

Triangle resilience of the square of a Hamilton cycle in random graphs

Manuela Fischer* Nemanja Škorić* Angelika Steger* Miloš Trujić*,†

Abstract

Since first introduced by Sudakov and Vu in 2008, the study of resilience problems in random graphs received a lot of attention in probabilistic combinatorics. Of particular interest are resilience problems of spanning structures. It is known that for spanning structures which contain many triangles, local resilience cannot prevent an adversary from destroying all copies of the structure by removing a negligible amount of edges incident to every vertex. In this paper we generalise the notion of local resilience to H -resilience and demonstrate its usefulness on the containment problem of the square of a Hamilton cycle. In particular, we show that there exists a constant $C > 0$ such that if $p \geq C \log^3 n / \sqrt{n}$ then w.h.p. in every subgraph G of a random graph $G_{n,p}$ there exists the square of a Hamilton cycle, provided that every vertex of G remains on at least a $(4/9 + o(1))$ -fraction of its triangles from $G_{n,p}$. The constant $4/9$ is optimal and the value of p slightly improves on the best-known appearance threshold of such a structure and is optimal up to the logarithmic factor.

1 Introduction

One of the central questions of extremal graph theory concerns determining sufficient conditions for the containment of (spanning) structures. Some of the most influential examples, dating back to the middle of the previous century, include Turán’s theorem [38] and Dirac’s theorem [13]. The former states that having more than $\lfloor n^2/4 \rfloor$ edges in a graph with n vertices is sufficient in order for a triangle to exist, while the latter states that a graph with minimum degree $\lfloor n/2 \rfloor$ is Hamiltonian. Several years later, first Pósa [14], and then Seymour [34], conjectured that for any integer $k \geq 2$, a graph G with n vertices and minimum degree $\delta(G) \geq kn/(k+1)$ contains the k -th power of a Hamilton cycle. For a cycle C and an integer $k \in \mathbb{N}$, the k -th power of a cycle (k -cycle for short) is obtained by including an edge between all pairs of vertices with distance on C of at most k . The second power of a cycle is also called the *square of a cycle*. It required the development of powerful tools, most notably Szemerédi’s regularity lemma and the blow-up lemma, before this conjecture was finally proven by Komlós, Sárközy, and Szemerédi [22], at least for all sufficiently large values of n .

Theorem 1.1 ([22]). *For any $k \in \mathbb{N}$, there exists an $n_0 \in \mathbb{N}$ such that if G has order n with $n \geq n_0$ and $\delta(G) \geq kn/(k+1)$, then G contains the k -th power of a Hamilton cycle.*

For more history on the problem and similar embedding questions we refer the reader to the literature, cf. e.g. [9, 10, 16, 23, 21, 27] and the survey [25].

*Institute of Theoretical Computer Science, ETH Zürich, 8092 Zürich, Switzerland

Email: {manuela.fischer|nskoric|steger|mttrujic}@inf.ethz.ch

†author was supported by grant no. 200021 169242 of the Swiss National Science Foundation.

Generalising the type of problems considered in the above theorem, we arrive at the following Dirac-type question: given a graph property $\mathcal{P}$, what is the minimum number α such that every graph G on n vertices and minimum degree at least αn satisfies $G \in \mathcal{P}$? This leads to the notion of *local resilience* that we now introduce formally.

Definition 1.2 (Local resilience). Let $G = (V, E)$ be a graph and $\mathcal{P}$ a monotone increasing graph property. The *local resilience* of G with respect to $\mathcal{P}$ is defined as:

$$r(G, \mathcal{P}) := \min\{r : \exists \tilde{G} \subseteq G \text{ such that each } v \in V \text{ satisfies} \\ \deg_{\tilde{G}}(v) \leq r \cdot \deg_G(v) \text{ and } G - \tilde{G} \text{ does not have } \mathcal{P}\}.$$

Looking back at the aforementioned results, Dirac's theorem implies that the local resilience of the complete graph K_n with respect to Hamiltonicity is at least $n/2$ and the theorem of Komlós, Sárközy, and Szemerédi implies that the local resilience of 'containment of the k -th power of a Hamilton cycle' is at least $n/(k+1)$. Moreover, it is not too difficult to construct examples that show that both of these results are optimal (consider, for example, $k=2$ and a complete 3-partite graph with two parts of size $(n-1)/3$ and one of size $(n+2)/3$).

In this paper we study how Theorem 1.1 can be transferred to the setting of random graphs. Such transference results recently received considerable attention including several breakthrough results by Balogh, Morris, and Samotij [5], Conlon and Gowers [11], Conlon, Gowers, Samotij, and Schacht [12], Saxton and Thomason [32], and Schacht [33].

We denote by $G_{n,p}$ the probability space of all graphs with vertex set $[n] := \{1, \dots, n\}$ where each edge appears randomly with probability $p := p(n) \in (0, 1)$, independently of all other edges. A systematic study of local resilience in random graphs was initiated by Sudakov and Vu [37] and already led to many beautiful and deep results, see e.g. [3, 1, 8, 24, 26, 28, 29] and the recent surveys [7, 36]. Inspired by other transference results (such as the ones mentioned above as well as many more) from dense graphs to the random setting, one may be tempted to guess that *with high probability*¹ a random graph is such that every subgraph with minimum degree roughly $(2/3 + o(1))np$ contains the square of a Hamilton cycle.

On second thoughts, however, one easily sees that this cannot hold for $p \gg \log n / \sqrt{n}$. An adversary can remove all the edges with both endpoints lying in the neighbourhood of an arbitrary vertex v , thus preventing v from being in a triangle (which implies in particular that v cannot be contained in any square of a cycle); note that the deletion of these edges changes the degree of every other vertex only by $o(np)$. In fact, Huang, Lee, and Sudakov [19] and Balogh, Lee, and Samotij [4] showed that an adversary can always prevent as many as $\Omega(p^{-2})$ vertices from being in triangles by deleting $o(np)$ edges touching each vertex, as long as $1/\sqrt{n} \ll p \ll 1$. The former result shows the claim even when p is a fixed constant, independent of n .

In this paper we overcome the obstacles that the notion of local resilience encounters with respect to containment of spanning structures (that contain triangles). For this we generalise the notion of local resilience. More precisely, we restrict the adversary to only remove a fraction of certain substructures touching each vertex. In the classic definition of local resilience these substructures correspond to edges. For obtaining the square of a Hamilton cycle it turns out that one should replace edges by triangles. This then motivates the following question:

How many triangles at a vertex does an adversary have to destroy in order to obtain a graph without the square of a Hamilton cycle?

¹We say that an event holds with high probability (w.h.p. for short), if the probability that it holds tends to 1 as n tends to infinity.

We capture this question under the notion of K_3 -resilience, or more generally H -resilience as given in the following definition.

Definition 1.3. Let H be a fixed graph and let $\mathcal{P}$ be a monotone increasing graph property. For a graph G , the H -resilience of G with respect to $\mathcal{P}$ is defined as

$$r_H(G, \mathcal{P}) := \min\{r : \exists \tilde{G} \subseteq G \text{ such that the removal of } \tilde{G} \text{ destroys at most an } r\text{-fraction of copies of } H \text{ in } G \text{ at every vertex and } G - \tilde{G} \text{ does not have } \mathcal{P}\}.$$

In the main result of this paper we show that the above definition can be used in order to determine the resilience of $G_{n,p}$ with respect to the containment of the square of a Hamilton cycle.

Theorem 1.4. *The K_3 -resilience of $G_{n,p}$ w.r.t. the containment of the square of a Hamilton cycle is w.h.p. $5/9 \pm o(1)$, provided that $p \gg n^{-1/2} \log^3 n$.*

In other words, the above theorem shows that w.h.p. the adversary needs to delete more than a $(5/9)$ -fraction of the triangles lying on each vertex in order to destroy all copies of the square of a Hamilton cycle in $G_{n,p}$. The density value p is optimal up to the logarithmic factor, as a simple application of the first moment method shows that for $p \ll n^{-1/2}$ a random graph $G_{n,p}$ w.h.p. does not contain the square of a Hamilton cycle. Additionally, this result marginally improves upon the current *appearance threshold* for the square of a Hamilton cycle in $G_{n,p}$ by Nenadov and the second author [30], by a $\log n$ factor in the density p .

The second result of the paper rephrases the above theorem in slightly different terms. From Theorem 1.1 we know that in the dense case it is sufficient to require that the minimum degree is at least $2n/3$. Although the analogous statement cannot be true in the case of random graphs, we prove that w.h.p. every spanning subgraph which satisfies the correct minimum degree condition and the additional property that each edge is contained in αnp^2 triangles, contains the square of a Hamilton cycle. Before we can state this result precisely, we need a definition.

Definition 1.5. Let Γ be a graph on n vertices. We denote by $\mathcal{G}(\Gamma, n, \alpha, p)$ the family of all spanning subgraphs $G \subseteq \Gamma$ that satisfy the following properties:

1. for every $v \in V(G)$: $\deg_G(v) \geq (2/3 + \alpha)np$, and
2. for every $\{u, v\} \in E(G)$: $|N_G(u) \cap N_G(v)| \geq \alpha np^2$.

With this at hand we can state the second result of the paper.

Theorem 1.6. *For every $\alpha > 0$ there exists a positive constant $C := C(\alpha)$, such that a random graph $\Gamma \sim G_{n,p}$ w.h.p. has the following property, provided that $p \geq Cn^{-1/2} \log^3 n$. Each member of $\mathcal{G}(\Gamma, n, \alpha, p)$ contains the square of a Hamilton cycle.*

As in the first result of the paper, the value of p is almost optimal. Furthermore, the constant $2/3$ in the definition of the class $\mathcal{G}(\Gamma, n, \alpha, p)$ cannot be improved, as the same counterexample as in the dense case works in this scenario as well. Even though the type of conditions in two theorems above look quite different at first sight, we prove Theorem 1.4 by a reduction to Theorem 1.6.

The proof of Theorem 1.6 uses the so-called *absorbing method*. In particular, we make use of a strategy paved by Nenadov and the second author [30]. This method is discussed in Section 6. In Section 2 we introduce some notation and probabilistic tools, and state several useful lemmas about properties of (random) graphs, culminating in Lemma 2.9 about edge expansion properties.

In Section 3 we give the proof of Theorem 1.4 by a reduction to Theorem 1.6. We also show in this section that the constant $5/9$ in Theorem 1.4 is best possible. In Section 4 we introduce several classes and definitions of graphs which we rely on throughout the paper. In Section 5 we give the proof of Theorem 1.6 modulo several lemmas. Each of the subsequent Sections 6–7.3 are dedicated to the proof of one of the technical lemmas and/or claims. Finally, we conclude by discussing some related open problems in Section 8.

2 Tools and preliminaries

Our graph theoretic notation is standard (see, e.g. [6]). In particular, for a graph $G = (V, E)$ we denote by $N_G(v)$ the neighbourhood of a vertex $v \in V$ and by $\deg_G(v)$ its size, i.e. $\deg_G(v) = |N_G(v)|$. Similarly, for $X \subseteq V$ we write $N_G(X)$ for the union of neighbourhoods of vertices in X , that is $N_G(X) := \{u : \{v, u\} \in E \text{ and } v \in X\}$. Furthermore, for $X, Y \subseteq V$, we let $N_G(X, Y)$ denote $N_G(X) \cap Y$ and if X consists of a single vertex we abbreviate $N_G(\{x\}, Y)$ to $N_G(x, Y)$. If $X, Y \subseteq V$ are disjoint subsets of vertices we write $e_G(X, Y)$ for the number of edges with one endpoint in X and the other in Y . We use a set of edges $I \subseteq E$ interchangeably as a set of edges and a (sub)graph. In particular, we write $\deg_I(v)$ to denote the number of edges from I that are incident to a vertex v and $e_I(X, Y)$ for the number of edges in I that have an endpoint in each of the subsets X and Y . We omit the subscript G (resp. I) whenever it is clear from the context to which graph G we refer to. For $k, \ell \in \mathbb{N}$ and a cycle C_ℓ with ℓ vertices, we let C_ℓ^k denote the k -th power of C_ℓ , that is a graph obtained by adding an edge between any two vertices of C_ℓ which are at distance at most k . Given two graphs H and G , and a function $f: V(H) \rightarrow V(G)$, we say that f is an *embedding* of H into G if it is an injection and for all $\{v, u\} \in E(H)$ we have $\{f(u), f(v)\} \in E(G)$.

For an integer $k \geq 2$ and a set V we write $\binom{V}{k}$ for the family of all subsets of V with cardinality exactly k . We write V^k to denote the family of all ordered k -tuples of V whose entries are pairwise different, that is $V^k := \{(v_1, \dots, v_k) : v_i \in V \text{ for all } i \in [k] \text{ and } v_i \neq v_j \text{ for all } 1 \leq i < j \leq k\}$. An element of V^k is usually denoted by a lower case bold letter. Given $\mathbf{w} = (v_1, \dots, v_k) \in V^k$, we write $\overline{\mathbf{w}}$ to denote the tuple obtained by reversing the order in $\mathbf{w}$, i.e. $\overline{\mathbf{w}} = (v_k, \dots, v_1)$. Moreover, for two ordered tuples $\mathbf{w}_1$ and $\mathbf{w}_2$, the tuple $\mathbf{w}_1\mathbf{w}_2$ is an ordered tuple obtained by concatenation of $\mathbf{w}_1$ and $\mathbf{w}_2$. For a function g applicable to the elements of a tuple $(x_1, \dots, x_n)$ we for convenience shorten $(g(x_1), \dots, g(x_n))$ to $g(x_1, \dots, x_n)$.

For an integer $n \in \mathbb{N}$ we write $[n] := \{1, \dots, n\}$. Given $a, b, c, x \in \mathbb{R}$ we write $x \in (a \pm b)c$ to denote $(a - b)c \leq x \leq (a + b)c$. We make use of the standard asymptotic notation, o , O , ω , Ω , and Θ . For two functions a and b , we write $a \ll b$ to denote $a = o(b)$ and similarly $a \gg b$ for $a = \omega(b)$. All logarithms are with respect to base e . We omit floors and ceilings whenever they are not of importance. Lastly, we write $C_{5.1}$ to indicate that the constant $C_{5.1}$ is given by Theorem/Lemma/Claim 5.1.

The following statement about r -star-matchings is an easy corollary of Hall's matching theorem [17]. A star of size r (r -star, for short) is a complete bipartite graph $K_{1,r}$ with the vertex adjacent to all others being the *centre*.

Lemma 2.1 (r -star-matching). *Let $r \geq 1$ be an integer and let $G = (A \cup B, E)$ be a bipartite graph. If for every subset $A' \subseteq A$ it holds that $|N(A')| \geq r|A'|$, then G contains a collection of pairwise disjoint r -stars, such that the centres of these stars cover all vertices in A .*

Proof. Consider the ‘blow-up’ of G in which each vertex in A is replaced by r copies that are connected to the same vertices as the original vertex. Then this new graph satisfies Hall's

condition and thus contains a matching that saturates all copies of the vertices in A . The corollary follows by contracting the copies of each vertex. $\square$

We also make use of a generalised version of Hall's theorem due to Haxell which has recently seen a surge of applications in embedding spanning structures into random graphs, especially in the resilience setting.

Theorem 2.2 ([18]). *Let $\mathcal{H} = (A \cup B, E)$ be an r -uniform hypergraph such that $|A \cap e| = 1$ and $|B \cap e| = r - 1$ for every edge $e \in E$. If for every $A' \subseteq A$ and $B' \subseteq B$ such that $|B'| \leq (2r - 3)|A'|$ there is an edge $e \in E$ intersecting A' but not B' , then $\mathcal{H}$ contains an A -saturating matching (that is, a collection of vertex-disjoint hyperedges whose union contains A).*

We repeatedly make use of the following two standard tail estimates used in random graph theory, cf. e.g. [2, 15].

Lemma 2.3 (Chernoff's inequality). *Let $X \sim \text{Bin}(n, p)$ and let $\mu := \mathbb{E}[X]$. Then for all $0 < a < 1$:*

- $\Pr[X < (1 - a)\mu] < e^{-a^2\mu/2}$, and
- $\Pr[X > (1 + a)\mu] < e^{-a^2\mu/3}$.

Moreover, the inequalities above also hold if X has the hypergeometric distribution with the same mean.

Theorem 2.4 (Janson's inequality). *Let $p \in (0, 1)$ and consider a family $\{H_i\}_{i \in \mathcal{I}}$ of subgraphs of the complete graph on the vertex set $[n]$. Let $\Gamma \sim G_{n,p}$. For each $i \in \mathcal{I}$, let X_i denote the indicator random variable for the event $\{H_i \subseteq \Gamma\}$ and, for each ordered pair $(i, j) \in \mathcal{I} \times \mathcal{I}$ with $i \neq j$, write $H_i \sim H_j$ if $E(H_i) \cap E(H_j) \neq \emptyset$. Let*

$$\begin{aligned} X &:= \sum_{i \in \mathcal{I}} X_i, \\ \mu &:= \mathbb{E}[X] = \sum_{i \in \mathcal{I}} p^{e(H_i)}, \\ \Delta &:= \sum_{\substack{(i,j) \in \mathcal{I} \times \mathcal{I} \\ H_i \sim H_j}} \mathbb{E}[X_i X_j] = \sum_{\substack{(i,j) \in \mathcal{I} \times \mathcal{I} \\ H_i \sim H_j}} p^{e(H_i) + e(H_j) - e(H_i \cap H_j)}. \end{aligned}$$

Then for all $0 < \gamma < 1$ we have

$$\Pr[X < (1 - \gamma)\mu] \leq e^{-\frac{\gamma^2 \mu^2}{2(\mu + \Delta)}}.$$

Next, we collect several facts about random graphs mostly concerning the number of edges and triangles between certain subsets, as well as a simple edge expansion property.

Lemma 2.5. *For every $\varepsilon \in (0, 1/100)$ and $p := p(n) \in (0, 1)$, the random graph $\Gamma \sim G_{n,p}$ w.h.p. satisfies the following.*

Let $s \geq \varepsilon^{-10} \log n / p^2$ be an integer. Then there are at least $(1 - n^{-6}) \binom{n}{s}$ subsets $W \subseteq V(\Gamma)$ of size s such that the following holds. For every $W' \subseteq W$ of size $|W'| \geq \varepsilon |W|$ and every family of pairs $\mathcal{P} \subseteq \binom{V(\Gamma) \setminus W}{2}$ of size $|\mathcal{P}| \geq \varepsilon^{-10} \log n / p^2$ and such that no vertex of Γ appears in more than $1 + 1/p$ pairs from $\mathcal{P}$, we have

$$\sum_{\{u,v\} \in \mathcal{P}} |N(u, W') \cap N(v, W')| \leq (1 + \varepsilon) |\mathcal{P}| |W'| p^2.$$

Proof. An easy application of Chernoff's inequality (Lemma 2.3) and the union bound shows that w.h.p. every pair of vertices $u, v \in V(\Gamma)$ have a common neighbourhood which satisfies $|N(u) \cap N(v)| \leq (1 + \varepsilon^6)np^2$. For a set $W'' \subseteq V(\Gamma)$ and a triple $(u, v, w) \in (V(\Gamma) \setminus W'') \times (V(\Gamma) \setminus W'') \times W''$, we define a random variable

$$X_{u,v,w} = \begin{cases} 1, & \text{if } \{u, v\} \in \mathcal{P} \text{ and } \{u, v\} \subseteq N(w), \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that w.h.p. for all $\mathcal{P}$ as in the statement of the lemma and all $W'' \subseteq V(\Gamma)$, $|W''| \geq \varepsilon^{-7} \log n/p^2$, the following holds:

$$\sum_{\{u,v\} \in \mathcal{P}} \sum_{w \in W''} X_{u,v,w} \geq (1 - \varepsilon^3)|\mathcal{P}||W''|p^2. \quad (1)$$

Condition on the fact that $\Gamma \sim G_{n,p}$ satisfies these two properties, which happens with high probability. We claim that this is sufficient in order to show the lemma.

Let $W \subseteq V(\Gamma)$ be a subset of size $s \geq \varepsilon^{-10} \log n/p^2$ chosen uniformly at random among all such subsets. For a fixed pair $u, v \in V(\Gamma)$, from Chernoff's inequality we have

$$\Pr[|N(u, W) \cap N(v, W)| > (1 + \varepsilon^3)|W|p^2] \leq e^{-(\varepsilon^3/3)|W|p^2} \leq e^{-\varepsilon^{-6} \log n} < 1/n^9,$$

and so by the union bound we conclude that with probability at least $1 - o(n^{-6})$ for every two vertices $u, v \in V(\Gamma)$ it holds that

$$|N(u, W) \cap N(v, W)| \leq (1 + \varepsilon^3)|W|p^2. \quad (2)$$

Let $W' \subseteq W$ be of size $|W'| \geq \varepsilon|W|$, let $X = \sum_{\{u,v\} \in \mathcal{P}} \sum_{w \in W'} X_{u,v,w}$, and note that X counts exactly the quantity we are interested in. Take $W'' := W \setminus W'$. Then, trivially,

$$X = \sum_{\{u,v\} \in \mathcal{P}} \sum_{w \in W'} X_{u,v,w} = \sum_{\{u,v\} \in \mathcal{P}} \sum_{w \in W} X_{u,v,w} - \sum_{\{u,v\} \in \mathcal{P}} \sum_{w \in W''} X_{u,v,w}.$$

The first term on the right hand side of the previous equation is by (2) bounded from above by $(1 + \varepsilon^3)|\mathcal{P}||W|p^2$. If $|W''| \geq \varepsilon^3|W| \geq \varepsilon^{-7} \log n/p^2$, then we can use the lower bound from (1) to obtain

$$X \leq (1 + \varepsilon^3)|\mathcal{P}||W|p^2 - (1 - \varepsilon^3)|\mathcal{P}||W''|p^2 \leq (1 + \varepsilon^3)|\mathcal{P}||W'|p^2 + 2\varepsilon^3|\mathcal{P}||W|p^2 \leq (1 + \varepsilon)|\mathcal{P}||W'|p^2,$$

where the last inequality holds because $|W'| \geq \varepsilon|W|$ and $\varepsilon < 1/100$. In case $|W''| < \varepsilon^3|W|$ we have $|W'| \geq (1 - \varepsilon^3)|W|$ and thus from (2) we have

$$X \leq (1 + \varepsilon^3)|\mathcal{P}||W|p^2 \leq \frac{1 + \varepsilon^3}{1 - \varepsilon^3}|\mathcal{P}||W'|p^2 \leq (1 + \varepsilon)|\mathcal{P}||W'|p^2,$$

since $\varepsilon < 1/100$. In conclusion, W satisfies the assertion of the lemma with probability at least $1 - o(n^{-6})$ which implies the desired statement. It remains to show that (1) is indeed true.

