Received: 23 March 2018 | Accepted: 16 November 2018 | Published on: 29 April 2019

DOI: 10.1002/rsa.20851

RESEARCH ARTICLE

WILEY

Regularity inheritance in pseudorandom graphs

Peter Allen[1] | **Julia Böttcher**[1] | **Jozef Skokan**[1,2] | **Maya Stein**[3,4]

[1] Department of Mathematics, London School of Economics, London, United Kingdom
[2]Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA
[3]Departamento de Ingeniería Matemática, Universidad de Chile, Santiago, Chile
[4]Centro de Modelamiento Matemático, Universidad de Chile, UMI 2807 CNRS, Santiago, Chile

Correspondence
Julia Böttcher, Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom.
Email: J.boettcher@lse.ac.uk

Funding information
This research was supported by the Fondecyt; 1183080, CONICYT + PIA/Apoyo a centros científicos y tecnológicos de excelencia con financiamiento Basal, Código AFB170001 and by Fondecyt Regular Grant 1183080, EPSRC; EP/P032125/1, EPSRC; EP/R00532X/1, National Science Foundation; DMS-1500121.

Abstract

Advancing the sparse regularity method, we prove one-sided and two-sided regularity inheritance lemmas for subgraphs of bijumbled graphs, improving on results of Conlon, Fox, and Zhao. These inheritance lemmas also imply improved H-counting lemmas for subgraphs of bijumbled graphs, for some H.

KEYWORDS

counting lemma, pseudorandom graphs, regularity inheritance, sparse regularity lemma, Szemerédi's regularity lemma

1 | INTRODUCTION

Over the past 40 years, the regularity method has developed into a powerful tool in discrete mathematics, with applications in combinatorial geometry, additive number theory and theoretical computer science (see [12, 15, 18, 21] for surveys).

The regularity method relies on Szemerédi's celebrated regularity lemma [25] and a corresponding counting lemma. Roughly speaking, the regularity lemma states that each graph can (almost) be partitioned into a bounded number of regular pairs. More precisely, a pair (U, W) of disjoint sets of vertices in a graph G is *ε-regular* if, for all $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$, we have $|d(U', W') - d(U, W)| \leq \varepsilon$, where $d(U, W) := e(U, W)/(|U||W|)$ is the *density* of the pair

(U,W) and $e(U,W)$ is the number of edges between U and W in G. The regularity lemma then says that every graph G has a vertex partition $V_1\dot\cup\dots\dot\cup V_m$ into almost equal-sized sets such that all but at most εm^2 pairs (V_i,V_j) are ε-regular and m is bounded by a function depending on ε but not on G.

The counting lemma complements the regularity lemma and states that in systems of regular pairs the number of copies of any fixed graph H is roughly as predicted by the densities of the regular pairs. In particular, if H is a graph with vertex set $V(H)=[m]:=\{1,\dots,m\}$ and G is an m-partite graph with partition $V_1\dot\cup\dots\dot\cup V_m$ of $V(G)$ such that (V_i,V_j) is ε-regular whenever $ij\in E(H)$, then the number of (labeled) copies of H in G with vertex i in V_i for each $i\in V(H)$ is $\prod_{ij\in E(H)}\big(d(V_i,V_j)\pm\gamma\big)\cdot\prod_{i\in[m]}|V_i|$, as long as ε is sufficiently small.

Such a counting lemma can easily be proved with the help of the fact that neighborhoods in dense regular pairs are large and therefore "inherit" regularity. More precisely, if (X,Y), (Y,Z), and (X,Z) are ε-regular and have density $d\gg\varepsilon$ then for most vertices $x\in X$ it is true that $|N(x)\cap Y|=(d\pm\varepsilon)|Y|$ and $|N(x)\cap Z|=(d\pm\varepsilon)|Z|$. Hence one can easily deduct from ε-regularity that the pair $\big(N(x)\cap Y,Z\big)$ is ε'-regular (this is called *one-sided inheritance*) and the pair $\big(N(x)\cap Y,N(x)\cap Z\big)$ is ε'-regular (this is called *two-sided inheritance*) for some ε'. Using this regularity inheritance, the counting lemma follows by induction on the number of vertices m of H.

For sparse graphs G, that is, G with n vertices and $o(n^2)$ edges, the error term in the definition of ε-regularity is too coarse, and hence the regularity method is, as such, not useful for such graphs. There are, however, sparse analogs of the regularity lemma, which "rescale" the error term and hence are meaningful for sparse graphs.

Definition 1 (sparse regularity)**.** Let $p>0$ and G be a graph. Let $U,W\subseteq V(G)$ be disjoint. The *p-density* of (U,W) is $d_p(U,W):=e(U,W)/(p|U||W|)$. The pair (U,W) is (ε,p)*-regular* if, for all $U'\subseteq U$ and $W'\subseteq W$ with $|U'|\ge\varepsilon|U|$ and $|W'|\ge\varepsilon|W|$, we have

$$\big|d_p(U',W')-d_p(U,W)\big|\le\varepsilon.$$

It is (ε,d,p)-regular if, moreover, $d_p(U,W)\ge d-\varepsilon$. An (ε)-regular pair (U,W) is an (ε,p)-regular pair with density $d(U,W)=p$.

The sparse regularity lemma (see [13,23]) states that the vertex set of any graph can be partitioned into sets, most pairs of which are (ε)-regular. However, a corresponding counting lemma for (ε)-regular pairs is not true in general: one can construct, say, balanced 4-partite graphs such that every pair of parts induces an (ε,d,p)-regular pair with $\varepsilon\ll d$, but which do not contain a single copy of K_4 (see, eg, [7, p. 11]).

Nevertheless, counting lemmas are known for sparse graphs G with additional structural properties. In the case that G is a subgraph of a random graph establishing such a counting lemma was a famous open problem, the so-called KŁR-Conjecture [14], which was settled only recently [4,6,22]. Proving an analogous result for subgraphs G of pseudorandom graphs has been another central problem in the area. The study of pseudorandom graphs was initiated by Thomason [26,27] (see also [19] for more background information on pseudorandom graphs), who considered a notion of pseudorandomness very closely related to that of bijumbledness.

Definition 2 (bijumbled)**.** A pair (U,V) of disjoint sets of vertices in a graph Γ is called (p,γ)*-bijumbled in* Γ if, for all pairs (U',V') with $U'\subseteq U$ and $V'\subseteq V$, we have

$$\big|e(U',V')-p|U'||V'|\big|\le\gamma\sqrt{|U'||V'|}.$$

A graph Γ is said to be (p,γ)-bijumbled if all pairs of disjoint sets of vertices in Γ are (p,γ)-bijumbled in Γ. A bipartite graph Γ with partition classes U and V is (p,γ)-bijumbled if the pair (U,V) is (p,γ)-bijumbled in Γ.

After partial results were obtained in [16], Conlon, Fox, and Zhao [7] recently proved a general counting lemma for subgraphs of bijumbled graphs. This counting lemma has various interesting applications for subgraphs of bijumbled graphs, including a removal lemma, Turán-type results, and Ramsey-type results.

For obtaining counting lemmas for sparse graphs the most straightforward approach is to try to mimic the strategy for the proof of the dense counting lemma outlined above. The main obstacle here is that in sparse graphs it is no longer true that neighborhoods of vertices in regular pairs are typically large and therefore trivially induce regular pairs—they are of size $pn \ll \varepsilon n$. One can overcome this difficulty by establishing that, under certain conditions, typically these sparse neighborhoods nevertheless inherit sparse regularity. Inheritance lemmas of this type were first considered by Gerke, Kohayakawa, Rödl, and Steger [11]. Conlon, Fox, and Zhao [7] proved inheritance lemmas for subgraphs of bijumbled graphs. The main results of the present paper are inheritance lemmas which require weaker bijumbledness conditions. The first result establishes one-sided regularity inheritance.

Lemma 3 (One-sided inheritance lemma). *For each $\varepsilon', d > 0$ there are $\varepsilon, c > 0$ such that for all $0 < p < 1$ the following holds. Let $G \subseteq \Gamma$ be graphs and X, Y, Z be disjoint vertex sets in $V(\Gamma)$. Assume that*

- *(X, Y) is $(p, cp^{3/2}\sqrt{|X||Y|})$-bijumbled in Γ,*
- *(Y, Z) is $\big(p, cp^2(\log_2 \frac{1}{p})^{-1/2}\sqrt{|Y||Z|}\big)$-bijumbled in Γ, and*
- *(Y, Z) is (ε, d, p)-regular in G.*

Then, for all but at most at most $\varepsilon'|X|$ vertices x of X, the pair $\big(N_\Gamma(x) \cap Y, Z\big)$ is (ε', d, p)-regular in G.

Comparing this result with the analog by Conlon, Fox, and Zhao in [7, Proposition 5.1], we need Γ to be a factor $(p \log_2 \frac{1}{p})^{1/2}$ less jumbled when $|X| = |Y| = |Z|$. The second result establishes two-sided regularity inheritance under somewhat stronger bijumbledness conditions.

Lemma 4 (Two-sided inheritance lemma). *For each $\varepsilon', d > 0$ there are $\varepsilon, c > 0$ such that for all $0 < p < 1$ the following holds. Let $G \subseteq \Gamma$ be graphs and X, Y, Z be disjoint vertex sets in $V(\Gamma)$. Assume that*

- *(X, Y) is $(p, cp^2\sqrt{|X||Y|})$-bijumbled in Γ,*
- *(X, Z) is $(p, cp^3\sqrt{|X||Z|})$-bijumbled in Γ,*
- *(Y, Z) is $(p, cp^{5/2}\big(\log_2 \frac{1}{p}\big)^{-\frac{1}{2}}\sqrt{|Y||Z|})$-bijumbled in Γ, and*
- *(Y, Z) is (ε, d, p)-regular in G.*

Then, for all but at most $\varepsilon'|X|$ vertices x of X, the pair $\big(N_\Gamma(x) \cap Y, N_\Gamma(x) \cap Z\big)$ is (ε', d, p)-regular in G.

Here Γ needs to be a factor p less jumbled when $|X| = |Y| = |Z|$ than in [7, Proposition 1.13]. We remark that the bijumbledness conditions in our results imply that these implicitly are statements about sufficiently large graphs (see Lemma 7). Our proofs use the counting lemma for C_4 of Conlon, Fox, and Zhao [7] as a fundamental ingredient.

1.1 | Applications

Blow-up lemmas

Blow-up lemmas are an important tool in the regularity method, which make it possible to derive results about large or even spanning subgraphs in certain graph classes (see, eg, [20]). In [2] a blow-up lemma which works relative to sparse jumbled graphs is proved. The proof of this lemma relies on our regularity inheritance lemmas, Lemmas 3 and 4.

Resilience theorems in jumbled graphs

As an application of the blow-up lemma for jumbled graphs in [1] *resilience problems* for jumbled graphs with respect to certain spanning subgraphs are considered. The study of such problems dates back to [3] where the name *fault-tolerance* was used, but lately the term *resilience* has come into vogue, following Sudakov and Vu [24].

In [1] Lemmas 3 and 4 together with the blow-up lemma for jumbled graphs are used to derive the following sparse version of the bandwidth theorem (proved for dense graphs in [5]).

Theorem 5 [1] *For each $\varepsilon > 0$, $\Delta \geq 2$, and $k \geq 1$, there exists a constant $c > 0$ such that the following holds for any $p > 0$. Given $\gamma \leq cp^{\max(4,(3\Delta+1)/2)}n$, suppose Γ is a $\left(p,\gamma\right)$-bijumbled graph, G is a spanning subgraph of Γ with $\delta(G) \geq \left(\frac{k-1}{k}+\varepsilon\right)pn$, and H is a k-colorable graph on n vertices with $\Delta(H) \leq \Delta$ and bandwidth at most cn. Suppose further that there are at least $c^{-1}p^{-6}\gamma^2n^{-1}$ vertices in $V(H)$ that are not contained in any triangles of H. Then G contains a copy of H.*

Note that the bijumbledness requirement implicitly places a lower bound on p. It is necessary to insist on some vertices of H not being in any triangles of H, but the number $c^{-1}p^{-6}\gamma^2n^{-1}$ comes from the requirements of Lemma 4, and improvement there would immediately improve this statement.[1] This is a very general resilience result, covering for example Hamilton cycles, clique factors, and much more. Note that although a Hamilton cycle might not be 2-colorable, in [1] a more complicated variant of the above statement is proved which allows occasional vertices to receive a $(k+1)^{\text{st}}$ color.

Counting lemmas

The most obvious application of our inheritance lemmas is to prove stronger counting lemmas than those in [7]. The results we obtain are not much stronger than those in [7], so we do not regard this as a main contribution of this paper. However, we feel it is worth providing the stronger results for future use, and that the (rather different to that in [7]) approach we take is worth highlighting.

Recall that for a dense graph G and fixed H the counting lemma provides matching upper and lower bounds on the number of copies of H in G. By contrast, when G is a subgraph of a sparse bijumbled graph Γ, we formulate two separate counting lemmas. The one-sided counting lemma gives only a lower bound on the number of copies of H in G, while the two-sided counting lemma gives in addition a matching upper bound.[2] The motivation for formulating two separate lemmas is that for many graphs H, the bijumbledness requirement on Γ to prove a one-sided counting lemma is significantly less than to prove a two-sided counting lemma, and for many applications the one-sided counting lemma suffices.

[1] As we discuss in Section 1.2, we believe that one can improve Lemma 4 in order to obtain $c^{-1}p^{-4}\gamma^2n^{-1}$ uncovered vertices. This is the best achievable using inheritance lemmas, but we are not sure whether the constructions giving a lower bound on inheritance lemmas can be modified to give a matching lower bound in this setting.

[2] Somewhat confusingly, the terms one-sided/two-sided refer to completely different aspects in the one-sided/two-sided counting lemmas and the one-sided/two-sided inheritance lemmas. Both are standard terminology.

The statements and proofs of our counting lemmas are quite technical, and we prefer to leave them as an appendix to this paper. Comparison with the results of [7] is unfortunately also not straightforward, in part because the two-sided counting lemma in [7] actually provides better performance than the one-sided counting lemma there in some important cases, such as for cliques. Briefly, our one-sided counting lemma always performs at least as well as either of [7, Thms. 1.12 and 1.14], and in some cases our results are better. For example, if H consists of 10 copies of K_3 sharing a single vertex, then our one-sided counting lemma requires (p, cp^3)-jumbledness to lower bound the number of copies of H, whereas the results in [7] require (p, cp^4)-bijumbledness. Our two-sided counting lemma sometimes performs better than [7, Thm. 1.12]. Again, for 10 copies of K_3 sharing a vertex, we require $\left(p, cp^{10.5}\right)$-bijumbledness while [7] requires $\left(p, cp^{12}\right)$-bijumbledness. In general, our results perform better when there are vertices of exceptionally high degree. For many interesting graphs (such as d-regular graphs for any $d \geq 3$) the performance is identical.

Of course, these counting lemmas can also be immediately applied in the (relatively straightforward) applications presented in [7]. For most of these applications what one requires is a one-sided counting lemma. In particular, by using the one-sided counting lemma resulting from our inheritance lemmas the bijumbledness requirements for the removal lemma [7, Thm. 1.1], the Turán result [7, Thm. 1.4], and the Ramsey result [7, Thm. 1.6] can always be matched, and in some cases be improved.

1.2 | Optimality

Our one-sided inheritance lemma is probably not optimal. In the case when H is a clique, Conlon, Fox, and Zhao [7] are able to obtain a one-sided counting lemma with a bijumbledness requirement matching ours by using a completely different strategy. In particular, when H is a triangle, these counting lemmas imply a triangle removal lemma for subgraphs of bijumbled graphs with $\beta = o(p^3 n)$. Such a result was obtained earlier already in [17], where it was also conjectured that this can be improved to $\beta = o(p^2 n)$. Conlon, Fox, and Zhao [7] conjecture the contrary. We sympathize with the former conjecture, and believe that it would be extremely interesting to resolve this question. We think it unlikely that Lemma 4 is optimal and believe there is room for improvement in our proof strategy; any improvement would disprove the conjecture of Conlon, Fox, and Zhao.

Organization

The remaining sections of this paper are devoted to the proofs of the inheritance lemmas. We start in Section 2 with an overview of these proofs. Section 3 collects necessary auxiliary results on bijumbled graphs and sparse regular pairs. In Sections 4 and 5 we prove various lemmas used in the proofs of the inheritance lemmas: Section 4 establishes lemmas on counting copies of C_4 in various bipartite graphs, and Section 5 concerns a classification of pairs of vertices in such graphs according to their codegrees. In Section 6 we prove Lemma 3 and in Section 7 Lemma 4.

