} \mid \mathcal{F}_i) &\le \sum_{e\in O_i(A,B)} \sum_{\alpha\in\mathcal{I}: e\in\alpha} p^{|\alpha|} \le \sum_{e\in O_i(A,B)} \Big(|Y_e(i)|\cdot p^2 + |X_e(i)|\cdot p^3\Big) \\
&\le q_i |O_0(A,B)| \cdot \big(2\pi_i q_i \sqrt{n}\cdot p^2 + q_i^2 n\cdot p^3\big) \le 3\sigma\cdot q_i|O_0(A,B)|p =: \mu^*_{i+1}.
\end{aligned}$$

Since $\mathcal{D}_{i+1}$ is a collection of edge-disjoint elements of $\mathcal{B}_{i+1}$ (and thus $\{\alpha \in \mathcal{D}_{i+1} : \alpha \cap \beta \neq \varnothing\} = \{\beta\}$ for all $\beta \in \mathcal{D}_{i+1}$), using $E(\mathcal{D}_{i+1}) = \bigcup_{\alpha\in\mathcal{D}_{i+1}} \alpha \subseteq \Gamma_{i+1} \subseteq O_i$, $|\alpha| \le 3$ and $T_i \subseteq E_i$ it is not difficult to check that

$$Y_{i+1} = \sum_{\alpha\in\mathcal{D}_{i+1}} |\alpha \cap O_i(A,B)| \le 3 \cdot \sum_{\alpha\in\mathcal{I}\cap\mathcal{D}_{i+1}} \mathbb{1}_{\{\alpha\in\Gamma_{i+1}\}} \le 3Z_1,$$

where Z_1 is defined as in Theorem 15. Applying inequality (50) with $C = 1$ and $\mu = \mu^*_{i+1}$ (in the probability space conditional on $\mathcal{F}_i$; cf. the beginning of Section 3.1), when $\mathfrak{X}_{\le i}$ holds it follows that, say,

$$\mathbb{P}(Y_{i+1} \ge y_{i+1} \mid \mathcal{F}_i) \le \mathbb{P}(Z_1 \ge y_{i+1}/3 \mid \mathcal{F}_i) \le \begin{cases} \Big(\frac{e\mu^*_{i+1}}{y_{i+1}/3}\Big)^{y_{i+1}/3} \le \sigma^{y_{i+1}/6} & \text{if } y_{i+1} \ge 9\mu^*_{i+1}/\sqrt{\sigma}, \\ 1 & \text{otherwise.} \end{cases} \tag{100}$$

Comparing the definition of $\sum_{0\le i<I} \mu^*_{i+1}$ with μ^-, using $\tau_i \ge \tau_I \ge 1/2$ (see (82)) and $\sigma \ll 1$ we see that

$$\sum_{\substack{0\le i<I: \\ y_{i+1} \le 9\mu^*_{i+1}/\sqrt{\sigma}}} y_{i+1} \le 9/\sqrt{\sigma} \cdot \sum_{0\le i<I} \mu^*_{i+1} \le 9/\sqrt{\sigma}\cdot 6\sigma\mu^- \ll \delta^2\mu^-/9 = y.$$

So, inserting (100) into (99), using (97) and the definition of s it follows that $y/\log y = \Omega(\sqrt{n}) \gg I$ and

$$\mathbb{P}\big(Y \ge \delta^2\mu^-/9 \text{ and } \mathfrak{X}_{\le I}\big) \le \sum_{\substack{(y_1,\dots,y_I)\in\mathbb{N}^I \\ \sum_{0\le i<I} y_{i+1} = \lceil y \rceil}} \sigma^{\lceil y \rceil/6 - o(y)} \le (y+2)^I \cdot \sigma^{y/7} \le e^{-\omega(\delta^2\mu^-)} \le n^{-\omega(s)},$$

completing the proof together with the probability estimate (98). □

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A note on Ramsey numbers. *J. Combin. Theory Ser. A* **29** (1980), 354–360.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. A dense infinite Sidon sequence. *European J. Combin.* **2** (1981), 1–11.

[3] N. Alon, J.H. Kim, and J. Spencer. Nearly perfect matchings in regular simple hypergraphs. *Israel J. Math.* **100** (1997), 171–187.