Denote the left hand side of (1) by X'' . By linearity of expectation we have $\mu := \mathbb{E}[X''] = |\mathcal{P}||W''|p^2$. For two triples (u_1, v_1, w_1) and (u_2, v_2, w_2) we write $(u_1, v_1, w_1) \sim (u_2, v_2, w_2)$ if their corresponding random variables are dependent. Note that (u_1, v_1, w_1) and (u_2, v_2, w_2) can only be dependent if $|\{u_1, v_1\} \cap \{u_2, v_2\}| = 1$ and $w_1 = w_2$. Thus

$$\Delta := \sum_{(u_1, v_1, w_1) \sim (u_2, v_2, w_2)} p^3 \leq |\mathcal{P}||W''| \frac{2}{p} p^3 = 2\mu.$$

By Janson's inequality (Theorem 2.4) it follows that

$$\Pr[X'' < (1 - \varepsilon^3)|\mathcal{P}||W''|p^2] \leq e^{-\varepsilon^6\mu^2/(6\mu)} = e^{-\varepsilon^6|\mathcal{P}||W''|p^2/6}.$$

Let us denote $\varepsilon^{-10} \log n/p^2$ by t . By the union bound over all choices of $\mathcal{P}$ and W'' and by using standard bounds on binomial coefficients, we get

$$\begin{aligned} \sum_{x=t}^{n^2} \binom{n^2}{x} \sum_{y=\varepsilon^3 t}^n \binom{n}{y} e^{-\frac{\varepsilon^6}{6} xy p^2} &\leq \sum_{x=t}^{n^2} \sum_{y=\varepsilon^3 t}^n e^{2x \log n} \cdot e^{y \log n} \cdot e^{-\frac{\varepsilon^9}{6} xy p^2} \\ &\leq \sum_{x=t}^{n^2} \sum_{y=\varepsilon^3 t}^n e^{(2x+y) \log n - \max\{x,y\} \cdot \frac{1}{6\varepsilon} \log n} \\ &\leq \sum_{x=t}^{n^2} \sum_{y=\varepsilon^3 t}^n e^{-10 \max\{x,y\} \log n} \leq n^3 \cdot e^{-10\varepsilon^{-10} \log^2 n/p^2} < 1/n^8. \end{aligned}$$

This implies that with probability at least $1 - o(n^{-5})$ we have

$$\sum_{\{u,v\} \in \mathcal{P}} \sum_{w \in W''} X_{u,v,w} \geq (1 - \varepsilon^3)|\mathcal{P}||W''|p^2,$$

for all permissible $\mathcal{P}$ and W'' , as required. $\square$

Lemma 2.6. *For every $\varepsilon \in (0, 1/300)$ and $p := p(n) \in (0, 1)$, the random graph $\Gamma \sim G_{n,p}$ w.h.p. satisfies the following.*

Let $s \geq \varepsilon^{-3} \log n/p^2$ be an integer. Then there are at least $(1 - n^{-6}) \binom{n}{s}$ subsets $W \subseteq V(\Gamma)$ of size s such that the following holds. For every family of pairs $\mathcal{P} \subseteq \binom{V(\Gamma) \setminus W}{2}$ such that no vertex of Γ appears in more than $1 + \varepsilon/p$ pairs from $\mathcal{P}$ and $|\mathcal{P}| \leq 1 + \varepsilon/p^2$, we have

$$\left| \bigcup_{\{u,v\} \in \mathcal{P}} (N(u, W) \cap N(v, W)) \right| = (1 \pm 5\varepsilon)|\mathcal{P}||W|p^2.$$

Proof. Fix a $\mathcal{P}$ as in the statement and let U denote the set of vertices of Γ not appearing in $\mathcal{P}$; note that $|U| = (1 - o(1))n$. Let q denote the probability for a vertex $w \in U$ to be in the common neighbourhood of u and v , for some $\{u, v\} \in \mathcal{P}$, and let $\mathcal{E}_{u,v}$ denote this event for fixed u and v . By the union bound we have that $q \leq |\mathcal{P}|p^2$. As for the lower bound on q , we use the inclusion-exclusion principle (Bonferroni's inequality) to get

$$\begin{aligned} q &\geq \sum_{\{u,v\} \in \mathcal{P}} \Pr[\mathcal{E}_{u,v}] - \sum_{\{u,v\} \cap \{u',v'\} < 2} \Pr[\mathcal{E}_{u,v} \wedge \mathcal{E}_{u',v'}] \\ &= \sum_{\{u,v\} \in \mathcal{P}} \Pr[\mathcal{E}_{u,v}] - \sum_{|\{u,v\} \cap \{u',v'\}|=0} \Pr[\mathcal{E}_{u,v} \wedge \mathcal{E}_{u',v'}] - \sum_{|\{u,v\} \cap \{u',v'\}|=1} \Pr[\mathcal{E}_{u,v} \wedge \mathcal{E}_{u',v'}]. \end{aligned}$$

Using the fact that no vertex of Γ appears in more than $1 + \varepsilon/p$ pairs from $\mathcal{P}$,

$$q \geq |\mathcal{P}|p^2 - |\mathcal{P}|(|\mathcal{P}| - 1)p^4 - |\mathcal{P}|\frac{\varepsilon}{p}p^3 \geq (1 - 2\varepsilon)|\mathcal{P}|p^2.$$

Let us denote $\bigcup_{\{u,v\} \in \mathcal{P}} (N(u, U) \cap N(v, U))$ by Z . Observe that the expected size of Z is $|U|q$ and thus by Chernoff's inequality and our estimates for q we get

$$\Pr[|Z| \leq (1 - 4\varepsilon)|\mathcal{P}|np^2] \leq \Pr[|Z| \leq (1 - \varepsilon)|U|q] \leq e^{-\frac{\varepsilon^2}{2}(1-2\varepsilon)|\mathcal{P}||U|p^2} \leq e^{-\frac{\varepsilon^2}{8}|\mathcal{P}|np^2}. \quad (3)$$

Similarly, we have

$$\Pr[|Z| \geq (1 + 4\varepsilon)|\mathcal{P}|np^2] \leq \Pr[|Z| \geq (1 + \varepsilon)|U|q] \leq e^{-\frac{\varepsilon^2}{3}(1-2\varepsilon)|\mathcal{P}||U|p^2} \leq e^{-\frac{\varepsilon^2}{12}|\mathcal{P}|np^2}. \quad (4)$$

Since $np^2 \geq \varepsilon^{-3} \log n$, by combining (3) and (4) together with the union bound over all choices for $\mathcal{P}$ we get that with high probability for all such sets $\mathcal{P}$ and $u, v \in \mathcal{P}$

$$\left| \bigcup_{\{u,v\} \in \mathcal{P}} (N(u, U) \cap N(v, U)) \right| = (1 \pm 4\varepsilon)|\mathcal{P}|np^2. \quad (5)$$

Condition on $\Gamma \sim G_{n,p}$ satisfying this. Let $W \subseteq V(\Gamma)$ be a set of size $s \geq \varepsilon^{-3} \log n/p^2$ chosen uniformly at random among all such sets. Let $\mathcal{P}$ be as in the statement of the lemma, and note that (5) is fulfilled for $\mathcal{P}$ and the corresponding set U . By Chernoff's inequality, we have

$$\left| \bigcup_{\{u,v\} \in \mathcal{P}} (N(u, W) \cap N(v, W)) \right| = (1 \pm 5\varepsilon)|\mathcal{P}||W|p^2,$$

with probability at least

$$1 - e^{-\frac{2\varepsilon^2}{3}|\mathcal{P}||W|p^2} \geq 1 - e^{-\frac{2}{3\varepsilon} \log n} \geq 1 - n^{-6},$$

as required. $\square$

The following definition is used at various places throughout the paper. It captures essential properties of a random graph $\Gamma \sim G_{n,p}$ and its subgraphs $G \in \mathcal{G}(\Gamma, \alpha, \varepsilon, p)$ which are used in order to prove the main result. Some of the properties follow easily from others; we list them all separately for ease of reference later on.

Definition 2.7. Let $\varepsilon, \alpha > 0$, $p \in (0, 1)$, and $n \in \mathbb{N}$. Let Γ be a graph on n vertices and $G \in \mathcal{G}(\Gamma, n, 2\alpha, p)$. We say that a subset $W \subseteq V(G)$ is (α, ε, p) -good with respect to Γ and G if the following properties hold:

(G1) For every two disjoint subsets $X \subseteq W$, $Y \subseteq V(\Gamma)$ of sizes $|X|, |Y| \geq \varepsilon^{-3} \log n/p$ we have

$$e_\Gamma(X, Y) = (1 \pm \varepsilon)|X||Y|p.$$

(G2) For every $v \in V(\Gamma)$ we have $\deg_\Gamma(v, W) = (1 \pm \varepsilon)|W|p$.

(G3) For every $v \in V(G)$ we have $\deg_G(v, W) = (1 \pm \varepsilon) \deg_G(v) \frac{|W|}{n}$.

(G4) For every $v \in V(G)$ we have $\deg_G(v, W) \geq (2/3 + \alpha)|W|p$.

(G5) For every $\{u, v\} \in E(G)$ we have $|N_G(u, W) \cap N_G(v, W)| \geq \alpha|W|p^2$.

(G6) For all subsets $W' \subseteq W$ of size $|W'| \geq \varepsilon|W|$ and all $\mathcal{P} \subseteq \binom{V(G) \setminus W}{2}$ of size $|\mathcal{P}| \geq \varepsilon^{-10} \log n/p^2$ such that no vertex of G appears in more than $1 + 1/p$ pairs from $\mathcal{P}$, we have

$$\sum_{\{u,v\} \in \mathcal{P}} |N_G(u, W') \cap N_G(v, W')| \leq (1 + \varepsilon)|\mathcal{P}||W'|p^2.$$

(G7) For every set of edges $\mathcal{P} \subseteq E(G)$ avoiding W such that no vertex of G appears in more than $1 + \varepsilon/p$ edges from $\mathcal{P}$ and $|\mathcal{P}| \leq 1 + \varepsilon/p^2$, we have

$$\left| \bigcup_{\{u,v\} \in \mathcal{P}} (N_G(u, W) \cap N_G(v, W)) \right| \geq \alpha|\mathcal{P}||W|p^2.$$

In order to be precise, whenever speaking about (α, ε, p) -good sets, one would always need to specify graphs Γ and G as well. However, for a cleaner exposition, we omit it as these two graphs are always clear from the context and almost always are the random graph $\Gamma \sim G_{n,p}$ and its subgraph $G \in \mathcal{G}(\Gamma, \alpha, \varepsilon, p)$. The following proposition shows that w.h.p. an overwhelming majority of sets W are $(\alpha/2, \varepsilon, p)$ -good with respect to $\Gamma \sim G_{n,p}$ and its subgraph G which belongs to the class $\mathcal{G}(\Gamma, \alpha, \varepsilon, p)$, for a certain choice of parameters. Namely, a random choice of W is typically good.

Proposition 2.8. *For every $\alpha > 0$ there exists a positive constant $\varepsilon := \varepsilon(\alpha)$ such that the random graph $\Gamma \sim G_{n,p}$ w.h.p. satisfies the following for every $G \in \mathcal{G}(\Gamma, n, \alpha, p)$.*

Let $s \geq \varepsilon^{-10} \log n / p^2$ be an integer. Then there are at least $(1 - n^{-5}) \binom{n}{s}$ subsets $W \subseteq V(G)$ such that W is $(\alpha/2, \varepsilon, p)$ -good.

Proof. Choose $\varepsilon > 0$ sufficiently small so that the arguments below follow through. One can easily show with Chernoff's inequality and the union bound that w.h.p. $\Gamma \sim G_{n,p}$ is such that

- (density) $e_\Gamma(X, Y) = (1 \pm \varepsilon)|X||Y|p$, for every $X, Y \subseteq V(\Gamma)$ with $|X|, |Y| \geq \varepsilon^{-3} \log n / p$;
- (degree) $\deg_\Gamma(v) = (1 \pm \varepsilon/2)np$, for every $v \in V(\Gamma)$; and
- (codegree) $|N_\Gamma(u) \cap N_\Gamma(v)| \leq (1 + \varepsilon/2)np^2$, for every $u, v \in V(\Gamma)$.

Furthermore, with high probability, Γ satisfies both the conclusion of Lemma 2.5 and that of Lemma 2.6. Condition on these five events from now on.

Let $G \subseteq \Gamma$ be a member of $\mathcal{G}(\Gamma, n, \alpha, p)$ and let $H := \Gamma - G$. We choose a set W uniformly at random among all subsets of size s and aim to show that with probability at least $1 - o(n^{-5})$ such a set satisfies all (G1)–(G7), which would imply the desired statement. Property (G1) follows directly from the density event we conditioned on above. Properties (G2)–(G5) follow from the degree event we conditioned on above, the definition of $\mathcal{G}(\Gamma, n, \alpha, p)$, and then the fact that W is chosen u.a.r., Chernoff's inequality, and a simple union bound. Moreover, (G6) holds for W with probability at least $1 - n^{-6}$ by the conclusion of Lemma 2.5, since $G \subseteq \Gamma$.

Finally, let us look at (G7). By the conclusion of Lemma 2.6, with probability at least $1 - n^{-6}$, W is one of the subsets for which the following holds: for any subset of edges $\mathcal{P} \subseteq E(G)$ avoiding W such that $|\mathcal{P}| \leq 1 + \varepsilon/p^2$ and no vertex of G appearing in more than $1 + \varepsilon/p$ edges from $\mathcal{P}$, we have

$$\left| \bigcup_{\{u,v\} \in \mathcal{P}} (N_\Gamma(u, W) \cap N_\Gamma(v, W)) \right| \geq (1 - 5\varepsilon)|\mathcal{P}||W|p^2. \quad (6)$$

From the codegree event we conditioned on above and the definition of the class $\mathcal{G}(\Gamma, n, \alpha, p)$ we further get that

$$|N_H(u) \cap N_H(v)| \leq (1 + \varepsilon/2 - \alpha)np^2 \leq (1 - 3\alpha/4)np^2,$$

for every edge $\{u, v\} \in E(G)$. Moreover, as W is chosen uniformly at random, by Chernoff's inequality and the union bound we also have that with probability at least $1 - o(n^{-5})$, every $\{u, v\} \in E(G)$ satisfies

$$|N_H(u, W) \cap N_H(v, W)| \leq (1 - 3\alpha/4 + \varepsilon)|W|p^2. \quad (7)$$

By combining (6) and (7) we get

$$\begin{aligned} \left| \bigcup_{\{u,v\} \in \mathcal{P}} (N_G(u, W) \cap N_G(v, W)) \right| &\geq \left| \bigcup_{\{u,v\} \in \mathcal{P}} (N_\Gamma(u, W) \cap N_\Gamma(v, W)) \right| - \sum_{\{u,v\} \in \mathcal{P}} |N_H(u, W) \cap N_H(v, W)| \\ &\geq (1 - 5\varepsilon)|\mathcal{P}||W|p^2 - |\mathcal{P}|(1 - 3\alpha/4 + \varepsilon)|W|p^2 \\ &\geq (\alpha/2)|\mathcal{P}||W|p^2, \end{aligned}$$

where the last inequality follows for small enough $\varepsilon > 0$. This concludes the proof. $\square$

2.1 Edge expansion and triangles

In this section we provide some expansion tools that we make use of in our proof. Before we state them, we give some motivational background. A standard approach for showing that two vertices a and b are connected by a path of length, say, ℓ , is to inductively prove a lower bound on the number of vertices that can be reached by paths of length i , starting from each a and b . To prove such a bound one usually relies on expansion properties of vertices in certain sets. In our case, we do not want to find paths, but *square paths* (see Section 4 for details), where each new vertex is given by a triangle lying on a previous edge. In particular, we build such paths by starting from an edge and determining how many edges (instead of vertices) we can reach starting from this edge. Correspondingly, we need expansion properties of edges instead of vertices. The goal of this section is to provide such expansion properties.

More precisely, we are trying to understand the following setup. Suppose we are given three disjoint subsets of vertices W_1, W_2, W_3 of a graph G . Assume further that $F_{12} \subseteq E_G(W_1, W_2)$ is some set of edges from G between W_1 and W_2 . What we are interested in is the set of edges

$$F_{23} := \{\{w_2, w_3\} \in E_G(W_2, W_3) : \exists w_1 \in W_1 \text{ s.t. } \{w_1, w_2\} \in F_{12} \text{ and } \{w_1, w_3\} \in E(G)\}$$

that ‘extend’ an edge from F_{12} to the set W_3 via a triangle. Namely, we are aiming at providing some bound on $|F_{23}|$ in terms of $|F_{12}|$.

Lemma 2.9. *For every $\alpha > 0$ there exists a positive constant $\varepsilon := \varepsilon(\alpha)$ such that for sufficiently large n and all $p := p(n) \in (0, 1)$, every graph Γ on n vertices and $G \in \mathcal{G}(\Gamma, n, 2\alpha, p)$ satisfy the following.*

Let $\tilde{n} \geq \varepsilon^{-22} \log^2 n / p^2$ be an integer. Let $W_1, W_2, W_3 \subseteq V(G)$ be three disjoint (α, ε, p) -good sets of size $\tilde{n}$ each, let $X \subseteq W_3$ be of size $|X| \leq \varepsilon^4 \tilde{n}$, and $F_{12} \subseteq E_G(W_1, W_2)$ be a subset of edges. Then the following statements hold, where $U \subseteq W_2$ is the set of vertices incident to the edges in F_{12} , and F_{23} is as defined above.

- (1) *If $|U| \geq |X| / \log n$ and $\deg_{F_{12}}(v) \leq \varepsilon/p$ for all $v \in W_1 \cup W_2$, then there exists a subset $U' \subseteq U$ of size $|U'| = (1 - \varepsilon) \min\{|U|, \varepsilon/p^2\}$ and an $(\alpha \tilde{n} p^2 / 2)$ -star-matching in F_{23} saturating U' and avoiding X in W_3 .*
- (2) *If $|F_{12}| \geq \varepsilon^{-17} \tilde{n} \log^2 n$ and $\deg_{F_{12}}(v, W_1) \leq \frac{\varepsilon^{-4} \log n}{p}$ for all $v \in U$, then $e_{F_{23}}(U, W_3 \setminus X) \geq \varepsilon^{-4} |F_{12}|$.*
- (3) *If $|F_{12}| \geq \varepsilon^{-5} \tilde{n} \log n$ and $\deg_{F_{12}}(v, W_1) \geq \frac{\varepsilon^{-4} \log n}{p}$ for all $v \in U$, then $e_{F_{23}}(U, W_3 \setminus X) \geq \frac{\alpha \tilde{n} p}{4} |U|$.*
- (4) *If $|F_{12}| \geq \frac{\varepsilon^{-5} \log n}{p} \tilde{n}$ and $\frac{\varepsilon^{-4} \log n}{p} \leq \deg_{F_{12}}(v, W_1) \leq \frac{\tilde{n} p}{3}$ for all $v \in U$, then $e_{F_{23}}(U, W_3 \setminus X) \geq (1 + \alpha/4) |F_{12}|$.*
- (5) *If $|F_{12}| \geq \frac{\varepsilon^{-10} \log n}{p} \tilde{n}$ and $\deg_{F_{12}}(v, W_1) \geq \frac{\tilde{n} p}{3}$ for all $v \in U$, then $e_{F_{23}}(U, W_3 \setminus X) \geq (1 - \varepsilon^2) e_G(U, W_3)$ and there exists a subset $L \subseteq U$ of size $|L| \geq (1 - 3\varepsilon^3) |U|$, such that for every $u \in L$ we have $\deg_{F_{23}}(u, W_3 \setminus X) \geq (1 - 2\varepsilon^3) \deg_G(u, W_3)$.*
- (6) *If $|U| \geq \frac{2}{3} \tilde{n}$ and $\deg_{F_{12}}(v, W_1) \geq \frac{\tilde{n} p}{3}$ for all $v \in U$, then there exists $L \subseteq W_3 \setminus X$ of size $|L| \geq (1 - \varepsilon) \tilde{n}$, such that for every $u \in L$ we have $\deg_{F_{23}}(u, W_2) \geq (1/3 + \alpha/2) \tilde{n} p$.*
- (7) *If $|F_{12}| \geq \varepsilon^{-18} \tilde{n} \log^2 n$ and $e_{F_{23}}(W_2, W_3 \setminus X) < \varepsilon^{-3} |F_{12}|$, then there exists a subset $L \subseteq W_2$ such that for every $v \in L$ we have $\deg_{F_{12}}(v, W_1) \geq \varepsilon^{-4} \log n / p$ and $e_{F_{12}}(W_1, L) \geq (1 - \varepsilon) |F_{12}|$.*
- (8) *If $|F_{12}| \geq \varepsilon \tilde{n}^2 p$ and $e_{F_{23}}(W_2, W_3 \setminus X) < (1 + \mu \alpha / 8) |F_{12}|$, then there exists a subset $L \subseteq W_2$ such that for every $v \in L$ we have $\deg_{F_{12}}(v, W_1) > \tilde{n} p / 3$ and $e_{F_{12}}(W_1, L) \geq (1 - \mu) |F_{12}|$,*

for every $\mu \in (32\varepsilon/\alpha, 1)$.

(9) If $|F_{12}| \geq \varepsilon \tilde{n}^2 p$, then $e_{F_{23}}(W_2, W_3 \setminus X) \geq (1 - \sqrt{\varepsilon})|F_{12}|$.

Let $\varepsilon := \varepsilon(\alpha) > 0$ be a small enough constant such that all the arguments follow through. We prove the statements one by one.

Proof of (1) in Lemma 2.9. Fix an arbitrary set $\tilde{U} \subseteq U$ of size $|\tilde{U}| = \min\{|U|, \varepsilon/p^2\}$. Every subset $A \subseteq \tilde{U}$, by using property (G7) applied for a single edge from F_{12} incident to each $a \in A$ (as $\mathcal{P}$), satisfies

$$|N_{F_{23}}(A)| \geq \alpha|A|\tilde{n}p^2.$$

Lemma 2.1 thus implies there exists an $(\alpha\tilde{n}p^2)$ -star-matching M in F_{23} saturating $\tilde{U}$. Let $\tilde{U}_X$ be the largest subset of $\tilde{U}$ such that for each $v \in \tilde{U}_X$ at least half of its edges in M are incident to vertices in X . It must be that $|\tilde{U}_X| \leq \varepsilon|\tilde{U}|$ as otherwise we have

$$\begin{aligned} |X| &\geq |\tilde{U}_X| \cdot (\alpha/2)\tilde{n}p^2 \geq \varepsilon \min\{|U|, \varepsilon/p^2\} \cdot (\alpha/2)\tilde{n}p^2 \\ &\geq \min\{\varepsilon^{-20}|U| \log^2 n, \varepsilon^3 \tilde{n}\} \geq \min\{|X| \log n, \varepsilon^3 \tilde{n}\}, \end{aligned}$$

which contradicts our assumptions on $|X|$. Thus, by setting $U' := \tilde{U} \setminus \tilde{U}_X$, we have for each vertex from U' that half of its edges from the matching M avoid X and that suffices. $\square$

Proof of (2) in Lemma 2.9. Let $F'_{12} \subseteq F_{12}$ be a largest subset of the edges from F_{12} obtained by keeping at most ε/p edges incident to every vertex in U . Note that

$$|F'_{12}| \geq |F_{12}| \cdot \frac{\varepsilon/p}{\varepsilon^{-4} \log n/p} \geq \frac{\varepsilon^5 |F_{12}|}{\log n} \geq \varepsilon^{-12} \tilde{n} \log n. \quad (8)$$

We define a sequence $J_1, \dots, J_t$ of disjoint subsets of F'_{12} as follows. Let J_1 be a largest subset of F'_{12} such that no vertex of W_1 is incident to more than ε/p edges from J_1 . Assume we have defined $J_1, \dots, J_i$ for some $i \geq 1$. We then define J_{i+1} as a largest subset of $F'_{12} \setminus (J_1 \cup \dots \cup J_i)$ such that no vertex from W_1 is incident to more than ε/p edges from J_{i+1} . We set t to be the smallest integer such that $|J_1| + \dots + |J_t| \geq |F'_{12}|/2$. Note that from (G2) it follows that for every $i \in [t]$

$$|J_i| \geq \frac{|F'_{12}|}{2} \cdot \frac{\varepsilon/p}{(1+\varepsilon)\tilde{n}p} \geq \frac{\varepsilon|F'_{12}|}{4\tilde{n}p^2} \stackrel{(8)}{\geq} \frac{\varepsilon^{-10} \log n}{p^2}.$$

By using property (G6) for J_i (as $\mathcal{P}$), X (as W'), and the previous inequality, we get that for every $i \in [t]$ it holds that $e_{F_{23}(J_i)}(U, X) \leq 2\varepsilon^4 |J_i| \tilde{n}p^2$, where $F_{23}(J_i)$ denotes the edges in F_{23} obtained by extending *only* the edges belonging to J_i . This further shows that

$$e_{F_{23}(J^*)}(U, X) \leq 2\varepsilon^4 |J^*| \tilde{n}p^2, \quad (9)$$

where $J^* := J_1 \cup \dots \cup J_t$.