Notation

For a graph $G = (V, E)$ we also write $V(G)$ for the vertex set and $E(G)$ for the edge set of G. We write $e(G)$ for the number of edges of G. For vertices $v, v' \in V$ and a set $U \subseteq V$ we write $N_G(v; U)$ and $N_G(v, v'; U)$ for the G-neighborhood of v in U and common G-neighborhood of v and v' in U, respectively. Similarly, $\deg_G(v; U) := |N_G(v; U)|$ and $\deg_G(v, v'; U) := |N_G(v, v'; U)|$. If $U = V$ we may omit U and, if G is clear from the context, we may also omit G.

For disjoint vertex sets $U, W \subseteq V$ the graph $G[U, W]$ is the bipartite subgraph of G containing exactly all edges of G with one end in U and the other in W. We write $e(U, W)$ for the number of edges in $G[U, W]$.

2 | PROOF OVERVIEW

We sketch the proof of Lemma 3 first. We label the pairs in Y as "typical," "heavy," or "bad," according to whether their G-common neighborhood in Z is not significantly larger than one would expect, or so large as to be unexpected even in Γ, or intermediate. By using the bijumbledness of (Y, Z) in Γ we can show that the heavy pairs are so few that one can ignore them (Lemma 16).

Now suppose that $x \in X$ is such that $\big(N_\Gamma(x; Y), Z\big)$ is either too dense or is not sufficiently regular. In either case, by several applications of the defect Cauchy-Schwarz inequality, we conclude that $\big(N_\Gamma(x; Y), Z\big)$ contains noticeably more copies of C_4 in G than one would expect if (Y, Z) were a random bipartite graph of the same density (Lemma 13). In particular, the average pair of vertices in $N_\Gamma(x; Y)$ has noticeably more G-common neighbors in Z than one would expect. It follows that a substantial fraction of the pairs y, y' in $N_\Gamma(x; Y)$ are bad or heavy. Since there are few heavy pairs, we see that there are many bad pairs (Lemma 17).

On the other hand, because (Y, Z) is regular, we can count copies of C_4 in G crossing the pair (Lemma 12, which is taken from [7]). A further application of the defect Cauchy-Schwarz inequality tells us that a very small fraction of the pairs in Y are bad, and using the bijumbledness of (X, Y) we conclude that there are few triples (x, y, y') such that xy and xy' are edges of Γ and (y, y') is bad (Lemma 18).

Putting these two statements together, we conclude that there are few $x \in X$ such that $\big(N_\Gamma(x; Y), Z\big)$ is either too dense or is not sufficiently regular. By averaging, if there are few dense pairs there are also few pairs which are too sparse. This completes the proof of Lemma 3.

The proof of Lemma 4 is very similar. We have to additionally classify the pairs in Y as typical, heavy, or bad with respect to $x \in X$, which we do according to their G-common neighborhood in $N_\Gamma(x; Z)$. Now Lemma 17 as before tells us that if $x \in X$ is such that $\big(N_\Gamma(x; Y), Z\big)$ is either too dense or is not sufficiently regular, then a substantial fraction of the pairs (y, y') in $N_\Gamma(x; Y)$ are bad with respect to x. Lemma 18 continues to tell us that there are few triples (x, y, y') such that xy and xy' are edges of Γ and (y, y') is bad, and Lemma 16 continues to tell us that we can ignore the heavy pairs. To complete the argument as before, it remains to show that if (y, y') is a typical pair, then there are few x such that $xy, xy' \in \Gamma$ and (y, y') is bad with respect to x. To prove this we do not use the requirement $xy, xy' \in \Gamma$, but simply bound, using bijumbledness of (X, Z), the number of x with abnormally many neighbors in $N_G(y, y'; Z)$. This step is where we require most bijumbledness. We believe it is wasteful, but were not able to find a more efficient way.

3 | PRELIMINARIES

3.1 | Bijumbledness

One consequence of a pair (U, V) being (p, γ)-bijumbled is that most vertices in U have about $p|V|$ neighbors in V.

Lemma 6 *Let $k \geq 1$, $c' > 0$, and $0 < p < 1$, and let (U, V) be a $(p, c'p^k\sqrt{|U||V|})$-bijumbled pair in a graph Γ. Then, for any $\gamma > 0$, we have*

$$\big|\{u \in U : \ \deg_\Gamma(u; V) \neq (1 \pm \gamma)p|V|\}\big| \leq 2(c')^2 p^{2k-2}\gamma^{-2}|U|.$$

Proof Let $U^+ := \{u \in U : \deg_\Gamma(u;V) > (1+\gamma)p|V|\}$. By bijumbledness applied to the pair (U^+, V) we have

$$(1+\gamma)p|U^+||V| < e(U^+, V) \le p|U^+||V| + c'p^k\sqrt{|U||V|}\sqrt{|U^+||V|}.$$

Simplifying this gives $|U^+| \le (c')^2p^{2k-2}\gamma^{-2}|U|$. A similar calculation for the set U^- of vertices in U with fewer than $(1-\gamma)p|V|$ neighbors in V yields the same bound on $|U^-|$, and the result follows. ■

Moreover, nontrivial bijumbled graphs cannot be very small.

Lemma 7 *Let $0 < c' \le \frac{1}{4}$, $0 < p \le \frac{1}{4}$ and $k \ge 1$. Let Γ be a graph, and let (U,V) be $(p, c'p^k\sqrt{|U||V|})$-bijumbled in Γ. Then we have*

$$|U|, |V| \ge \tfrac{1}{8}(c')^{-2}p^{1-2k}.$$

Proof By Lemma 6, the number of vertices in U with more than $2p|V|$ neighbors in V is at most $2(c')^2p^{2k-2}|U| \le \frac{1}{2}|U|$. It follows that we can take a set $U' \subseteq U$ of $\min\left\{\frac{1}{4}p^{-1}, \frac{1}{2}|U|\right\}$ vertices, each with degree at most $2p|V|$. The union of their neighborhoods covers by definition at most $\frac{1}{4}p^{-1}\cdot 2p|V| = \frac{1}{2}|V|$ vertices of V, so we can let V' be a subset of $\frac{1}{2}|V|$ vertices in V with no edges between U' and V'. Applying bijumbledness to the pair (U', V'), we have

$$0 = e\big(U', V'\big) \ge p|U'||V'| - c'p^k\sqrt{|U||V|}\sqrt{|U'||V'|},$$

which implies $(c')^2p^{2k}|U||V| \ge p^2|U'||V'| = p^2\min\left\{\frac{1}{4}p^{-1}, \frac{1}{2}|U|\right\}\frac{1}{2}|V|$. Hence, we obtain

$$|U| \ge \tfrac{1}{2}(c')^{-2}p^{2-2k}\min\left\{\tfrac{1}{4}p^{-1}, \tfrac{1}{2}|U|\right\}.$$

The inequality $|U| \ge \frac{1}{4}(c')^{-2}p^{2-2k}|U|$ is false for all $U \ne \emptyset$ by our choice of c', p, and k, so we conclude that

$$|U| \ge \tfrac{1}{8}(c')^{-2}p^{1-2k}.$$

The same bound applies to $|V|$. ■

Remark 8 Erdős and Spencer [8] (see also Theorem 5 in [9]) observed that there exists $c > 0$ such that every m-vertex graph with density p contains two disjoint sets X and Y for which $\big|e(X,Y) - p|X||Y|\big| \ge c\sqrt{pm}\sqrt{|X||Y|}$, as long as $p(1-p) \ge 1/m$. One can also recover Lemma 7 using this result. (See also Remark 6 in [16].)

3.2 | Sparse regularity

The slicing lemma, Lemma 9, states that large subpairs of regular pairs remain regular. Its proof, which we omit, follows directly from Definition 1.

Lemma 9 (slicing lemma). *For any $0 < \varepsilon < \gamma$ and any $p > 0$, any (ε, p)-regular pair (U, W) in G, and any $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \ge \gamma|U|$ and $|W'| \ge \gamma|W|$, the pair (U', W') is $(\varepsilon/\gamma, p)$-regular in G with p-density $d(U,W) \pm \varepsilon$.*

In the other direction, the following lemma shows that, under certain conditions, adding a few vertices to either side of a regular pair cannot destroy regularity completely.

Lemma 10 *Let $0 < \varepsilon < \frac{1}{10}$ and $c \leq \frac{1}{10}\varepsilon^3$. Let G be a spanning subgraph of a graph Γ, let (U', V') be a pair of disjoint sets in $V(\Gamma)$, and let $U \subseteq U'$ and $V \subseteq V'$. Assume (U', V') is $(p, cp\sqrt{|U||V|})$-bijumbled in Γ and (U, V) is (ε, d, p)-regular in G.*

If $|U'| \leq \left(1 + \frac{1}{10}\varepsilon^3\right)|U|$ and $|V'| \leq \left(1 + \frac{1}{10}\varepsilon^3\right)|V|$, then (U', V') is $(2\varepsilon, d, p)$-regular in G.

Proof Let $X \subseteq U'$ with $|X| \geq 2\varepsilon|U'|$ and $Y \subseteq V'$ with $|Y| \geq 2\varepsilon|V'|$ be arbitrary. Using $(p, cp\sqrt{|U||V|})$-bijumbledness of (U', V') in Γ we have

$$e(X \cap U, Y \setminus V) \leq e_\Gamma(X \cap U, Y \setminus V) \leq p|U| \cdot \tfrac{\varepsilon^3}{10}|V| + cp\sqrt{|U||V|}\sqrt{|U| \cdot \tfrac{\varepsilon^3}{10}|V|} \leq \tfrac{1}{5}\varepsilon^3 p|U||V|.$$

Similarly, we have

$$e(X \setminus U, Y) \leq p\tfrac{\varepsilon^3}{10}|U| \cdot \left(1 + \tfrac{\varepsilon^3}{10}\right)|V| + cp\sqrt{|U||V|}\sqrt{\tfrac{\varepsilon^3}{10}|U| \cdot \left(1 + \tfrac{\varepsilon^3}{10}\right)|V|} \leq \tfrac{1}{5}\varepsilon^3 p|U||V|.$$

Moreover, since (U, V) is (ε, d, p)-regular, $e(X \cap U, Y \cap V) = (d \pm \varepsilon)p|X \cap U||Y \cap V|$. Hence

$$\begin{aligned} e(X, Y) &= e(X \cap U, Y \cap V) + e(X \cap U, Y \setminus V) + e(X \setminus U, Y) \\ &= (d \pm \varepsilon)p|X \cap U||Y \cap V| \pm \tfrac{2}{5}\varepsilon^3 p|U||V| \\ &= \left(d \pm \tfrac{3}{2}\varepsilon\right)p|X \cap U||Y \cap V| \\ &= (d \pm 2\varepsilon)p|X||Y|. \end{aligned}$$

We conclude that (U', V') is $(2\varepsilon, d, p)$-regular in G. ■

3.3 | Cauchy-Schwarz

We use the following "defect" form of the Cauchy-Schwarz inequality. This inequality and a proof can be found in [10, Fact B].

Lemma 11 (Defect form of Cauchy-Schwarz). *Let $a_1, \dots, a_k$ be real numbers with average at least a. If for some $\delta \geq 0$ at least μk of them average at least $(1 + \delta)a$, then*

$$\sum_{i=1}^{k} a_i^2 \geq ka^2\left(1 + \tfrac{\mu\delta^2}{1-\mu}\right),$$

and the same bound is obtained if at least μk of the a_i average at most $(1 - \delta)a$.

4 | COUNTING COPIES OF C_4 IN REGULAR, IRREGULAR AND DENSE PAIRS

The following counting lemma for counting C_4 in (ε, d, p)-regular subgraphs of bijumbled graphs is as given by Conlon, Fox, and Zhao [7, Proposition 4.13]. We write $C_4(G)$ for the number of unlabeled copies of C_4 in G.

Lemma 12 (counting C_4 in regular pairs). *For any $\varepsilon > 0$, $c > 0$ and $d \in [0, 1]$ the following holds. If (U, V) is a $(p, cp^2\sqrt{|U||V|})$-bijumbled pair in a graph Γ, and G is a bipartite subgraph of Γ with parts U and V which forms an (ε, d, p)-regular pair, then $C_4(G) = \frac{1}{4}\left(d^4 \pm 100(c+\varepsilon)^{1/2}\right)p^4|U|^2|V|^2$.*

The next lemma gives a lower bound on the number of copies of C_4 in a bipartite graph of a given density. Moreover, if this bipartite graph is not (ε)-regular we obtain an even stronger lower bound. Observe that for this lemma we do *not* require that the pair is a subgraph of a pseudorandom graph.

Lemma 13 (counting C_4 in dense pairs and irregular pairs). *Let $0 < \varepsilon_{\mathrm{L}13} \leq 10^{-3}$, let G be a bipartite graph with vertex classes U and V of sizes $m \geq n \geq 2\varepsilon_{\mathrm{L}13}^{-9}$ respectively. Suppose that G has density $q \geq \varepsilon_{\mathrm{L}13}^{-10} n^{-1/2}$.*

(a) $C_4(G) \geq (1 - \varepsilon_{\mathrm{L}13}^{8}) q^4 \frac{1}{4} m^2 n^2$.
(b) If G is not $(\varepsilon_{\mathrm{L}13})$-regular, then we have $C_4(G) \geq (1 + \varepsilon_{\mathrm{L}13}^{13}) q^4 \frac{1}{4} m^2 n^2$.

Proof Assume G has density q. Clearly, we have

$$C_4(G) = \sum_{\{u,u'\} \in \binom{U}{2}} \binom{\deg(u,u')}{2}. \tag{1}$$

Hence, for bounding this quantity we will analyze common neighborhoods of vertices in U. Let us first bound the average

$$a := \binom{m}{2}^{-1} \sum_{u \neq u' \in U} \deg(u, u').$$

Observe that if $a \geq (1 + \varepsilon_{\mathrm{L}13}^{8}) q^2 n$ then, using Jensen's inequality and facts that $q \geq 2\varepsilon_{\mathrm{L}13}^{-4} n^{-1/2}$ and $m \geq n \geq 2\varepsilon_{\mathrm{L}13}^{-9}$, we get

$$C_4(G) \geq \binom{m}{2}\binom{(1+\varepsilon_{\mathrm{L}13}^{8})q^2 n}{2} \geq \frac{(1+\varepsilon_{\mathrm{L}13}^{8})(1+\varepsilon_{\mathrm{L}13}^{9})q^4 n^2}{2}\binom{m}{2} \geq (1+\varepsilon_{\mathrm{L}13}^{8}) q^4 \frac{1}{4} n^2 m^2,$$

and thus are done. Hence we may assume in the following that

$$a \leq (1 + \varepsilon_{\mathrm{L}13}^{8}) q^2 n. \tag{2}$$

For obtaining a corresponding lower bound on a note that the average degree of the vertices in V is qm. Hence by Jensen's inequality we have

$$\sum_{v \in V} \binom{\deg(v)}{2} \geq n \binom{qm}{2} = n \frac{qm(qm-1)}{2} \geq n \frac{(1-\varepsilon_{\mathrm{L}13}^{20}) q^2 m^2}{2},$$

where the second inequality uses $q \geq q^2 \geq \varepsilon_{\mathrm{L}13}^{-20} m^{-1}$. Therefore

$$\sum_{\{u,u'\} \in \binom{U}{2}} \deg(u,u') = \sum_{v \in V} \binom{\deg(v)}{2} \geq n \frac{(1-\varepsilon_{\mathrm{L}13}^{20}) q^2 m^2}{2} \geq (1 - \varepsilon_{\mathrm{L}13}^{20}) q^2 \binom{m}{2} n.$$

This gives

$$a \geq (1 - \varepsilon_{\mathrm{L13}}^{20}) q^2 n. \tag{3}$$

Moreover, we obtain from (1) and (2) that

$$C_4(G) \geq \frac{1}{2} \sum_{\{u,u'\} \in \binom{U}{2}} \deg(u,u')^2 - (1 + \varepsilon_{\mathrm{L13}}^{8}) n q^2 \binom{m}{2}. \tag{4}$$

For estimating the sum of squares in this inequality, we will use the defect form of Cauchy-Schwarz (Lemma 11).

Let us first establish the first part of Lemma 13. We apply Lemma 11 with $k = \binom{m}{2}$, $\mu = \delta = 0$ (so actually without defect) to obtain that

$$\sum_{\{u,u'\} \in \binom{U}{2}} \deg(u,u')^2 \geq \binom{m}{2} a^2 \overset{(3)}{\geq} \binom{m}{2} (1 - \varepsilon_{\mathrm{L13}}^{20})^2 q^4 n^2.$$

Hence, by (4), we have

$$\begin{aligned} C_4(G) &\geq \frac{1}{2} \binom{m}{2} (1 - \varepsilon_{\mathrm{L13}}^{20})^2 q^4 n^2 - (1 + \varepsilon_{\mathrm{L13}}^{8}) n q^2 \binom{m}{2} \\ &\geq (1 - 2\varepsilon_{\mathrm{L13}}^{20}) q^4 \frac{n^2}{2} \binom{m}{2} - n \binom{m}{2} q^4 \left(\varepsilon_{\mathrm{L13}}^{-10} n^{-1/2}\right)^{-2} \\ &\geq (1 - \varepsilon_{\mathrm{L13}}^{8}) q^4 \frac{1}{4} m^2 n^2 , \end{aligned}$$

as desired, where we used $q \geq \varepsilon_{\mathrm{L13}}^{-10} n^{-1/2}$ in the second inequality.