[4] J. Beck. On size Ramsey number of paths, trees, and circuits. I. *J. Graph Theory* **7** (1983), 115–129.
[5] T. Bohman. The triangle-free process. *Adv. Math.* **221** (2009), 1653–1677.
[6] T. Bohman and P. Keevash. The early evolution of the H-free process. *Invent. Math.* **181** (2010), 291–336.
[7] T. Bohman and P. Keevash. Dynamic concentration of the triangle-free process. Preprint (2013). `arXiv:1302.5963`.
[8] S. A. Burr, P. Erdős, and L. Lovász. On graphs of Ramsey type. *Ars Combinatoria* **1** (1976), 167–190.
[9] D. Conlon, J. Fox, and B. Sudakov. Recent developments in graph Ramsey theory. In *Surveys in Combinatorics 2015*, pp. 49–118. Cambridge Univ. Press, Cambridge (2015).
[10] P. Erdős. Some remarks on the theory of graphs. *Bull. Amer. Math. Soc.* **53** (1947), 292–294.
[11] P. Erdős. Graph theory and probability. II. *Canad. J. Math.* **13** (1961), 346–352.
[12] P. Erdős, R.J. Faudree, C.C. Rousseau, and R.H. Schelp. The size Ramsey number. *Period. Math. Hungar.* **9** (1978), 145–161.
[13] P. Erdős, S. Suen, and P. Winkler. On the size of a random maximal graph. *Rand. Struct. & Algor.* **6** (1995), 309–318.
[14] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.* **2** (1935), 463–470.
[15] G. Fiz Pontiveros, S. Griffiths, and R. Morris. The triangle-free process and $R(3, k)$. *Memoirs of the Am. Math. Soc.*, to appear.
[16] J. Fox, A. Grinshpun, A. Liebenau, Y. Person, and T. Szabó. On the minimum degree of minimal Ramsey graphs for multiple colours. *J. Combin. Theory Ser. B* **120** (2016), 64–82.
[17] J. Fox and K. Lin. The minimum degree of Ramsey-minimal graphs. *J. Graph Theory* **54** (2007), 167–177.
[18] T.E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.* **56** (1960), 13–20.
[19] J. Kahn. Asymptotically good list-colorings. *J. Combin. Theory Ser. A* **73** (1996), 1–59.
[20] J.H. Kim. The Ramsey number $R(3, t)$ has order of magnitude $t^2/\log t$. *Rand. Struct. & Algor.* **7** (1995), 173–207.
[21] M. Krivelevich. Bounding Ramsey numbers through large deviation inequalities. *Rand. Struct. & Algor.* **7** (1995), 145–155.
[22] A. Liebenau. *Orientation Games and Minimal Ramsey Graphs.* PhD thesis, FU Berlin (2013).
[23] C. McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics 1989*, pp. 148–188. Cambridge Univ. Press, Cambridge (1989).
[24] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, pp. 195–248. Springer, Berlin (1998).
[25] D. Osthus and A. Taraz. Random maximal H-free graphs. *Rand. Struct. & Algor.* **18** (2001), 61–82.
[26] Y. Person. Personal communication (RSA 2013 conference in Poznań), 2013.
[27] M. Picollelli. The diamond-free process. *Rand. Struct. & Algor.* **45** (2014), 513–551.
[28] F.P. Ramsey. On a Problem of Formal Logic. *Proc. Lond. Math. Soc.* **30** (1930), 264–286.
[29] V. Rödl and M. Siggers. On Ramsey minimal graphs. *SIAM J. Discrete Math.* **22** (2008), 467–488.
[30] V. Rödl and E. Szemerédi. On size Ramsey numbers of graphs with bounded degree. *Combinatorica* **20** (2000), 257–262.
[31] J. Spencer. Asymptotic lower bounds for Ramsey functions. *Discrete Math.* **20** (1977), 69–76.
[32] J. Spencer. Maximal triangle-free graphs and Ramsey $R(3, t)$. Unpublished manuscript (1995). `http://cs.nyu.edu/spencer/papers/ramsey3k.pdf`.
[33] T. Szabó, P. Zumstein, and S. Zürcher. On the minimum degree of minimal Ramsey graphs. *J. Graph Theory* **64** (2010), 150–164.
[34] L. Warnke. When does the K_4-free process stop? *Rand. Struct. & Algor.* **44** (2014), 355–397.
[35] L. Warnke. The C_ℓ-free process. *Rand. Struct. & Algor.* **44** (2014), 490–526.
[36] L. Warnke. On the method of typical bounded differences. *Combin. Probab. Comput.* **25** (2016), 269–299.
[37] L. Warnke. Upper tails for arithmetic progressions in random subsets. *Israel J. Math.* **221** (2017), 317–365.
[38] G. Wolfovitz. Triangle-free subgraphs in the triangle-free process. *Rand. Struct. & Algor.* **39** (2011), 539–543.