Next, consider an arbitrary $v \in U$. As every such v is incident to at most ε/p edges in J^* by definition, we have by (G7) applied for edges of J^* incident to v (as $\mathcal{P}$) that $\deg_{F_{23}}(v, W_3) \geq \alpha \deg_{J^*}(v, W_1) \tilde{n}p^2$. Thus, it follows that

$$e_{F_{23}(J^*)}(U, W_3) \geq \alpha |J^*| \tilde{n}p^2. \quad (10)$$

From the previous inequality and the fact that $|J^*| \geq |F'_{12}|/2$, we obtain

$$\begin{aligned} e_{F_{23}}(U, W_3 \setminus X) &\geq e_{F_{23}(J^*)}(U, W_3) - e_{F_{23}(J^*)}(U, X) \stackrel{(9),(10)}{\geq} \alpha |J^*| \tilde{n}p^2 - 2\varepsilon^4 |J^*| \tilde{n}p^2 \\ &\geq (\alpha/2) |J^*| \tilde{n}p^2 \geq (\alpha/4) |F'_{12}| \tilde{n}p^2 \stackrel{(8)}{\geq} \frac{(\varepsilon^5 \alpha/4) \tilde{n}p^2}{\log n} |F_{12}| \geq \varepsilon^{-4} |F_{12}|, \end{aligned}$$

where the last inequality follows from the assumption on $\tilde{n}$. $\square$

Proof of (3) in Lemma 2.9. From (G2) we have that $|U| \geq |F_{12}|/((1+\varepsilon)\tilde{n}p) \geq \varepsilon^{-4} \log n/p$. Thus, by (G1) and the fact that $|X| \leq \varepsilon^4 \tilde{n}$, we get

$$e_G(U, X) \leq 2\varepsilon^4 |U| \tilde{n}p. \quad (11)$$

Consider an arbitrary $v \in U$ and let $N_1 := N_{F_{12}}(v, W_1)$ and $N_3 := N_{F_{23}}(v, W_3)$. From the fact that $|N_1| \geq \varepsilon^{-3} \log n/p$, (G5), and (G1) applied for N_1 (as X) and N_3 (as Y), we know

$$|N_1| \cdot \alpha \tilde{n}p^2 \leq e_G(N_1, N_3) \leq (1+\varepsilon)|N_1| \max\{\varepsilon^{-3} \log n/p, |N_3|\}p.$$

By the assumption on $\tilde{n}$ we cannot have $|N_1| \alpha \tilde{n}p^2 \leq (1+\varepsilon)|N_1| \varepsilon^{-3} \log n$, hence it must be $|N_3| \geq (\alpha/2)\tilde{n}p$. Together with (11) we get

$$e_{F_{23}}(U, W_3 \setminus X) = e_{F_{23}}(U, W_3) - e_{F_{23}}(U, X) \geq (\alpha/2)|U| \tilde{n}p - 2\varepsilon^4 |U| \tilde{n}p \geq (\alpha/4)|U| \tilde{n}p,$$

completing the proof. $\square$

Proof of (4) in Lemma 2.9. Let $U' \subseteq U$ be a subset of vertices in U defined as

$$U' := \{v \in U : \deg_{F_{23}}(v, W_3) < (1/3 + \alpha/2)\tilde{n}p\}.$$

Assume for the moment that we can show $e_{F_{12}}(U', W_1) \leq \varepsilon |F_{12}|$. Then we would have

$$(1-\varepsilon)|F_{12}| \leq e_{F_{12}}(U \setminus U', W_1) \leq |U \setminus U'| \cdot \frac{\tilde{n}p}{3}, \quad (12)$$

where the upper bound follows from the assumption on $\deg_{F_{12}}(v, W_1)$. From this we get that $|U \setminus U'| \geq \varepsilon^{-4} \log n/p$ (with room to spare). Therefore, by the definition of U' and (G1) applied for X and $U \setminus U'$ (as Y) we obtain

$$\begin{aligned} e_{F_{23}}(U, W_3 \setminus X) &\geq e_{F_{23}}(U \setminus U', W_3) - e_{F_{23}}(U \setminus U', X) \\ &\geq |U \setminus U'|((1/3 + \alpha/2)\tilde{n}p - (1+\varepsilon)|X|p) \geq |U \setminus U'|(1/3 + \alpha/2 - 2\varepsilon^4)\tilde{n}p, \end{aligned}$$

where the last step follows from $|X| \leq \varepsilon^4 \tilde{n}$. This, together with (12) shows

$$e_{F_{23}}(U, W_3 \setminus X) \geq |U \setminus U'| \cdot (1/3 + \alpha/4)\tilde{n}p \geq 3(1-\varepsilon)|F_{12}|(1/3 + \alpha/4) \geq (1 + \alpha/4)|F_{12}|,$$

as desired. It remains to prove $e_{F_{12}}(U', W_1) \leq \varepsilon |F_{12}|$. Towards a contradiction assume this is not the case.

For every $v \in U'$ let $L_v := N_{F_{12}}(v, W_1)$ and $R_v := N_{G-F_{23}}(v, W_3)$. Observe that by (G4) and the definition of U' we have

$$|R_v| \geq (1/3 + \alpha/2)\tilde{n}p. \quad (13)$$

Moreover, the definition of R_v gives $e_G(L_v, R_v) = 0$. On the other hand, property (G1), as $|L_v|, |R_v| \geq \varepsilon^{-4} \log n/p$, states $e_\Gamma(L_v, R_v) = (1 \pm \varepsilon)|L_v||R_v|p$. Let

$$L'_v := \{w \in L_v : \deg_\Gamma(w, R_v) \geq (1-\varepsilon)|R_v|p\}. \quad (14)$$

It immediately follows from (G1) that $|L'_v| \geq (1-\varepsilon)|L_v|$. Hence, the assumption $e_{F_{12}}(U', W_1) \geq \varepsilon |F_{12}|$ implies

$$\sum_{v \in U'} |L'_v| \geq \sum_{v \in U'} (1-\varepsilon)|L_v| = (1-\varepsilon)e_{F_{12}}(U', W_1) \geq (1-\varepsilon) \cdot \varepsilon |F_{12}| \geq \frac{\varepsilon^{-3} \log n}{p} \cdot \tilde{n}.$$

Consequently, by averaging over vertices in W_1 , there is a vertex $w \in W_1$ and a set $U_w \subseteq U'$ such that $|U_w| \geq \varepsilon^{-3} \log n/p$ and $w \in L'_v$ for all $v \in U_w$.

Set $T := \bigcup_{v \in U_w} N_\Gamma(w, R_v)$. Note that all vertices in T are connected to w in Γ , however none is connected to w in G due to the way R_v 's are defined. For all $v \in U_w$ we have

$$\deg_G(v, T) \geq |R_v \cap T| \geq |N_\Gamma(w, R_v)| \stackrel{(14)}{\geq} (1 - \varepsilon)|R_v|p \stackrel{(13)}{\geq} (1/3 + \alpha/4)\tilde{n}p^2.$$

Therefore,

$$e_\Gamma(U_w, T) \geq e_G(U_w, T) \geq |U_w| \cdot (1/3 + \alpha/4)\tilde{n}p^2.$$

Since $|U_w| \geq \varepsilon^{-3} \log n/p$, we get from (G1) applied for U_w (as X) and T (as Y) that

$$|T| \geq \frac{(1/3 + \alpha/4)}{1 + \varepsilon} \tilde{n}p \geq (1/3 + \alpha/8)\tilde{n}p.$$

Lastly, as $T \subseteq N_{\Gamma-G}(w, W_3)$, (G2) and (G4) provide the desired contradiction. $\square$

Proof of (5) in Lemma 2.9. From (G2) we get that $|U| \geq |F_{12}|/((1 + \varepsilon)\tilde{n}p) \geq \varepsilon^{-9} \log n/p^2$. Let $U' \subseteq U$ be defined as

$$U' := \{v \in U : \deg_{F_{23}}(v, W_3) < (1 - \varepsilon^3) \deg_G(v, W_3)\}. \quad (15)$$

Assume for the moment that we can show $e_G(U', W_3) \leq \varepsilon^3 \cdot e_G(U, W_3)$. Then

$$|U'| \cdot (2/3 + \alpha/2)\tilde{n}p \stackrel{(G4)}{\leq} e_G(U', W_3) \leq \varepsilon^3 \cdot e_G(U, W_3) \stackrel{(G1)}{\leq} \varepsilon^3 \cdot (1 + \varepsilon)|U|\tilde{n}p.$$

This implies $|U'| \leq 2\varepsilon^3|U|$. Since $|X| \leq \varepsilon^4\tilde{n}$, we deduce from (G1) that there can be at most $\varepsilon^{-3} \log n/p \leq \varepsilon^3|U|$ vertices in U with degree into X larger than $2\varepsilon^4\tilde{n}p$. Let L be the vertices in $U \setminus U'$ that do not have this property. Therefore, $|L| \geq |U| - |U'| - \varepsilon^3|U| \geq (1 - 3\varepsilon^3)|U|$ and for all $v \in L$

$$\deg_{F_{23}}(v, W_3 \setminus X) \stackrel{(15)}{\geq} (1 - \varepsilon^3) \deg_G(v, W_3) - 2\varepsilon^4\tilde{n}p \geq (1 - 2\varepsilon^3) \deg_G(v, W_3).$$

Moreover, by definition of U' and (G1) applied for X and $U \setminus U'$ (as Y), since $|X| \leq \varepsilon^4\tilde{n}$, we obtain

$$\begin{aligned} e_{F_{23}}(U, W_3 \setminus X) &\geq e_{F_{23}}(U \setminus U', W_3) - e_{F_{23}}(U \setminus U', X) \\ &\geq \left(\sum_{v \in U \setminus U'} (1 - \varepsilon^3) \deg_G(v, W_3) \right) - 2\varepsilon^4|U \setminus U'|\tilde{n}p \stackrel{(G4)}{\geq} \sum_{v \in U \setminus U'} (1 - 2\varepsilon^3) \deg_G(v, W_3). \end{aligned}$$

This together with the assumption $e_G(U', W_3) \leq \varepsilon^3 \cdot e_G(U, W_3)$ further shows

$$\begin{aligned} e_{F_{23}}(U, W_3 \setminus X) &\geq \sum_{v \in U \setminus U'} (1 - 2\varepsilon^3) \deg_G(v, W_3) \geq (1 - 2\varepsilon^3)e_G(U \setminus U', W_3) \\ &\geq (1 - 2\varepsilon^3)(1 - \varepsilon^3)e_G(U, W_3) \geq (1 - \varepsilon^2)e_G(U, W_3), \end{aligned}$$

as desired. It remains to prove $e_G(U', W_3) \leq \varepsilon^3 \cdot e_G(U, W_3)$. Towards a contradiction assume this is not the case.

For every $v \in U'$ let $L_v := N_{F_{12}}(v, W_1)$ and $R_v := N_{G-F_{23}}(v, W_3)$. Observe that by (G4) and the definition of U' we have $|R_v| \geq \varepsilon^3 \cdot \deg_G(v, W_3) \geq (2\varepsilon^3/3)\tilde{n}p$. Moreover, the definition of R_v gives $e_G(L_v, R_v) = 0$. On the other hand, property (G1), as $|L_v|, |R_v| \geq \varepsilon^{-4} \log n/p$, states $e_\Gamma(L_v, R_v) = (1 \pm \varepsilon)|L_v||R_v|p$. Let

$$R'_v := \{w \in R_v : \deg_\Gamma(w, L_v) \geq (1 - \varepsilon)|L_v|p\}. \quad (16)$$

It immediately follows from (G1) that $|R'_v| \geq (1 - \varepsilon)|R_v|$. Hence, the assumption $e_G(U', W_3) > \varepsilon^3 \cdot e_G(U, W_3)$ implies

$$\sum_{v \in U'} |R'_v| \geq \sum_{v \in U'} (1 - \varepsilon)|R_v| \geq \sum_{v \in U'} (1 - \varepsilon) \cdot \varepsilon^3 \cdot \deg_G(v, W_3) \geq \frac{\varepsilon^3}{2} e_G(U', W_3) > \frac{\varepsilon^6}{2} e_G(U, W_3).$$

Next, using (G3), we know that $(1 + \varepsilon)e_G(U, W_3) \geq (1 - \varepsilon)e_G(U, W_1)$ and thus

$$\sum_{v \in U'} |R'_v| \geq \frac{\varepsilon^6}{4} e_G(U, W_1) \geq \frac{\varepsilon^6}{4} |F_{12}| \geq \frac{\varepsilon^{-3} \log n}{p} \cdot \tilde{n}.$$

Consequently, by averaging over vertices in W_3 , there is a vertex $w \in W_3$ and a set $U_w \subseteq U'$ such that $|U_w| \geq \varepsilon^{-3} \log n/p$ and $w \in R'_v$ for all $v \in U_w$.

Set $T := \bigcup_{v \in U_w} N_\Gamma(w, L_v)$. Note that all vertices in T are connected to w in Γ , however none is connected to w in G . For all $v \in U'_w$ we now have

$$\deg_G(v, T) \geq |L_v \cap T| \geq |N_\Gamma(w, L_v)| \stackrel{(16)}{\geq} (1 - \varepsilon)|L_v|p \geq (1/3 - \varepsilon/3)\tilde{n}p^2,$$

where the last inequality follows from the assumption $\deg_{F_{12}}(v, W_1) \geq \tilde{n}p/3$. Therefore,

$$e_\Gamma(U_w, T) \geq e_G(U_w, T) \geq |U_w| \cdot (1/3 - \varepsilon/3)\tilde{n}p^2.$$

Since $|U_w| \geq \varepsilon^{-3} \log n/p$ we get from (G1) applied for U_w (as X) and T (as Y) that

$$|T| \geq \frac{1/3 - \varepsilon/3}{1 + \varepsilon} \tilde{n}p \geq (1/3 - 2\varepsilon)\tilde{n}p.$$

Lastly, as $T \subseteq N_{\Gamma-G}(w, W_1)$, (G2) and (G4) provide the desired contradiction. $\square$

Proof of (6) in Lemma 2.9. Note that $|F_{12}| \geq |U| \cdot \tilde{n}p/3 \geq \varepsilon^{-10}\tilde{n} \log n/p$. By part (5) of this lemma, we get

$$e_{F_{23}}(U, W_3 \setminus X) \geq (1 - \varepsilon^2)e_G(U, W_3). \quad (17)$$

Let

$$L := \{u \in W_3 \setminus X : \deg_{F_{23}}(u, W_2) \geq (1/3 + \alpha/2)\tilde{n}p\}.$$

We aim to show that $|L| \geq (1 - \varepsilon)\tilde{n}$. Assume towards a contradiction that this is not the case. Together with (G4) and the assumption $|X| \leq \varepsilon^4 \tilde{n}$ this implies that there exists a set $Q \subseteq W_3 \setminus (X \cup L)$ of size $|Q| \geq (\varepsilon/2)\tilde{n}$ such that for each $v \in Q$ we have

$$\deg_{G-F_{23}}(v, W_2) \geq (1/3 + \alpha/2)\tilde{n}p. \quad (18)$$

Let $Q' := \{v \in Q : \deg_{G-F_{23}}(v, U) \geq (\alpha/4)\tilde{n}p\}$. If $|Q'| \geq (\varepsilon/4)\tilde{n}$ then $e_{G-F_{23}}(U, Q') \geq (\varepsilon\alpha/16)\tilde{n}^2p$, which together with (G1) applied for U (as X) and Q' (as Y) contradicts (17). Therefore, $|Q'| \leq (\varepsilon/2)\tilde{n}$. From (G1) we then get

$$e_{G-F_{23}}(W_2 \setminus U, Q \setminus Q') \leq (1 + \varepsilon) \max\{\varepsilon^{-3} \log n/p, |W_2 \setminus U|\} \cdot |Q \setminus Q'|p. \quad (19)$$

On the other hand, from (18) and the definition of Q and Q' we have

$$e_{G-F_{23}}(W_2 \setminus U, Q \setminus Q') \geq (1/3 + \alpha/4)|Q \setminus Q'|\tilde{n}p.$$

Together with (19) we deduce

$$\max\{\varepsilon^{-3} \log n/p, |W_2 \setminus U|\} \geq \frac{1/3 + \alpha/4}{1 + \varepsilon} \tilde{n} \geq (1/3 + \alpha/8)\tilde{n},$$

which contradicts the assumption $|U| \geq (2/3)\tilde{n}$. $\square$

Proof of (7) in Lemma 2.9. Let $S := \{v \in U : \deg_{F_{12}}(v, W_1) < \varepsilon^{-4} \log n/p\}$ and let $F_S \subseteq F_{12}$ be the subset of edges that are incident to S . Observe that if $|F_S| < \varepsilon|F_{12}|$ we are done. So assume otherwise. By using part (2) of this lemma with F_S (as F_{12}) we get

$$e_{F_{23}}(W_2, W_3 \setminus X) \geq e_{F_{23}}(S, W_3 \setminus X) \geq \varepsilon^{-4}|F_S| \geq \varepsilon^{-3}|F_{12}|,$$

contradicting the starting assumption $e_{F_{23}}(W_2, W_3 \setminus X) < \varepsilon^{-3}|F_{12}|$. $\square$

Proof of (8) in Lemma 2.9. Define sets $S, M, L \subseteq W_2$ (which may be thought of as ‘small’, ‘medium’, and ‘large’) as

$$\begin{aligned} S &= \{v \in W_2 : 1 \leq \deg_{F_{12}}(v, W_1) \leq \varepsilon^{-4} \log n/p\}, \\ M &= \{v \in W_2 : \varepsilon^{-4} \log n/p \leq \deg_{F_{12}}(v, W_1) \leq \tilde{n}p/3\}, \\ L &= \{v \in W_2 : \tilde{n}p/3 < \deg_{F_{12}}(v, W_1)\}, \end{aligned}$$

and denote by F_S, F_M , and F_L the subsets of edges in F_{12} incident to S, M , and L , respectively. Note that F_S, F_M , and F_L partition the set F_{12} . We claim that

- (i) if $|F_S| \geq \varepsilon^4|F_{12}|$, then $e_{F_{23}}(S, W_3 \setminus X) \geq (1 + \alpha/4)|F_S|$,
- (ii) if $|F_M| \geq \varepsilon^4|F_{12}|$, then $e_{F_{23}}(M, W_3 \setminus X) \geq (1 + \alpha/4)|F_M|$, and
- (iii) if $|F_L| \geq \varepsilon^4|F_{12}|$, then $e_{F_{23}}(L, W_3 \setminus X) \geq (1 - 3\varepsilon)|F_L|$.

To see this assume first that $|F_S| \geq \varepsilon^4|F_{12}|$. Then $|F_S| \geq \varepsilon^5 \tilde{n}^2 p \geq \varepsilon^{-17} \tilde{n} \log^2 n$ and we can thus apply part (2) of this lemma to S (as U) and F_S (as F_{12}) to get

$$e_{F_{23}}(S, W_3 \setminus X) \geq \varepsilon^{-4}|F_S| \geq (1 + \alpha/4)|F_S|.$$

Next, assume $|F_M| \geq \varepsilon^4|F_{12}|$. Then again $|F_M| \geq \varepsilon^{-5} \tilde{n} \log n/p$ and we can apply part (4) of this lemma to M (as U) and F_M (as F_{12}) to get

$$e_{F_{23}}(M, W_3 \setminus X) \geq (1 + \alpha/4)|F_M|.$$

Lastly, assume $|F_L| \geq \varepsilon^4|F_{12}|$. Then $|F_L| \geq \varepsilon^{-10} \tilde{n} \log n/p$ and we can apply part (5) of this lemma to L (as U) and F_L (as F_{12}) to get

$$e_{F_{23}}(L, W_3 \setminus X) \geq (1 - \varepsilon^2)e_G(L, W_3). \quad (20)$$

Moreover, from (G3) applied for all $v \in L$ and W_1, W_3 (as W), we have $(1 + \varepsilon)e_G(L, W_3) \geq (1 - \varepsilon)e_G(L, W_1)$. Therefore, together with (20):

$$e_{F_{23}}(L, W_3 \setminus X) \geq (1 - \varepsilon^2)e_G(L, W_3) \geq (1 - \varepsilon^2) \frac{1 - \varepsilon}{1 + \varepsilon} e_G(L, W_1),$$

from which the third property follows as, trivially, $e_G(L, W_1) \geq |F_L|$.

Having these three properties at hand, we are ready to prove the lemma. If $|F_L| \geq (1 - \mu)|F_{12}|$ we are done, so assume the contrary. Observe that this implies that at least one of $|F_S|, |F_M|$ has size at least $\varepsilon^4|F_{12}|$. If $|F_L|$ is strictly smaller than $\varepsilon^4|F_{12}|$, then either at least one of $|F_S|$ and $|F_M|$ has size at least $(1 - 2\varepsilon^4)|F_{12}|$ or both have size at least $\varepsilon^4|F_{12}|$. Thus by (i) and (ii) we get

$$e_{F_{23}}(W_2, W_3 \setminus X) \geq \max\{(1 + \alpha/4)(1 - 2\varepsilon^4)|F_{12}|, (1 + \alpha/4)(1 - \varepsilon^4)|F_{12}|\} > (1 + \alpha/8)|F_{12}|,$$

which is a contradiction to the assumption of the lemma. Therefore, $|F_L| \geq \varepsilon^4|F_{12}|$ and at least one of $|F_S|$ and $|F_M|$ has size at least $|F_{12}| - |F_L| - \varepsilon^4|F_{12}|$ or both have size at least $\varepsilon^4|F_{12}|$. Thus, again by (i) and (ii),

$$\begin{aligned} e_{F_{23}}(W_2, W_3 \setminus X) &\geq \max\{(1 + \alpha/4)(|F_{12}| - \varepsilon^4|F_{12}| - |F_L|), (1 + \alpha/4)(|F_{12}| - |F_L|)\} + (1 - 3\varepsilon)|F_L| \\ &\geq (1 + \alpha/4)(1 - \varepsilon^4)|F_{12}| - (3\varepsilon + \alpha/4)|F_L|. \end{aligned}$$

Using our assumption $|F_L| < (1 - \mu)|F_{12}|$ and $\mu \geq 32\varepsilon/\alpha$ this implies $e_{F_{23}}(W_2, W_3 \setminus X) \geq (1 + \mu\alpha/8)|F_{12}|$, again contradicting the assumption of the lemma. $\square$

Proof of (9) in Lemma 2.9. Clearly, if $e_{F_{23}}(W_2, W_3 \setminus X) \geq (1 + 4\varepsilon)|F_{12}|$ we are done. Let us assume the opposite. By applying part (8) of this lemma with $32\varepsilon/\alpha$ (as μ) we obtain a set $L \subseteq W_2$ such that for each $v \in L$ it holds that $\deg_{F_{12}}(v, W_1) > \tilde{n}p/3$ and

$$e_{F_{12}}(W_1, L) \geq (1 - 32\varepsilon/\alpha)|F_{12}| \geq (1 - \varepsilon^{2/3})|F_{12}|. \quad (21)$$

Next, we use part (5) of this lemma with L (as U) and $E_{F_{12}}(W_1, L)$ (as F_{12}) to conclude

$$e_{F_{23}}(L, W_3 \setminus X) \geq (1 - \varepsilon^2)e_G(L, W_3). \quad (22)$$

By taking together (21) and (22), we finally get

$$\begin{aligned} e_{F_{23}}(L, W_3 \setminus X) &\geq (1 - \varepsilon^2)e_G(L, W_3) \stackrel{(G3)}{\geq} (1 - \varepsilon^2)(1 - 2\varepsilon)e_G(W_1, L) \\ &\geq (1 - 3\varepsilon)e_{F_{12}}(W_1, L) \stackrel{(21)}{\geq} (1 - 3\varepsilon)(1 - \varepsilon^{2/3})|F_{12}| \geq (1 - \sqrt{\varepsilon})|F_{12}|, \end{aligned}$$

and the assertion follows. $\square$

3 Proof of Theorem 1.4

The proof of Theorem 1.4 is split into two natural parts. In Theorem 3.1 we show that the K_3 -resilience of $G_{n,p}$ w.r.t. the containment of C_n^2 is w.h.p. at most $5/9 + \alpha$, for any $\alpha > 0$. Next, in Theorem 3.2 we show that the K_3 -resilience is w.h.p. at least $5/9 - \alpha$, for any $\alpha > 0$. Both of the theorems rely on the following fact: for every $\varepsilon > 0$ a random graph $\Gamma \sim G_{n,p}$ w.h.p. has the property that each vertex is contained in $(1 \pm \varepsilon)\binom{n}{2}p^3$ triangles, provided that $p \gg n^{-2/3} \log^{1/3} n$ (see, e.g. [2, Theorem 8.5.4]).