For the second part of the lemma, we will use a similar calculation, but we will apply Lemma 11 with $\mu, \delta > 0$. So we need to find a subset $\tilde{U} \subseteq U$ of vertices whose average pair degrees differ significantly from a.

The following definition will be useful. For a set $\tilde{U} \subseteq U$, let

$$a(\tilde{U}) := \binom{|\tilde{U}|}{2}^{-1} \sum_{\{u,u'\} \in \binom{\tilde{U}}{2}} \deg(u,u') = \binom{|\tilde{U}|}{2}^{-1} \sum_{v \in V} \binom{\deg(v,\tilde{U})}{2}. \tag{5}$$

Claim 14 *If G is not $(\varepsilon_{\mathrm{L13}})$-regular, then there is a set $\tilde{U} \subseteq U$ with $|\tilde{U}| \geq \varepsilon_{\mathrm{L13}} m$ such that*

$$a(\tilde{U}) \geq (1 + 2\varepsilon_{\mathrm{L13}}^{5}) q^2 n \geq (1 + \varepsilon_{\mathrm{L13}}^{5}) a,$$

where the second inequality follows from (2).

Before we prove this claim, let us show how it implies the second part of our lemma. For this, assume that G is not $(\varepsilon_{\mathrm{L13}})$-regular, and let $\tilde{U}$ be the set guaranteed by Claim 14. Since $|\tilde{U}| \geq \varepsilon_{\mathrm{L13}} m$, there are at least $\binom{\varepsilon_{\mathrm{L13}} m}{2} \geq \frac{1}{2}\varepsilon_{\mathrm{L13}}^2\binom{m}{2}$ pairs of vertices in $\tilde{U}$. Thus we can use Lemma 11 with $k := \binom{m}{2}$, $\mu = \varepsilon_{\mathrm{L13}}^2/2$ and $\delta = \varepsilon_{\mathrm{L13}}^5$ to infer that

$$\begin{aligned}\sum_{\{u,u'\}\in\binom{U}{2}} \deg(u,u')^2 &\geq \binom{m}{2} a^2\left(1+\frac{\varepsilon_{\mathrm{L13}}^{12}}{2}\right)\\ &\overset{(3)}{\geq} \binom{m}{2}(1-\varepsilon_{\mathrm{L13}}^{20})^2 q^4 n^2\left(1+\frac{\varepsilon_{\mathrm{L13}}^{12}}{2}\right)\\ &\geq \binom{m}{2}\left(1+\frac{\varepsilon_{\mathrm{L13}}^{12}}{4}\right) q^4 n^2.\end{aligned}$$

Together with (4) this gives the desired

$$\begin{aligned}C_4(G) &\geq \frac{1}{2}\binom{m}{2}\left(1+\frac{\varepsilon_{\mathrm{L13}}^{12}}{4}\right) q^4 n^2 - (1+\varepsilon_{\mathrm{L13}}^{8}) n q^2 \binom{m}{2}\\ &\geq \frac{1}{2}\binom{m}{2}\left(1+\frac{\varepsilon_{\mathrm{L13}}^{12}}{5}\right) q^4 n^2 \geq (1+\varepsilon_{\mathrm{L13}}^{13}) q^4 \tfrac{1}{4} n^2 m^2\,,\end{aligned}$$

where again we used $q \geq \varepsilon_{\mathrm{L13}}^{-10} n^{-1/2}$ in the second inequality.

It remains to prove the claim.

Proof of Claim 14 Since G is not $(\varepsilon_{\mathrm{L13}})$-regular there are sets $U' \subseteq U$ and $V' \subseteq V$ with $|U'| = \varepsilon_{\mathrm{L13}} m$ and $|V'| = \varepsilon_{\mathrm{L13}} n$ such that either

$$d(U', V') > (1+\varepsilon_{\mathrm{L13}})q \quad \text{or} \quad d(U', V') < (1-\varepsilon_{\mathrm{L13}})q. \tag{6}$$

Now we distinguish three cases.

First suppose that $d(U', V) \geq \left(1+\frac{\varepsilon_{\mathrm{L13}}^3}{10}\right)q =: \tilde{q}$. Then, using again Jensen's inequality, we have

$$\begin{aligned}a(U') &= \binom{\varepsilon_{\mathrm{L13}} m}{2}^{-1} \sum_{v\in V}\binom{\deg(v,U')}{2} \geq \frac{2}{\varepsilon_{\mathrm{L13}}^2 m^2}\cdot n\frac{\tilde{q}\varepsilon_{\mathrm{L13}} m(\tilde{q}\varepsilon_{\mathrm{L13}} m - 1)}{2}\\ &\geq \frac{2}{\varepsilon_{\mathrm{L13}}^2 m^2}\cdot n\frac{(1-\varepsilon_{\mathrm{L13}}^7)\tilde{q}^2\varepsilon_{\mathrm{L13}}^2 m^2}{2}\\ &= (1-\varepsilon_{\mathrm{L13}}^7)\tilde{q}^2 n \geq \left(1+\frac{\varepsilon_{\mathrm{L13}}^3}{5}\right) q^2 n\,,\end{aligned}$$

where the second inequality uses $\tilde{q} \geq q \geq \varepsilon_{\mathrm{L13}}^{-8}/m$ and the last inequality uses $\varepsilon_{\mathrm{L13}} \leq 10^{-3}$. Hence we can choose U' as $\tilde{U}$.

Secondly, suppose that $d(U', V) \leq \left(1-\frac{\varepsilon_{\mathrm{L13}}^3}{10}\right)q$ and let $U'' := U \setminus U'$. Then

$$d(U'', V) = \frac{d(U,V)nm - d(U',V)\varepsilon_{\mathrm{L13}} nm}{(1-\varepsilon_{\mathrm{L13}})nm} \geq \frac{q - (1-\frac{\varepsilon_{\mathrm{L13}}^3}{10})q\varepsilon_{\mathrm{L13}}}{1-\varepsilon_{\mathrm{L13}}} \geq \left(1+\frac{\varepsilon_{\mathrm{L13}}^4}{10}\right)q.$$

Using an analogous calculation as in the previous case we obtain $a(U'') \geq (1+2\varepsilon_{\mathrm{L13}}^5)q^2n$ and thus can choose U'' as $\tilde{U}$.

Finally, suppose $\left(1-\frac{\varepsilon_{\mathrm{L13}}^3}{10}\right)q < d(U',V) < \left(1+\frac{\varepsilon_{\mathrm{L13}}^3}{10}\right)q$. In this case we will use $\tilde{U} := U'$ and apply Lemma 11 to bound

$$a(U') \geq \frac{1}{\varepsilon_{\mathrm{L13}}^2 m^2}\Big(\sum_{v\in V}\deg(v,U')^2 - \sum_{v\in V}\deg(v,U')\Big). \tag{7}$$

For this observe that

$$b := \frac{1}{n}\sum_{v\in V}\deg(v,U') = \frac{1}{n}d(U',V)\varepsilon_{\mathrm{L13}}mn = \left(1\pm\frac{\varepsilon_{\mathrm{L13}}^3}{10}\right)q\varepsilon_{\mathrm{L13}}m. \tag{8}$$

On the other hand,

$$b(V') := \frac{1}{\varepsilon_{\mathrm{L13}}n}\sum_{v\in V'}\deg(v,U') = \frac{1}{\varepsilon_{\mathrm{L13}}n}d(U',V')\varepsilon_{\mathrm{L13}}^2 mn$$

and thus, by (6), we obtain that either

$$b(V') > (1+\varepsilon_{\mathrm{L13}})q\varepsilon_{\mathrm{L13}}m \geq \left(1+\frac{\varepsilon_{\mathrm{L13}}}{2}\right)b$$

or

$$b(V') < (1-\varepsilon_{\mathrm{L13}})q\varepsilon_{\mathrm{L13}}m \leq \left(1-\frac{\varepsilon_{\mathrm{L13}}}{2}\right)b.$$

Therefore, Lemma 11 applied with $k := n$, $\delta := \varepsilon_{\mathrm{L13}}/2$, $\mu := \varepsilon_{\mathrm{L13}}$, and with b instead of a implies that

$$\sum_{v\in V}\deg(v,U')^2 \geq n\left(1-\frac{\varepsilon_{\mathrm{L13}}^3}{10}\right)^2 q^2\varepsilon_{\mathrm{L13}}^2 m^2\left(1+\frac{\varepsilon_{\mathrm{L13}}^3}{4(1-\varepsilon_{\mathrm{L13}})}\right) \geq \left(1+\frac{\varepsilon_{\mathrm{L13}}^3}{100}\right)q^2n\cdot\varepsilon_{\mathrm{L13}}^2m^2.$$

Together with (7) and (8) this gives

$$a(U') \geq \left(1+\frac{\varepsilon_{\mathrm{L13}}^3}{100}\right)q^2n - \left(1+\frac{\varepsilon_{\mathrm{L13}}^3}{10}\right)\frac{qn}{\varepsilon_{\mathrm{L13}}m} \geq \left(1+\frac{\varepsilon_{\mathrm{L13}}^3}{1000}\right)q^2n$$

as desired, where we used $q \geq 400\varepsilon_{\mathrm{L13}}^{-4}/m$. ∎

5 | TYPICAL PAIRS, BAD PAIRS, HEAVY PAIRS

The proofs of our inheritance lemmas rely on estimating the number of copies of C_4 which use certain types of vertex pairs in one part of a regular pair which is a subgraph of a bijumbled graph. We will consider vertex pairs that are atypical for the regular pair, which we call bad, and vertex pairs which are even atypical for the underlying bijumbled graph, which we call heavy.

Definition 15 (bad, heavy pairs)**.** Let G be a graph and U and V be disjoint vertex sets in G. Let $q \in [0,1]$ and $\delta > 0$. We say that a pair uu' of distinct vertices in U is (V, q, δ)-*bad* in G if

$$\deg_G(u, u'; V) \geq (1+\delta)q^2|V|.$$

Moreover, uu' is (V, q)-*heavy* in G if

$$\deg_G(u, u'; V) \geq 4q^2|V|.$$

Pairs which are neither heavy nor bad (with certain parameters) will usually be called *typical*.

In a bijumbled graph we can establish good bounds on the number of copies of C_4 which use heavy pairs.

Lemma 16 (C_4-copies using heavy pairs). *Let Γ be a bipartite graph with partition classes U and V that is $\left(p, c'p^{3/2}(\log_2 \frac{1}{p})^{-1/2}\sqrt{|U||V|}\right)$-bijumbled. Assume further that for all $u \in U$ we have* $\deg_\Gamma(u; V) \leq 2p|V|$.

Then the number of copies of C_4 in Γ which use a pair in U which is (V, p)-heavy in Γ is less than $64(c')^2p^4|U|^2|V|^2$.

Proof We first fix $u \in U$ and count the number of copies of C_4 in Γ which use a pair that contains u and is (V, p)-heavy. Let $W_u \subseteq U \setminus \{u\}$ be the set of vertices $u' \in U \setminus \{u\}$ such that uu' is a (V, p)-heavy pair. We now split W_u according to the number of common neighbors the vertices of W_u have with u. Since $4p^2|V| \leq \deg(u, u') \leq 2p|V|$ for all $u' \in W_u$, we can partition W_u into $W_u = S_1 \dot\cup \ldots \dot\cup S_{\lfloor \log_2 \frac{1}{p} \rfloor}$ with

$$S_t = \left\{u' \in W_u : 2^{t-1} \cdot 4p^2|V| \leq |N(u, u')| < 2^t \cdot 4p^2|V|\right\}$$

for $t = 1, 2, \ldots, \lfloor \log_2 \frac{1}{p} \rfloor$. Since $\deg_\Gamma(u; V) \leq 2p|V|$, we can take a superset $N_u \subseteq V$ of $N(u)$ of size $2p|V|$. Applying Lemma 6 to (U, N_u) with c' replaced by $c'(\log_2 \frac{1}{p})^{-1/2}$, $k = 1$ and $\gamma = 2^{t-1}$, we see that the number of vertices in U with at least $(1+\gamma)p|N_u| = (1+\gamma) \cdot 2p^2|V|$ neighbors in N_u is at most

$$2(c')^2(\log_2 \tfrac{1}{p})^{-1}2^{2-2t}|U|.$$

Since each vertex of S_t has at least $4 \cdot 2^{t-1}p^2|V| \geq (1+\gamma) \cdot 2p^2|V|$ neighbors in $N(u) \subseteq N_u$, we conclude that

$$|S_t| \leq 2^{3-2t}(c')^2|U|(\log_2 \tfrac{1}{p})^{-1}.$$

For a fixed u', the number of copies of C_4 using u and u' is $\binom{|N(u,u')|}{2} \leq \frac{1}{2}|N(u, u')|^2$. Hence, the total number of copies of C_4 using u and any vertex of S_t is at most

$$|S_t|\frac{1}{2}\left(2^t \cdot 4p^2|V|\right)^2 \leq 64(c')^2p^4|U||V|^2(\log_2 \tfrac{1}{p})^{-1}.$$

Summing over the at most $\log_2 \frac{1}{p}$ values of t, we conclude that the total number of copies of C_4 in Γ using u and some $u' \in W_u$ is at most $64(c')^2p^4|U||V|^2$.

Finally, summing over all $u \in U$, the total number of copies of C_4 in Γ using (V, p)-heavy pairs in U is at most $64(c')^2p^4|U|^2|V|^2$ as desired. ∎

Using this lemma we obtain a good lower bound on the number of bad pairs in subgraphs of bijumbled graphs which are irregular or exceed a certain density.

Lemma 17 (many bad pairs). *Given $d \in (0,1)$ and $\varepsilon^* \le 10^{-3}$, if $\delta \le (\varepsilon^*)^{14}/10$, $\varepsilon \le (\varepsilon^*)^{14} d/100$ and $c' \le d^2\varepsilon^{10}/100$ then for any $p \in (0,1/2)$ the following holds.*
Let Γ be a graph and let G be a bipartite subgraph of Γ with vertex classes U and V. Assume further that Γ is $\big(p, c'p^{3/2}(\log_2 \frac{1}{p})^{-1/2}\sqrt{|U||V|}\big)$-bijumbled and $\deg_\Gamma(u;V) \le 2p|V|$ for all $u \in U$. If

- *(i) (U,V) has density at least $(d-\varepsilon)p$ and is not (ε^*,p)-regular in G, or*
- *(ii) (U,V) has density at least $(d+\varepsilon^*)p$ in G,*

then at least $(\varepsilon^*)^{15} d^4 |U|^2$ pairs $uu' \in \binom{U}{2}$ are (V, dp, δ)-bad in G.

Proof Let P_h be the set of (V,p)-heavy pairs in Γ and P_b be the set of (V,dp,δ)-bad pairs in G which are not in P_h. Let $P_t := \binom{U}{2} \setminus (P_b \cup P_h)$. Denote by C_4^h the number of those copies of C_4 in G that use a pair in P_h, and define C_4^b and C_4^t similarly.

We claim that, if (i) or (ii) are satisfied, then

$$C_4(G) \ge \big(1 + (\varepsilon^*)^{14}\big) d^4 p^4 \tfrac{1}{4}|U|^2|V|^2. \tag{9}$$

Indeed, Lemma 7 implies that U and V are of size at least $\frac{1}{8}\big(c'(\log_2 \frac{1}{p})^{-1/2}p\big)^{-2} \ge \frac{1}{8}(c'p)^{-2}$, and hence $v_0 := \big(\min\{|U|,|V|\}\big)^{-1/2} \le \sqrt{8}c'p$. Now assume first that (i) holds. Then G has density at least

$$(d-\varepsilon)p \ge (\varepsilon^*)^{-10}\sqrt{8}c'p \ge (\varepsilon^*)^{-10} v_0$$

because $c' \le d^2\varepsilon^{10}/100$ and $\varepsilon \le \varepsilon^*$. But this is the condition we require to apply Lemma 13 with $\varepsilon_{\mathrm{L13}} = \varepsilon^*$ and $q = (d-\varepsilon)p$ to G. By Lemma 13(b) we have

$$C_4(G) \ge \big(1 + (\varepsilon^*)^{13}\big)(d-\varepsilon)^4 p^4 \tfrac{1}{4}|U|^2|V|^2 \ge (1 + (\varepsilon^*)^{14}) d^4 p^4 \tfrac{1}{4}|U|^2|V|^2,$$

where we used $\varepsilon \le (\varepsilon^*)^{14} d/100$ in the second inequality. Hence (9) holds in this case. If, on the other hand, (ii) holds, then G has density at least $(d+\varepsilon^*)p \ge dp \ge \varepsilon^{-10}\sqrt{8}c'p \ge \varepsilon^{-10}v_0$, where we used $c' \le d^2\varepsilon^{10}/100$ in the second inequality. Hence Lemma 13(a) applied with $\varepsilon_{\mathrm{L13}} = \varepsilon$ and $q = (d+\varepsilon^*)p$ gives

$$C_4(G) \ge (1-\varepsilon^8)(d+\varepsilon^*)^4 p^4 \tfrac{1}{4}|U|^2|V|^2 \ge (1+(\varepsilon^*)^{14}) d^4 p^4 \tfrac{1}{4}|U|^2|V|^2,$$

because $\varepsilon \le (\varepsilon^*)^{14} d/100$, which means we also get (9) in this case.