A Appendix

Proof of Theorem 12. We may assume that $\mathcal{I} = \{1, \ldots, |\mathcal{I}|\}$. Recalling $X = f\big((\xi_i)_{i\in\mathcal{I}}\big)$, we define

$$D_i := \mathbb{E}(X \mid \xi_1, \ldots, \xi_{i-1},\ \xi_i = 1) - \mathbb{E}(X \mid \xi_1, \ldots, \xi_{i-1},\ \xi_i = 0) \in [-c_i, 0],$$

where $D_i \leq 0$ follows from the assumption that f is decreasing, and $|D_i| \leq c_i$ follows, as usual, from the assumed discrete Lipschitz property of f. Analogous to, e.g., the proof of [36, Theorem 1.3], writing

$p_i = \mathbb{P}(\xi_i = 1)$ it is routine to check that

$$\Delta_i := \mathbb{E}(X \mid \xi_1, \ldots, \xi_i) - \mathbb{E}(X \mid \xi_1, \ldots, \xi_{i-1}) = D_i(1-p_i)\mathbb{1}_{\{\xi_i=1\}} - D_i p_i \mathbb{1}_{\{\xi_i=0\}}.$$

Since $1 + x \le e^x$ for $x \in \mathbb{R}$ and $e^x \le 1 + x + x^2/2$ for $x \le 0$, for $\theta \ge 0$ it follows easily that

$$\begin{aligned}\mathbb{E}\big(e^{\theta \Delta_i} \mid \xi_1, \ldots, \xi_{i-1}\big) &= (1-p_i) \cdot e^{-\theta D_i p_i} + p_i \cdot e^{\theta D_i (1-p_i)} = e^{-\theta D_i p_i}(1 - p_i + p_i e^{\theta D_i}) \\ &\le e^{-\theta D_i p_i + p_i(e^{\theta D_i}-1)} \le e^{\theta^2 D_i^2 p_i/2} \le e^{\theta^2 c_i^2 p_i/2}.\end{aligned}$$

Hence $\mathbb{E}\big(e^{\theta \sum_{i \in \mathcal{I}} \Delta_i}\big) \le e^{\theta^2 \lambda/2}$, where $\lambda = \sum_{i\in\mathcal{I}} c_i^2 p_i$. Noting $X - \mathbb{E}\, X = \sum_{i \in \mathcal{I}} \Delta_i$, we deduce

$$\mathbb{P}(X \ge \mathbb{E}\, X + t) = \mathbb{P}\big(e^{\theta \sum_{i\in\mathcal{I}} \Delta_i} \ge e^{\theta t}\big) \le \mathbb{E}\big(e^{\theta \sum_{i\in\mathcal{I}}\Delta_i}\big)e^{-\theta t} \le e^{\theta^2\lambda/2 - \theta t} = e^{-t^2/(2\lambda)}$$

by choosing $\theta = t/\lambda$, completing the proof of (48). □

Proof of Lemma 17. Note that the ODE $\Psi'(x) = e^{-\Psi^2(x)}$ and $\Psi(0) = 0$ has the implicit solution

$$x = \int_0^{\Psi(x)} e^{t^2}\, dt. \tag{101}$$

For $x \ge 0$ it follows that $\Psi(x)$ is strictly increasing, so that $\Psi'(x) \ge 0$ is strictly decreasing. Recalling $q_i = \Psi'(i\sigma)$, we deduce $q_i \ge q_{i+1}$ and $0 \le q_i \le q_0 = 1$ for all $i \ge 0$.

To facilitate our upcoming calculations, we first prove the auxiliary claim that, for all $i \ge 0$,

$$\pi_i - \Psi(i\sigma) \in [\sigma, 2\sigma]. \tag{102}$$

Indeed, using $\Psi(0) = 0$ and monotonicity of Ψ' (for the first two inequalities) together with $\Psi'(0) = 1$ and $\Psi' \ge 0$ (for the last inequality) it follows that

$$0 \le \Bigg(\sum_{0 \le j \le i-1} \sigma \Psi'(j\sigma)\Bigg) - \Psi(i\sigma) \le \sigma(\Psi'(0) - \Psi'(i\sigma)) \le \sigma,$$

which establishes (102) by the definition (37) of π_i and $\Psi'(j\sigma) = q_j$.