The proof of the upper bound of K_3 -resilience stems from a simple construction and an application of Janson's inequality. We actually show that w.h.p. there exists a subgraph G of $G_{n,p}$ such that each vertex is contained in $(4/9 - \gamma)\binom{n}{2}p^3$ triangles and G does not contain a family of more than $(1 - \gamma)n/3$ vertex-disjoint triangles. This is sufficient since C_n^2 contains $\lfloor n/3 \rfloor$ vertex-disjoint triangles.

Theorem 3.1. *For every $\gamma > 0$, there exists a positive constant $C := C(\gamma)$ such that for all $p \geq Cn^{-2/3} \log^{1/3} n$ a random graph $\Gamma \sim G_{n,p}$ w.h.p. contains a spanning subgraph $G \subseteq \Gamma$ in which each vertex is contained in at least $(4/9 - \gamma)\binom{n}{2}p^3$ triangles and such that G does not contain a family of more than $(1 - \gamma)n/3$ vertex-disjoint triangles.*

Proof. Let $V(G) = V_1 \cup V_2$ be a partition of the vertex set of G such that

$$|V_1| = \left\lceil \left(\frac{1}{3} + \frac{2\gamma}{3} \right) n \right\rceil \quad \text{and} \quad |V_2| = n - |V_1|.$$

Observe that $|V_2| = (2/3 - 2\gamma/3 - o(1))n$. Furthermore, let G be the graph obtained from Γ by removing all edges with both endpoints in V_1 . For a vertex $v \in V(G)$ let $\mathcal{T}_v$ denote the family of all triangles in K_n which contain v and do not have more than one vertex in V_1 . Set

$$X = \sum_{T \in \mathcal{T}_v} X_T, \quad \mu = \mathbb{E}[X], \quad \text{and} \quad \Delta = \sum_{\substack{T_1, T_2 \in \mathcal{T}_v \\ T_1 \sim T_2}} \mathbb{E}[X_{T_1} X_{T_2}],$$

where X_T is an indicator random variable for the event $\{T \subseteq \Gamma\}$. Note that X is a random variable counting the number of triangles in G that contain v . We aim to show that

$$\Pr \left[X < (4/9 - \gamma) \binom{n}{2} p^3 \right] \leq e^{-2 \log n}, \quad (23)$$

for every $v \in V(G)$.

Let us estimate μ and Δ for an arbitrary vertex $v \in V_1$. We have

$$\mu = \sum_{T \in \mathcal{T}_v} \mathbb{E}[X_T] = \binom{|V_2|}{2} p^3 \geq (2/3 - 2\gamma/3 - o(1)) \frac{2n^2 p^3}{2} \geq (4/9 - 8\gamma/9) \binom{n}{2} p^3. \quad (24)$$

Note that if two triangles $T_1, T_2 \in \mathcal{T}_v$ do not share an edge, they are independent and thus $T_1 \not\sim T_2$. Therefore, we can bound Δ as follows:

$$\Delta = \sum_{\substack{T_1, T_2 \in \mathcal{T}_v \\ T_1 \sim T_2}} \mathbb{E}[X_{T_1} X_{T_2}] \leq |V_2|^3 p^5 \leq \frac{5\mu^2}{|V_2|p}. \quad (25)$$

Let us choose ε such that $(1-\varepsilon)(4/9 - 8\gamma/9) \geq 4/9 - \gamma$ and apply Janson's inequality (Theorem 2.4) with ε (as γ) to obtain

$$\begin{aligned} \Pr \left[X < (4/9 - \gamma) \binom{n}{2} p^3 \right] &\leq \Pr[X < (1 - \varepsilon)\mu] \leq e^{-\frac{\varepsilon^2 \mu^2}{2(\mu + \Delta)}} \leq e^{-\frac{\varepsilon^2 \mu^2}{4 \max\{\mu, \Delta\}}} \\ &\leq \max\{e^{-\varepsilon^2 \mu/4}, e^{-\varepsilon^2 |V_2|p/20}\}, \end{aligned}$$

where the last step follows from (24) and (25). Since $|V_2|p \gg \log n$ and $\mu \geq (4/9 - 8\gamma/9 - o(1))C^3 \log n$, by choosing C large enough with respect to ε and γ we obtain (23) for every vertex in V_1 . The proof of (23) for the case when $v \in V_2$ follows analogously and is omitted. By (23) and the union bound over all vertices we get that w.h.p. each vertex $v \in V(G)$ is contained in at least $(4/9 - \gamma) \binom{n}{2} p^3$ triangles in G .

Let $\mathcal{F}$ be the largest family of vertex-disjoint triangles in G . Since G does not contain an edge with both endpoints in V_1 , there is no triangle in $\mathcal{F}$ with more than one vertex in V_1 . This implies that $|V(\mathcal{F}) \cap V_2| \geq 2|V(\mathcal{F}) \cap V_1|$, which further shows

$$|V(\mathcal{F})| \leq 3/2 \cdot |V_2| \leq 3/2 \cdot (2/3 - 2\gamma/3)n \leq (1 - \gamma)n,$$

completing the proof. $\square$

The following lower bound on the K_3 -resilience is proven by a reduction to Theorem 1.6.

Theorem 3.2. *For every $\gamma > 0$, there exists a positive constant $C := C(\gamma)$ such that a random graph $\Gamma \sim G_{n,p}$ w.h.p. satisfies the following, provided that $p \geq Cn^{-1/2} \log^3 n$. Every spanning subgraph $G \subseteq \Gamma$ in which each vertex is contained in at least $(4/9 + \gamma) \binom{n}{2} p^3$ triangles contains the square of a Hamilton cycle.*

Proof. Choose $\varepsilon > 0$ such that $(1 + \varepsilon)(4/9 + 5\gamma/12) < 4/9 + \gamma/2$ and $\varepsilon \leq \gamma/4$, and set $C = 10/\varepsilon$. Let $G \subseteq \Gamma$ be an arbitrary spanning subgraph of Γ such that each vertex of G is contained in at least $(4/9 + \gamma) \binom{n}{2} p^3$ triangles. Let $G' \subseteq G$ be a subgraph obtained by removing each edge of G which is contained in fewer than εnp^2 triangles. We aim to show that G' has minimum degree at least $(2/3 + \gamma/4)np$. If this is the case then by Theorem 1.6 we are done.

First, we show that by removing the edges which are contained in only a few triangles, we did not significantly change the overall number of triangles each vertex is in. As $np = \omega(\log n)$,

w.h.p. we have that every $v \in V(\Gamma)$ satisfies $\deg_\Gamma(v) \leq (1 + \varepsilon)np$. Moreover, each edge in $E(G) \setminus E(G')$ is contained in at most εnp^2 triangles, which implies that we did not remove more than $\varepsilon np^2 \cdot (1 + \varepsilon)np \leq 2\varepsilon n^2 p^3$ triangles from G touching a single vertex. Since G has the property that each vertex is contained in at least $(4/9 + \gamma)\binom{n}{2}p^3$ triangles, the previous observation and the choice of ε show that in G' each vertex is in at least $(4/9 + \gamma/2)\binom{n}{2}p^3$ triangles.

In order to finish the argument we use the following claim, whose proof follows below.

Claim 3.3. *The following holds w.h.p. For every vertex $v \in V(\Gamma)$ and every subset $S \subseteq N_\Gamma(v)$ of size $|S| \geq (2/3)np$, we have $e(S) \leq (1 + \varepsilon)\binom{|S|}{2}p$.*

With this claim we can easily complete the proof of the theorem. Suppose for contradiction that $v \in V(G')$ is a vertex with degree smaller than $(2/3 + \gamma/4)np$. By the claim above and the choice of ε we have that v is contained in at most

$$(1 + \varepsilon)(2/3 + \gamma/4)^2 \frac{n^2 p^3}{2} \leq (1 + \varepsilon)(4/9 + 5\gamma/12) \frac{n^2 p^3}{2} < (4/9 + \gamma/2) \binom{n}{2} p^3,$$

triangles—a contradiction. $\square$

Proof of Claim 3.3. It suffices to show that the claim holds for a fixed vertex v with probability at least $1 - e^{-\omega(\log n)}$, as the claim then follows by the union bound over all vertices. Let v be a vertex from Γ and let $S \subseteq N_\Gamma(v)$. Recall, we have that $\deg_\Gamma(v) \leq (1 + \varepsilon)np$ with probability at least $1 - e^{-\omega(\log n)}$. Similarly we have

$$\Pr \left[e(\Gamma[S]) > (1 + \varepsilon) \binom{|S|}{2} p \right] \leq e^{-\varepsilon^2 \binom{|S|}{2} p/3}. \quad (26)$$

By using (26) and the union bound, the probability that the assertion of the claim fails is at most

$$\begin{aligned} \sum_{s=(2/3)np}^{(1+\varepsilon)np} \binom{(1+\varepsilon)np}{s} e^{-\varepsilon^2 s(s-1)p/6} &\leq \sum_{s=(2/3)np}^{(1+\varepsilon)np} \left(\frac{2enp}{s} \right)^s e^{-\varepsilon^2 s(s-1)p/6} \\ &\leq \sum_{s=(2/3)np}^{(1+\varepsilon)np} (3e)^{np} e^{-\varepsilon^2 s^2 p/12} \leq \sum_{s=(2/3)np}^{(1+\varepsilon)np} (e^{3 \log 3 - \varepsilon^2 sp/12})^s \\ &\leq \sum_{s=(2/3)np}^{(1+\varepsilon)np} e^{-2s} \leq n \cdot e^{-\Omega(np)}, \end{aligned}$$

where in the second to last inequality we used the fact that $\varepsilon^2 sp/12 \geq \varepsilon^2 np^2/18 \geq 3 \log 3$. Finally, since $np = \omega(\log n)$, the claim follows. $\square$

4 Definitions of some graphs

The following graphs are used often throughout the paper and we thus give their definitions here, for easier reference later on. We note that most of these come from or were inspired by similar definitions in [30].

An ℓ -**square-path**, denoted by P_ℓ^2 , is a graph defined on a vertex set $\{v_1, \dots, v_\ell\}$ such that v_i and v_j are connected by an edge if $1 \leq i < j \leq i + 2$ (see Figure 1).

Given a graph $G = (V, E)$ and $\mathbf{a}, \mathbf{b} \in V^2$, we say that G contains a square-path *connecting* $\mathbf{a}$ to $\mathbf{b}$, if there exists an $\ell \in \mathbb{N}$ and an embedding $g: V(P_\ell^2) \rightarrow V(G)$ such that $\mathbf{a} = (g(v_1), g(v_2))$ and $\mathbf{b} = (g(v_{\ell-1}), g(v_\ell))$. Note that due to the fact that $\mathbf{a}$ and $\mathbf{b}$ are (ordered) pairs of vertices, a

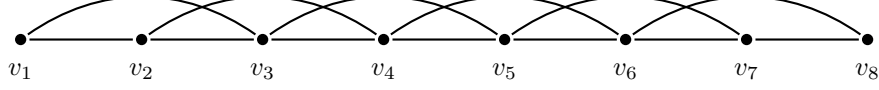


Figure 1: The 8-square-path P_8^2 .

path connecting $\mathbf{a}$ to $\mathbf{b}$ is *not* the same as a path connecting $\mathbf{b}$ to $\mathbf{a}$. However, a path connecting $\mathbf{a}$ to $\mathbf{b}$ is also a path connecting $\bar{\mathbf{b}}$ to $\bar{\mathbf{a}}$. It is easy to see that one can connect two square-paths in order to get a longer square-path.

Proposition 4.1. *Let G be a graph and $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V(G)^2$ disjoint pairs of vertices. Suppose that in G there exists a square-path connecting $\mathbf{a}$ to $\mathbf{b}$ and a square-path connecting $\mathbf{b}$ to $\mathbf{c}$ (and thus also $\bar{\mathbf{c}}$ to $\bar{\mathbf{b}}$) such that these paths are internally vertex-disjoint. Then the union of these two paths is a square-path that connects $\mathbf{a}$ to $\mathbf{c}$. $\square$*

A (b, ℓ) -**pseudo-path** S_ℓ^b , where $b \in \{1, 2\}$, is a graph defined on the vertex set $\{u_1, \dots, u_\ell\}$ with the edge set

$$E(S_\ell^b) := \bigcup_{i=2}^{\ell} \{u_{i-1}, u_i\} \cup \bigcup_{\substack{i \in \{3, \dots, \ell\} \\ i \text{ is odd}}} \{u_{i-2}, u_i\} \cup \bigcup_{\substack{i \in \{4, \dots, \ell\} \\ i \text{ is even}}} \{u_{i-1-b}, u_i\}.$$

Observe that a $(1, \ell)$ -pseudo-path is isomorphic to an ℓ -square-path; a $(2, \ell)$ -pseudo-path is depicted in Figure 2.

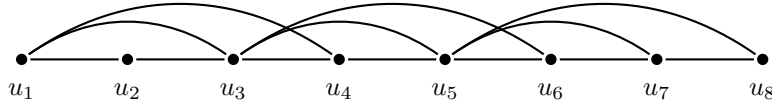


Figure 2: The $(2, 8)$ -pseudo-path S_8^2

The notion of a (b, ℓ) -pseudo-path connecting $\mathbf{a}$ to $\mathbf{b}$ is defined in a natural way, similarly as above.

An ℓ -**backbone-path** B_ℓ , is a graph defined on the vertex set

$$W_\ell = \bigcup_{i \in [\ell]} \{w_{i,1}, w_{i,2}, w_{i,3}, w_{i,4}\}.$$

We set $\mathbf{w}_i^a = (w_{i,1}, w_{i,2})$ and $\mathbf{w}_i^b = (w_{i,3}, w_{i,4})$, for every $i \in [\ell]$. The edge set of B_ℓ is given by the union of following graphs (see Figure 3):

- edges $\{w_{1,1}, w_{1,2}\}$ and $\{w_{1,3}, w_{1,4}\}$;
- the 4-square-path $(\mathbf{w}_i^a, \mathbf{w}_i^b)$ for every $2 \leq i \leq \ell$;
- the 4-square-path $(\mathbf{w}_1^a, \bar{\mathbf{w}}_2^a)$;
- the 4-square-path $(\bar{\mathbf{w}}_i^b, \bar{\mathbf{w}}_{i+2}^a)$ for every $1 \leq i \leq \ell - 2$;
- the 4-square-path $(\bar{\mathbf{w}}_{\ell-1}^b, \mathbf{w}_\ell^b)$.

Given a graph G and $\mathbf{a}, \mathbf{b} \in V^2$, we say that a backbone-path *connecting $\mathbf{a}$ to $\mathbf{b}$* is an embedding $g: V(B_\ell) \rightarrow V(G)$, for an appropriate $\ell \in \mathbb{N}$, such that $\mathbf{a} = (g(w_{1,2}), g(w_{1,1}))$ and $\mathbf{b} = (g(w_{1,4}), g(w_{1,3}))$.

The connection between backbone-paths and pseudo-paths is given by the following proposition.

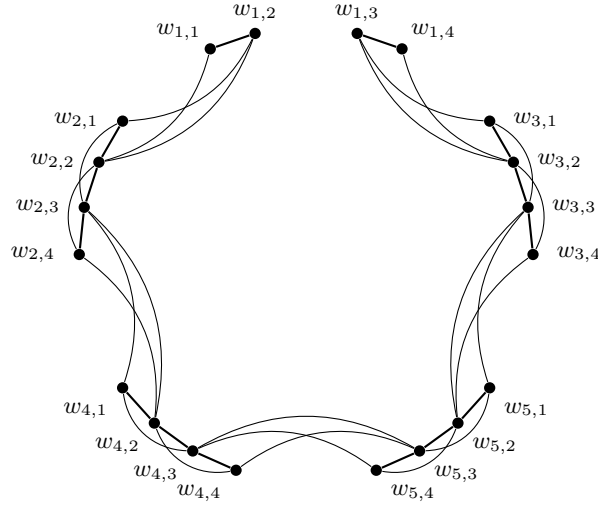
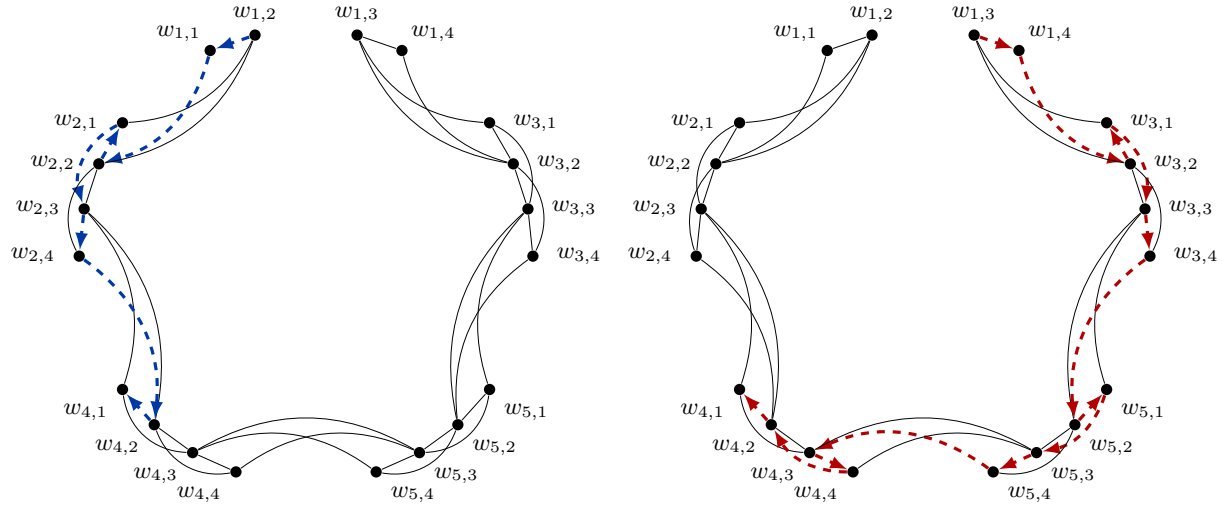


Figure 3: The graph B_5

Proposition 4.2. Let G be a graph, $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V(G)^2$ disjoint pairs of vertices, and $\ell_1, \ell_2 \in \mathbb{N}$ are such that both ℓ_1 and $\ell_1 + \ell_2 - 2$ are divisible by four. Suppose that in G there exists a $(2, \ell_1)$ -pseudo-path connecting $\mathbf{a}$ to $\mathbf{b}$ and a $(2, \ell_2)$ -pseudo-path connecting $\bar{\mathbf{c}}$ to $\mathbf{b}$ such that these paths are internally vertex-disjoint. Then the union of these two paths is a backbone-path that connects $\mathbf{a}$ to $\mathbf{c}$.

Proof. One easily verifies that Figure 4 describes an embedding of the two pseudo-paths whose union is a backbone-path. We omit the details. $\square$



(a) Blue arrows indicate the order of the vertices mapped by g_1 . (b) Red arrows indicate the order of the vertices mapped by g_2 .

Figure 4: An embedding of a graph B_5 by combining two 2-pseudo-paths.

The reason behind a rather complex looking definition of a backbone-path should become more apparent once we make use of it as a building block for absorbers later on (see Figure 5).

An (b, ℓ) -**connecting-path** C_ℓ^b , for $b \in \{1, 2\}$ and ℓ divisible by four, is a graph on ℓ vertices

defined as

$$C_\ell^b := \begin{cases} P_\ell^2, & \text{if } b = 1, \\ B_{\ell/4}, & \text{if } b = 2. \end{cases}$$

5 Proof of Theorem 1.6

Our proof strategy uses the *absorbing method*, in particular following a variant used by Nenadov and the second author in [30]. Let A be a graph and $\mathbf{a}, \mathbf{b} \in V(A)^2$ disjoint pairs of vertices of A . Given a subset $X \subseteq V(A)$, we say that A is an $(\mathbf{a}, \mathbf{b}, X)$ -*absorber* if for every subset $X' \subseteq X$ there exists a square-path $P \subseteq A$ connecting $\mathbf{a}$ to $\mathbf{b}$ such that $V(P) = V(A) \setminus X'$.

The following lemma shows that one can find an absorber in a member of $\mathcal{G}(\Gamma, n, \alpha, p)$ for a large subset X and $\Gamma \sim G_{n,p}$. The proof of the Absorbing Lemma is deferred to Section 6.

Lemma 5.1 (Absorbing Lemma). *For every $\alpha > 0$, there exists a positive constant $C := C(\alpha)$ such that w.h.p. for a random graph $\Gamma \sim G_{n,p}$ every $G \in \mathcal{G}(\Gamma, n, \alpha, p)$ has the following property.*

Let $s \geq C \log^4 n / p^2$ be an integer. Then there are at least $(1 - n^{-3}) \binom{n}{s}$ subsets $W \subseteq V(G)$ of size s satisfying the following. For every subset $X \subseteq V(G) \setminus W$ of size $|X| \leq |W| / (C \log^2 n)$ there exists an $(\mathbf{a}, \mathbf{b}, X)$ -absorber A in G such that $V(A) \setminus X \subseteq W$, where $\mathbf{a}, \mathbf{b} \in W^2$ are two disjoint pairs of vertices.

In order to construct absorbers one typically resorts to what is usually called a *Connecting Lemma*. Intuitively, it allows us to connect certain pairs of vertices by vertex-disjoint copies of a fixed graph F through a reservoir of vertices W . For ease of reference, we now define such a notion formally.

Definition 5.2. Let $t, \ell \in \mathbb{N}$, let G be a graph, and let $W \subseteq V(G)$ be a subset of vertices of G . Given a family $\mathcal{I} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \in [t]} \subseteq V(G)^4$ of pairwise disjoint 4-tuples and $b \in \{1, 2\}$, we say that a collection $\{F_i \subseteq G\}_{i \in [t]}$ of subgraphs of G forms an $(\mathcal{I}, b, \ell)$ -*matching* in W if the following holds:

- F_i is a copy of a (b, ℓ) -connecting-path connecting $\mathbf{x}_i$ to $\mathbf{y}_i$, for every $i \in [t]$,
- $V(F_i) \setminus \{\mathbf{x}_i, \mathbf{y}_i\} \subseteq W$, and
- $V(F_i) \cap V(F_j) = \emptyset$ for all distinct $i, j \in [t]$.

In other words, a $(b, \mathcal{I}, \ell)$ -matching ‘connects’ prescribed tuples of vertices from $\mathcal{I}$ with copies of (b, ℓ) -connecting-paths. The Connecting Lemma shows that under certain conditions such a matchings exist.

Lemma 5.3 (Connecting Lemma). *For every $b \in \{1, 2\}$ and every $\alpha > 0$, there exist positive constants $\varepsilon := \varepsilon(\alpha)$ and $C := C(\alpha)$ such that w.h.p. for a random graph $\Gamma \sim G_{n,p}$ every $G \in \mathcal{G}(\Gamma, n, \alpha, p)$ has the following property.*

Let $s \geq C \log^4 n / p^2$ be an integer. Then there are at least $(1 - n^{-4}) \binom{n}{s}$ subsets $W \subseteq V(G)$ of size s satisfying the following. For every family of disjoint 4-tuples $\{\mathbf{x}_i, \mathbf{y}_i\}_{i \in [t]} \subseteq (V(G) \setminus W)^4$, such that $t \log n \leq \varepsilon^6 |W|$, there exists an $(\{\mathbf{x}_i, \mathbf{y}_i\}_{i \in [t]}, b, 4 \log n)$ -matching in W .

In [30] the authors rely on Janson’s inequality in order to show such a statement. As we are working with a subgraph of a random graph, we cannot apply this technique here. The proof of the Connecting Lemma thus becomes a much more challenging task and requires a detailed analysis of ‘expansion of the edges’ in certain subsets. We defer it to Section 7.