Our next goal is to obtain a lower bound for C_4^b. For this purpose observe that since $\deg_\Gamma(u;V) \le 2p|V|$ for all $u \in U$ Lemma 16 applies and we obtain $C_4^h \le 64(c')^2 p^4 |U|^2|V|^2$. Moreover, each pair $uu' \in P_t$ lies in at most $\binom{\deg_G(u,u';V)}{2} \le \binom{(1+\delta)d^2p^2|V|}{2} \le \frac{1}{2}(1+\delta)^2 d^4 p^4 |V|^2$ copies of C_4 in G by definition of P_t. Hence

$$C_4^t \le \binom{|U|}{2} \cdot \frac{1}{2}(1+\delta)^2 d^4 p^4 |V|^2 \le \frac{1}{4}(1+3\delta) d^4 p^4 |U|^2|V|^2.$$

Thus we conclude from (9) that

$$C_4^b = C_4(G) - C_4^h - C_4^t \ge \Big(\frac{(\varepsilon^*)^{14} d^4}{4} - 64(c')^2 - \frac{3\delta d^4}{4}\Big) p^4 |U|^2|V|^2$$

$$\geq 8(\varepsilon^*)^{15} d^4 p^4 |U|^2 |V|^2 ,$$

where we use $\varepsilon^* \leq 10^{-3}$, $c' \leq d^2\varepsilon^{10}/100$, $\delta \leq (\varepsilon^*)^{14}/10$, and $\varepsilon \leq \varepsilon^*$.

Now observe that each pair uu' in P_b is in at most

$$\binom{\deg_\Gamma(u,u';V)}{2} \leq \binom{4p^2|V|}{2} < 8p^4|V|^2$$

copies of C_4 in G by definition of P_b. It follows that

$$|P_b| \geq \frac{8(\varepsilon^*)^{15} d^4 p^4 |U|^2 |V|^2}{8p^4|V|^2} = (\varepsilon^*)^{15} d^4 |U|^2 ,$$

as desired. ■

The next lemma provides an upper bound for the number of bad pairs in neighborhoods.

Lemma 18 (few bad pairs). *Let $d, \delta > 0$, let $c' \leq \varepsilon \leq 10^{-10}\delta^6 d^8$, and $p \in (0,1)$. Let $G \subseteq \Gamma$ and let $U, V, W \subseteq V(\Gamma)$ be disjoint sets such that*

(i) (U,V) is $(p, c'p^{3/2}\sqrt{|U||V|})$-bijumbled in Γ,
(ii) (V,W) is $(p, c'p^2\sqrt{|V||W|})$-bijumbled in Γ, and (ε, d, p)-regular in G, and
(iii) each $v \in V$ has $\deg_\Gamma(v;U) = (1 \pm \varepsilon)p|U|$.

Then, for the sets $P_b(u)$ of pairs $vv' \in \binom{N_\Gamma(u;V)}{2}$ which are (W, dp, δ)-bad in G, we have $\sum_{u\in U} |P_b(u)| \leq \delta p^2 |U||V|^2$.

Proof Let P_b be the set of all pairs $vv' \in \binom{V}{2}$ which are (W, dp, δ)-bad in G. Our first step is to obtain an upper bound on $|P_b|$.

Claim 19 $|P_b| \leq \frac{1}{2}\delta\binom{|V|}{2}$.

Proof of Claim 19 We conclude from Lemma 7 applied to (V,W) that $|V|, |W| \geq \frac{1}{8}(c')^{-2}p^{-3} \geq \frac{1}{8}p^{-2}\varepsilon^{-2}$. This implies

$$|V| - 1 \geq (1-\varepsilon)|V| \quad \text{and} \quad (d-\varepsilon)p|V| - 1 \geq (1-\varepsilon)(d-\varepsilon)p|V| , \tag{10}$$

which we will use to estimate binomial coefficients.

Let μ be such that $|P_b| = \mu\binom{|V|}{2}$. Our goal is to get an upper bound on μ. For this purpose we shall first use the defect form of Cauchy-Schwarz, Lemma 11, to get a lower bound on the number of C_4-copies in (V,W) in terms of μ. Then we combine this bound with the upper bound on the number of C_4-copies in regular pairs provided by Lemma 12.

For the application of Lemma 11 set $a_{vv'} := \deg_G(v,v';W)$ for each $vv' \in \binom{V}{2}$, and define

$$a' := \binom{|V|}{2}^{-1} \sum_{vv' \in \binom{V}{2}} a_{vv'} = \binom{|V|}{2}^{-1} \sum_{w\in W} \binom{\deg_G(w;V)}{2}$$

to be the average of the $a_{vv'}$. Let us now first establish some bounds on a'. Since (V, W) is (ε, d, p)-regular in G, all but at most $\varepsilon|W|$ vertices of W have at least $(d-\varepsilon)p|V|$ neighbors in V. This gives

$$a' \geq \binom{|V|}{2}^{-1}(1-\varepsilon)|W|\binom{(d-\varepsilon)p|V|}{2} \overset{(10)}{\geq} (1-\varepsilon)^2(d-\varepsilon)^2p^2|W| =: a. \tag{11}$$

On the other hand, by Lemma 6, the number of vertices $w \in W$ with $\deg_\Gamma(w; V) > 2p|V|$ is at most $2(c'p)^2|W|$. We conclude that

$$\begin{aligned} a' &\leq \binom{|V|}{2}^{-1}\left(|W|\binom{2p|V|}{2} + 2(c'p)^2|W|\binom{|V|}{2}\right) \\ &\leq \left(4+2(c')^2\right)p^2|W| \leq 5p^2|W|. \end{aligned} \tag{12}$$

Now we apply Lemma 11 with $k = \binom{|V|}{2}$ and a, δ and μ as given. By (11) the $a_{vv'}$ average at least a. Moreover, by definition all μk pairs $vv' \in P_b$ are (W, dp, δ)-bad in G, that is, $a_{vv'} \geq (1+\delta)d^2p^2|W| \geq (1+\delta)a$ by (11). Lemma 11 thus guarantees that

$$\begin{aligned} \sum_{vv' \in \binom{V}{2}} a_{vv'}^2 &\geq ka^2\left(1+\frac{\mu\delta^2}{1-\mu}\right) \geq \binom{|V|}{2}(1-\varepsilon)^4(d-\varepsilon)^4p^4|W|^2(1+\mu\delta^2) \\ &\overset{(10)}{\geq} \tfrac{1}{2}(1-\varepsilon)^5(d-\varepsilon)^4(1+\mu\delta^2)p^4|V|^2|W|^2 \\ &\geq \tfrac{1}{2}(d-3\varepsilon)^4(1+\mu\delta^2)p^4|V|^2|W|^2, \end{aligned}$$

since $(1-\varepsilon)^5(d-\varepsilon)^4 \geq (d-3\varepsilon)^4$.

Hence the number of copies of C_4 in $G[V, W]$ is

$$\begin{aligned} \sum_{vv' \in \binom{V}{2}} \binom{a_{vv'}}{2} &= \tfrac{1}{2}\sum_{vv' \in \binom{V}{2}} a_{vv'}^2 - \tfrac{1}{2}\sum_{vv' \in \binom{V}{2}} a_{vv'} \overset{(12)}{\geq} \tfrac{1}{4}(d-3\varepsilon)^4(1+\mu\delta^2)p^4|V|^2|W|^2 - |V|^2\tfrac{5}{4}p^2|W| \\ &\geq \tfrac{1}{4}(d-4\varepsilon)^4(1+\mu\delta^2)p^4|V|^2|W|^2, \end{aligned}$$

where we used $|W| \geq \frac{1}{8}p^{-2}\varepsilon^{-2}$ in the last inequality. On the other hand, since (V, W) is (ε, d, p)-regular in G, and $(p, c'p^2\sqrt{|V||W|})$-bijumbled in Γ, Lemma 12 implies that the number of copies of C_4 in $G[V, W]$ is at most $\frac{1}{4}\left(d^4+100(c'+\varepsilon)^{1/2}\right)p^4|V|^2|W|^2$. Putting these two inequalities together we obtain

$$(d-4\varepsilon)^4(1+\mu\delta^2) \leq d^4 + 100(c'+\varepsilon)^{1/2}.$$

Using the assumption that $c' \leq \varepsilon \leq 10^{-10}\delta^6d^8$, we deduce that $\mu \leq \frac{1}{2}\delta$ as desired. ■

Now for each $v \in V$, let $V_v := \{v' \in V : vv' \in P_b\}$. Note that $vv' \in P_b(u)$ if and only if $vv' \in P_b$ and $u \in N_\Gamma(v; U)$ and $uv' \in E(\Gamma)$. It follows that

$$\sum_{u \in U} |P_b(u)| = \tfrac{1}{2}\sum_{v \in V} e_\Gamma\left(V_v, N_\Gamma(v; U)\right).$$

Since (U,V) is $(p, c'p^{3/2}\sqrt{|U||V|})$-bijumbled in Γ, we have for each $v \in V$ that

$$\begin{aligned} e_\Gamma\big(V_v, N_\Gamma(v;U)\big) &\le p|V_v|\deg_\Gamma(v;U) + c'p^{3/2}\sqrt{|U||V|}\sqrt{|V_v|\deg_\Gamma(v;U)} \\ &\le (1+\varepsilon)p^2|V_v||U| + c'p^{3/2}\sqrt{|U||V|}\sqrt{(1+\varepsilon)p|U||V|} \\ &= \big((1+\varepsilon)|V_v| + c'\sqrt{1+\varepsilon}|V|\big)p^2|U|\,, \end{aligned}$$

where we use assumption (iii) for the second inequality. We therefore obtain

$$\begin{aligned} \sum_{u\in U}|P_b(u)| &\le \tfrac{1}{2}\sum_{v\in V}\big((1+\varepsilon)|V_v| + c'\sqrt{1+\varepsilon}|V|\big)p^2|U| \\ &\le (1+\varepsilon)p^2|P_b||U| + c'p^2|V|^2|U| \\ &\le \tfrac{1}{2}(1+\varepsilon)\delta p^2\binom{|V|}{2}|U| + c'p^2|V|^2|U| \\ &\le \delta p^2|U||V|^2, \end{aligned}$$

as desired, where in the third inequality we use Claim 19. ■

6 | ONE-SIDED INHERITANCE

To prove Lemma 3 we combine Lemmas 17 and 18. The former asserts that any vertex x such that $\big(N_\Gamma(x;Y), Z\big)$ is not (ε', d, p)-regular in G creates many pairs in $N_\Gamma(x,Y)$ which are bad in $(Y,Z) \subseteq G$, whereas the latter upper bounds the sum over $x \in X$ of the number of such bad pairs.

Proof of Lemma 3 We may assume without loss of generality that $0 < \varepsilon' < 10^{-4}$. Given in addition $d > 0$ set

$$\delta = 10^{-10}(\varepsilon')^{20}d^4\,, \quad \varepsilon = 10^{-16}(\varepsilon')^{22}d^{16}\delta^6 \quad \text{and} \quad c = 10^{-4}\varepsilon^{10}d^4\delta^2. \tag{13}$$

As a preparation we first "clean up" the partition classes X, Y, Z as follows. We let $Y' \subseteq Y$ be the set of vertices y of Y with

$$\begin{aligned} \deg_\Gamma(y;X) = (1\pm\varepsilon)p|X|\,, \quad \deg_\Gamma(y;Z) &= (1\pm\varepsilon)p|Z|\,, \quad \text{and} \\ \deg_G(y;Z) &= (d\pm\varepsilon)p|Z|. \end{aligned} \tag{14}$$

Observe that by Lemma 6 and by (ε, d, p)-regularity of (X, Y) in G we have

$$|Y\setminus Y'| \le 2c^2p\varepsilon^{-2}|Y| + 2c^2p^2\varepsilon^{-2}|Y| + 2\varepsilon|Y| \overset{(13)}{\le} 3\varepsilon|Y|. \tag{15}$$

Hence, (X, Y') is $(p, \frac{3}{2}cp^{3/2}\sqrt{|X||Y'|})$-bijumbled in Γ. We then let $X' \subseteq X$ be the set of vertices x of X with

$$\deg_\Gamma(x;Y') = (1\pm\varepsilon)p|Y'| \quad \text{and} \quad \deg_\Gamma(x;Y\setminus Y') \le 4\varepsilon p|Y|. \tag{16}$$

Similarly as before, we apply Lemma 6 once to (X, Y') with $\gamma = \varepsilon$ and once to the pair $(X, Y \setminus Y')$ in (X, Y) with $\gamma = \frac{1}{3}$ and use (13) and (16) to obtain

$$|X \setminus X'| \leq 2(\tfrac{3}{2}c)^2 p\varepsilon^{-2}|X| + 2\big(3c\big)^2 p|X| \leq \varepsilon p|X|. \tag{17}$$

By Lemma 9 and because of (14) and (17) it follows that

$$(Y', Z) \text{ is } (2\varepsilon, d, p)\text{-regular in } G \qquad \text{and} \qquad \deg_\Gamma(y; X') = (1 \pm 3\varepsilon)p|X'| \tag{18}$$

for each $y \in Y'$. Moreover, (Y', Z) is $\big(p, \frac{3}{2}cp^2(\log_2 \frac{1}{p})^{-1/2}\sqrt{|Y'||Z|}\big)$-bijumbled in Γ. Thus for each $x \in X'$, because $|Y'| \leq \deg_\Gamma(x; Y')/(p(1-\varepsilon))$ by (16),

$$(Y', Z) \text{ is } \big(p, 2cp^{3/2}(\log_2 \tfrac{1}{p})^{-1/2}\sqrt{|N_\Gamma(x; Y')||Z|}\big)\text{-bijumbled in } \Gamma. \tag{19}$$

Finally, let X^* be the set of vertices in X' such that $\big(N_\Gamma(x; Y'), Z\big)$ is not $\big(\frac{\varepsilon'}{2}, d, p\big)$-regular in G.

We claim that $\big(N_\Gamma(x; Y), Z\big)$ is (ε', d, p)-regular in G for all $x \in X' \setminus X^*$. In order to show this we apply Lemma 10 with $\varepsilon_{\mathrm{L10}} = \frac{1}{2}\varepsilon'$ and $c_{\mathrm{L10}} = 2c$, and with $U = N_\Gamma(x; Y')$, $U' = N_\Gamma(x, Y)$ and $V = V' = Z$. This is possible by (19), the definition of X^*, and because

$$\begin{aligned} |U' \setminus U| &\leq \deg_\Gamma(x; Y \setminus Y') \overset{(16)}{\leq} 4\varepsilon p|Y| \overset{(13)}{\leq} \frac{1}{10}\left(\frac{\varepsilon'}{2}\right)^3 (1-3\varepsilon)p|Y| \\ &\leq \frac{1}{10}\left(\frac{\varepsilon'}{2}\right)^3 \deg_\Gamma(x; Y) \leq \frac{1}{10}\left(\frac{\varepsilon'}{2}\right)^3 |U|, \end{aligned} \tag{20}$$

where for the second to last inequality we use (15) and (16). We conclude that indeed $\big(N_\Gamma(x; Y), Z\big)$ is (ε', d, p)-regular in G. Therefore, by (17) it suffices to show that $|X^*| \leq \frac{1}{2}\varepsilon'|X|$ to complete the proof.

For this purpose, we define for each $x \in X'$

$$P_b(x) := \left\{ yy' \in \binom{N_\Gamma(x; Y')}{2} : yy' \text{ is } (Z, dp, \delta)\text{-bad in } G \right\}.$$

and determine a lower bound on $\sum_{x \in X'} |P_b(x)|$ in terms of $|X^*|$ with the help of Lemma 17 and an upper bound in terms of $|X'|$ with the help of Lemma 18.

