THE REGULARITY METHOD FOR GRAPHS WITH FEW 4-CYCLES

DAVID CONLON, JACOB FOX, BENNY SUDAKOV, AND YUFEI ZHAO

Abstract. We develop a sparse graph regularity method that applies to graphs with few 4-cycles, including new counting and removal lemmas for 5-cycles in such graphs. Some applications include:

- Every n-vertex graph with no 5-cycle can be made triangle-free by deleting $o(n^{3/2})$ edges.
- For $r \geq 3$, every n-vertex r-graph with girth greater than 5 has $o(n^{3/2})$ edges.
- Every subset of $[n]$ without a nontrivial solution to the equation $x_1 + x_2 + 2x_3 = x_4 + 3x_5$ has size $o(\sqrt{n})$.

1. Introduction

Szemerédi's regularity lemma [52] is a rough structure theorem that applies to all graphs. The lemma originated in Szemerédi's proof of his celebrated theorem that dense sets of integers contain arbitrarily long arithmetic progressions [51] and is now considered one of the most useful and important results in combinatorics. Among its many applications, one of the earliest was the influential triangle removal lemma of Ruzsa and Szemerédi [40], which says that any n-vertex graph with $o(n^3)$ triangles can be made triangle-free by removing $o(n^2)$ edges. Surprisingly, this simple sounding statement is already sufficient to imply Roth's theorem, the special case of Szemerédi's theorem for 3-term arithmetic progressions, and a generalization known as the corners theorem.

Most applications of the regularity lemma, including the triangle removal lemma, rely on also having an associated *counting lemma*. Such a lemma roughly says that the number of embeddings of a fixed graph H into a pseudorandom graph G can be estimated by pretending that G is a random graph. This combined application of the regularity lemma and a counting lemma is often referred to as the *regularity method* and has had important applications in graph theory, combinatorial geometry, additive combinatorics, and theoretical computer science. For surveys on the regularity method and its applications, we refer the interested reader to [11, 30, 38].

The original version of the regularity lemma is only meaningful for dense graphs. However, many interesting and challenging combinatorial problems concern sparse graphs, so it would be extremely valuable to develop a regularity method that also applies to these graphs. The first step in this direction was already taken in the 1990's by Kohayakawa [28] and Rödl (see [24]), who proved an analogue of Szemerédi's regularity lemma for sparse graphs (see also [48]). The problem of proving an associated sparse counting lemma has been a more serious challenge, but one that has seen substantial progress in recent years.

It is known that one needs to make some nontrivial assumptions about a sparse graph in order for a counting lemma to hold. Usually, that has meant that the graph is assumed to be a subgraph of another well-behaved sparse graph, such as a random or pseudorandom graph. For subgraphs of random graphs, proving such a counting lemma (or, more accurately in this context, embedding lemma) was a famous problem, known as the KŁR conjecture [29], which has only been resolved very recently [3, 10, 46] as part of the large body of important work (see also [9, 47]) extending

Conlon is supported by NSF Award DMS-2054452 and in part by ERC Starting Grant 676632.

Fox is supported by a Packard Fellowship and by NSF Award DMS-1855635.

Sudakov is supported in part by SNSF grant 200021_196965.

Zhao is supported by NSF Award DMS-1764176, the MIT Solomon Buchsbaum Fund, and a Sloan Research Fellowship.

classical combinatorial theorems such as Turán's theorem and Szemerédi's theorem to subsets of random sets.

In the pseudorandom setting, the aim is again to prove analogues of combinatorial theorems, but now for subsets of pseudorandom sets. For instance, the celebrated Green–Tao theorem [26], that the primes contain arbitrarily long arithmetic progressions, may be viewed in these terms. Indeed, the main idea in their work is to prove an analogue of Szemerédi's theorem for dense subsets of pseudorandom sets and then to show that the primes form a dense subset of a pseudorandom set of "almost primes" so that their relative Szemerédi theorem can be applied. For graphs, a sparse counting lemma, which easily enables the transference of combinatorial theorems from dense to sparse graphs, was first developed in full generality by Conlon, Fox, and Zhao [13] and then extended to hypergraphs in [14], where it was used to give a simplified proof of a stronger relative Szemerédi theorem, valid under weaker pseudorandomness assumptions than in [26].