5.1 Proof of the main result

Let us first briefly give an overview of the various steps of the proof. The first step is to partition the graph uniformly at random into sets U , W , and X such that $|U| = |W| = \Theta(n)$ and $|X| = \Theta(n/\log^2 n)$. Next, we find an $(\mathbf{a}, \mathbf{b}, X)$ -absorber A for some $\mathbf{a}, \mathbf{b} \in W^2$, such that $V(A) \setminus X \subseteq W$. Let W' denote the vertices of W which are not part of the absorber, and let $U' = U \cup W'$. In the third step, we construct t vertex-disjoint square-paths $P_1, \dots, P_t$ in U' such that

$$\left| U' \setminus \bigcup_{i \in [t]} V(P_i) \right| \ll |X|/\log n \quad \text{and} \quad t \ll |X|/\log n.$$

Let us denote the set of vertices from U' not contained in any P_i by Q . Using the Connecting Lemma with X as the ‘reservoir’ (set W in Lemma 5.3) we connect $P_1, \dots, P_t$, vertices from Q , and pairs $\mathbf{a}$ and $\mathbf{b}$ into a square-path P such that P connects $\mathbf{b}$ to $\mathbf{a}$ and $V(G) \setminus X \subseteq V(P)$. Let X' be the set of vertices from X contained in P . By the definition of the absorber A there exists a square-path P' connecting $\mathbf{a}$ to $\mathbf{b}$ such that $V(P') = V(A) \setminus X'$. By combining P and P' we obtain the square of a Hamilton cycle. In the remainder of the section we formalise this argument.

Let $C_1 = C_{5.1}(\alpha)$, $C_2 = C_{5.3}(\alpha)$, $\varepsilon_1 = \varepsilon_{2.8}(\alpha)$, $\varepsilon_2 = \varepsilon_{5.3}(\alpha)$, $\varepsilon = \min\{\varepsilon_1, \varepsilon_2, 1/(4C_1)\}$, and $C = \max\{\varepsilon^{-2}C_1C_2, 100\}$. Let G be a member of $\mathcal{G}(\Gamma, n, \alpha, p)$ and let $V(G) = X_1 \cup X_2 \cup W \cup U$ be a partition of $V(G)$ chosen uniformly at random, where $X := X_1 \cup X_2$, such that

$$|X| = \left\lfloor \frac{\varepsilon n}{2C_1 \log^2 n} \right\rfloor, \quad |X_1| = \lfloor \varepsilon |X| \rfloor, \quad |W| = \lfloor \varepsilon n \rfloor, \quad \text{and} \quad |U| = n - |X| - |W|.$$

Note that w.h.p. W satisfies the conclusion of the Absorbing Lemma (Lemma 5.1), X_2 the conclusion of the Connecting Lemma (Lemma 5.3), and all four sets are $(\alpha/2, \varepsilon, p)$ -good by Proposition 2.8. From now on we fix such a choice of subsets.

We apply Lemma 5.1 with W and X to obtain an $(\mathbf{a}_0, \mathbf{b}_0, X)$ -absorber A for some $\mathbf{a}_0, \mathbf{b}_0 \in W^2$, such that $V(A) \setminus X \subseteq W$. We can indeed do this, as $|W| \geq \varepsilon n - 1 \geq C_1 \log^4 n / p^2$ and $|X| \leq |W| / (C_1 \log^2 n)$. Let us denote by W' the subset of vertices from W , which are not contained in A . Furthermore, let $U' := U \cup W'$. Next, we use the following claim whose proof is given at the end of the section.

Claim 5.4 (Covering Claim). *For every $\varepsilon > 0$, there exists a positive constant $K := K(\varepsilon)$ such that w.h.p. the following holds. The induced subgraph $G[U']$ contains $t \leq K \log \log n$ vertex-disjoint square-paths $P_1, \dots, P_t$ such that $|U' \setminus \bigcup_{i \in [t]} V(P_i)| \leq \varepsilon |U'| / \log^3 n$.*

By applying the Covering Claim with ε^{10} (as ε) we get that there is a constant $K = K_{5.4}(\varepsilon^{10})$ and that w.h.p. $G[U']$ contains $t \leq K \log \log n$ vertex-disjoint square-paths $P_1, \dots, P_t$ which contain all but at most

$$\frac{\varepsilon^{10} |U'|}{\log^3 n} \leq \frac{\varepsilon^{10} n}{\log^3 n} \leq \frac{\varepsilon^8 |X|}{\log n} \quad (27)$$

vertices from U' . We denote the set of uncovered vertices by Q and the end-pairs of P_i by $\mathbf{a}_i$ and $\mathbf{b}_i$, for every $i \in [t]$.

We next show that there exists a matching between Q and X_1 which saturates Q by verifying Hall’s condition. Recall, X_1 is $(\alpha/2, \varepsilon, p)$ -good. Let $Q_1 \subseteq Q$ be an arbitrary subset of Q and let us denote $N_G(Q_1, X_1)$ by Z . If $|Q_1| \leq \varepsilon^{-3} \log n / p$, then for a vertex $v \in Q_1$

$$|Z| \geq \deg_G(v, X_1) \geq (2/3 + \alpha/2) |X_1| p \geq \frac{\varepsilon^2 n p}{4C_1 \log^2 n} \geq \frac{\varepsilon^{-3} \log n}{p} \geq |Q_1|, \quad (28)$$

where the second to last inequality follows from the bound on p . On the other hand, if $|Q_1| \geq \varepsilon^{-3} \log n/p$ then by (G1) we have

$$|Q_1|(2/3 + \alpha/2)|X_1|p \leq e_G(Q_1, Z) \leq (1 + \varepsilon)|Q_1| \max\{|Z|, \varepsilon^{-3} \log n/p\}p,$$

which implies $|Z| \geq |X_1|/4$ (with room to spare) as otherwise we get $|X_1| \leq 4\varepsilon^{-3} \log n/p$, which is not true, again by the bound on p . Since $|Z| \geq |X_1|/4$ we have by (28) that $|Z| \geq |Q_1|$ and thus by Hall's theorem (Lemma 2.1) there exists a matching between Q and X_1 which saturates Q . Let us denote the edges of the matching by $\mathbf{m}_i$ for all $i \in \{1, \dots, |Q|\}$.

As the last step, we apply the Connecting Lemma (Lemma 5.3) with X_2 (as W) and with the following family of 4-tuples (as $\mathcal{I}$):

$$\{(\mathbf{b}_i, \mathbf{a}_{i+1})\}_{i \in \{0, \dots, t-1\}} \cup (\mathbf{b}_t, \mathbf{m}_1) \cup \{(\mathbf{m}_i, \mathbf{m}_{i+1})\}_{i \in [|Q|-1]} \cup (\mathbf{m}_{|Q|}, \mathbf{a}_0),$$

in order to obtain a square path P which connects $\mathbf{b}_0$ to $\mathbf{a}_0$ and contains all vertices from $V(G) \setminus V(A)$ and possibly some vertices from X_2 . Note that we can indeed apply the lemma as

$$|X_2| \geq (1 - \varepsilon) \frac{\varepsilon n}{2C_1 \log^2 n} - 2 \geq \frac{C_2 \log^4 n}{p^2}$$

and by (27) we have

$$t + |Q| + 1 \leq K \log \log n + \frac{\varepsilon^8 |X|}{\log n} + 1 \leq \frac{\varepsilon^6 |X_2|}{\log n}.$$

Finally, let $X' \subseteq X$ be the set of vertices from X contained in P . By the definition of the $(\mathbf{a}_0, \mathbf{b}_0, X)$ -absorber A there exists a path P' connecting $\mathbf{a}_0$ to $\mathbf{b}_0$ such that $V(P') = V(A) \setminus X'$. By combining P and P' we obtain the square of a Hamilton cycle in G , which concludes the proof of Theorem 1.6.

In the next subsection we provide the missing proof of the Covering Claim.

5.2 Proof of the Covering Claim

The goal of this subsection is to show that $G[U']$ contains $o(|X|/\log n)$ vertex-disjoint square-paths which contain all but at most $o(|X|/\log n)$ vertices from U' . In an earlier paper by some of the authors [35] we proved that w.h.p. any subgraph of $G_{n,p}$ contains the square of a Hamilton cycle on $(1 - o(1))n$ vertices, provided that $p \gg (\log n/n)^{1/2}$. The proof of Claim 5.4 relies on this result which we thus state precisely.

Theorem 5.5 (Škorić, Steger, Trujić [35]). *For every $\varepsilon, \alpha > 0$ there exist positive constants $C := C(\varepsilon, \alpha)$ and $b := b(\varepsilon, \alpha)$, such that if $p \geq C(\log n/n)^{1/2}$ then the random graph $\Gamma \sim G_{n,p}$ has the following property with probability at least $1 - e^{-bn^2 p / \log^2 n}$. Every spanning subgraph of Γ with minimum degree at least $(2/3 + \alpha)np$, contains the square of a cycle on at least $(1 - \varepsilon)n$ vertices.*

As a corollary we get the following statement.

Corollary 5.6. *For every $\varepsilon, \alpha > 0$ there exists a positive constant $C := C(\varepsilon, \alpha)$, such that if $p \geq C(\log^4 n/n)^{1/2}$ then the random graph $\Gamma \sim G_{n,p}$ w.h.p. has the following property. Every subgraph $G \subseteq \Gamma$ of size $v(G) \geq \varepsilon n / \log^3 n$ with minimum degree at least $(2/3 + \alpha)v(G)p$, contains the square of a cycle on at least $(1 - \varepsilon)v(G)$ vertices.*

Proof. Let $C' = C_{5.5}(\varepsilon, \alpha)$, $b = b_{5.5}(\varepsilon, \alpha)$, and let $C = \max\{1000/(\varepsilon b), C'/\sqrt{\varepsilon}\}$. Note that for all subgraphs $G \subseteq \Gamma$ of size s , such that $s \geq \varepsilon n / \log^3 n$, we have

$$p \geq C \left(\frac{\log^4 n}{n} \right)^{1/2} \geq C \left(\frac{\varepsilon \log^4 n}{s \log^3 n} \right)^{1/2} \geq C' \left(\frac{\log s}{s} \right)^{1/2}.$$

Thus, for a fixed subgraph G of size s with the required minimum degree we have by Theorem 5.5 that with probability at least $1 - e^{-bs^2p/\log^2 s}$, G contains the square of a cycle on at least $(1 - \varepsilon)s$ vertices. Since

$$\frac{bs^2p}{\log^2 s} \geq s \cdot \frac{b\varepsilon np}{\log^5 n} \geq s \cdot \frac{b\varepsilon C \cdot \sqrt{n}}{\log^2 n} \geq 1000s \log n,$$

we may additionally do the union bound over all $s \geq \varepsilon n / \log^3 n$ and all subsets of size s . $\square$

With Corollary 5.6 at hand we are ready to give the proof of Claim 5.4 by using a bootstrapping technique developed by Nenadov and the second author [31].

Proof of Claim 5.4. Without loss of generality we assume that ε is sufficiently small w.r.t. α . Recall that U is $(\alpha/2, \varepsilon, p)$ -good and $|U| \geq (1 - 2\varepsilon)|U'|$, so for each $v \in U'$ we have by (G4)

$$\deg_G(v, U') \geq \deg_G(v, U) \geq (2/3 + \alpha/2)|U|p \geq (1 - 2\varepsilon)(2/3 + \alpha/2)|U'|p \geq (2/3 + \alpha/4)|U'|p.$$

Let q be the largest integer such that $|U'|/2^{q-1} \geq \lceil n/\log^3 n \rceil$ and note that $q = O(\log \log n)$. Consider a uniformly at random chosen partition $U' = V_1 \cup \dots \cup V_q$ such that $V_i = \lfloor |U'|/2^i \rfloor$ for all $i \in [q-1]$ and

$$|V_q| = |U'| - (|V_1| + \dots + |V_{q-1}|) \geq |U'| - |U'| \sum_{i=1}^{q-1} 2^{-i} = |U'|/2^{q-1}.$$

Similarly, $|V_q| \leq |U'|/2^{q-1} + q$. Since $p \gg n^{-1/2} \log^3 n$ we have

$$|V_i| \geq |U'|/2^{q-1} - 1 \geq \frac{n}{2 \log^3 n} \geq \frac{\log^2 n}{p}. \quad (29)$$

Thus, as V_i is a random subset of U' , by a simple application of Chernoff's inequality for a hypergeometric distribution we get

$$\Pr[\deg_G(v, V_i) < (2/3 + \alpha/8)|V_i|p] = e^{-\Omega(|V_i|p)} < 1/n^2,$$

where the last inequality follows from (29). Using the union bound over all $v \in U'$ and $i \in [q]$ we have that w.h.p. for each $v \in U'$ and each $i \in [q]$

$$\deg_G(v, V_i) \geq (2/3 + \alpha/8)|V_i|p.$$

Since $p \gg n^{-1/2} \log^3 n$ we have that w.h.p. Corollary 5.6 holds when applied with $\varepsilon/8$ (as ε) and $\alpha/16$ (as α). Having this in mind, we prove by induction on $i \in [q]$ that $G[V_1 \cup \dots \cup V_i]$ contains i square-paths which cover all but at most $(\varepsilon/4)|V_i|$ vertices from $V_1 \cup \dots \cup V_i$. Since $q = O(\log \log n)$ and

$$|V_q| \leq |U'|/2^{q-1} + q \leq \frac{2n}{\log^3 n} + q \leq \frac{4n}{\log^3 n},$$

we have that by setting $i = q$ the induction implies the claim.

By Corollary 5.6 we directly get that there exists a square-path in $G[V_1]$ which covers all but at most $(\varepsilon/8)|V_1|$ vertices from V_1 , settling the base case. Assume now that the hypothesis

holds for some $1 \leq i < q$ and let $P_1, \dots, P_i$ be the square-paths guaranteed by the hypothesis. Furthermore, let $Q \subseteq V_1 \cup \dots \cup V_i$ denote the subset of vertices not contained in any of the P_i 's. Then $|Q| \leq (\varepsilon/8)|V_i| \leq \varepsilon|V_{i+1}|$ and for every $v \in U'$ we have

$$\begin{aligned} \deg_G(v, Q \cup V_{i+1}) &\geq \deg_G(v, V_{i+1}) \geq (2/3 + \alpha/8)|V_{i+1}|p \\ &\geq (2/3 + \alpha/8) \frac{|Q| + |V_{i+1}|}{1 + \varepsilon} p \geq (2/3 + \alpha/16)|Q \cup V_{i+1}|p, \end{aligned} \quad (30)$$

where the last inequality follows from the assumption on ε . Therefore, by Corollary 5.6 we know that $G[Q \cup V_{i+1}]$ contains a square-path P_{i+1} which covers all but at most $(\varepsilon/8)|Q \cup V_{i+1}| \leq (\varepsilon/4)|V_{i+1}|$ vertices. Observe that we can indeed use Corollary 5.6 since $|V_{i+1}| \geq n/(2 \log^3 n)$ and by (30) we have that $\delta(G[Q \cup V_{i+1}]) \geq (2/3 + \alpha/16)|Q \cup V_{i+1}|p$. As the vertices from $(V_1 \cup \dots \cup V_i) \setminus Q$ are already contained in $V(P_1) \cup \dots \cup V(P_i)$, this shows that the hypothesis holds for $i + 1$. $\square$

6 Proof of the Absorbing Lemma

Our strategy for constructing an absorber for a set X consists of two steps. In the first step we find an $(\mathbf{a}_i, \mathbf{b}_i, \{x_i\})$ -absorber A_i (a *single-vertex* absorber) for each $x_i \in X$, such that they are pairwise disjoint. In the second step, by using the Connecting Lemma, we find a square-path from $\mathbf{b}_i$ to $\mathbf{a}_{i+1}$, for every $1 \leq i \leq m - 1$, such that they are pairwise disjoint and also disjoint from A_i 's. It is easy to see that this gives an $(\mathbf{a}_1, \mathbf{b}_m, X)$ -absorber: given $X' \subseteq X$, for every $x_i \in X$ we choose a square-path in A_i depending on whether $x_i \in X'$ or $x_i \notin X'$.

We use the following construction for a single-vertex absorber.

Claim 6.1. *Let A_x be a graph obtained as the union of following graphs and edges*

$$B_\ell \cup \{w_{1,2}, w_{1,3}\} \cup \bigcup_{i \in [4]} \{w_{1,i}, x\} \cup \bigcup_{i \in [\ell]} U_i,$$

where U_i is a square-path connecting $\mathbf{w}_i^b$ to $\mathbf{w}_{i+1}^a$ for every $1 \leq i < \ell$, such that all the square-paths are pairwise vertex-disjoint and also disjoint from B_ℓ (except for the pairs of vertices they connect). Furthermore, vertex x is not contained in either B_ℓ or any U_i . Then A_x is a $(\mathbf{w}_1^a, \mathbf{w}_\ell^b, \{x\})$ -absorber.

Proof. There are only two cases we need to consider: $X' = \emptyset$ and $X' = \{x\}$. We specify the desired square-path from $\mathbf{w}_1^a$ to $\mathbf{w}_\ell^b$ in each case by giving the ordering in which we traverse the vertices of such a path (see Figure 5):

- $X' = \emptyset$: $\mathbf{w}_1^a, x, \mathbf{w}_1^b, U_1, \mathbf{w}_2^a, \mathbf{w}_2^b, U_2, \mathbf{w}_3^a, \mathbf{w}_3^b, U_3, \mathbf{w}_4^a, \dots, \mathbf{w}_{\ell-1}^b, U_{\ell-1}, \mathbf{w}_\ell^a, \mathbf{w}_\ell^b$,
- $X' = \{x\}$: $\mathbf{w}_1^a, \overline{\mathbf{w}}_2^a, \overline{U}_1, \overline{\mathbf{w}}_1^b, \overline{\mathbf{w}}_3^a, \overline{U}_2, \overline{\mathbf{w}}_2^b, \overline{\mathbf{w}}_4^a, \overline{U}_3, \overline{\mathbf{w}}_3^b, \dots, \overline{\mathbf{w}}_\ell^a, \overline{U}_{\ell-1}, \overline{\mathbf{w}}_{\ell-1}^b, \mathbf{w}_\ell^b$.

$\square$

Now, we are ready to present the proof of the Absorbing Lemma. For the convenience of the reader we first restate the lemma.

Lemma 5.1 (Absorbing Lemma). *For every $\alpha > 0$, there exists a positive constant $C := C(\alpha)$ such that w.h.p. for a random graph $\Gamma \sim G_{n,p}$ every $G \in \mathcal{G}(\Gamma, n, \alpha, p)$ has the following property.*

Let $s \geq C \log^4 n / p^2$ be an integer. Then there are at least $(1 - n^{-3}) \binom{n}{s}$ subsets $W \subseteq V(G)$ of size s satisfying the following. For every subset $X \subseteq V(G) \setminus W$ of size $|X| \leq |W| / (C \log^2 n)$ there exists an $(\mathbf{a}, \mathbf{b}, X)$ -absorber A in G such that $V(A) \setminus X \subseteq W$, where $\mathbf{a}, \mathbf{b} \in W^2$ are two disjoint pairs of vertices.

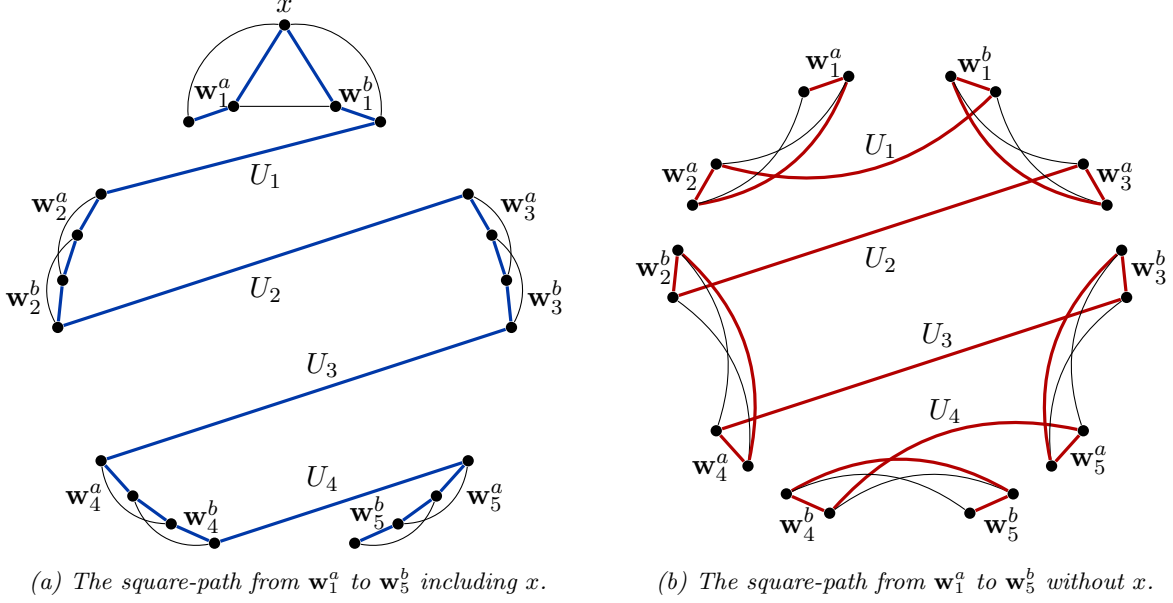


Figure 5: A single-vertex absorber A_x . Note that both square-paths use all vertices of A_x .

Proof of Lemma 5.1. Let $\varepsilon' = \varepsilon_{5.3}(\alpha)$, $\varepsilon'' = \varepsilon_{2.8}(\alpha)$, $\varepsilon = \min\{\varepsilon', \varepsilon'', 1/100\}$, $C' = C_{5.3}(\varepsilon', \alpha)$, and $C = 40 \max\{C', \varepsilon^{-6}\}$. Let $W = W_1 \cup \dots \cup W_7$ be such that W_i 's are pairwise vertex-disjoint, all of size $|W_i| \geq \lfloor s/10 \rfloor$, and chosen uniformly at random from G among all sets of prescribed size. By the union bound, with probability at least $1 - o(n^{-3})$ it holds that for all $i \in [7]$, W_i is $(\alpha/2, \varepsilon, p)$ -good by Proposition 2.8, and satisfies the conclusion of the Connecting Lemma (Lemma 5.3). From now on fix such a choice of sets W_i (and thus W) and note that it is sufficient to show the conclusion of the lemma for this particular W .

Let $X = \{x_1, \dots, x_m\} \subseteq V(G) \setminus W$ be a subset of vertices such that $m \leq s/(C \log^2 n)$ and observe that $4|X| \log^2 n \leq \varepsilon^6 |W_i|$. Furthermore, let S_{x_i} be the subgraph of A_{x_i} (as defined in Claim 6.1) induced on the vertex set $\{w_{1,1}, w_{1,2}, w_{1,3}, w_{1,4}, x_i\}$. We aim to construct a vertex-disjoint collection $\{S_{x_i}\}_{i \in [m]}$ in G , such that each S_{x_i} contains x_i and no other vertex from X . We do this in four steps.

Step 1. First, we show that there exists a matching M_1 between X and W_1 saturating X . Let $X' \subseteq X$ be a subset of X and let us denote $N_G(X', W_1)$ by Z . If $|X'| \leq \varepsilon^{-3} \log n/p$ then for a vertex $x_i \in X'$

$$|Z| \geq |N_G(x_i, W_1)| \geq (2/3 + \alpha/2)|W_1|p \geq (2C/30) \log^4 n/p \geq |X'|,$$

where the second inequality follows from (G4). If we assume $|X'| > \varepsilon^{-3} \log n/p$ (and hence $|Z| \geq \varepsilon^{-3} \log n/p$ by analysis from above) then by (G1) and (G4) we have

$$|X'|(2/3 + \alpha/2)|W_1|p \leq e_G(X', Z) \leq (1 + \varepsilon)|X'||Z|p,$$

which implies $|Z| \geq |W_1|/2 \geq |X'|$. Thus, Hall's condition is satisfied and the desired matching M_1 exists. Let us denote the matched vertex of some $x_i \in X$ in M_1 by $M_1(x_i)$.