For the lower bound, fix $x \in X^*$. By (14) the density of $\big(N_\Gamma(x; Y'), Z\big)$ in G is at least $(d - \varepsilon)p$. Hence, by (13), (14), (19) and the definition of X^* we may apply Lemma 17 with parameters $d, \varepsilon^* = \frac{\varepsilon'}{2}$, δ, $\varepsilon_{\mathrm{L17}} = 2\varepsilon$, $c' = 2c$ and p to the pair $\big(N_\Gamma(x; Y'), Z\big)$ in G, in the bijumbled graph (Y, Z) in Γ, using condition (i) of this lemma. We obtain $|P_b(x)| \geq \big(\frac{\varepsilon'}{2}\big)^{15} d^4 \deg_\Gamma(x; Y')^2$, and therefore

$$\sum_{x \in X'} |P_b(x)| \geq \sum_{x \in X^*} |P_b(x)| \overset{(16)}{\geq} |X^*| \cdot \left(\frac{\varepsilon'}{2}\right)^{15} d^4 (1-\varepsilon)^2 p^2 |Y'|^2. \tag{21}$$

For the upper bound we use Lemma 18 with input δ, $\varepsilon_{\mathrm{L18}} = 3\varepsilon$, and $c' = 2c$, and setting $U = X'$, $V = Y'$ and $W = Z$, which we may do by (13), (15), (17), and (18). The conclusion is that $\sum_{x \in X'} |P_b(x)| \leq \delta p^2 |X'||Y'|^2$. Together with (21) this gives $\big(\frac{\varepsilon'}{2}\big)^{15} d^4 (1-\varepsilon)^2 |X^*| \leq \delta|X'| \leq \delta|X|$ and therefore by (13) we indeed have $|X^*| \leq \frac{1}{2}\varepsilon'|X|$. ∎

7 | TWO-SIDED INHERITANCE

The proof of Lemma 4 follows a similar pattern to that of Lemma 3.

Proof of Lemma 4 Assume without loss of generality that $0 < \varepsilon' < 10^{-4}$. Given $d > 0$, we set

$$\begin{aligned} \varepsilon^* &= 10^{-20}(\varepsilon')^{14}d\,, & \delta &= 10^{-20}d^4(\varepsilon^*)^{31}\,, \\ \varepsilon &= 10^{-20}(\varepsilon^*)^{30}\delta^6 d^8 \quad \text{and} & c &= 10^{-3}d^2\varepsilon^{10}\delta. \end{aligned} \tag{22}$$

We now "clean up" the partition classes X, Y, Z as follows. First, let $Y' \subseteq Y$ be the set of vertices $y \in Y$ with

$$\deg_\Gamma(y;Z) = (1 \pm \varepsilon)p|Z|\,, \quad \text{and} \quad \deg_\Gamma(y;X) = (1 \pm \varepsilon)p|X|. \tag{23}$$

By Lemma 6 and (22) we have

$$|Y \setminus Y'| \le 2c^2(\log_2 \tfrac{1}{p})^{-1}p^3\varepsilon^{-2}|Y| + 2c^2p^2\varepsilon^{-2}|Y| \le \varepsilon|Y|. \tag{24}$$

We let $X' \subseteq X$ be the set of vertices $x \in X$ with

$$\begin{aligned} &\deg_\Gamma(x;Y') = (1 \pm \varepsilon)p|Y'|\,, \quad \deg_\Gamma(x;Y \setminus Y') \le 2\varepsilon p|Y|\,, \quad \text{and} \\ &\deg_\Gamma(x;Z) = (1 \pm \varepsilon)p|Z|. \end{aligned} \tag{25}$$

Again, by Lemma 6 and (22) we have

$$|X \setminus X'| \le 2 \cdot 8c^2p^2\varepsilon^{-2}|X| + 2c^2p^4\varepsilon^{-2}|X| \le \varepsilon p|X|. \tag{26}$$

By Lemma 9, by (23) and by (26), we obtain

$$(Y', Z) \text{ is } (2\varepsilon, d, p)\text{-regular in } G \quad \text{and} \quad \deg_\Gamma(y;X') = (1 \pm 3\varepsilon)p|X'| \tag{27}$$

for each $y \in Y'$. Moreover, (Y', Z) is $\big(p, 2cp^{5/2}\big(\log_2 \frac{1}{p}\big)^{-1/2}\sqrt{|Y'||Z|}\big)$-bijumbled in Γ. Thus for each $x \in X'$, because $|Y'| \le \deg_\Gamma(x;Y')/(p(1-\varepsilon))$ and $|Z| \le \deg_\Gamma(x;Z)/(p(1-\varepsilon))$ by (25),

$$(Y', Z) \text{ is } \big(p, 2cp^{3/2}(\log_2 \tfrac{1}{p})^{-1/2}\sqrt{|N_\Gamma(x;Y')||N_\Gamma(x;Z)|}\big)\text{-bijumbled in } \Gamma. \tag{28}$$

For $x \in X'$ let

$$Y_x := N_\Gamma(x;Y') \quad \text{and} \quad Z_x := N_\Gamma(x;Z).$$

Define

$$\begin{aligned} X_1^* &:= \left\{x \in X' : d_G(Y_x, Z_x) \ge (d - \varepsilon^*)p \text{ and } (Y_x, Z_x)_G \text{ is not } \big(\tfrac{\varepsilon'}{2}, d, p\big)\text{-regular}\right\}, \\ X_2^* &:= \left\{x \in X' : d_G(Y_x, Z_x) \ge \big(d + (\varepsilon^*)^2\big)p\right\}, \end{aligned}$$

and let $X^* := X_1^* \cup X_2^*$. Finally, let X^{**} be the set of $x \in X' \setminus X^*$ such that $\big(Y_x, Z_x\big)$ has density less than $(d - \varepsilon^*)p$ in G.

We claim that $\big(N_\Gamma(x;Y),Z_x\big)$ is (ε',d,p)-regular in G for all $x\in X'\setminus(X^*\cup X^{**})$. This again follows from Lemma 10, which we apply with $\varepsilon_{\mathrm{L10}}=\frac{1}{2}\varepsilon'$ and $c_{\mathrm{L10}}=2c$, and with $U=Y_x$, $U'=N_\Gamma(x,Y)$, $V=V'=Z_x$. This is possible by (28), because $\big(Y_x,Z_x\big)$ is $(\frac{1}{2}\varepsilon',d,p)$-regular in G by the definition of X^* and X^{**}, and because $|U'|\le\big(1+\frac{1}{10}(\frac{1}{2}\varepsilon')^3\big)|U|$ by a calculation analogous to (20). We conclude that indeed $\big(N_\Gamma(x;Y),Z_x\big)$ is (ε',d,p)-regular in G. Therefore, by (26) it suffices to show that $|X^*|\le\frac{1}{3}\varepsilon'|X|$ and $|X^{**}|\le\frac{1}{3}\varepsilon'|X|$ to complete the proof.

We start with the former. For each $x\in X'$, let

$$P_b^*(x):=\left\{yy'\in\binom{Y_x}{2}:\ yy'\text{ is }\big(Z_x,dp,\delta\big)\text{-bad in }G\right\}.$$

To bound $|X^*|$, we will again estimate $\sum_{x\in X'}|P_b^*(x)|$ in two different ways. The first part is given by the following claim.

Claim 20 $\sum_{x\in X'}|P_b^*(x)|\ge\sum_{x\in X^*}|P_b^*(x)|\ge 10^{-10}(\varepsilon^*)^{30}d^4p^2|X^*||Y'|^2$.

Proof This bound will follow from Lemma 17. We first need to "clean up" the pairs (Y_x,Z_x) for the application of this lemma. Let $Y_x'\subseteq Y_x$ consist of the vertices $y\in Y_x$ with $\deg_\Gamma(y;Z_x)\le 2p|Z_x|$. The pair (Y_x,Z_x) is $(p,2cp^{3/2}\big(\log_2\frac{1}{p}\big)^{-1/2}\sqrt{|Y_x||Z_x|})$-bijumbled since $|Y_x||Z_x|=(1\pm\varepsilon)^2p^2|Y||Z|$ by (28). So Lemma 6 and (22) imply $|Y_x\setminus Y_x'|\le 8c^2\big(\log_2\frac{1}{p}\big)^{-1}p|Y_x|\le\varepsilon p|Y_x|$. Moreover,

$$(Y_x',Z_x)\text{ is }\big(p,4cp^{3/2}\big(\log_2\tfrac{1}{p}\big)^{-1/2}\sqrt{|Y_x'||Z_x|}\big)\text{-bijumbled.}\tag{29}$$

We now first consider vertices $x\in X_1^*$. To bound $|P_b^*(x)|$ we want to apply Lemma 17 to (Y_x',Z_x), using condition (i) of Lemma 17. For this purpose we will first show that (Y_x',Z_x) is also not $\big(\frac{\varepsilon'}{4},d,p\big)$-regular in G. Indeed, by (22) and (29) we can apply the contrapositive of Lemma 10 with $\varepsilon_{\mathrm{L10}}=\frac{\varepsilon'}{4}$ and $c_{\mathrm{L10}}=4c$, and with $U=Y_x'$, $U'=Y_x$, $V=V'=Z_x$ because

$$|Y_x\setminus Y_x'|\le\varepsilon p|Y_x|\le 2\varepsilon p|Y_x'|\le\frac{1}{10}(\varepsilon'/4)^3|Y_x'|.$$

Since (Y_x,Z_x) is not $\big(\frac{\varepsilon'}{2},d,p\big)$-regular in G, this lemma implies that (Y_x',Z_x) is also not $\big(\frac{\varepsilon'}{4},d,p\big)$-regular in G as claimed. Hence, by (22), (29), the definition of X_1^* and the definition of Y_x' we may apply Lemma 17 to (Y_x',Z_x) with input d, $\varepsilon^*_{\mathrm{L17}}=\frac{\varepsilon'}{4}$, δ, $\varepsilon_{\mathrm{L17}}=\varepsilon^*$, $c'=4c$ and p. We conclude that

$$\big|P_b^*(x)\big|\ge\big(\tfrac{\varepsilon'}{4}\big)^{15}d^4|Y_x'|^2\overset{(22),(25)}{\ge}10^{-10}(\varepsilon^*)^{30}d^4p^2|Y'|^2.$$

It remains to consider $x\in X_2^*$. In this case we want to use Lemma 17(ii). To obtain the required density condition, observe that

$$\begin{aligned}e_\Gamma(Y_x\setminus Y_x',Z_x)&\le p\cdot\varepsilon p|Y_x||Z_x|+2cp^{3/2}\big(\log_2\tfrac{1}{p}\big)^{-1/2}\sqrt{|Y_x||Z_x|}\sqrt{\varepsilon p|Y_x||Z_x|}\\&\overset{(22)}{\le}\tfrac{1}{3}(\varepsilon^*)^2p^2|Y_x||Z_x|.\end{aligned}$$

Since $d_G(Y_x, Z_x) \geq \big(d + (\varepsilon^*)^2\big)p$, it follows that $d_G(Y'_x, Z_x) \geq \big(d + \frac{1}{3}(\varepsilon^*)^2\big)p$ by (22). Hence, by (22) and (29) we can apply Lemma 17 to (Y'_x, Z_x) with input d, $\varepsilon^*_{\mathrm{L17}} = \frac{1}{3}(\varepsilon^*)^2$, δ, $\varepsilon_{\mathrm{L17}} = \varepsilon$ and $c' = 4c$ and conclude that

$$P_b^*(x) \geq \big(\tfrac{(\varepsilon^*)^2}{3}\big)^{15} d^4 |Y'_x|^2 \overset{(22)}{\geq} 10^{-10}(\varepsilon^*)^{30} d^4 p^2 |Y'|^2.$$

Summing over all $x \in X^* = X_1^* \cup X_2^*$, the claim follows. ∎

The next claim establishes a complementing upper bound for $\sum_{x\in X'} |P_b^*(x)|$.

Claim 21 $\sum_{x\in X'} |P_b^*(x)| \leq \delta p^2 |X||Y'|^2$.

Proof In order to estimate $\sum_{x\in X'} |P_b^*(x)|$ we will distinguish between the contribution made to this sum by the pairs

$$P_b := \left\{ yy' \in \binom{Y'}{2} : \ yy' \text{ is } \big(Z, dp, \tfrac{\delta}{2}\big)\text{-bad in } G \right\}$$

and that made by the pairs $P_t := \binom{Y'}{2} \setminus P_b$.

For the former let $P_b(x) := \{yy' \in \binom{Y_x}{2} : \ yy' \text{ is } (Z, dp, \frac{\delta}{2})\text{-bad in } G\}$. We use the very rough bound

$$\sum_{x\in X'} |P_b^*(x) \cap P_b| \leq \sum_{x\in X'} |P_b(x)| \,,$$

which holds since $P_b^*(x) \cap P_b \subseteq P_b(x)$ for all $x \in X'$. By Lemma 18 applied to X', Y', Z with parameters d, $\delta_{\mathrm{L18}} = \frac{1}{2}\delta$, $c' = 2c$, $\varepsilon_{\mathrm{L18}} = 3\varepsilon$, which we can do by (22), (27) and since (Y', Z) is $(2\varepsilon, d, p)$-regular in G, we thus have

$$\sum_{x\in X'} |P_b^*(x) \cap P_b| \leq \sum_{x\in X'} |P_b(x)| \leq \tfrac{1}{2}\delta p^2 |X'||Y'|^2 \leq \tfrac{1}{2}\delta p^2 |X||Y'|^2. \tag{30}$$

For the contribution of P_t on the other hand, define

$$V_b(yy') := \{x \in X' : \ \deg_\Gamma\big(x, N_G(y, y'; Z)\big) \geq (1+\delta) d^2 p^2 |Z_x|\}$$

for $yy' \in \binom{Y'}{2}$. Observe that we have $yy' \in P_b^*(x)$ for some $x \in X'$ if and only if $x \in N_\Gamma(y, y'; X')$ and $x \in V_b(yy')$. It follows that

$$\sum_{x\in X'} |P_b^*(x) \cap P_t| \leq \sum_{yy'\in P_t} |V_b(yy')|.$$

Now let $yy' \in P_t$ be fixed. We have $\deg_G(y, y'; Z) \leq \big(1 + \frac{1}{2}\delta\big) d^2 p^2 |Z|$ by definition of P_t. Let $Z_{yy'}$ be a superset of $N_G(y, y'; Z)$ of size $\big(1 + \frac{1}{2}\delta\big) d^2 p^2 |Z|$. By assumption, (X, Z) is $(p, cp^3\sqrt{|X||Z|})$-bijumbled, and so $(X, Z_{yy'})$ is $(p, cd^{-1}p^2\sqrt{|X||Z_{yy'}|})$-bijumbled. Lemma 6, with parameters $\gamma = \varepsilon$, $c' = cd^{-1}$, $k = 2$, then gives

$$\big|\{x \in X' : \ \deg_\Gamma(x; Z_{yy'}) \geq (1+\varepsilon) p |Z_{yy'}|\}\big| \leq 2c^2 d^{-2} p^2 \varepsilon^{-2} |X|. \tag{31}$$

Since

$$(1+\delta)d^2p^2|Z_x| \overset{(25)}{\geq} (1+\delta)d^2p^2(1-\varepsilon)p|Z| \overset{(22)}{\geq} (1+\varepsilon)p|Z_{yy'}|$$

and by the choice of $Z_{yy'}$, the left-hand side of (31) is at least $|V_b(yy')|$. Summing over all $yy' \in P_t$ we conclude

$$\sum_{x\in X'} |P_b^*(x) \cap P_t| \leq \sum_{yy'\in P_t} |V_b(yy')| \leq 2c^2d^{-2}p^2\varepsilon^{-2}|X||Y'|^2 \overset{(22)}{\leq} \tfrac{1}{2}\delta p^2|X||Y'|^2.$$

Together with (30) this proves the claim. ∎

Claims 20 and 21 imply $10^{-10}(\varepsilon^*)^{30}d^4p^2|X^*||Y'|^2 \leq \delta p^2|X||Y'|^2$ and hence

$$|X^*| \leq 10^{10}\delta d^{-4}(\varepsilon^*)^{-30}|X| \leq \varepsilon^*|X| \leq \tfrac{1}{3}\varepsilon'|X|\,, \tag{32}$$

by (22). It remains to bound $|X^{**}|$. Let $X'' := X' \setminus X^*$.

Claim 22 $|X^{**}| \leq 2\varepsilon^*|X''| \leq \frac{1}{3}\varepsilon'|X|$.