In this paper, we develop the sparse regularity method in another direction, without any assumption that our graph is contained in a sufficiently pseudorandom host. Instead, our only assumption will be that the graph has few 4-cycles and our main contribution will be a counting lemma that lower bounds the number of 5-cycles in such graphs.[1] Unlike the previous results on sparse regularity, the method developed here has natural applications in extremal and additive combinatorics with few hypotheses about the setting. We begin by exploring these applications.

Asymptotic notation. For positive functions f and g of n, we write $f = O(g)$ or $f \lesssim g$ to mean that $f \le Cg$ for some constant $C > 0$; we write $f = \Omega(g)$ or $f \gtrsim g$ to mean that $f \ge cg$ for some constant $c > 0$; we write $f = o(g)$ to mean that $f/g \to 0$; and we write $f = \Theta(g)$ or $f \asymp g$ to mean that $g \lesssim f \lesssim g$.

1.1. Sparse graph removal lemmas.

The famous triangle removal lemma of Ruzsa and Szemerédi [40] states that:

> *An n-vertex graph with $o(n^3)$ triangles can be made triangle-free by deleting $o(n^2)$ edges.*

One of the main applications of our sparse regularity method is a removal lemma for 5-cycles in C_4-free graphs. Since a C_4-free graph on n vertices has $O(n^{3/2})$ edges, a removal lemma in such graphs is only meaningful if the conclusion is that we can remove $o(n^{3/2})$ edges to achieve our goal, in this case that the graph should also be C_5-free. We show that such a removal lemma holds if our n-vertex C_4-free graph has $o(n^{5/2})$ C_5's.

Theorem 1.1. *An n-vertex C_4-free graph with $o(n^{5/2})$ C_5's can be made C_5-free by removing $o(n^{3/2})$ edges.*

This theorem is a special case of the following more general result.

Theorem 1.2 (Sparse C_3–C_5 removal lemma)**.** *An n-vertex graph with $o(n^2)$ C_4's and $o(n^{5/2})$ C_5's can be made $\{C_3, C_5\}$-free by deleting $o(n^{3/2})$ edges.*

Remark. Let us motivate the exponents that appear in this theorem. It is helpful to compare the quantities with what is expected in an n-vertex random graph with edge density p. Provided that $pn \to \infty$, the number of C_k's in $G(n,p)$ is typically on the order of $p^k n^k$ for each fixed k. Moreover, if $p \gtrsim n^{-1/2}$, then a second-moment calculation shows that $G(n,p)$ typically contains on the order of pn^2 edge-disjoint triangles and so cannot be made triangle-free by removing $o(pn^2)$ edges. Hence, the random graph $G(n,p)$ with $p \asymp n^{-1/2}$ shows that Theorem 1.2 becomes false if we only assume that there are $O(n^{5/2})$ C_5's and $O(n^2)$ C_4's.

[1] Longer cycles, as well as several other families of graphs, can also be counted using our techniques. We hope to return to this point in a future paper.

For a different example, we note that the polarity graph of Brown [6] and Erdős–Rényi–Sós [16] has n vertices, $\Theta(n^{3/2})$ edges, no C_4's, $\Theta(n^{5/2})$ C_5's, and every edge is contained in exactly one triangle (see [33] for the proof of this latter property). Thus, Theorem 1.2 is false if we relax the hypothesis on the number of C_5's from $o(n^{5/2})$ to $O(n^{5/2})$. For another example, showing that the hypothesis on the number of C_4's also cannot be entirely dropped, we refer the reader to Proposition 1.5 below.

We state two additional corollaries of Theorem 1.2, the second being an immediate consequence of the first. See Section 5 for the short deductions.