For (57), note that by (102) and $I = \lceil n^\beta \rceil \gg 1$ it suffices to show $\sqrt{\log x} - 1 \le \Psi(x) \le \sqrt{\log x} + 1$ for $x \ge e$ (with room to spare). The upper bound follows from $\int_0^{\sqrt{\log x}+1} e^{t^2}\, dt \ge x$ and (101). Using the inequality $(y-1)e^{-2y+1} \le 1$ with $y = \sqrt{\log x}$, the lower bound follows from $\int_0^{\sqrt{\log x}-1} e^{t^2}\, dt \le x$ and (101).

Turning to (54), note that the above calculations for (57) imply $\Psi'(x) = e^{-\Psi^2(x)} = x^{-1+o(1)}$ as $x \to \infty$, so that $q_I = n^{-\beta + o(1)}$. Together with $q_i \ge q_I$, it then is routine to see that (54) holds for $\beta < \beta_0 = 1/14$.

Now we focus on (53). As a warm-up, note that $\pi_i \le \pi_I$ for $0 \le i \le I$ by the definition (37) of π_i, and that $\pi_I \le \sqrt{\log(I\sigma)} + 2 \ll \log n = \sigma^{-1/2}$ by (57), so that $\sqrt{\sigma}\pi_i \le 1$. Next, using (102) together with the simple inequalities $e^{-x^2} x \le 1/2$ and $e^{-x^2}x^2 \le 1/2$, we also infer that

$$q_i \pi_i \le e^{-\Psi^2(i\sigma)}\big(\Psi(i\sigma) + 2\sigma\big) \le 1, \tag{103}$$

$$q_i \pi_i^2 \le e^{-\Psi^2(i\sigma)}\big(\Psi^2(i\sigma) + 4\sigma\Psi(i\sigma) + 4\sigma^2\big) \le 1. \tag{104}$$

Combined with $q_i \le 1$ this implies $q_i \pi_i^j \le 1$ for all $j \in \{0, 1, 2\}$, completing the proof of (53).

Turning to (55), note that $\Psi((i+1)\sigma) \le \pi_{i+1} - \sigma \le \pi_i$ by (102), (37) and $q_i \le 1$. Since $\Psi \ge 0$ is increasing and $\Psi' \ge 0$ is decreasing, using $q_j = \Psi'(j\sigma)$ together with $\Psi''(x) = -2\Psi'(x)^2\Psi(x)$ and (103) it follows that

$$|q_i - q_{i+1}| \le \sigma \max_{i\sigma \le \xi \le (i+1)\sigma} |\Psi''(\xi)| \le \sigma \cdot 2\Psi'(i\sigma)^2 \cdot \Psi((i+1)\sigma) \le \sigma \cdot 2q_i^2\pi_i \le \sigma \cdot 2\min\{q_i, q_i\pi_i\}. \tag{105}$$

Noting that (105) also implies $q_i \sim q_{i+1}$, this completes the proof of (55) since $q_i \ge q_{i+1}$.

Finally, for (56) it suffices to show $|q_i - q_{i+1} - 2\sigma q_i^2 \pi_i| \le 8\sigma^2 q_i^2$. Since $q_i = \Psi'(i\sigma)$, it follows that

$$\big|q_i - q_{i+1} + \sigma\Psi''(i\sigma)\big| \le \tfrac{\sigma^2}{2} \max_{i\sigma \le \xi \le (i+1)\sigma} |\Psi'''(\xi)|.$$

As $\Psi'(x) = e^{-\Psi^2(x)}$, it is routine to check that $\Psi'''(x) = 2\Psi'(x)^3\big(4\Psi^2(x) - 1\big)$. Since $\Psi \geq 0$ is increasing and $\Psi' \geq 0$ is decreasing, using $\Psi((i+1)\sigma) \leq \pi_i$ (as above), (104) and $q_i \leq 1$ we infer

$$\max_{i\sigma \leq \xi \leq (i+1)\sigma} |\Psi'''(\xi)| \leq 2\Psi'(i\sigma)^3 \cdot \max\big\{4\Psi^2((i+1)\sigma),\, 1\big\} \leq 2q_i^3 \max\big\{4\pi_i^2,\, 1\big\} \leq 8q_i^2.$$

Furthermore, since $\Psi''(x) = -2\Psi'(x)^2\Psi(x)$, using (102) we deduce

$$\big|\Psi''(i\sigma) - (-2q_i^2\pi_i)\big| = \big|-2q_i^2\Psi(i\sigma) + 2q_i^2\pi_i\big| \leq 4\sigma q_i^2,$$

which completes the proof of (56). $\square$