Proof Let $\mu := |X^{**}||X''|^{-1}$. We bound μ by considering the number T of triples xyz with $x \in X''$, $y \in Y'$, $z \in Z$ which are such that $xy, xz \in E(\Gamma)$ and $yz \in E(G)$. Observe that $T = \sum_{x\in X''} e_G\big(Y_x, Z_x\big)$ and

$$e_G(Y_x,Z_x) = d_G(Y_x,Z_x)\deg_\Gamma(x;Y')\deg_\Gamma(x;Z) \overset{(25)}{\leq} d_G\big(Y_x,Z_x\big)(1+\varepsilon)^2p^2|Y'||Z|$$

for each $x \in X''$. Since $X^* \cap X'' = \emptyset$ we get by the definition of X^* and X^{**}

$$\begin{aligned} T &\leq \Big(|X^{**}|(d-\varepsilon^*)p + |X''\setminus X^{**}|(d+(\varepsilon^*)^2p)\Big)(1+\varepsilon)^2p^2|Y'||Z| \\ &= \Big(\mu(d-\varepsilon^*) + (1-\mu)(d+(\varepsilon^*)^2)\Big)(1+\varepsilon)^2p^3|X''||Y'||Z| \\ &\leq \big(d - \mu\varepsilon^* + (\varepsilon^*)^2 + 3\varepsilon\big)p^3|X''||Y'||Z|. \end{aligned} \tag{33}$$

For obtaining a lower bound on T, let $Y'' \subseteq Y'$ be the set of vertices $y \in Y'$ with

$$\deg_G(y;Z) \geq (d-\varepsilon)p|Z| \quad \text{and} \quad \deg_\Gamma(y;X'') \geq (1-\varepsilon)p|X''|.$$

Since (Y,Z) is (ε,d,p)-regular in G and (X'',Y) is $(p, 2cp^2\sqrt{|X''||Y|})$-bijumbled in Γ, applying Lemma 6 we obtain $|Y'\setminus Y''| \leq |\varepsilon|Y| + 8c^2p^2\varepsilon^{-2}|Y| \leq 2\varepsilon|Y|$ by (22). Now, each $y \in Y''$ contributes at least $T(y) := e_\Gamma\big(N_\Gamma(y,X''), N_G(y,Z)\big)$ triples to T. As (X,Z) is $(p, cp^3\sqrt{|X||Z|})$-bijumbled the definition of Y'' thus implies that

$$\begin{aligned} T(y) &\geq p\cdot(1-\varepsilon)p|X''|(d-\varepsilon)p|Z| - cp^3\sqrt{|X||Z|}\sqrt{(1-\varepsilon)p|X''|(d-\varepsilon)p|Z|} \\ &\overset{(22)}{\geq} (d-3\varepsilon)p^3|X''||Z|\,, \end{aligned}$$

for each $y \in Y''$. We conclude that

$$T \geq \sum_{y\in Y''} T(y) \geq \big(|Y'| - 2\varepsilon|Y|\big)(d-3\varepsilon)p^3|X''||Z| \geq (d-10\varepsilon)p^3|X''||Y'||Z|.$$

Together with (33) this gives $d - \mu\varepsilon^* + (\varepsilon^*)^2 + 3\varepsilon \geq d - 10\varepsilon$ and so $\mu \leq \varepsilon^* + 13\varepsilon(\varepsilon^*)^{-1} \leq 2\varepsilon^*$ by (22) as desired. ∎

Claim 22 and (32) prove the lemma. ∎

REFERENCES

1. P. Allen, J. Böttcher, J. Ehrenmüller, and A. Taraz, *The bandwidth theorem for sparse graphs*. arXiv: 1612.00661.
2. P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Blow-up lemmas for sparse graphs*. arXiv: 1612.00622.
3. N. Alon, M. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński, and E. Szemerédi, *Universality and tolerance (extended abstract)*. 41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000), IEEE Computer Society Press, Los Alamitos, CA, 2000, pp. 14–21.
4. J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), no. 3, 669–709.
5. J. Böttcher, M. Schacht, and A. Taraz, *Proof of the bandwidth conjecture of Bollobás and Komlós*, Math. Ann. **343** (2009), 175–205.
6. D. Conlon, W.T. Gowers, W. Samotij, and M. Schacht, *On the KŁR conjecture in random graphs*, Israel J. Math. **203** (2014a), no. 1, 535–580.
7. D. Conlon, J. Fox, and Y. Zhao, *Extremal results in sparse pseudorandom graphs*, Adv. Math. **256** (2014b), 206–290.
8. P. Erdös and J. Spencer, *Imbalances in k-colorations*, Networks. **1** (1971/72), 379–385.
9. P. Erdös, M. Goldberg, J. Pach, and J. Spencer, *Cutting a graph into two dissimilar halves*, J. Graph Theory. **12** (1988), no. 1, 121–131.
10. P. Frankl and V. Rödl, *Extremal problems on set systems*, Random Structures Algorithms. **20** (2002), no. 2, 131–164.
11. S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger, *Small subsets inherit sparse ϵ-regularity*, J. Combin. Theory Ser. B. **97** (2007), no. 1, 34–56.
12. S. Gerke and A. Steger, *The sparse regularity lemma and its applications*. Surveys in Combinatorics 2005, London Math. Soc. Lecture Note Ser., Cambridge University Press, Cambridge, Vol. 327, 2005, pp. 227–258.
13. Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*. Foundations of Computational Mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230.
14. Y. Kohayakawa, T. Łuczak, and V. Rödl, *On K^4-free subgraphs of random graphs*, Combinatorica. **17** (1997), no. 2, 173–213.
15. Y. Kohayakawa and V. Rödl, *Szemerédi's regularity lemma and quasi-randomness*. Recent Advances in Algorithms and Combinatorics, CMS Books in Mathematics/Ouvrages de mathématiques de la SMC, Springer, New York, 2003, pp. 289–351.
16. Y. Kohayakawa, V. Rödl, M. Schacht, P. Sissokho, and J. Skokan, *Turán's theorem for pseudo-random graphs*, J. Combin. Theory Ser. A. **114** (2007), no. 4, 631–657.
17. Y. Kohayakawa, V. Rödl, Mathias Schacht, and Jozef Skokan, *On the triangle removal lemma for subgraphs of sparse pseudorandom graphs*.
18. J. Komlós, A. Shokoufandeh, M. Simonovits, and Endre Szemerédi, *The regularity lemma and its applications in graph theory*. Theoretical Aspects of Computer Science (Tehran, 2000), Lecture Notes in Computer Science, Springer, Berlin, Vol. 2292, 2002, pp. 84–112.
19. M. Krivelevich and B. Sudakov, *Pseudo-random graphs*. More Sets, Graphs and Numbers, Bolyai Society Mathematical Studies, Springer, Berlin, Vol. 15, 2006, pp. 199–262.
20. D. Kühn and D. Osthus, *Embedding large subgraphs into dense graphs*. Surveys in Combinatorics 2009, London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, Vol. 365, 2009, pp. 137–167.
21. V. Rödl and M. Schacht, *Regularity lemmas for graphs*. Fete of Combinatorics and Computer Science, Bolyai Society Mathematical Studies, János Bolyai Math. Soc., Budapest, Vol. 20, 2010, pp. 287–325.
22. D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), no. 3, 925–992.
23. A. Scott, *Szemerédi's regularity lemma for matrices and sparse graphs*, Combin. Probab. Comput. **20** (2011), no. 3, 455–466.
24. B. Sudakov and V.H. Vu, *Local resilience of graphs*, Random Structures Algorithms. **33** (2008), no. 4, 409–433.

25. E. Szemerédi, *Regular partitions of graphs*. Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, CNRS, Paris, Vol. 260, 1978, pp. 399–401.
26. A. Thomason, *Pseudorandom graphs*. Random graphs ’85 (Poznań, 1985), North-Holland Mathematics Studies, Amsterdam, North-Holland, Vol. 144, 1987, pp. 307–331.
27. A. Thomason, *Random graphs, strongly regular graphs and pseudorandom graphs*. Surveys in Combinatorics 1987 (New Cross, 1987), London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, Vol. 123, 1987, pp. 173–195.

How to cite this article: Allen P, Böttcher J, Skokan J, Stein M. Regularity inheritance in pseudorandom graphs. *Random Struct Alg.* 2020;56:306–338. https://doi.org/10.1002/rsa.20851

APPENDIX A. | COUNTING LEMMAS

In this appendix we formulate a sparse one-sided counting lemma and a sparse two-sided counting lemma (requiring stronger bijumbledness), which both follow from our inheritance lemmas.

Given a graph H with $V(H) = [m]$, a graph G, and vertex subsets $V_1, \dots, V_m$ of $V(G)$, we write $n(H;G)$ for the number of labeled copies of H in G with i in V_i for each i. Observe that the quantity $n(H;G)$ depends on the choice of the sets $V_1, \dots, V_m$, but this choice will always be clear from the context. Given $0 < p \le 1$, we write

$$d(H;G) := \prod_{ij\in E(H)} d_p(V_i, V_j).$$

Again, this quantity depends on the choice of $V_1, \dots, V_m$, and again this will always be clear from the context.

Still for any given graph H with vertex set $[m]$, which we think of as having order $1, \dots, m$, and given $u, v \in [m]$, we define

$$\begin{aligned} N^+(v) &:= \left\{ w \in N_H(v) : \ w > v \right\}, \\ N^-(v) &:= \left\{ w \in N_H(v) : \ w < v \right\}, \\ N^{<u}(v) &:= \left\{ w \in N_H(v) : \ w < u \right\}. \end{aligned}$$

Finally, we let $k_{\mathrm{reg}}(H)$ be the smallest number with the following properties for each $1 \le i \le m$. For each $j \ge i$ such that $ij \in E(H)$

$$k_{\mathrm{reg}}(H) \ge \tfrac{1}{2}|N^-(i)| + \tfrac{1}{2}|N^{<i}(j)| + \begin{cases} \frac{3}{2} & \text{if } \exists k > i : \ jk \in E(H) \\ 2 & \text{if } \exists k > i : \ jk, ik \in E(H) \\ 3 & \text{if } \exists k > i : \ jk, ik \in E(H) \text{ and} \\ & \qquad |N^{<i}(k)| \le |N^{<i}(j)| \\ 1 & \text{otherwise} \end{cases}$$

and for each $j, j' \geq i$ such that $ij, jj' \in E(H)$

$$k_{\mathrm{reg}}(H) \geq \tfrac{1}{2}|N^{<i}(j)| + \tfrac{1}{2}|N^{<i}(j')| + \begin{cases} 2.501 & \text{if } ij' \in E(H)\,, \\ 2.001 & \text{otherwise.} \end{cases}$$

Informally, the idea is that $(p, cp^{k_{\mathrm{reg}}(H)})$-bijumbledness is enough to use Lemmas 3 and 4 to find copies of H in G one vertex at a time, in the natural order $1, \dots, m$. The following lemma formalizes this.

Lemma 23 (One-sided counting lemma). *For every graph H with $V(H) = [m]$ and every $\gamma > 0$, there exist $\varepsilon, c > 0$ such that the following holds. Let G and Γ be graphs with $G \subseteq \Gamma$, and let $V_1, \dots, V_m$ be subsets of $V(G)$. Suppose that for each edge $ij \in H$, the sets V_i and V_j are disjoint, and the pair (V_i, V_j) is (ε, p)-regular in G and $(p, cp^{k_{\mathrm{reg}}(H)}\sqrt{|V_i||V_j|})$-bijumbled in Γ. Then we have*

$$n(H; G) \geq \big(d(H; G) - \gamma\big)p^{e(H)} \prod_{i \in V(H)} |V_i|.$$

The proof of this lemma is similar to the proof of [7, Lemma X]. It is also contained in the proof of Lemma 24, so we omit the details.

The jumbledness requirement in our two-sided counting lemma depends on another graph parameter, which is also different from the parameter in the two-sided counting lemma in [7] and may appear somewhat exotic at first sight. We shall later compare this parameter to other more common graph parameters.

Let H be given with vertex set $[m]$, which again we think of as having the order $1, \dots, m$. For each $v \in V(H)$, let τ_v be any ordering of $N^+(v)$ such that $|N^{<v}(w)|$ is decreasing. We define

$$\tilde{\mathrm{d}}(H) := \max_{v \in V(H)} \Big(|N^-(v)| + \max_{w \in N^+(v)} \big(\tau_v(w) + |N^{<v}(w)|\big)\Big).$$

The idea is that this parameter controls the bijumbledness we require in order to prove an upper bound on the number copies of H in Γ. In order to count in G, we need both to be able to do this and to use our inheritance lemmas, and we need to consider the same order on $V(H)$ for both.

Lemma 24 (Two-sided counting lemma). *For every graph H with $V(H) = [m]$ and every $\gamma > 0$, there exist $\varepsilon, c > 0$ such that the following holds. We set*

$$\beta = cp^{\max\left(k_{\mathrm{reg}}(H), \frac{1}{2} + \frac{1}{2}\tilde{\mathrm{d}}(H)\right)}. \tag{A.1}$$

Let G and Γ be graphs with $G \subseteq \Gamma$, and let $V_1, \dots, V_m$ be subsets of $V(G)$. Suppose that for each edge $ij \in H$, the sets V_i and V_j are disjoint, and the pair (V_i, V_j) is (ε, p)-regular in G and $(p, \beta\sqrt{|V_i||V_j|})$-bijumbled in Γ. Then we have

$$n(H; G) = \big(d(H; G) \pm \gamma\big)p^{e(H)} \prod_{i \in V(H)} |V_i|.$$

As with Lemma 23, in applications one should choose the order on $V(H)$ so that the resulting β is as large as possible.

For comparison to more standard graph parameters, observe in an optimal order we have

$$\tfrac{\Delta(H)+1}{2} \le \tfrac{1}{2}\tilde{\mathrm{d}}(H) + \tfrac{1}{2} \le \tfrac{\Delta(H)+\mathrm{degen}(H)}{2}\,,$$

where $\mathrm{degen}(H) = \min\{d : \forall H' \subseteq H, \delta(H') \le d\}$ is the *degeneracy* of H. To see that the former inequality is true, observe that for any $v \in V(H)$ we have

$$|N^-(v)| + \max_{w \in N^+(v)} \big(\tau_v(w) + |N^{<v}(w)|\big) \ge |N^-(v)| + |N^+(v)| = d(v)\,,$$

and thus, $\tilde{d}(H) \ge \Delta(H)$. For the latter, consider a degeneracy order on H, that is, an order in which each vertex has at most $\mathrm{degen}(H)$ neighbors preceding it. For such an order, for any v and any $w \in N^+(v)$, we have $|N^{<v}(w)| \le \mathrm{degen}(H) - 1$, since $N^{<v}(w)$ contains neighbors of w preceding w, but not including v. We thus have, for each $v \in V(H)$,

$$|N^-(v)| + \max_{w \in N^+(v)} \big(\tau_v(w) + |N^{<v}(w)|\big) \le |N^-(v)| + |N^+(v)| + \mathrm{degen}(H) - 1.$$

In the similar two-sided counting result of [7], the exponent of p in bijumbledness is

$$\min\big(\tfrac{\Delta(L(H))+4}{2}, \tfrac{\mathrm{degen}(L(H))+6}{2}\big)\,,$$

where $L(H)$ is the *line graph* of H, namely the graph with vertex set $E(H)$ in which two vertices are adjacent if they are incident as edges of H. It is easy to check that this parameter is bounded between $\frac{\Delta(H)+3}{2}$ and $\frac{\Delta(H)+\mathrm{degen}(H)+4}{2}$ (and both bounds can be sharp).

We now briefly outline the proof of Lemma 24. We prove this statement by induction. We count the number of copies of H in G, which is a subgraph of the bijumbled Γ, by embedding H one vertex at a time, and bounding the number of choices at each step. Most of the time, we will choose to embed to vertices which maintain regularity, and thus we can accurately estimate the number of choices. This part of the proof is very similar to the usual proof of the counting lemma for dense graphs, except that we use Lemmas 3 and 4 to argue that regularity is usually maintained rather than this being a triviality. To deal with the exceptional event that we embed to a vertex and regularity is lost, we require an upper bound on H-copies in Γ. This is the content of the following Lemma 25.

Given H with $V(H) = [m]$ in an order realizing $\tilde{\mathrm{d}}(H)$, and $x \in V(H)$, let

$$H^{\ge x} := H\big[\{y \in V(H) : y \ge x\}\big].$$

Lemma 25 *Given H with vertex set $[m]$ and $0 < p < \frac{1}{10}$, suppose*

$$\beta \le \tfrac{1}{2}\varepsilon(50\Delta(H))^{-\Delta(H)} p^{\frac{1}{2}+\frac{1}{2}\tilde{\mathrm{d}}(H)}.$$

Let $V_1, \ldots, V_m$ be subsets of $V(\Gamma)$, and suppose (V_i, V_j) is $(p, \beta\sqrt{|V_i||V_j|})$-bijumbled in Γ for each $ij \in E(H)$. Let $x \in V(H)$, and for each $y \in V\big(H^{\ge x}\big)$, let $W_y \subseteq V_y$ satisfy $|W_y| \ge \varepsilon p^{|N^{<x}(y)|}|V_y|$. Then the number of copies of $H^{\ge x}$ in Γ with y in the set W_y for each y is at most

$$(4p)^{e(H_\pi^{\ge x})} \prod_{x \le y \le m} |W_y|.$$

We now show how this implies Lemma 24

Proof of Lemma 24 Suppose that Γ, G, and H are as in the lemma statement. We will prove by induction the following statement (†).