Corollary 1.3. *An n-vertex graph with $o(n^2)$ C_5's can be made triangle-free by deleting $o(n^{3/2})$ edges.*

Corollary 1.4. *An n-vertex C_5-free graph can be made triangle-free by deleting $o(n^{3/2})$ edges.*

We do not know if the exponent $3/2$ in Corollary 1.4 is best possible, but the next statement, whose proof can be found in Section 7, shows that the hypothesis on the number of C_5's in Corollary 1.3 cannot be relaxed from $o(n^2)$ to $o(n^{5/2})$.

Proposition 1.5. *There exist n-vertex graphs with $o(n^{2.442})$ C_5's that cannot be made triangle-free by deleting $o(n^{3/2})$ edges.*

We also state a 5-partite version of the sparse 5-cycle removal lemma. This statement will be used in our arithmetic applications.

Theorem 1.6 (Sparse removal lemma for 5-cycles in 5-partite graphs)**.** *For every $\epsilon > 0$, there exists $\delta > 0$ such that if G is a 5-partite graph on vertex sets $V_1, \dots, V_5$ with $|V_1| = \cdots = |V_5| = n$, all edges of G lie between V_i and V_{i+1} for some i (taken mod 5), and*

(a) *(Few 4-cycles between two parts) G has at most δn^2 copies of C_4 whose vertices lie in two different parts V_i,*
(b) *(Few 5-cycles) G has at most $\delta n^{5/2}$ copies of C_5,*

then G can be made C_5-free by removing at most $\epsilon n^{3/2}$ edges.

We note that the exponent $3/2$ in the conclusion above is tight, as shown by the next statement, whose proof can be found in Section 7.

Proposition 1.7. *For every n, there exists a 5-partite graph on vertex sets $V_1, \dots, V_5$ with $|V_1| = \cdots = |V_5| = n$, where all edges lie between V_i and V_{i+1} for some i (taken mod 5), such that the graph is C_4-free, every edge lies in exactly one 5-cycle, and there are $e^{-O(\sqrt{\log n})} n^{3/2}$ edges.*

Related results. An earlier application of sparse regularity to C_4-free (and, more generally, $K_{s,t}$-free) graphs may be found in [1], where it was used to study a conjecture of Erdős and Simonovits [17] in extremal graph theory. For instance, they show that if $s = 2$ and $t \geq 2$ or if $s = t = 3$, then the maximum number of edges in an n-vertex graph with no copy of $K_{s,t}$ and no copy of C_k for some odd $k \geq 5$ is asymptotically the same as the maximum number of edges in a bipartite n-vertex graph with no copy of $K_{s,t}$.

1.2. **Extremal results in hypergraphs.** In an r-graph (i.e., an r-uniform hypergraph), a (v, e)-*configuration* is a subgraph with e edges and at most v vertices. A central problem in extremal combinatorics is to estimate $f_r(n, v, e)$, the maximum number of edges in an n-vertex r-graph without a (v, e)-configuration. For brevity, we drop the subscript when $r = 3$, simply writing $f(n, v, e) := f_3(n, v, e)$.

The systematic study of this function was initiated almost five decades ago by Brown, Erdős, and Sós [7, 50]. A famous conjecture that arose from their work [18, 15] asks whether $f(n, e+3, e) = o(n^2)$ for any fixed $e \geq 3$. For $e = 3$, this problem was resolved by Ruzsa and Szemerédi [40]. In fact, this

$(6,3)$-theorem, rather than the triangle removal lemma, was their original motivation for studying such problems. Their result has been extended in many directions, but the problem of showing that $f(n, e+3, e) = o(n^2)$ remains open for all $e \geq 4$.

Our methods give the following new bound for a problem of this type, which turns out to be equivalent to Corollary 1.4.

Corollary 1.8. $f(n, 10, 5) = o(n^{3/2})$.