Step 2. In the next step, we want to find a family of m vertex-disjoint triangles $\mathcal{T}_1$, such that each triangle contains exactly one edge from M_1 and intersects W_2 in exactly one vertex. We achieve this again with the help of Hall's matching theorem (Lemma 2.1). Let $X' \subseteq X$ and let us denote

$$\bigcup_{x_i \in X'} (N_G(x_i, W_2) \cap N_G(M_1(x_i), W_2))$$

by Z . By applying (G7) to the edges of M_1 incident to the vertices of X' or its arbitrary subset of size ε/p^2 in case $|X'| > \varepsilon/p^2$ (as $\mathcal{P}$) we have

$$|Z| \geq \alpha \cdot \min\{\varepsilon/p^2, |X'|\} |W_2| p^2 \geq \alpha \cdot \min\{\varepsilon |W_2|, |X'| \cdot (C/10) \log^4 n\} \geq |X'|,$$

and by Lemma 2.1 we conclude that the desired collection $\mathcal{T}_1$ exists. For $x_i \in X$ let us denote by $\mathbf{u}_i = (u_{i,1}, u_{i,2})$ the two vertices sharing the triangle with x_i from $\mathcal{T}_1$, where $u_{i,1} \in W_1$ and $u_{i,2} \in W_2$.

Step 3. In a manner analogous to that seen in the second step, we find a collection $\mathcal{T}_2$ of m vertex-disjoint triangles such that each triangle in $\mathcal{T}_2$ contains exactly one vertex from W_3 and vertices x_i and $u_{i,2}$, for every $i \in [m]$. Let us denote by $v_{i,1}$ the third vertex in the triangle from $\mathcal{T}_2$ which contains x_i and $u_{i,2}$.

Step 4. In the last step, we find a collection $\mathcal{T}_3$ of m vertex disjoint triangles such that each triangle in $\mathcal{T}_3$ contains exactly one vertex from W_4 and vertices x_i and $v_{i,1}$, for every $i \in [m]$. Let us denote by $v_{i,2}$ the third vertex in the triangle from $\mathcal{T}_3$ which contains x_i and $v_{i,1}$ and $\mathbf{v}_i = (v_{i,1}, v_{i,2})$.

This completes the first part of the embedding scheme as we have constructed the vertex-disjoint collection $\{S_{x_i}\}_{i \in [m]}$ containing vertices from X .

The rest of the proof consists of three consecutive applications of the Connecting Lemma (Lemma 5.3). First, by applying it with $b = 2$, W_5 (as W), and $\{(\bar{\mathbf{u}}_i, \bar{\mathbf{v}}_i)\}_{i \in [m]}$ (as $\{(\mathbf{x}_i, \mathbf{y}_i)\}_i$) we conclude that there exists a

$$(\{(\bar{\mathbf{u}}_i, \bar{\mathbf{v}}_i)\}_{i \in [m]}, 2, 4 \log n)\text{-matching}$$

in W_5 . We can apply the Connecting Lemma as $|W_5| \geq (C/10) \log^4 n / p^2$ and $m \log n \leq \varepsilon^6 |W_5|$. Let $\ell = \log n$ and let g_i be the embedding of the backbone-path B_ℓ given by the above matching, for each $i \in [m]$. Next, consider the family of 4-tuples $\{(g_i(\mathbf{w}_j^b), g_i(\mathbf{w}_{j+1}^a))\}_{i \in [m], j \in [\ell-1]}$. We apply the Connecting Lemma with $b = 1$, W_6 (as W), and $\{(g_i(\mathbf{w}_j^b), g_i(\mathbf{w}_{j+1}^a))\}_{i \in [m], j \in [\ell-1]}$ (as $\{(\mathbf{x}_i, \mathbf{y}_i)\}_i$) to conclude that there exists a

$$(\{(g_i(\mathbf{w}_j^b), g_i(\mathbf{w}_{j+1}^a))\}_{i \in [m], j \in [\ell-1]}, 1, 4 \log n)\text{-matching}$$

in W_6 . We can do that as $|W_6| \geq (C/10) \log^4 n / p^2$ and $4m \log^2 n \leq \varepsilon^6 |W_6|$. Let us denote pairs $g_i(\mathbf{w}_1^a)$ and $g_i(\mathbf{w}_\ell^b)$ by $\mathbf{a}_i$ and $\mathbf{b}_i$, for every $i \in [m]$. By Claim 6.1 we conclude that the set $X \cup W_1 \cup \dots \cup W_6$ contains an $(\mathbf{a}_i, \mathbf{b}_i, \{x_i\})$ -absorber A_i for each $i \in [m]$, such that all A_i 's are pairwise vertex-disjoint.

Lastly, using the vertices in W_7 we connect all A_i 's into a single absorber for the set X . Consider the family of 4-tuples $\{(\mathbf{b}_i, \mathbf{a}_{i+1})\}_{i \in [m-1]}$. By applying the Connecting Lemma with $b = 1$, W_7 (as W), and $\{(\mathbf{b}_i, \mathbf{a}_{i+1})\}_{i \in [m-1]}$ (as $\{(\mathbf{x}_i, \mathbf{y}_i)\}_i$), there exists an

$$(\{(\mathbf{b}_i, \mathbf{a}_{i+1})\}_{i \in [m-1]}, 1, 4 \log n)\text{-matching}$$

in W_7 . We can do that as $|W_7| \geq (C/10) \log^4 n / p^2$ and $m \log n \leq \varepsilon^6 |W_7|$. If we denote the square-paths connecting $\mathbf{b}_i$ to $\mathbf{a}_{i+1}$ (given by the last application of the Connecting Lemma) by $Q_1, Q_2, \dots, Q_{m-1}$, then Proposition 4.1 implies that

$$A = \bigcup_{i \in [m]} A_i \cup \bigcup_{i \in [m-1]} Q_i$$

is an $(\mathbf{a}_1, \mathbf{b}_m, X)$ -absorber: consider some subset $X' \subseteq X$ and for each $i \in [m]$ let $P_i \subseteq A_i$ be the square-path from $\mathbf{a}_i$ to $\mathbf{b}_i$ which contains x_i if and only if $x_i \notin X'$ and, moreover, contains all

other vertices in A_i . Such a path exists as A_i is an $(\mathbf{a}_i, \mathbf{b}_i, \{x_i\})$ -absorber. Then

$$\begin{array}{ccccccccc} P_1 & & Q_1 & & P_2 & & Q_2 & & \dots & & Q_{m-1} & & P_m \\ \mathbf{a}_1 & \rightsquigarrow & \mathbf{b}_1 & \rightsquigarrow & \mathbf{a}_2 & \rightsquigarrow & \mathbf{b}_2 & \rightsquigarrow & \dots & \rightsquigarrow & \mathbf{a}_m & \rightsquigarrow & \mathbf{b}_m \end{array}$$

is a square-path from $\mathbf{a}_1$ to $\mathbf{b}_m$ which contains all vertices in A except those in X' . This concludes the proof of the lemma. $\square$

7 Proof of the Connecting Lemma

In the remainder of the paper we give the proof of the Connecting Lemma. For the convenience of the reader let us first restate the lemma.

Lemma 5.3 (Connecting Lemma). *For every $b \in \{1, 2\}$ and every $\alpha > 0$, there exist positive constants $\varepsilon := \varepsilon(\alpha)$ and $C := C(\alpha)$ such that w.h.p. for a random graph $\Gamma \sim G_{n,p}$ every $G \in \mathcal{G}(\Gamma, n, \alpha, p)$ has the following property.*

Let $s \geq C \log^4 n / p^2$ be an integer. Then there are at least $(1 - n^{-4}) \binom{n}{s}$ subsets $W \subseteq V(G)$ of size s satisfying the following. For every family of disjoint 4-tuples $\{\mathbf{x}_i, \mathbf{y}_i\}_{i \in [t]} \subseteq (V(G) \setminus W)^4$, such that $t \log n \leq \varepsilon^6 |W|$, there exists an $(\{\mathbf{x}_i, \mathbf{y}_i\}_{i \in [t]}, b, 4 \log n)$ -matching in W .

The proof relies on the following lemma whose proof we defer to the next subsection. The idea behind it is that even after removal of a not too large subset X from the reservoir W , we can find a copy of a $(b, 4 \log n)$ -connecting path, or in the phrasing of the lemma above—a ‘matching’, connecting at least one pair $\mathbf{x}_i$ to the corresponding pair $\mathbf{y}_i$.

Lemma 7.1. *For every $b \in \{1, 2\}$ and every $\alpha > 0$, there exist positive constants $\varepsilon := \varepsilon(\alpha)$ and $C := C(\alpha)$ such that w.h.p. for a random graph $\Gamma \sim G_{n,p}$ every $G \in \mathcal{G}(\Gamma, n, 2\alpha, p)$ has the following property.*

Let $s \geq C \log^4 n / p^2$ be an integer. Then there are at least $(1 - n^{-4}) \binom{n}{s}$ subsets $W \subseteq V(\Gamma)$ of size s satisfying the following. For every family of disjoint 4-tuples $\{\mathbf{x}_i, \mathbf{y}_i\}_{i \in [t]} \subseteq (V(G) \setminus W)^4$, such that $t \log n \leq \varepsilon^6 |W|$, and every subset $X \subseteq W$ of size $|X| \leq 8t \log n$, there exists an $i \in [t]$ and an $(\{\mathbf{x}_i, \mathbf{y}_i\}, b, 4 \log n)$ -matching in $W \setminus X$.

Proof of Lemma 5.3. For given α , let $\varepsilon = \varepsilon_{7.1}(\alpha/2)$ and $C = C_{7.1}(\varepsilon, \alpha/2)$. Set $\ell = 4 \log n$ and $\mathcal{I} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \in [t]}$. Let W be one of the $(1 - n^{-4}) \binom{n}{s}$ subsets satisfying the conclusion of Lemma 7.1.

We define an ℓ -uniform hypergraph $\mathcal{H}$ on the vertex set $\mathcal{I} \cup W$ whose edge set is defined as follows. For every 4-tuple $(\mathbf{x}_i, \mathbf{y}_i)$ and every set $Y \subseteq W$ of size $\ell - 4$, we add an edge $(\mathbf{x}_i, \mathbf{y}_i) \cup Y$ if and only if G contains a (b, ℓ) -connecting-path connecting $\mathbf{x}_i$ to $\mathbf{y}_i$ and its internal vertices belong to Y . Clearly, if there is an $\mathcal{I}$ -saturating matching in $\mathcal{H}$, then there is an $(\mathcal{I}, b, \ell)$ -matching in W . We use Haxell’s criteria (Theorem 2.2) in order to show this.

Let $\mathcal{I}' \subseteq \mathcal{I}$ and $X \subseteq W$ be arbitrary subsets such that $|X| \leq 2|\mathcal{I}'| \cdot \ell$. It is enough to show that for some $(\mathbf{x}_i, \mathbf{y}_i) \in \mathcal{I}'$ there is a (b, ℓ) -connecting-path connecting $\mathbf{x}_i$ to $\mathbf{y}_i$ whose internal vertices are completely contained in the set $W \setminus X$. This in turn implies that $\mathcal{H}$ contains an edge intersecting $\mathcal{I}'$ and not intersecting X and the condition of Theorem 2.2 is satisfied. Applying Lemma 7.1 with $\mathcal{I}'$ (as $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i \in [t]}$) gives us exactly that. Namely, we may apply the lemma since $|\mathcal{I}'| \log n \leq t \log n \leq \varepsilon^6 |W|$ and $|X| \leq 2|\mathcal{I}'| \cdot \ell \leq 8t \log n$. $\square$

7.1 Proof of Lemma 7.1

Let us set $\varepsilon = \min\{1/2^{300}, \alpha^{10}, \varepsilon_{2.8}(2\alpha)\}$, $C = 5\varepsilon^{-27}$, and take $m = 4\log n - 4$. Let G be an arbitrary member of $\mathcal{G}(\Gamma, n, 2\alpha, p)$. Let $s \geq C \log^4 n / p^4$ and $\tilde{n} := s / (5 \log n) \geq \varepsilon^{-27} \log^3 n / p^2$. Let W be a set of size s chosen uniformly at random among all such sets. As by Proposition 2.8 w.h.p. Γ is such that for every $G \in \mathcal{G}(\Gamma, n, 2\alpha, p)$ at least $(1 - n^{-5}) \binom{n}{\tilde{n}}$ sets of size $\tilde{n}$ are (α, ε, p) -good, it follows by Chernoff's inequality for hypergeometrically distributed random variables that with high probability at least $(1 - n^{-4}) \binom{s}{\tilde{n}}$ subsets $W' \subseteq W$ of size $\tilde{n}$ are (α, ε, p) -good. In particular, w.h.p. Γ is such that for every $G \in \mathcal{G}(\Gamma, n, 2\alpha, p)$, there are at least $(1 - n^{-4}) \binom{n}{s}$ sets $W \subseteq V(G)$ of size s which in turn contain at least $(1 - n^{-4}) \binom{s}{\tilde{n}}$ subsets of size $\tilde{n}$ which are (α, ε, p) -good. Condition on this event and let W be one of those $(1 - n^{-4}) \binom{n}{s}$ sets of size s .

Fix an arbitrary $X \subseteq W$ of size $|X| \leq 8t \log n$. We now show the existence of disjoint subsets $W_1, \dots, W_m$ of W which for all $i \in [m]$ satisfy:

$$(W1) \quad |W_i| = \tilde{n}, \text{ for } \tilde{n} := \left\lceil \frac{|W|}{5 \log n} \right\rceil \geq \frac{\varepsilon^{-27} \log^3 n}{p^2},$$

$$(W2) \quad W_i \text{ is } (\alpha, \varepsilon, p)\text{-good, and}$$

$$(W3) \quad |X \cap W_i| \leq \varepsilon^5 \tilde{n}.$$

Consider a uniformly at random chosen collection $W_1, \dots, W_m$ of disjoint subsets of W satisfying (W1). We claim that such a random collection satisfies (W2) and (W3) with positive probability. First, as each W_i is u.a.r. chosen from W which contains $(1 - n^{-4}) \binom{s}{\tilde{n}}$ subsets which are good, (W2) follows the union bound over all $i \in [m]$. As for (W3), observe first that $|X| \leq 8t \log n \leq 8\varepsilon^6 |W|$. If $|X| \leq \varepsilon^{-3} \log^2 n$ then the claim holds vacuously. Otherwise another application of Chernoff's inequality and the union bound implies that with probability at least

$$1 - 4 \log n \cdot e^{-\varepsilon^2 |X| / (6 \log n)} \geq 1 - 4 \log n \cdot e^{-\log n / (6\varepsilon)} \geq 1 - n^{-5},$$

we have

$$|X \cap W_i| \leq (1 + \varepsilon) |X| \frac{\tilde{n}}{|W|} \leq (1 + \varepsilon) \cdot 8\varepsilon^6 \tilde{n} \leq \varepsilon^5 \tilde{n},$$

for every $i \in [m]$. As $\max\{\varepsilon^{-3} \log^2 n, \varepsilon^5 \tilde{n}\} = \varepsilon^5 \tilde{n}$, it holds that $|X \cap W_i| \leq \varepsilon^5 \tilde{n}$. From now on we thus assume that we have disjoint subsets $W_1, \dots, W_m$ that satisfy (W1)–(W3). For convenience, we also set $X_i := X \cap W_i$ and $\tilde{W}_i := W_i \setminus X_i$, for every $i \in [m]$.

The goal of the remainder of the proof is to show that there exists an embedding g of a $(b, m+4)$ -pseudo-path connecting $\mathbf{x}_i$ to $\mathbf{y}_i$, for some $i \in [t]$. In order to do this, we first introduce a couple of definitions. Let $f: \{0, \dots, m\} \rightarrow \mathbb{Z}$ be a function defined as:

$$f(i) = \begin{cases} i - 1, & \text{if } i \text{ is even,} \\ i - b, & \text{otherwise.} \end{cases}$$

Note that this function can be used to describe the left neighbour other than u_{i-1} of a vertex u_i , $i \geq 3$, in a (b, ℓ) -pseudo-path. Indeed, the two left neighbours are $u_{f(i-1)}$ and u_{i-1} . Next, we define a graph that is the union of all (b, ℓ) -pseudo-paths that start in a set of given edges.

Definition 7.2 (Projection graph). Let $\pi: [m] \rightarrow [m]$ be a permutation of the set $[m]$ and let $\{(a_i, b_i)\}_{i \in [t]} \subseteq (V(G) \setminus W)^2$, denoted by $\mathcal{I}$, be a set of t disjoint ordered pairs. We define an $(\mathcal{I}, \pi)$ -projection graph F on the vertex set

$$V(F) = W_{-1} \cup W_0 \cup W_{\pi(1)} \cup \dots \cup W_{\pi(m)},$$

where $W_{-1} = \{a_1, \dots, a_t\}$ and $W_0 = \{b_1, \dots, b_t\}$. The edge set of F is defined inductively as follows. Let E_0 be the edges between the sets W_{-1} and W_0 , i.e. all edges in $\bigcup_{i \in [t]} \{a_i, b_i\}$. Then for all $j \in [m]$ we let

$$E_j := \{\{u, v\}, \{w, v\} : v \in \tilde{W}_{\pi(j)} \text{ and } \{u, w\} \in E_{j-1}(W_{\pi(f(j-1))}, W_{\pi(j-1)}), \{w, v\}, \{u, v\} \in E\}.$$

Lastly, we set the edge set of F to $E_0 \cup \dots \cup E_m$ (Figure 6).

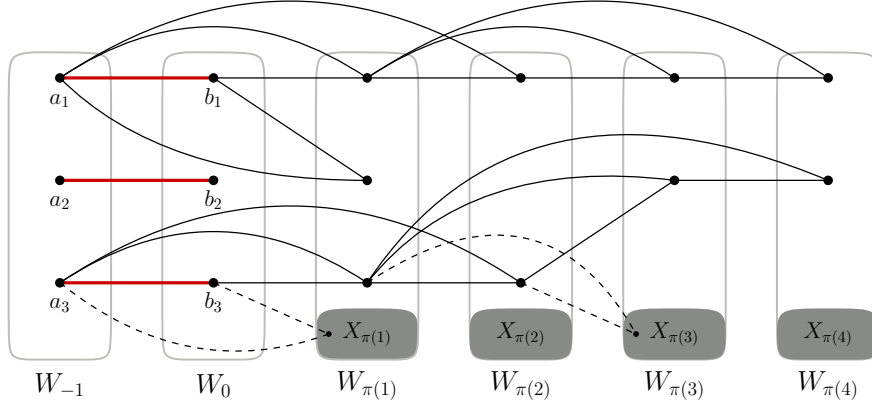


Figure 6: An example of a projection graph F where $m = 4$ and $b = 2$. Dashed edges do not belong to the projection graph F .

To understand this definition, observe that E_{j-1} are exactly those edges for which an edge expansion into $\tilde{W}_{\pi(j)}$ ‘extends’ the pseudo-path constructed so far by a vertex from $\tilde{W}_{\pi(j)}$. Crucially, even though the vertex set of F is defined as the union of the sets W_i , the edge set consists *only* of edges that run between $\tilde{W}_i$ ’s.

The next proposition thus follows immediately.

Proposition 7.3. *Let $b \in \{1, 2\}$, let $\pi: [m] \rightarrow [m]$ be a permutation of the set $[m]$, and let $\{(a_i, b_i)\}_{i \in [t]} \subseteq (V(G) \setminus W)^2$, denoted by $\mathcal{I}$, be a family of t disjoint ordered pairs. Furthermore, let F be an $(\mathcal{I}, \pi)$ -projection graph. Then for each j , where $1 \leq j \leq m/2$, and each edge $\{v, w\} \in E_F(W_{\pi(2j-1)}, W_{\pi(2j)})$, there exists an $i \in [t]$ and an embedding of a $(b, 2j + 2)$ -pseudo-path in F connecting (a_i, b_i) to (v, w) that contains exactly one vertex from each set W_i , $i \in [m]$, and no vertex from the set X . $\square$*

The following claim is the main tool in the proof of Lemma 7.1. The proof of the claim is technical and quite involved and thus it is presented in the next section. In the remainder of this section, we show how the claim implies Lemma 7.1.

Claim 7.4. *Let $t' \in \mathbb{N}$ be such that $t/3 \leq t' \leq t$. Let $\pi: [m] \rightarrow [m]$ be a permutation of the set $[m]$ and let $\{(a_i, b_i)\}_{i \in [t']} \subseteq (V(G) \setminus W)^2$, denoted by $\mathcal{I}$, be a set of t' disjoint edges in G . Then there exists $(a_i, b_i) \in \mathcal{I}$ such that*

$$e_F(W_{\pi(m/2-1)}, W_{\pi(m/2)}) \geq \frac{2}{3} \tilde{n}^2 p \quad \text{and} \quad e_F(W_{\pi(m/2+1)}, W_{\pi(m/2+2)}) \geq \frac{2}{3} \tilde{n}^2 p,$$

where F is the $(\{(a_i, b_i)\}, \pi)$ -projection graph.

Having the previous claim at hand, we finish the proof of Lemma 7.1. Let π_1 be the identity permutation of the set $[m]$ and let π_2 be a permutation of $[m]$ defined as $\pi_2(i) = m - i + 1$. Let $\mathcal{I}_x$ be the largest subset of $\{\mathbf{x}_i\}_{i \in [t]}$ such that for every $\mathbf{x}_i \in \mathcal{I}_x$ it holds that

$$e_{F'}(W_{m/2-1}, W_{m/2}) \geq \frac{2}{3} \tilde{n}^2 p,$$

where F' is the $(\mathcal{I}_x, \pi_1)$ -projection graph. If $|\mathcal{I}_x| \leq t/2$ then by applying Claim 7.4 with $\{\mathbf{x}_i\}_{i \in [t]} \setminus \mathcal{I}_x$ (as $\mathcal{I}$) and π_1 (as π) we get a contradiction with the maximality of $\mathcal{I}_x$. Thus $|\mathcal{I}_x| > t/2$. Similarly, let $\mathcal{I}_y$ be the largest subset of $\{\bar{\mathbf{y}}_i\}_{i \in [t]}$ such that for every $\bar{\mathbf{y}}_i \in \mathcal{I}_y$ it holds that

$$e_{F''}(W_{\pi_2(m/2+1)}, W_{\pi_2(m/2+2)}) \geq \frac{2}{3}\tilde{n}^2 p,$$

where F'' is the $(\mathcal{I}_y, \pi_2)$ -projection graph. As in the case of $\mathcal{I}_x$ it must be that $|\mathcal{I}_y| > t/2$. The fact that both $\mathcal{I}_x$ and $\mathcal{I}_y$ are larger than $t/2$ implies that there must be a single integer $i^* \in [t]$ such that

$$e_{F_x}(W_{m/2-1}, W_{m/2}) \geq \frac{2}{3}\tilde{n}^2 p \quad \text{and} \quad e_{F_y}(W_{\pi_2(m/2+1)}, W_{\pi_2(m/2+2)}) \geq \frac{2}{3}\tilde{n}^2 p, \quad (31)$$

where F_x and F_y are the $(\{\mathbf{x}_{i^*}\}, \pi_1)$ -projection graph and the $(\{\bar{\mathbf{y}}_{i^*}\}, \pi_2)$ -projection graph, respectively.