For every $\gamma' > 0$ there exist $\varepsilon', c' > 0$ with the following property. Given $x \in V(H)$, for each $y \in V(H^{\geq x})$, let $W_y \subseteq V_y$ satisfy $|W_y| \geq \varepsilon' p^{|N^{<x}(y)|}|V_y|$. Suppose that for each $ij \in E(H)$ with $i,j \geq x$, the pair (W_i, W_j) has p-density $d_p(W_i, W_j) \geq \gamma'$, is $(\varepsilon', d_p(V_i, V_j), p)$-regular in G, and $(p, \beta\sqrt{|V_i||V_j|})$-bijumbled in Γ with $\beta = c' p^{\max\left(k_{\mathrm{reg}}(H), \frac{1}{2}+\frac{1}{2}\tilde{\mathrm{d}}(H)\right)}$. Then the number of copies of $H^{\geq x}$ in G with $y \in W_y$ for each y is

$$\left(d(H^{\geq x}; G) \pm \gamma'\right) p^{e(H^{\geq x})} \prod_{x \leq i \leq m} |W_i|.$$

Before we prove this, we show that it implies the statement of Lemma 24. To that end, we assume (†) holds with input $\gamma' = \frac{\gamma}{4}$ and $x = 1$, returning constants c' and ε'. We set $c = c'$ and $\varepsilon = \frac{1}{2}\varepsilon'$, and use $W_y := V_y$ for each $y \in V(H)$.

If for each $ij \in E(H)$ we have $d_p(V_i, V_j) \geq \frac{\gamma}{4}$, then by (†) the lemma statement follows. We may therefore assume that there is some $ij \in E(H)$ such that $d_p(V_i, V_j) < \frac{\gamma}{4}$, and thus $d(H; G) < \frac{\gamma}{2}$. To establish the lemma statement it thus suffices to show

$$n(H; G) \leq \gamma p^{e(H)} \prod_{1 \leq i \leq m} |V_i|.$$

We generate a graph G' with $G \subseteq G' \subseteq \Gamma$ by adding edges of Γ to each (ε, p)-regular pair (V_i, V_j) with $d_p(V_i, V_j; G) < \frac{\gamma}{4}$. We do this by choosing edges of Γ in such pairs uniformly at random with probability $\frac{3\gamma}{8} - d_p(V_i, V_j)$. We claim that the result is that any such pair (V_i, V_j) is $(2\varepsilon, p)$-regular in G' with density between $\frac{\gamma}{4}p$ and $\frac{\gamma}{2}p$. The proof of this claim is a standard application of the Chernoff bound, and we omit it. Since there is a pair in G' with density less than $\frac{\gamma}{2}p$, we have $d(H; G') \leq \frac{3\gamma}{4}$. By construction we have $n(H; G) \leq n(H; G')$, and by (†) the lemma statement follows.

We now prove (†) by induction on x. The base case $x = m$ is trivial, with $\varepsilon'' = c'' = 1$, so we assume $x < m$. Given γ', we set $\gamma'' = \frac{\gamma'}{2}$. Let ε'' and c'' be returned by (†) for input γ'' and $x + 1$. Without loss of generality we assume $\varepsilon'' < 4^{-m^2}\gamma'/(24qm)$. We set $\varepsilon' > 0$ small enough for Lemmas 3 and 4 with input $\frac{1}{2}\gamma'\varepsilon''$ and γ', and such that $(1 + \gamma'^{-1}\varepsilon')^m < 1 + \frac{\gamma'}{8}$. We suppose that $c' \leq \min\left(\varepsilon'^3, c''\right)$ is small enough for both these applications and for Lemma 25. When $N_{H^{\geq x}}(x) = \emptyset$, statement (†) for x and γ' follows trivially by (†) with input $x + 1$ and γ'. Thus, suppose $N_{H^{\geq x}}(x) = \{y_1, \dots, y_q\}$ for some $q \geq 1$.

If $v \in W_x$ fails to satisfy any of the following conditions, we say v is bad.

(a) For each i we have $\deg_G(v, W_{y_i}) = (d_p(V_x, V_{y_i}) \pm \varepsilon')p|W_{y_i}|$.
(b) For each i we have $\deg_\Gamma(v, W_{y_i}) = (1 \pm \varepsilon')p|W_{y_i}|$.
(c) For each i and $z > x$ such that $y_i z \in E(H^{\geq x})$, the pair $\left(N_G(v, W_{y_i}), W_z\right)$ is $\left(\varepsilon'', d_p(V_{y_i}, V_z), p\right)$-regular in G.
(d) For each $i \neq j$ such that $y_i y_j \in E(H^{\geq x})$ the pair $\left(N_G(v, W_{y_i}), N_G(v, W_{y_j})\right)$ is $\left(\varepsilon'', d_p(V_{y_i}, V_{y_j}), p\right)$-regular in G.

Let $B \subseteq W_x$ be the set of bad vertices. We now show that B is much smaller than W_x. Since for each $1 \leq i \leq q$ the pair $\left(W_x, W_{y_i}\right)$ is $\left(\varepsilon', d_p(V_1, V_{y_i}), p\right)$-regular, there are at most $2q\varepsilon'|W_x|$ vertices in W_x for which (a) fails.

For the remaining estimates, it is convenient to estimate the bijumbledness of pairs (W_i, W_j) for $i,j \geq x$. Specifically, since (V_i, V_j) is $\big(p, \beta\sqrt{|V_i||V_j|}\big)$-bijumbled, and since $|W_i| \geq \varepsilon' p^{|N^{<x}(i)|}|V_i|$, and $|W_j| \geq \varepsilon' p^{|N^{<x}(j)|}|V_j|$, we conclude that

$$(W_i, W_j) \text{ is } \Big(p, \varepsilon'^{-1}p^{-\frac{1}{2}|N^{<x}(i)|-\frac{1}{2}|N^{<x}(j)|}\beta\sqrt{|W_i||W_j|}\Big)\text{-bijumbled.} \tag{A.2}$$

Next, let $Z \subseteq W_x$ be the set of vertices with at least $(1+\varepsilon')p|W_{y_i}|$ neighbors in W_{y_i}. By (A.2), we have

$$(1+\varepsilon')p|W_{y_i}||Z| - p|W_{y_i}||Z| \leq \varepsilon'^{-1}p^{-\frac{1}{2}|N^{<x}(y_i)|-\frac{1}{2}|N^{<x}(x)|}\beta\sqrt{|W_x||W_{y_i}||W_{y_i}||Z|}.$$

Since $\beta \leq c' p^{k_{\mathrm{reg}}(H)}$, and by definition of k_{reg}, we conclude $|Z| \leq \varepsilon'|W_x|$. A similar argument applies to the set of vertices in W_x with at most $(1-\varepsilon')p|W_{y_i}|$ neighbors in W_{y_i} in Γ, so there are at most $2q\varepsilon'|W_x|$ vertices in W_x for which (b) fails.

We move on to the regularity statements. By (A.2) and definition of k_{reg}, the bijumbledness requirements of Lemma 3 are satisfied, so since $d_p(V_{y_i}, V_z) \geq \gamma'$ and $\big(V_{y_i}, V_z\big)$ is (ε', p)-regular, the number of vertices $v \in W_x$ such that $\big(N_\Gamma(v, V_{y_i}), V_z\big)$ is not $\big(\frac{1}{2}\gamma'\varepsilon'', d_p(V_{y_i}, V_z), p\big)$-regular is at most $qm\varepsilon''|W_x|$. Similarly, the bijumbledness requirements of Lemma 4 are satisfied. Again, the number of vertices $v \in W_x$ such that $\big(N_\Gamma(v, V_{y_i}), N_\Gamma(v, V_{y_j})\big)$ is not $\big(\frac{1}{2}\gamma'\varepsilon'', d_p(V_{y_i}, V_{y_j}), p\big)$-regular is at most $qm\varepsilon''|W_x|$. Now suppose that v satisfies (a) and (b). By choice of ε' and by Lemma 9, if $\big(N_\Gamma(v, V_{y_i}), V_z\big)$ is $\big(\frac{1}{2}\gamma'\varepsilon'', d_p(V_{y_i}, V_z), p\big)$-regular then $\big(N_G(v, V_{y_i}), V_z\big)$ is $\big(\varepsilon'', d_p(V_{y_i}, V_z), p\big)$-regular, if $\big(N_\Gamma(v, V_{y_i}), N_\Gamma(v, V_{y_j})\big)$ is $\big(\frac{1}{2}\gamma'\varepsilon'', d_p(V_{y_i}, V_{y_j}), p\big)$-regular then $\big(N_G(v, V_{y_i}), N_G(v, V_{y_j})\big)$ is $\big(\varepsilon'', d_p(V_{y_i}, V_{y_j}), p\big)$-regular. We conclude that in total at most $2qm\varepsilon''|W_x|$ vertices of W_x which satisfy (a) and (b) fail either (c) or (d), so $|B| \leq 2qm(2\varepsilon' + \varepsilon'')|W_x|$.

Now given any $v \in W_x \setminus B$, we wish to estimate the number of copies of $H^{\geq x}$ in G such that x is mapped to v and i is in W_i for each $x+1 \leq i \leq m$. In other words, we need to know the number of copies of $H^{\geq x+1}$ such that i is in W'_i for each $x+1 \leq i \leq m$, where $W'_i = W_i$ if $xi \notin E(H)$ and $W'_i = N_G(v) \cap W_i$ if $xi \in E(H)$. Because v satisfies (c) and (d), for each $ij \in E(H^{\geq x+1})$ the pair (W'_i, W'_j) is (ε'', p)-regular in G. We now use the induction hypothesis. By choice of c', we can apply (†) with input $x+1$ and $\frac{\gamma'}{2}$ to obtain

$$\begin{aligned} n(H^{\geq x+1}; G) &= \big(d(H^{\geq x+1}; G) \pm \tfrac{\gamma'}{2}\big)p^{e(H^{\geq x+1})} \prod_{x+1\leq i\leq m} |W'_i| \\ &= \big(d(H^{\geq x}; G) \pm \tfrac{\gamma'}{2}\big)p^{e(H^{\geq x})}(1 \pm \varepsilon'\gamma'^{-1})^q \prod_{x+1\leq i\leq m} |W_i|, \end{aligned}$$

where the second line uses the fact that (W'_i, W'_j) is $\big(\varepsilon'', d_p(V_i, V_j), p\big)$-regular for each $ij \in E(H^{\geq x+1})$ by (c) and (d), and the fact $|W'_i| = \big(d_p(V_x, V_i) \pm \varepsilon'\big)p|W_i|$ for each i such that $xi \in E(H^{\geq x})$. We conclude that the number of copies of $H^{\geq x}$ in G with x mapped to $W_x \setminus B$ and i mapped to W_i for each $x+1 \leq i \leq m$ is

$$\begin{aligned} \big(|W_x| - |B|\big)&\big(d(H^{\geq x}; G) \pm \tfrac{\gamma'}{2}\big)p^{e(H^{\geq x})}(1 \pm \varepsilon'\gamma'^{-1})^q \prod_{x+1\leq i\leq m} |W_i| \\ &= \big(d(H^{\geq x}; G) \pm \tfrac{3\gamma'}{4}\big)p^{e(H^{\geq x})} \prod_{x\leq i\leq m} |W_i|, \end{aligned}$$

where the second line follows by choice of ε'. This gives the desired lower bound; it only remains to complete the upper bound by showing that the number of copies of $H^{\geq x}$ in G with x in B and i in W_i for each $x+1 \leq i \leq m$ is at most

$$\frac{\gamma'}{4} p^{e(H^{\geq x})} \prod_{x \leq i \leq m} |W_i|.$$

We may assume $|B| = 6qm\varepsilon''|W_x|$ by, if necessary, adding arbitrary vertices of W_x. By choice of c' and (A.1), the jumbledness requirements of Lemma 25 are satisfied, so by that lemma the number of copies of $H^{\geq x}$ in Γ with x in B and i in W_i for each $x+1 \leq i \leq m$ is at most

$$(4p)^{e(H^{\geq x})}|B| \prod_{x \leq i \leq m} |W_i| \leq \frac{\gamma'}{4} p^{e(H^{\geq x})} \prod_{x \leq i \leq m} |W_i|,$$

where the inequality is true by choice of ε''. ∎

It remains to prove Lemma 25. In this proof we will need to optimize over certain configurations. This optimization problem is captured in the following lemma, where it is phrased as the problem of bounding a certain sum of real numbers.

Lemma 26 *Let $0 < p \leq \frac{1}{10}$ be real, $q \geq 1$ be an integer and $b_1 \geq \cdots \geq b_q$ be nonnegative integers. Let $P := \lfloor \log(p^{-1}) \rfloor$ and $C := \max_{1 \leq i \leq q}(b_i + i)$. Let $A := [0, P]^q \setminus \{\mathbf{0}\}$ be the set of nonzero q-dimensional vectors with integer entries between 0 and P. Then*

$$\sum_{\boldsymbol{\alpha} \in A} \frac{2^{\alpha_1 + \cdots + \alpha_q}}{\max_{i:\alpha_i \neq 0} 2^{2\alpha_i} p^{b_i}} \leq (50q)^q p^{1-C}. \tag{A.3}$$

Proof Given $\boldsymbol{\alpha} \in A$ let

$$M(\boldsymbol{\alpha}) := \frac{2^{\alpha_1 + \cdots + \alpha_q}}{\max_{i:\alpha_i \neq 0} 2^{2\alpha_i} p^{b_i}}. \tag{A.4}$$

We first establish the following bounds on $M(\boldsymbol{\alpha})$.

Claim 27 *For each $\boldsymbol{\alpha} \in A$ one of the following holds.*

(i) $M(\boldsymbol{\alpha}) \leq 2^{-\alpha_1 - \cdots - \alpha_q} p^{-b_1}$.
(ii) $M(\boldsymbol{\alpha}) \leq p^{-C + \frac{5}{4}}$.

Proof of Claim 27 It simplifies our arguments to pass to an optimization over real-valued variables: Let $\tilde{A}$ be the set of nonzero vectors from $\mathbb{R}^q$ with entries between 0 and P. Then, $M(\boldsymbol{\alpha})$ for $\boldsymbol{\alpha} \in \tilde{A}$ is defined as in (A.4).

We prove the claim by finding for each $\boldsymbol{\alpha} \in A$ an $\boldsymbol{\alpha}' \in \tilde{A}$ of simple structure such that $M(\boldsymbol{\alpha})$ and $M(\boldsymbol{\alpha}')$ are related suitably. We construct $\boldsymbol{\alpha}'$ by applying the following three operations successively until no further operation is possible. Each of these operations takes a vector $\tilde{\boldsymbol{\alpha}} \in \tilde{A}$ and returns a new vector $\tilde{\boldsymbol{\alpha}}' \in \tilde{A}$.

The first operation applies when there are two entries $\tilde{\alpha}_i$ and $\tilde{\alpha}_j$ such that $\tilde{\alpha}_i$ does not realize the maximum in $M(\tilde{\boldsymbol{\alpha}})$ (by this we mean the term $\max_{i:\alpha_i \neq 0} 2^{2\alpha_i} p^{b_i}$), but $\tilde{\alpha}_j$ does, and moreover, we have $2^{2\tilde{\alpha}_i} p^{b_i} < 2^{2\tilde{\alpha}_j} p^{b_j}$ (which might not be the case if $\tilde{\alpha}_i = 0$). Then we can increase $\tilde{\alpha}_i$ until it contributes the same value to the maximum as $\tilde{\alpha}_j$ or hits P. More precisely, if there exist i and j such that $\tilde{\alpha}_i < P$, such that $\tilde{\alpha}_j > 0$ and such that $2^{2\tilde{\alpha}_i}p^{b_i} < 2^{2\tilde{\alpha}_j}p^{b_j} = \max_{k:\alpha_k\neq 0} 2^{2\alpha_k}p^{b_k}$, then we can take $\tilde{\boldsymbol{\alpha}}'$ equal to $\tilde{\boldsymbol{\alpha}}$ at all entries except the ith, and

$$\tilde{\alpha}_i' = \min\left(P, \tfrac{1}{2}\log(2^{2\tilde{\alpha}_j}p^{b_j-b_i})\right).$$

Then the maximum in $M(\tilde{\boldsymbol{\alpha}}')$ is equal to that in $M(\tilde{\boldsymbol{\alpha}})$ because

$$2^{2\tilde{\alpha}_i'}p^{b_i} \leq 2^{2\tilde{\alpha}_j}p^{b_j-b_i}p^{b_i} = 2^{2\tilde{\alpha}_j}p^{b_j} = 2^{2\tilde{\alpha}_j'}p^{b_j}.$$

Moreover $\tilde{\alpha}_i' > \tilde{\alpha}_i$, so $\tilde{\alpha}_1' + \cdots + \tilde{\alpha}_q' \geq \tilde{\alpha}_1 + \cdots + \tilde{\alpha}_q$ and $M(\tilde{\boldsymbol{\alpha}}') > M(\tilde{\boldsymbol{\alpha}})$.