Let $\ell_1 = m/2$. Note that $\{\pi_2(m - \ell_1 + 2), \pi_2(m - \ell_1 + 1)\} = \{\ell_1 - 1, \ell_1\}$. Hence, from (31) we have

$$e_{F_x}(W_{\ell_1-1}, W_{\ell_1}) \geq \frac{2}{3}\tilde{n}^2 p \quad \text{and} \quad e_{F_y}(W_{\ell_1-1}, W_{\ell_1}) \geq \frac{2}{3}\tilde{n}^2 p.$$

This implies, by (G1), that there must exist an edge $e = \{u, v\}$ such that

$$e \in E_{F_x}(W_{\ell_1-1}, W_{\ell_1}) \cap E_{F_y}(W_{\ell_1-1}, W_{\ell_1}).$$

By Proposition 7.3 and the definitions of π_1 and π_2 we get that there exist two embeddings g_1 and g_2 of a $(b, \ell_1 + 2)$ -pseudo-path and a $(b, \ell_1 + 4)$ -pseudo-path such that:

- $(g_1(u_1), g_1(u_2)) = \mathbf{x}_{i^*}$ and $(g_2(u_1), g_2(u_2)) = \bar{\mathbf{y}}_{i^*}$,
- $g_1(u_i) \in W_{i-2}$, for every $i \in \{3, \dots, \ell_1\}$,
- $g_2(u_i) \in W_{m-i+3}$, for every $i \in \{3, \dots, \ell_1 + 2\}$,
- $g_1(u_{\ell_1+1}) = u$, $g_1(u_{\ell_1+2}) = v$, and
- $(g_2(u_{\ell_1+3}), g_2(u_{\ell_1+4})) = \begin{cases} (v, u), & \text{if } b = 1, \\ (u, v), & \text{if } b = 2. \end{cases}$

Using Proposition 4.1 and Propositions 4.2 we conclude that there exists an $(\{(\mathbf{x}_i, \mathbf{y}_i)\}, b, m + 4)$ -matching in $W \setminus X$, as desired. This concludes the proof of Lemma 7.1. It remains to prove Claim 7.4.

7.2 Proof of Claim 7.4

In this subsection we give the proof of Claim 7.4. As the choice of the permutation π does not play a role in the proof, we assume π is the identity permutation and we completely omit it from the definition of the projection graph. Thus, throughout the section when we write $\mathcal{I}$ -projection graph we mean $(\mathcal{I}, \pi)$ -projection graph, where π is the identity permutation. For a projection graph F , we refer to a pair of bipartite graphs $F[W_{f(i)}, W_i]$ and $F[W_{f(i+1)}, W_{i+1}]$ as the i -th step of F . Next, we introduce some terminology and define when a step is *expanding* or *non-expanding*.

Definition 7.5. Let F be a projection graph. Let $C \geq 1$ be a real number and let $i \geq 1$. We say that the i -th step of F is C -*expanding* if

$$e_F(W_{f(i+1)}, W_{i+1}) \geq C \cdot e_F(W_{f(i)}, W_i).$$

Otherwise, it is C -*non-expanding*.

Due to the asymmetry in the definition of pseudo-paths (for $b = 2$), it is easier to not consider every step, but to group steps into *blocks of two*.

Definition 7.6. Let F be a projection graph. Let $C \geq 1$ be a real number and let $i \geq 1$. We say that the i -th block is C -*expanding* in F , if at least one of the $2i$ -th and the $(2i + 1)$ -st step of F are C -expanding. Otherwise, it is C -*non-expanding*.

To understand how the two definitions relate, it is helpful to observe that for the edges between blocks, i.e. the red edges in Figure 7, we have $(W_{2i-1}, W_{2i}) = (W_{f(2i)}, W_{2i})$, by the definition of f .

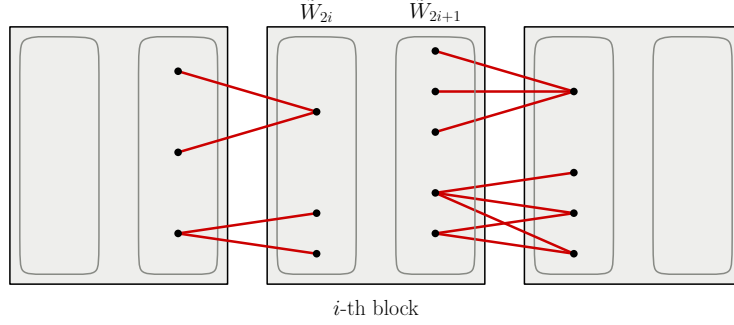


Figure 7: If i -th block is expanding, then the right set of red edges is larger than the left set of red edges. (Note: this expansion does not follow from the definition, but requires some proofs.)

The intuition behind expanding blocks is that the number of edges in F ‘after’ an expanding block should be larger than the number of edges in F ‘before’ it (see Figure 7). The following claim makes this precise.

Claim 7.7. Let F be a projection graph and $1 \leq i \leq m/2 - 2$.

(i) Suppose $e_F(W_{2i-1}, W_{2i}) \geq \varepsilon^{-21} \tilde{n} \log^2 n$. If the i -th block is $(1/\varepsilon)$ -expanding, then

$$e_F(W_{2i+1}, W_{2i+2}) \geq (1/\sqrt{\varepsilon}) \cdot e_F(W_{2i-1}, W_{2i}),$$

and otherwise

$$e_F(W_{2i+1}, W_{2i+2}) \geq \frac{\alpha^2}{256} \tilde{n}^2 p.$$

(ii) Suppose $e_F(W_{2i-1}, W_{2i}) \geq 2\varepsilon \tilde{n}^2 p$. If the i -th block is $(1 + 4\sqrt{\varepsilon})$ -expanding, then

$$e_F(W_{2i+1}, W_{2i+2}) \geq (1 + 2\sqrt{\varepsilon}) \cdot e_F(W_{2i-1}, W_{2i}),$$

and otherwise

$$e_F(W_{2i+3}, W_{2i+4}) \geq (2/3 + \alpha/2) \tilde{n}^2 p.$$

The proof of the claim is quite technical and relies mostly on properties given by Lemma 2.9 about expansion of edges and triangles; we defer it to the next section. With this claim at hand we are ready to give the proof of Claim 7.4.

Proof of Claim 7.4. As mentioned earlier, we assume w.l.o.g. that π is the identity permutation as the actual choice of π does not play a role in the proof. The proof comprises of four natural steps:

- (1) starting from the edges in $\mathcal{I}$ show that $e_F(W_{2i-1}, W_{2i}) \geq \varepsilon^{-21} \tilde{n} \log^2 n$, for all $m/80 \leq i \leq m/2$;

- (2) starting from the edges obtained in the previous step show that $e_F(W_{2i-1}, W_{2i}) \geq 2\varepsilon\tilde{n}^2p$, for all $m/40 \leq i \leq m/2$;
- (3) knowing that starting from all edges in $\mathcal{I}$ we can reach $2\varepsilon\tilde{n}^2p$ edges in some number of steps, show that there is at least one edge $e \in \mathcal{I}$ such that $e_{F'}(W_{2i-1}, W_{2i}) \geq 2\varepsilon\tilde{n}^2p$, for all $m/10 \leq i \leq m/2$, where F' is the e -projection graph;
- (4) starting from the edges obtained in the previous step show that $e_{F'}(W_{2i-1}, W_{2i}) \geq (2/3)\tilde{n}^2p$, for all $m/5 \leq i \leq m/2$.

Step (1). Let us first deal with the trivial case in which $p > \varepsilon^2$. For $j \geq 0$ and a vertex $v \in W_j$ with $\deg_F(v, W_{f(j)}) \geq 1$, by (G5) and as $|X_{j+1}| \leq \varepsilon^5\tilde{n}$ (recall, (W3)), it holds that

$$\deg_F(v, W_{j+1}) \geq \alpha\tilde{n}p^2 - |X_{j+1}| \geq \alpha\varepsilon^4\tilde{n} - \varepsilon^5\tilde{n} \geq \varepsilon^5\tilde{n},$$

and thus also $e_F(W_{2i-1}, W_{2i}) \geq \varepsilon^5\tilde{n}^2 \gg \tilde{n} \log^2 n$, for all $i \geq 1$.

In the remainder of the proof we assume $p \leq \varepsilon^2$. We first show by induction that the following invariant is true for every $j \in \{0, 1, \dots, m\}$: there exists $E_j \subseteq E_F(W_{f(j)}, W_j)$ such that

- (i) $|E_j \cap W_j| \geq \min\{\varepsilon^2\tilde{n}, |\mathcal{I}|(1/\varepsilon)^j\}$ and $|E_j \cap W_{f(j)}| \geq (1 - j\varepsilon + f(j)\varepsilon) \min\{\varepsilon^2\tilde{n}, |\mathcal{I}|(1/\varepsilon)^{f(j)}\}$, where $E_j \cap W_j$ and $E_j \cap W_{f(j)}$ denote the set of vertices from W_j and $W_{f(j)}$ incident to the edges in E_j , respectively,

- (ii) for each $v \in W_{f(j)} \cup W_j$ we have $\deg_{E_j}(v) \leq \varepsilon/p$.

The base of the induction $j = 0$ holds trivially as, by assumption, the starting set of edges is a matching of size $|\mathcal{I}|$. Thus, let $j \geq 1$ and let us assume the hypothesis holds for all values smaller than j . Let $E_{j-1} \subseteq E_F(W_{f(j-1)}, W_{j-1})$ be as given by the induction hypothesis for $j - 1$. Consider first the case $f(j) = j - 2$. Note that then $f(j - 1) = j - 2$ as well and thus $|E_{j-1} \cap W_{f(j-1)}| \geq (1 - \varepsilon) \min\{\varepsilon^2\tilde{n}, |\mathcal{I}|(1/\varepsilon)^{j-2}\}$. We apply Lemma 2.9 (1) with W_{j-1}, W_{j-2}, W_j (as W_1, W_2, W_3) and E_{j-1} (as F_{12}) to conclude that there exists a subset $U' \subseteq E_{j-1} \cap W_{j-2}$ of size

$$|U'| \geq (1 - \varepsilon) \min\{|E_{j-1} \cap W_{j-2}|, \varepsilon/p^2\} \geq (1 - 2\varepsilon) \min\{\varepsilon^2\tilde{n}, |\mathcal{I}|(1/\varepsilon)^{j-2}, \varepsilon/p^2\}$$

and an $(\alpha\tilde{n}p^2/2)$ -star-matching M saturating U' . We may indeed apply the lemma by (ii) and since

$$|X_j| \leq |X| \leq 4t \log n \leq 12|\mathcal{I}| \log n.$$

As $\alpha\tilde{n}p^2/2 \geq 1/\varepsilon^4$, $E_j := M$ satisfies all the required properties. In case $f(j) = j - 1$ we apply Lemma 2.9 (1) with W_{j-2}, W_{j-1}, W_j (as W_1, W_2, W_3). Observing that $|E_{j-1} \cap W_{j-1}| \geq \min\{\varepsilon^2\tilde{n}, |\mathcal{I}|(1/\varepsilon)^{j-1}\}$, and doing the same analysis as in the previous case shows that the invariant holds also in this case.

With these preparations at hand we can now finish the proof. For $j \geq \frac{m}{200} \geq \frac{\log n}{200}$, (i) implies

$$|E_j \cap W_j| \geq (1 - 2\varepsilon) \min\{\varepsilon^2\tilde{n}, |\mathcal{I}|(1/\varepsilon)^{\log n/200}\} \geq \varepsilon^3\tilde{n},$$

as ε is chosen in order for $\log(1/\varepsilon) > 200$ to hold. Next, for every $\frac{m}{100} \leq i \leq \frac{m}{2}$, we apply Lemma 2.9 (1) with $W_{f(2i-1)}, W_{2i-1}, W_{2i}$ (as W_1, W_2, W_3) and E_{2i-1} (as F_{12}) to conclude that for every subset $U \subseteq E_{2i-1} \cap W_{2i-1}$ of size $|U| = \min\{\varepsilon/p^2, |E_{2i-1} \cap W_{2i-1}|\}$ there exists a subset $U' \subseteq U$ of size $(1 - \varepsilon)|U|$ and an $(\alpha\tilde{n}p^2/2)$ -star-matching M saturating U' (note that degree assumption needed for Lemma 2.9 (1) holds by (ii)). This further implies

$$e_F(W_{2i-1}, W_{2i}) \geq (1 - \varepsilon)|E_{2i-1} \cap W_{2i-1}| \geq (1 - \varepsilon)\varepsilon^3\tilde{n} \cdot (\alpha/2)\tilde{n}p^2 \geq \varepsilon^4\tilde{n}^2p^2 \geq \varepsilon^{-21}\tilde{n} \log^2 n.$$

Step (2). Fix $i = m/80$. From Step (1) we know that $e_F(W_{2i-1}, W_{2i}) \geq \varepsilon^{-21} \tilde{n} \log^2 n$. From Claim 7.7 (i), and the fact that $(\alpha^2/256) \tilde{n}^2 p \geq \varepsilon^{-21} \tilde{n} \log^2 n$, we further have that for all $j \geq i + 1$

$$e_F(W_{2j-1}, W_{2j}) \geq \min \left\{ \varepsilon^{-(j-i)/2} \cdot e_F(W_{2i-1}, W_{2i}), \frac{\alpha^2}{256} \tilde{n}^2 p \right\}.$$

Recall that $m \geq \log n$ and that we have chosen ε sufficiently small so that $\varepsilon^{-(j-i)/2} \geq n^2$ for all $j \geq i + m/80$. The assertion in Step (2) then follows as $\alpha^2/256 \geq 2\varepsilon$.

Step (3). Fix $i = m/40$. From Step (2) we know that $e_F(W_{2i-1}, W_{2i}) \geq 2\varepsilon \tilde{n}^2 p$. Let $\mathcal{I} = \mathcal{I}_1 \cup \dots \cup \mathcal{I}_{\lceil \tilde{n}^{1/3} \rceil}$ be an arbitrary partition of $\mathcal{I}$ such that $|\mathcal{I}_j| \leq |\mathcal{I}|/\lceil \tilde{n}^{1/3} \rceil$ and let F_j be the $\mathcal{I}_j$ -projection graph, for every $j \in \{1, \dots, \lceil \tilde{n}^{1/3} \rceil\}$. Since

$$\sum_{j=1}^{\lceil \tilde{n}^{1/3} \rceil} e_{F_j}(W_{2i-1}, W_{2i}) \geq e_F(W_{2i-1}, W_{2i}),$$

there must be a $j^* \in \{1, \dots, \lceil \tilde{n}^{1/3} \rceil\}$ such that

$$e_{F_{j^*}}(W_{2i-1}, W_{2i}) \geq \frac{2\varepsilon \tilde{n}^2 p}{\lceil \tilde{n}^{1/3} \rceil} \geq \varepsilon^{-21} \tilde{n} \log^2 n,$$

where the second inequality holds as $\tilde{n}^{2/3} p \geq \varepsilon^{-21} \log^2 n$. By the same argument as in the Step (2), this time for F_{j^*} and i , we get that for $k = i + \frac{m}{80}$

$$e_{F_{j^*}}(W_{2k-1}, W_{2k}) \geq (\alpha^2/256) \tilde{n}^2 p,$$

implying that we can repeat the argument from above and partition $\mathcal{I}_{j^*}$ into $\lceil \tilde{n}^{1/3} \rceil$ parts. The claim follows by applying this argument successively at most two more times.

Step (4). Fix $i = m/10$ and let $e \in \mathcal{I}$ be the edge obtained in Step (3) and F' the e -projection graph. From Step (3) we know that $e_{F'}(W_{2i-1}, W_{2i}) \geq (\alpha^2/256) \tilde{n}^2 p$. Let us choose a constant $L := L(\alpha, \varepsilon)$ such that $(1 + 2\sqrt{\varepsilon})^{L-2} (\alpha^2/256) > 2$. Observe that not all blocks $i, \dots, i + L - 2$, can be $(1 + 4\sqrt{\varepsilon})$ -expanding as then Claim 7.7 (ii) would imply

$$e_{F'}(W_{2(i+L-2)-1}, W_{2(i+L-2)}) \geq (1 + 2\sqrt{\varepsilon})^{L-2} \cdot \frac{\alpha^2}{256} \tilde{n}^2 p \geq 2\tilde{n}^2 p,$$

which is a contradiction with (G1). Let $i^* \in [i, i + L - 2]$ be the smallest index such that the i^* -th block is $(1 + 4\sqrt{\varepsilon})$ -non-expanding. Then $e_{F'}(W_{2i^*+3}, W_{2i^*+4}) \geq (2/3 + \alpha/2) \tilde{n}^2 p$, by Claim 7.7 (ii).

If the $(i^* + 2)$ -nd block is $(1 + 4\sqrt{\varepsilon})$ -expanding then

$$e_{F'}(W_{2i^*+5}, W_{2i^*+6}) \geq (1 + 2\sqrt{\varepsilon}) \cdot e_{F'}(W_{2i^*+3}, W_{2i^*+4}) \geq (2/3 + \alpha/2) \tilde{n}^2 p.$$

If, on the other hand, the $(i^* + 2)$ -nd block is $(1 + 4\sqrt{\varepsilon})$ -non-expanding, then $e_{F'}(W_{2i^*+7}, W_{2i^*+8}) \geq (2/3 + \alpha/2) \tilde{n}^2 p$, by Claim 7.7 (ii). In addition, by applying Lemma 2.9 (9) to $W_{2i^*+3}, W_{2i^*+4}, W_{2i^*+5}$ (as W_1, W_2, W_3) and $E_{F'}(W_{2i^*+3}, W_{2i^*+4})$ (as F_{12}) we obtain by symmetry that

$$e_{F'}(W_{2i^*+3}, W_{2i^*+5}), e_{F'}(W_{2i^*+4}, W_{2i^*+5}) \geq (1 - \sqrt{\varepsilon}) e_{F'}(W_{2i^*+3}, W_{2i^*+4}).$$

Applying Lemma 2.9 (9) again, this time to $W_{f(2i^*+5)}, W_{2i^*+5}, W_{2i^*+6}$ (as W_1, W_2, W_3) and $E_{F'}(W_{f(2i^*+5)}, W_{2i^*+5})$ (as F_{12}), we obtain

$$e_{F'}(W_{2i^*+5}, W_{2i^*+6}) \geq (1 - \sqrt{\varepsilon})^2 e_{F'}(W_{2i^*+3}, W_{2i^*+4}) \geq (2/3) \tilde{n}^2 p,$$

by our choice of ε .

Repeating this argument, that is starting from $E_{F'}(W_{2i^*+5}, W_{2i^*+6})$ or $E_{F'}(W_{2i^*+7}, W_{2i^*+8})$ depending on whether the $(i^* + 2)$ -nd block was $(1 + 4\sqrt{\varepsilon})$ -expanding or not, shows that $e_{F'}(W_{2j-1}, W_{2j}) \geq (2/3) \tilde{n}^2 p$ for all $j \in [i^* + 2, m/2]$, and the claim follows since $i^* \leq m/5 - 2$ (recall that we had set $m = 4 \log n - 4$ and thus $m/2$ is even).

This completes the proof of Claim 7.4. $\square$

7.3 Proof of Claim 7.7

In this section we provide the proof of the assertions in Claim 7.7. We start with some general remarks. Recall that the i -th block consists of steps $2i$ and $2i + 1$. Recall also that step $2i$ extends edges between sets W_{2i-1} and W_{2i} into set $W_{2i+1} \setminus X_{2i+1}$ via triangles. Similarly, step $2i + 1$ extends edges between sets $W_{f(2i+1)}$ and W_{2i+1} into set $W_{2i+2} \setminus X_{2i+2}$. These extensions exactly mimic the setting covered by Lemma 2.9. However, the actual set $W_{f(2i+1)}$ depends on the value of b of the b -pseudo-path that we want to construct. In order to hide this difference we often use variables $t_1, \dots, t_4$ as follows: $t_4 = 2i + 2$, $t_3 = 2i + 1$, $t_2 = f(2i + 1)$, and t_1 is the unique element from $\{2i, 2i - 1\} \setminus \{t_2\}$. One easily checks that this implies that we can always apply Lemma 2.9 with $W_{t_1}, W_{t_2}, W_{t_3}$ to address the $2i$ -th step and to $W_{t_2}, W_{t_3}, W_{t_4}$ (as W_1, W_2, W_3) to address the $(2i + 1)$ -st step.

Proof of (i) in Claim 7.7: i -th block is $(1/\varepsilon)$ -expanding. If both steps $2i$ and $2i + 1$ are $(1/\varepsilon)$ -expanding then the claim follows directly from the definition of an expanding step together with the observation that $f(j) = j - 1$ whenever j is even. Thus, let us assume that one of the two steps is $(1/\varepsilon)$ -non-expanding. We aim to prove the following for every $j \geq 1$:

$$\text{If } e_F(W_{f(j)}, W_j) \geq \varepsilon^{-21} \tilde{n} \log^2 n \text{ then } e_F(W_{f(j+1)}, W_{j+1}) \geq \frac{\alpha}{16} e_F(W_{f(j)}, W_j). \quad (32)$$

Note that this implies the claim regardless of whether the non-expanding step is the first or the second step within the block, as then

$$e_F(W_{2i+1}, W_{2i+2}) \geq \frac{1}{\varepsilon} \cdot \frac{\alpha}{16} e_F(W_{2i-1}, W_{2i}) \geq \frac{1}{\sqrt{\varepsilon}} \cdot e_F(W_{2i-1}, W_{2i}),$$

which is what we wanted to prove.

We now prove (32). So assume $e_F(W_{f(j)}, W_j) \geq \varepsilon^{-21} \tilde{n} \log^2 n$, for some $j \geq 1$. Observe that the definition of the function f implies that (32) involves exactly three sets W_j . Indeed, by the definition of f we have $f(j + 1) \in \{j, f(j)\}$. Let $t_2 = f(j + 1)$ and t_1 be the unique element from $\{j, f(j)\} \setminus \{t_2\}$. Let $S \subseteq W_{t_2}$ be the set of vertices with the degree at most $\varepsilon^{-4} \log n/p$ in $E_F(W_{t_1}, W_{t_2})$ and set $M := W_{t_2} \setminus S$. Furthermore, let us denote $E_F(W_{t_1}, S)$ and $E_F(W_{t_1}, M)$ by I_S and I_M , respectively. If $|I_S| \geq e_F(W_{t_1}, W_{t_2})/2$ then by applying Lemma 2.9 (2) with $W_{t_1}, W_{t_2}, W_{j+1}$ (as W_1, W_2, W_3), S (as U), and I_S (as F_{12}) we get

$$e_F(S, W_{j+1}) \geq \frac{1}{\varepsilon^4} e_F(W_{t_1}, S) = \frac{1}{\varepsilon^4} |I_S| \geq \frac{1}{2\varepsilon^4} e_F(W_{f(j)}, W_j) \geq \frac{\alpha}{16} e_F(W_{f(j)}, W_j).$$

On the other hand, if $|I_M| \geq e_F(W_{t_1}, W_{t_2})/2$, by applying Lemma 2.9 (3) with $W_{t_1}, W_{t_2}, W_{j+1}$ (as W_1, W_2, W_3), M (as U), and I_M (as F_{12}) we get

$$e_F(M, W_{j+1}) \geq \frac{\alpha \tilde{n} p}{4} |M| \geq \sum_{v \in M} \frac{\alpha}{8} \deg_{I_M}(v, W_{t_1}) = \frac{\alpha}{8} |I_M| \geq \frac{\alpha}{16} e_F(W_{f(j)}, W_j),$$

where the second inequality follows from (G2). $\square$

Proof of (i) in Claim 7.7: i -th block is $(1/\varepsilon)$ -non-expanding. Set $t_4 = 2i + 2$, $t_3 = 2i + 1$, $t_2 = f(2i + 1)$, and let t_1 be the unique element from $\{2i, 2i - 1\} \setminus \{t_2\}$.