The second operation applies when there are two entries strictly between 0 and P. In this case we can increase both entries by the same amount until one hits P. More precisely, if there exist i and j with $0 < \tilde{\alpha}_i \leq \tilde{\alpha}_j < P$, then we can take $\tilde{\boldsymbol{\alpha}}'$ equal to $\tilde{\boldsymbol{\alpha}}$ at all entries except the ith and jth. We set $\tilde{\alpha}_j' = P$, and $\tilde{\alpha}_i' = \tilde{\alpha}_i + P - \tilde{\alpha}_j$. Then the maximum in $M(\tilde{\boldsymbol{\alpha}}')$ is greater than the one in $M(\tilde{\boldsymbol{\alpha}})$ by a factor of at most $\max\{2^{2(\tilde{\alpha}_j'-\tilde{\alpha}_j)}, 2^{2(\tilde{\alpha}_i'-\tilde{\alpha}_i)}\} = 2^{2P-2\tilde{\alpha}_j}$, while the sum of the entries of $\tilde{\boldsymbol{\alpha}}'$ is greater by $2P - 2\tilde{\alpha}_j$ than that of $\tilde{\boldsymbol{\alpha}}$. So again $\tilde{\alpha}_1' + \cdots + \tilde{\alpha}_q' \geq \tilde{\alpha}_1 + \cdots + \tilde{\alpha}_q$ and $M(\tilde{\boldsymbol{\alpha}}') \geq M(\tilde{\boldsymbol{\alpha}})$.

The third operation is the only operation that decreases coordinates: It decreases a coordinate if it is the only coordinate realizing the maximum in $M(\tilde{\boldsymbol{\alpha}})$ until it contributes as much to the maximum as some other coordinate or hits 0. More precisely, it applies if there are at least two nonzero entries in $\tilde{\boldsymbol{\alpha}}$, and j is the unique coordinate realizing $\max_{i:\tilde{\alpha}_i>0} 2^{2\tilde{\alpha}_i}p^{b_i}$. Let s be the second greatest value of $2^{\tilde{\alpha}_i}p^{b_i}$ over i such that $\tilde{\alpha}_i > 0$, and set

$$c = \min\left(\tilde{\alpha}_j, \tfrac{1}{2}\log(sp^{-b_j})\right).$$

Then let $\tilde{\boldsymbol{\alpha}}'$ be equal to $\tilde{\boldsymbol{\alpha}}$ in all coordinates except $\tilde{\alpha}_j' = \tilde{\alpha}_j - c$. The maximum in $M(\tilde{\boldsymbol{\alpha}}')$ is less than that in $M(\tilde{\boldsymbol{\alpha}})$ by a factor of at least 2^{2c} (if $c = \tilde{\alpha}_j$ the factor may be greater). We conclude that $\tilde{\alpha}_1' + \cdots + \tilde{\alpha}_q' = \tilde{\alpha}_1 + \cdots + \tilde{\alpha}_q - c$ and $M(\tilde{\boldsymbol{\alpha}}') \geq 2^c M(\tilde{\boldsymbol{\alpha}})$.

Now consider the result $\boldsymbol{\alpha}'$ of the successive application of these operations until none of them can be applied anymore. (Observe that eventually such a final $\boldsymbol{\alpha}'$ must be reached, since the way the three operations are defined prevents the process from going on indefinitely.) Clearly $M(\boldsymbol{\alpha}') \geq M(\boldsymbol{\alpha})$ since no operation decreases $M(\cdot)$. For the structure of $\boldsymbol{\alpha}'$ there are two possibilities. On the one hand, $\boldsymbol{\alpha}'$ could have exactly one nonzero coordinate α_ℓ. Then, since the third operation is the only operation that decreases coordinates, this operation was applied to each coordinate which was nonzero in $\boldsymbol{\alpha}$ but coordinate ℓ. As the first two operations increase the sum of coordinates of $\boldsymbol{\alpha}$ and increase the value of $M(\cdot)$, and the third operation increases the value of $M(\cdot)$ by at least a factor of 2^c, we conclude $M(\boldsymbol{\alpha}) \leq 2^{-\alpha_1-\cdots-\alpha_q+\alpha_\ell'}M(\boldsymbol{\alpha}')$. Clearly $M(\boldsymbol{\alpha}') = 2^{-\alpha_\ell'}p^{-b_\ell} \leq 2^{-\alpha_\ell'}p^{-b_1}$, and hence $M(\boldsymbol{\alpha}) \leq 2^{-\alpha_1-\cdots-\alpha_q}p^{-b_1}$, so (i) is satisfied.

On the other hand, it could be that $\boldsymbol{\alpha}'$ has at least two nonzero coordinates. Observe that at most one entry of $\boldsymbol{\alpha}'$ is not in $\{0, P\}$ by the second operation. By the third operation there are also at least two coordinates realizing the maximum in $M(\boldsymbol{\alpha}')$, one of which has value P. If there is a coordinate j such that $0 < \alpha_j' < P$ then this coordinate also realizes the maximum by the first operation. So it follows from $b_{i+1} \geq b_i$ that and the first operation that there is an index $\ell \geq 1$ such that for all $i \leq \ell$ we have $\alpha_i' = P$, for all $i > \ell + 1$ we have $\alpha_i' = 0$, and $\alpha_{\ell+1}' < P$. Now if $\alpha_{\ell+1}' > 0$, then it realizes the maximum and thus $2^{2\alpha_{\ell+1}'}p^{b_{\ell+1}} = 2^{2P}p^{b_\ell}$. Since b_ℓ and $b_{\ell+1}$ are integers, this equation can only be solved if $b_\ell - b_{\ell+1}$ is equal to 0, 1 or 2. In the first case we obtain $\alpha_{\ell+1}' = P$, contradicting the definition of ℓ. In both of the other two cases we have $\alpha'_{\ell+1} \le \frac{3}{4}P$. It follows that $\alpha'_{\ell+1} \le \frac{3}{4}P$. Clearly we have $\max_{i:\alpha_i\neq 0} 2^{2\alpha_i}p^{b_i} = 2^{2P}p^{b_\ell}$, and so

$$M(\boldsymbol{\alpha}) \le M(\boldsymbol{\alpha}') \le \frac{2^{\ell P+\frac{3}{4}P}}{2^{2P}p^{b_\ell}} = 2^{(\ell-\frac{5}{4})P}p^{-b_\ell} \le p^{-\ell+\frac{5}{4}-b_\ell} \le p^{-C+\frac{5}{4}},$$

because $C = \max_{1\le i\le q}(b_i + i)$. Hence (ii) holds. ∎

Now consider first all $\boldsymbol{\alpha} \in A$ for which Claim 27(i) holds. Since $M(\boldsymbol{\alpha}) \le 2^{-\alpha_1-\cdots-\alpha_q}p^{-b_1}$, the contribution to (A.3) of all such $\boldsymbol{\alpha}$ is at most $\sum_{K=1}^{qP}(K+q)^q 2^{-K}p^{-b_1}$, where we used that the number of vectors in A whose entries sum to K is at most $\binom{K+q-1}{q-1} < (K+q)^q$. Since $(1+\frac{K}{10q})^q \le \exp(K/10) \le 2^K$ we have $(K+q)^q 2^{-K} \le (10q)^q$ and hence the contribution to (A.3) is at most

$$\sum_{K=1}^{qP}(K+q)^q 2^{-K}p^{-b_1} \le q(10q)^q p^{-b_1}.$$

Finally, consider all $\boldsymbol{\alpha} \in A$ for which Claim 27(ii) holds. We require a preliminary estimate. For all $z > 1$ we have $1 + \log_e z - z < 0$, since this function is equal to zero at $z = 1$ and has first derivative $\frac{1}{z} - 1$ which is negative for all $z > 1$. It follows that for any $z > 1$, if $x \ge e^{4q}$ we have

$$\left(2\log_2(x^z)\right)^q x^{-z/4} \le \left(2\log x\right)^q x^{-1/4} z^q e^{q(1-z)} < \left(2\log x\right)^q x^{-1/4},$$

where we used $1 + \log_e z - z < 0$ to establish the second inequality. It follows that

$$(2\log_2 x)^q x^{-1/4} \le (16q)^q,$$

holds for all $x \ge 1$. To see that this is true, observe that the left hand side is trivially at most the claimed bound when $1 \le x \le e^{4q}$ (since the term $x^{-1/4}$ is at most one), and strictly decreasing for $x \ge e^{4q}$ by the previous calculation.

Since $M(\boldsymbol{\alpha}) \le p^{-C+\frac{5}{4}}$ the contribution to (A.3) of such $\boldsymbol{\alpha}$ is at most

$$(P+1)^q p^{-C+\frac{5}{4}} \le (2\log(p^{-1}))^q p^{-C+\frac{5}{4}} \le (16q)^q p^{-C+1},$$

where we used the above estimate for the second inequality. We obtain (A.3). ∎

With the help of Lemma 26 we can now prove Lemma 25.

Proof of Lemma 25 We prove the statement by induction on x. The base case $x = m$ is trivial, so suppose that $1 \le x \le m-1$, and for an induction hypothesis that the lemma statement holds for $x+1$.

In the case $N^+(x) = \emptyset$, the statement follows by applying the induction hypothesis and the same sets W_y for $y > x$. Thus we can assume $\left|N^+(x)\right| = q \ge 1$. Let $N^+(x) = \{y_1, \dots, y_q\}$ in an order such that $\left|N^{<x}(y_i)\right| \ge \left|N^{<x}(y_j)\right|$ whenever $i < j$.

For a fixed $v \in W_x$ we obtain the following estimate of $H^{\ge x}$ copies using this vertex. For $1 \le i \le q$ set $W'_{y_i} := N_\Gamma(v; W_{y_i})$, and possibly add some arbitrary vertices of W_{y_i} to W'_{y_i} until $|W'_{y_i}| \ge p|W_{y_i}|$. For all $y \notin N^+(x)$ with $y > x$ set $W'_y := W_y$. Then by induction, the number of copies of $H^{\ge x+1}$ in Γ with y

mapped to W'_y for each y is at most $(4p)^{e(H^{\geq x})-q}\prod_{y>x}|W'_y|$. It follows that the number of copies of $H^{\geq x}$ with x mapped to v and y mapped to W_y for each y is at most

$$(4p)^{e(H^{\geq x})-q}\Big(\prod_{1\leq i\leq q}\min\big\{p|W_{y_i}|,\deg_\Gamma(v;W_{y_i})\big\}\Big)\prod_{y>x,y\notin N^+(x)}|W_y|. \tag{A.5}$$

We next partition W_x as follows. Given $\boldsymbol{\alpha}\in[0,\lfloor\log p^{-1}\rfloor]^q$ with integer entries, we let $B_{\boldsymbol{\alpha}}$ be the set of vertices $v\in W_x$ such that for each $1\leq i\leq q$, either $\alpha_i=0$ and we have $\deg_\Gamma(v,W_{y_i})\leq 2p|W_{y_i}|$, or $\alpha_i>0$ and we have

$$2^{\alpha_i}p|W_{y_i}|<\deg_\Gamma(v,W_{y_i})\leq 2^{\alpha_i+1}p|W_{y_i}|.$$

Note that this is a partition because $2^{\lfloor\log p^{-1}\rfloor+1}p>1$.

Using $|B_{\mathbf{0}}|\leq|W_x|$ and (A.5) with $\deg_\Gamma(v;W_{y_i})\leq 2p|W_{y_i}|$ for each i, we can immediately bound the number of copies of $H^{\geq x}$ in Γ with x mapped to $B_{\mathbf{0}}$ and y mapped to W_y for each $y>x$ from above by

$$|W_x|(4p)^{e(H^{\geq x})-q}(2p)^q\prod_{y>x}|W_y|\leq\tfrac{1}{2}(4p)^{e(H^{\geq x})}\prod_{x\leq y\leq m}|W_y|. \tag{A.6}$$

It remains to establish an analogous bound for the sets $B_{\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha}\neq\mathbf{0}$. For this we use the jumbledness of Γ. Given $\boldsymbol{\alpha}$ and some $1\leq i\leq q$ such that $\alpha_i\neq 0$, we have $e\big(B_{\boldsymbol{\alpha}},W_{y_i}\big)\geq 2^{\alpha_i}p|B_{\boldsymbol{\alpha}}||W_{y_i}|$, and since $\alpha_i\geq 1$ it follows that

$$e\big(B_{\boldsymbol{\alpha}},W_{y_i}\big)-p\big|B_{\boldsymbol{\alpha}}\big|\big|W_{y_i}\big|\geq\tfrac{1}{2}\cdot 2^{\alpha_i}p|B_{\boldsymbol{\alpha}}||W_{y_i}|.$$

Since $\big(V_x,V_{y_i}\big)$ is $\big(p,\beta\sqrt{|V_x||V_{y_i}|}\big)$-bijumbled, this implies

$$\tfrac{1}{2}\cdot 2^{\alpha_i}p|B_{\boldsymbol{\alpha}}||W_{y_i}|\leq\beta\sqrt{|V_x||V_{y_i}|}\sqrt{|B_{\boldsymbol{\alpha}}||W_{y_i}|}.$$

Rearranging this we obtain

$$|B_{\boldsymbol{\alpha}}|\leq\frac{4\beta^2|V_x||V_{y_i}|}{2^{2\alpha_i}p^2|W_{y_i}|}.$$

Since this holds for each i with $\alpha_i>0$, using (A.5) with $\deg_\Gamma(v;W_{y_i})\leq 2^{\alpha_i+1}p|W_{y_i}|$ for each i, the number of ϕ-partite copies of $H^{\geq x}$ in Γ with x in $B_{\boldsymbol{\alpha}}$ and $y\in W_y$ for each $y>x$ is at most

$$\begin{aligned}&\Big(\min_{i:\alpha_i>0}\frac{4\beta^2|V_x||V_{y_i}|}{2^{2\alpha_i}p^2|W_{y_i}|}\Big)(4p)^{e(H^{\geq x})-q}\Big(\prod_{i=1}^{q}2^{\alpha_i+1}p\Big)\prod_{y>x}|W_y|\\&\leq\Big(\min_{i:\alpha_i>0}\frac{4\beta^2|V_x||V_{y_i}|}{2^{2\alpha_i}p^2|W_{y_i}|}\Big)\Big(\prod_{i=1}^{q}2^{\alpha_i+1}\Big)4^{-q}(4p)^{e(H^{\geq x})}\prod_{y>x}|W_y|.\end{aligned}$$

Since $|W_{y_i}|\geq\varepsilon p^{|N^{<x}(y_i)||V_{y_i}|}$ this is at most

$$\frac{2^{\alpha_1+\cdots+\alpha_q}}{\max_{i:\alpha_i>0}2^{2\alpha_i}p^{|N^{<x}(y_i)|}}\cdot 4\beta^2\varepsilon^{-1}|V_x|p^{-2}2^q\cdot 4^{-q}(4p)^{e(H^{\geq x})}\prod_{y>x}|W_y|.$$

By Lemma 26, with $b_i = |N^{<x}(y_i)|$ for each $1 \leq i \leq q$, letting $C = \max_{i=1}^{q}(b_i + i)$, the sum of these terms over all $B_{\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha} \neq \mathbf{0}$ is at most

$$(50q)^q p^{1-C} \cdot 4\beta^2 \varepsilon^{-1} |V_x| p^{-2} (4p)^{e(H^{\geq x})} \prod_{y>x} |W_y|.$$

We have $|V_x| \leq \varepsilon^{-1} p^{-|N^-(x)|} |W_x|$, and $C + |N^-(x)| \leq \tilde{\mathrm{d}}(H)$ by definition, so this is bounded above by

$$4\beta^2 \varepsilon^{-2} (50q)^q p^{-1-\tilde{\mathrm{d}}(H)} (4p)^{e(H^{\geq x})} \prod_{x \leq y \leq m} |W_y| \leq \frac{1}{2} (4p)^{e(H^{\geq x})} \prod_{x \leq y \leq m} |W_y|,$$

where the inequality is by choice of β. Together with (A.6) we obtain the claimed upper bound. ■