By assumption, the steps $2i$ and $2i + 1$ are both $(1/\varepsilon)$ -non-expanding. Thus, in particular $e_F(W_{t_2}, W_{t_3}) < (1/\varepsilon) e_F(W_{t_1}, W_{t_2})$. Additionally, since $e_F(W_{t_1}, W_{t_2}) \geq \varepsilon^{-21} \tilde{n} \log^2 n$, we can apply Lemma 2.9 (7) with $W_{t_1}, W_{t_2}, W_{t_3}$ (as W_1, W_2, W_3), and $E_F(W_{t_1}, W_{t_2})$ (as F_{12}) to get a set $L_2 \subseteq W_{t_2}$ such that for every $v \in L_2$ we have $\deg_F(v, W_{t_1}) \geq \varepsilon^{-4} \log n/p$ and

$$e_F(W_{t_1}, L_2) \geq (1 - \varepsilon) e_F(W_{t_1}, W_{t_2}). \quad (33)$$

The previous inequality implies $e_F(W_{t_1}, L_2) \geq \varepsilon^{-20} \tilde{n} \log n$ and thus by applying Lemma 2.9 (3) with $W_{t_1}, W_{t_2}, W_{t_3}$ (as W_1, W_2, W_3), L_2 (as U), and $E_F(W_{t_1}, L_2)$ (as F_{12}) we conclude

$$e_F(L_2, W_{t_3}) \geq \frac{\alpha \tilde{n} p}{4} |L_2| \stackrel{(G2)}{\geq} \frac{\alpha}{8} e_F(W_{t_1}, L_2) \stackrel{(33)}{\geq} \frac{\alpha}{10} e_F(W_{t_1}, W_{t_2}) \quad (34)$$

Since the i -th block is non-expanding we have $e_F(W_{t_3}, W_{t_4}) \leq (1/\varepsilon^2) e_F(W_{t_1}, W_{t_2})$ and thus

$$e_F(W_{t_3}, W_{t_4}) < \frac{1}{\varepsilon^2} e_F(W_{t_1}, W_{t_2}) \stackrel{(34)}{\leq} \frac{1}{\varepsilon^2} \frac{10}{\alpha} e_F(L_2, W_{t_3}) < \frac{1}{\varepsilon^3} e_F(L_2, W_{t_3}).$$

Hence, we can apply Lemma 2.9 (7) with $W_{t_2}, W_{t_3}, W_{t_4}$ (as W_1, W_2, W_3), and $E_F(W_{t_2}, W_{t_3})$ (as F_{12}) to get a set $L_3 \subseteq W_{t_3}$ such that for every $v \in L_3$ we have $\deg_F(v, W_{t_2}) \geq \varepsilon^{-4} \log n/p$ and

$$e_F(W_{t_2}, L_3) \geq (1 - \varepsilon) e_F(W_{t_2}, W_{t_3}) \geq (1 - \varepsilon) e_F(L_2, W_{t_3}). \quad (35)$$

Note that w.l.o.g. we may assume that L_3 contains all vertices $v \in W_{t_3}$ with $\deg_F(v, W_{t_2}) \geq \varepsilon^{-4} \log n/p$. As (34) and (35) imply $e_F(W_{t_2}, L_3) \geq \varepsilon^{-19} \tilde{n} \log n$, we can apply Lemma 2.9 (3) with L_3 (as U) to obtain

$$e_F(W_{t_3}, W_{t_4}) \geq e_F(L_3, W_{t_4}) \geq \frac{\alpha \tilde{n} p}{4} |L_3|.$$

If $|L_3| \geq (\alpha/32) \tilde{n}$ then we are done. In the remainder of the proof we show that such an assumption is actually true. Towards a contradiction assume that $|L_3| < (\alpha/32) \tilde{n}$. Observe that (recall, (W1) for W_{t_1} in order to apply (G2))

$$|L_2| \stackrel{(G2)}{\geq} \frac{e_F(W_{t_1}, L_2)}{(1 + \varepsilon) \tilde{n} p} \stackrel{(33)}{\geq} \frac{(1 - \varepsilon) e_F(W_{t_1}, W_{t_2})}{(1 + \varepsilon) \tilde{n} p} \geq \frac{(1 - \varepsilon) \varepsilon^{-21} \log^2 n}{(1 + \varepsilon) p} \geq \frac{\varepsilon^{-19} \log^2 n}{p}. \quad (36)$$

Let now $S \subseteq Y \subseteq W_{t_3}$ be sets defined as

$$Y := \{v \in W_{t_3} : \deg_F(v, L_2) > \varepsilon |L_2| p\} \quad \text{and} \quad S := Y \setminus L_3.$$

If $|Y| \geq (\alpha/16) \tilde{n}$, then $|S| \geq (\alpha/32) \tilde{n}$, as we assumed $|L_3| < (\alpha/32) \tilde{n}$, and thus

$$e_F(W_2, S) \geq e_F(L_2, S) \geq |S| \cdot \varepsilon |L_2| p \stackrel{(36)}{\geq} \frac{\varepsilon \alpha}{32} \tilde{n} \cdot \varepsilon^{-19} \log^2 n \geq \varepsilon^{-17} \tilde{n} \log^2 n.$$

Recall that all vertices in $W_{t_3} \setminus L_3$ have degree in F at most $\varepsilon^{-4} \log n/p$ into W_{t_2} . Therefore, Lemma 2.9 (2) applied with $W_{t_2}, W_{t_3}, W_{t_4}$ (as W_1, W_2, W_3), S (as U), and $E_F(W_2, S)$ (as F_{12}) shows

$$e_F(S, W_{t_4}) \geq \varepsilon^{-4} e_F(W_{t_2}, S) \geq \varepsilon^{-3} |S| |L_2| p \stackrel{(36)}{\geq} \varepsilon^{-3} \cdot \frac{\alpha}{16} (1 - 2\varepsilon) e_F(W_{t_1}, W_{t_2}) > \varepsilon^{-2} e_F(W_{t_1}, W_{t_2}),$$

which is a contradiction with our assumption that the i -th block is $(1/\varepsilon)$ -non-expanding. Therefore, $|Y| < (\alpha/16) \tilde{n}$. However, as by (G1) and (36) we know that there are at most $\varepsilon^{-3} \log n/p$ vertices $v \in W_{t_3}$ with $\deg_G(v, L_2) \geq 2|L_2|p$, we then get

$$e_F(L_2, W_{t_3}) \leq |Y| \cdot 2|L_2|p + \frac{\varepsilon^{-3} \log n}{p} \cdot 2\tilde{n}p + \tilde{n} \cdot \varepsilon |L_2|p \leq \left(\frac{\alpha}{8} + \varepsilon + \varepsilon \right) |L_2| \tilde{n} p < \frac{\alpha \tilde{n} p}{4} |L_2|,$$

which is a contradiction with the first inequality in (34). We conclude $|L_3| \geq (\alpha/32) \tilde{n}$ and the claim follows. $\square$

Proof of (ii) in Claim 7.7: i -th block is $(1 + 4\sqrt{\varepsilon})$ -expanding. Note that if both steps $2i$ and $2i + 1$ are $(1 + 4\sqrt{\varepsilon})$ -expanding, then the statement follows directly from the definition of an expanding step. Thus, let us assume one of the two steps is $(1 + 4\sqrt{\varepsilon})$ -non-expanding and let us denote

that step with $t \in \{2i, 2i + 1\}$. Furthermore, let $t_2 = f(t + 1)$ and let t_1 be the unique element from $\{f(t), t\} \setminus \{t_2\}$. By applying Lemma 2.9 (9) with $W_{t_1}, W_{t_2}, W_{t+1}$ (as W_1, W_2, W_3) and $E_F(W_{t_1}, W_{t_2})$ (as F_{12}) we get $e_F(W_{f(t+1)}, W_{t+1}) \geq (1 - \sqrt{\varepsilon})e_F(W_{f(t)}, W_t)$. From here we conclude

$$e_F(W_{2i+1}, W_{2i+2}) \geq (1 + 4\sqrt{\varepsilon})(1 - \sqrt{\varepsilon})e_F(W_{2i-1}, W_{2i}) \geq (1 + 2\sqrt{\varepsilon})e_F(W_{2i-1}, W_{2i}),$$

which is what we wanted to prove. $\square$

Proof of (ii) in Claim 7.7: i -th block is $(1 + 4\sqrt{\varepsilon})$ -non-expanding. Set $t_4 = 2i + 2$, $t_3 = 2i + 1$, $t_2 = f(2i + 1)$, and let t_1 be the unique element from $\{2i, 2i - 1\} \setminus \{t_2\}$. Let us define L_2 and L_3 as

$$L_2 = \{v \in W_{t_2} : \deg_F(v, W_{t_1}) > \tilde{n}p/3\} \quad \text{and} \quad L_3 = \{v \in W_{t_3} : \deg_F(v, W_{t_2}) > \tilde{n}p/3\}.$$

Our first goal is to show $|L_3| \geq (2/3 + \alpha/4)\tilde{n}$. As the i -th block is $(1 + 4\sqrt{\varepsilon})$ -non-expanding, we know that $e_F(W_{t_2}, W_{t_3}) < (1 + 4\sqrt{\varepsilon})e_F(W_{t_1}, W_{t_2})$ and thus by Lemma 2.9 (8) applied with $W_{t_1}, W_{t_2}, W_{t_3}$ (as W_1, W_2, W_3), $E_F(W_{t_1}, W_{t_2})$ (as F_{12}), and $32\sqrt{\varepsilon}/\alpha$ (as μ) we conclude

$$e_F(W_{t_1}, L_2) \geq (1 - 32\sqrt{\varepsilon}/\alpha)e_F(W_{t_1}, W_{t_2}) \geq (1 - \varepsilon^{1/3})e_F(W_{t_1}, W_{t_2}). \quad (37)$$

Next, by Lemma 2.9 (5) with $W_{t_1}, W_{t_2}, W_{t_3}$ (as W_1, W_2, W_3), L_2 (as U), and $E_F(W_{t_1}, L_2)$ (as F_{12}) we get

$$e_F(L_2, W_{t_3}) \geq (1 - \varepsilon^2)e_G(L_2, W_{t_3}) \stackrel{\text{(G3)}}{\geq} (1 - 3\varepsilon)e_G(W_{t_1}, L_2) \geq (1 - 3\varepsilon)e_F(W_{t_1}, L_2). \quad (38)$$

This together with (37) implies

$$e_F(W_{t_2}, W_{t_3}) \geq e_F(L_2, W_{t_3}) \geq (1 - 3\varepsilon)e_F(W_{t_1}, L_2) \geq (1 - 2\varepsilon^{1/3})e_F(W_{t_1}, W_{t_2}). \quad (39)$$

Once again using the fact that the i -th block is $(1 + 4\sqrt{\varepsilon})$ -non-expanding, we get

$$\begin{aligned} e_F(W_{t_3}, W_{t_4}) &< (1 + 4\sqrt{\varepsilon})^2 e_F(W_{t_1}, W_{t_2}) \leq (1 + 4\sqrt{\varepsilon})^2 (1 - 2\varepsilon^{1/3})^{-1} e_F(L_2, W_{t_3}) \\ &\leq (1 + 4\varepsilon^{1/3})e_F(L_2, W_{t_3}). \end{aligned} \quad (40)$$

Next, we apply Lemma 2.9 (8) with $W_{t_2}, W_{t_3}, W_{t_4}$ (as W_1, W_2, W_3), $E_F(L_2, W_{t_3})$ (as F_{12}), and $32\varepsilon^{1/3}/\alpha$ (as μ). We can do that since by (39) and (40) we know that $e_F(L_2, W_{t_3}) \geq \varepsilon\tilde{n}^2p$ and $e_F(W_{t_3}, W_{t_4}) < (1 + \mu\alpha/8)e_F(L_2, W_{t_3})$. Therefore, Lemma 2.9 (8) and (38) imply

$$e_F(L_2, L_3) \geq \left(1 - \frac{32\varepsilon^{1/3}}{\alpha}\right)e_F(L_2, W_{t_3}) \geq (1 - \varepsilon^{1/4})e_F(L_2, W_{t_3}) \stackrel{\text{(38)}}{\geq} (1 - 2\varepsilon^{1/4})e_G(L_2, W_{t_3}). \quad (41)$$

Furthermore, (41) and (G4) show

$$e_F(L_2, L_3) \geq (1 - 2\varepsilon^{1/4})e_G(L_2, W_{t_3}) \geq (1 - 2\varepsilon^{1/4}) \cdot |L_2|(2/3 + \alpha)\tilde{n}p \geq (2/3 + \alpha/2)|L_2|\tilde{n}p.$$

From (G1), the fact that $|L_2| \geq (\varepsilon/2)\tilde{n}$ (follows from (37) and (G2)), and $e_F(L_2, L_3) \geq (2/3 + \alpha/2)|L_2|\tilde{n}p$, we obtain $|L_3| \geq (2/3 + \alpha/4)\tilde{n}$.

Next, we define

$$\mathcal{L} = \{v \in W_{2i+1} : \deg_F(v, W_{2i+2}) > \tilde{n}p/3\} \quad \text{and} \quad \mathcal{L}' = \{v \in W_{2i+2} : \deg_F(v, W_{2i+1}) > \tilde{n}p/3\}.$$

We aim to show that $|\mathcal{L}|, |\mathcal{L}'| \geq (2/3)\tilde{n}$. Since $|L_3| \geq (2/3 + \alpha/4)\tilde{n}$, we can apply Lemma 2.9 (6) with $W_{f(2i+1)}, W_{2i+1}, W_{2i+2}$ (as W_1, W_2, W_3), and L_3 (as U) to conclude that $|\mathcal{L}'| \geq (1 - \varepsilon)\tilde{n}$, as

desired. Similarly, applying Lemma 2.9 (5) with $W_{f(2i+1)}, W_{2i+1}, W_{2i+2}$ (as W_1, W_2, W_3), L_3 (as U), and $E_F(W_{t_2}, L_3)$ (as F_{12}) we get that there exists a $L'_3 \subseteq L_3$ of size $|L'_3| \geq (1 - 3\varepsilon^3)|L_3|$ such that for all $v \in L'_3$ we have $\deg_F(v, W_{2i+2}) \geq (2/3 + \alpha/2)\tilde{n}p$. Clearly $L'_3 \subseteq \mathcal{L}$ and $|L'_3| \geq (2/3)\tilde{n}$.

Set $r_5 = 2i + 3$, $r_4 = f(2i + 3)$, and let r_3 be the unique element from $\{2i + 2, 2i + 1\} \setminus \{r_4\}$. Moreover, let L_4 and L_5 be defined as

$$L_4 = \{v \in W_{r_4} : \deg_F(v, W_{r_3}) > \tilde{n}p/3\} \quad \text{and} \quad L_5 = \{v \in W_{r_5} : \deg_F(v, W_{r_4}) > \tilde{n}p/3\}.$$

Note that, depending on $f(2i + 3)$, the set L_4 lies either in W_{2i+1} or W_{2i+2} . Thus, since $|\mathcal{L}|, |\mathcal{L}'| \geq (2/3)\tilde{n}$ and $L_4 \in \{\mathcal{L}, \mathcal{L}'\}$, we have $|L_4| \geq (2/3)\tilde{n}$ as well.

Having this at hand we can finally show $e_F(W_{2i+3}, W_{2i+4}) \geq (2/3 + \alpha/2)\tilde{n}^2p$. By applying Lemma 2.9 (6) with $W_{r_3}, W_{r_4}, W_{r_5}$ (as W_1, W_2, W_3), and L_4 (as U) we get that $|L_5| \geq (1 - \varepsilon)\tilde{n}$. Finally, we apply Lemma 2.9 (5) with $W_{r_4}, W_{r_5}, W_{2i+4}$ (as W_1, W_2, W_3), L_5 (as U), and $E_F(W_{r_4}, L_5)$ (as F_{12}) to obtain

$$e_F(W_{2i+3}, W_{2i+4}) \geq (1 - \varepsilon^2)e_G(L_5, W_{2i+4}) \stackrel{(G3)}{\geq} (1 - \varepsilon^2) \cdot |L_5|(2/3 + \alpha)\tilde{n}p \geq (2/3 + \alpha/2)\tilde{n}^2p.$$

This concludes the proof of Claim 7.7. $\square$

8 Concluding remarks

In this paper we introduce the notion of H -resilience which measures the fraction of H -copies touching a given vertex that an adversary may delete without destroying a certain given property. We demonstrate the usefulness of the definition by showing that the K_3 -resilience of $G_{n,p}$ w.r.t. the containment of the square of a Hamilton cycle is w.h.p. $5/9 \pm o(1)$. In other words, the adversary needs to delete more than a $(5/9)$ -fraction of the triangles lying on a vertex in order to destroy all copies of C_n^2 in $G_{n,p}$. Our result is optimal with respect to the constant $5/9$ and the density p up to logarithmic factors.

Having the notion of H -resilience at hand, one can ask for similar statements for other (spanning) graph properties. Of particular interest is the question of the K_3 -resilience of $G_{n,p}$ with respect to the containment of a triangle factor. Theorem 3.1 shows that also here the resilience is at most $5/9 + o(1)$. Moreover, as C_n^2 contains a triangle factor, provided $3 \mid n$, it follows that this is the correct one whenever $p \gg n^{-1/2} \log^3 n$. However, the threshold for the appearance of a triangle factor is significantly lower than the threshold for the appearance of a C_n^2 , cf. the seminal result of Johansson, Kahn, and Vu [20]. In light of this, we conjecture that the resilience variant of this result holds when p is close to the threshold for having a K_3 -factor.

An analogous construction as in Theorem 3.1 shows that the K_r -resilience for a K_r -factor is at most $1 - (1 - 1/r)^{r-1}$. It is thus tempting to conjecture that this value is also the K_r -resilience of $G_{n,p}$ w.r.t. containment of a K_r -factor, provided that $p \gg n^{-2/r}(\log n)^{1/e(K_r)}$, as well as C_n^{r-1} , provided that $p \gg n^{-1/r}$. The conjecture is true in the case when $p = 1$, as every graph with $(1 - 1/r)^{r-1} \binom{n}{r-1}$ copies of K_r at each vertex must have a minimum degree of at least $(r - 1)n/r$ and the statement thus follows from the theorem of Hajnal and Szemerédi [16] and Theorem 1.1.

References

- [1] P. Allen, J. Böttcher, J. Ehrenmüller, and A. Taraz. The bandwidth theorem in sparse graphs. *Advances in Combinatorics*, (6):1–60, 2020. 1
- [2] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Hoboken, NJ, fourth edition, 2016. 2, 3

- [3] J. Balogh, B. Csaba, and W. Samotij. Local resilience of almost spanning trees in random graphs. *Random Structures & Algorithms*, 38(1-2):121–139, 2011. [1](#)
- [4] J. Balogh, C. Lee, and W. Samotij. Corrádi and Hajnal’s theorem for sparse random graphs. *Combinatorics, Probability and Computing*, 21(1-2):23–55, 2012. [1](#)
- [5] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 28(3):669–709, 2015. [1](#)
- [6] J. A. Bondy and U. S. R. Murty. *Graph theory*, volume 244 of *Graduate Texts in Mathematics*. Springer, New York, 2008. [2](#)
- [7] J. Böttcher. Large-scale structures in random graphs. In *Surveys in combinatorics 2017*, pages 87–140. Cambridge University Press, Cambridge, 2017. [1](#)
- [8] J. Böttcher, Y. Kohayakawa, and A. Taraz. Almost spanning subgraphs of random graphs after adversarial edge removal. *Combinatorics, Probability and Computing*, 22(5):639–683, 2013. [1](#)
- [9] J. Böttcher, M. Schacht, and A. Taraz. Proof of the bandwidth conjecture of Bollobás and Komlós. *Mathematische Annalen*, 343(1):175–205, 2009. [1](#)
- [10] P. Châu, L. DeBiasio, and H. A. Kierstead. Pósa’s Conjecture for graphs of order at least 2×10^8 . *Random Structures & Algorithms*, 39(4):507–525, 2011. [1](#)
- [11] D. Conlon and W. T. Gowers. Combinatorial theorems in sparse random sets. *Annals of Mathematics. Second Series*, 184(2):367–454, 2016. [1](#)
- [12] D. Conlon, W. T. Gowers, W. Samotij, and M. Schacht. On the KLR conjecture in random graphs. *Israel Journal of Mathematics*, 203(1):535–580, 2014. [1](#)
- [13] G. A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society. Third Series*, 2:69–81, 1952. [1](#)
- [14] P. Erdős. Problem 9. *Theory of Graphs and its Applications (M. Fieldler ed.)*, Czech. Acad. Sci. Publ., Prague, page 159, 1964. [1](#)
- [15] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016. [2](#)
- [16] A. Hajnal and E. Szemerédi. Proof of a conjecture of P. Erdős. In *Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969)*, pages 601–623. North-Holland, Amsterdam, 1970. [1](#), [8](#)
- [17] P. Hall. On representatives of subsets. *Journal of the London Mathematical Society*, 1(1):26–30, 1935. [2](#)
- [18] P. E. Haxell. A condition for matchability in hypergraphs. *Graphs and Combinatorics*, 11(3):245–248, 1995. [2.2](#)
- [19] H. Huang, C. Lee, and B. Sudakov. Bandwidth theorem for random graphs. *Journal of Combinatorial Theory. Series B*, 102(1):14–37, 2012. [1](#)
- [20] A. Johansson, J. Kahn, and V. Vu. Factors in random graphs. *Random Structures & Algorithms*, 33(1):1–28, 2008. [8](#)
- [21] J. Komlós, G. N. Sárközy, and E. Szemerédi. On the Pósa-Seymour conjecture. *Journal of Graph Theory*, 29(3):167–176, 1998. [1](#)
- [22] J. Komlós, G. N. Sárközy, and E. Szemerédi. Proof of the Seymour conjecture for large graphs. *Annals of Combinatorics*, 2(1):43–60, 1998. [1](#), [1.1](#)
- [23] J. Komlós, G. N. Sárközy, and E. Szemerédi. Proof of the Alon-Yuster conjecture. *Discrete Mathematics*, 235(1-3):255–269, 2001. [1](#)
- [24] M. Krivelevich, C. Lee, and B. Sudakov. Resilient pancyclicity of random and pseudorandom graphs. *SIAM Journal on Discrete Mathematics*, 24(1):1–16, 2010. [1](#)
- [25] D. Kühn and D. Osthus. Embedding large subgraphs into dense graphs. In *Surveys in combinatorics 2009*, volume 365 of *London Math. Soc. Lecture Note Ser.*, pages 137–167. Cambridge University Press, Cambridge, 2009. [1](#)
- [26] C. Lee and B. Sudakov. Dirac’s theorem for random graphs. *Random Structures & Algorithms*, 41(3):293–305, 2012. [1](#)
- [27] I. Levitt, G. N. Sárközy, and E. Szemerédi. How to avoid using the regularity Lemma: Pósa’s

- conjecture revisited. *Discrete Mathematics*, 310(3):630–641, 2010. [1](#)
- [28] R. Montgomery. Hamiltonicity in random graphs is born resilient. *Journal of Combinatorial Theory. Series B*, 139:316–341, 2019. [1](#)
- [29] R. Nenadov, A. Steger, and M. Trujić. Resilience of perfect matchings and Hamiltonicity in random graph processes. *Random Structures & Algorithms*, 54(4):797–819, 2019. [1](#)
- [30] R. Nenadov and N. Škorić. Powers of Hamilton cycles in random graphs and tight Hamilton cycles in random hypergraphs. *Random Structures & Algorithms*, 54(1):187–208, 2019. [1](#), [1](#), [4](#), [5](#), [5](#)
- [31] R. Nenadov and N. Škorić. On Komlós’ tiling theorem in random graphs. *Combinatorics, Probability and Computing*, 29(1):113–127, 2020. [5.2](#)
- [32] D. Saxton and A. Thomason. Hypergraph containers. *Inventiones Mathematicae*, 201(3):925–992, 2015. [1](#)
- [33] M. Schacht. Extremal results for random discrete structures. *Annals of Mathematics. Second Series*, 184(2):333–365, 2016. [1](#)
- [34] P. Seymour. Problem section. In *Combinatorics: Proceedings of the British Combinatorial Conference*, pages 201–202, 1973. [1](#)
- [35] N. Škorić, A. Steger, and M. Trujić. Local resilience of an almost spanning k -cycle in random graphs. *Random Structures & Algorithms*, 53(4):728–751, 2018. [5.2](#), [5.5](#)
- [36] B. Sudakov. Robustness of graph properties. In *Surveys in combinatorics 2017*, volume 440 of *London Math. Soc. Lecture Note Ser.*, pages 372–408. Cambridge Univ. Press, Cambridge, 2017. [1](#)
- [37] B. Sudakov and V. H. Vu. Local resilience of graphs. *Random Structures & Algorithms*, 33(4):409–433, 2008. [1](#)
- [38] P. Turán. Eine Extremalaufgabe aus der Graphentheorie. *Matematikai és Fizikai Lapok*, 48:436–452, 1941. [1](#)