We next explain how to deduce Corollary 1.8 from Corollary 1.4. Suppose H is a 3-graph on n vertices without a $(10,5)$-configuration. We greedily delete vertices from H one at a time if they are in at most four edges. In total, this process deletes at most $4n$ edges. The resulting induced subgraph H' has the property that each vertex is in at least 5 edges. Furthermore, H' is linear, that is, any two edges intersect in at most one vertex. Indeed, if there are two edges sharing vertices u, v, then, by adding three additional edges touching v, we get a $(10,5)$-configuration. If we now let G be the underlying graph formed by converting all edges of H' to triangles, we see that G is a union of edge-disjoint triangles. Moreover, it is C_5-free, since otherwise it would contain a $(10,5)$-configuration. Hence, by Corollary 1.4, it has $o(n^{3/2})$ edges. But this then implies that H' and, therefore, H has $o(n^{3/2})$ edges. Conversely, to show that Corollary 1.8 implies Corollary 1.4, it suffices to observe that the 3-graph formed by a collection of edge-disjoint triangles in a C_5-free graph does not have a $(10,5)$-configuration.

A *(Berge) cycle* of length $k \geq 2$ (or simply a *k-cycle*) in a hypergraph is an alternating sequence of distinct vertices and edges $v_1, e_1, \ldots, v_k, e_k$ such that $v_i, v_{i+1} \in e_i$ for each i (where indices are taken modulo k). For example, a 2-cycle consists of a pair of edges intersecting in a pair of distinct vertices. The *girth* of an r-graph is the length of the shortest cycle.

Let $h_r(n, g)$ denote the maximum number of edges in an r-graph on n vertices of girth larger than g. The following observation, whose proof may be found in Appendix B, relates the Brown–Erdős–Sós problem to that of estimating $h_r(n, g)$.

Proposition 1.9. *For $r \geq 2$ and $e \geq 2$, there exists $n_0(r, e)$ such that $f_r(n, (r-1)e, e) = h_r(n, e)$ for all $n \geq n_0(r, e)$.*

By Corollary 1.8, we thus have the following result for $r = 3$. Note that the general result follows from the case $r = 3$. Indeed, if an r-graph has girth g, replacing each edge by a subset of size three, we get a 3-graph on the same set of vertices with the same number of edges and girth at least g.

Corollary 1.10. *Let $r \geq 3$. Then $h_r(n, 5) = o(n^{3/2})$, i.e., every r-graph on n vertices of girth greater than 5 has $o(n^{3/2})$ edges.*

Related results. Previously, upper bounds of the form $h_r(n, 5) \leq c_r n^{3/2}$ were known [33, 19]. In fact, Lazebnik and Verstraëte [33] showed that $f(n, 8, 4) = h_3(n, 4) = (1/6 + o(1))n^{3/2}$ for all sufficiently large n.[2] Bollobás and Győri [5] proved that the maximum number of edges in a 3-graph on n vertices with no 5-cycle is $\Theta(n^{3/2})$, which implies that the maximum number of triangles in an n-vertex C_5-free graph is $\Theta(n^{3/2})$. In contrast, Corollary 1.4 says that the maximum number of *edge-disjoint* triangles in a C_5-free graph on n vertices is $o(n^{3/2})$. See [2, 19, 20, 23] for further improvements and simplifications of the Bollobás–Győri result.

Given a family $\mathcal{F}$ of r-graphs, we say that an r-graph is *$\mathcal{F}$-free* if it contains no copy of any element of $\mathcal{F}$ as a subgraph. Define $\mathrm{ex}(n, \mathcal{F})$ to be the maximum number of edges in an $\mathcal{F}$-free

[2] Lazebnik and Verstraete [33] actually claim that $f(n, 8, 4) = h_3(n, 4)$ for all n. However, this is false for small n. For instance, it is easy to see that $f(6, 8, 4) = 3$, while $h_3(6, 4) = 2$. Nevertheless, their claim that $f(n, 8, 4) = (1/6 + o(1))n^{3/2}$ still stands by our Proposition 1.9 and their result that $h_3(n, 4) = (1/6 + o(1))n^{3/2}$.

r-graph on n vertices. It is easy to see that

$$2^{\mathrm{ex}(n,\mathcal{F})} \le \#\{\mathcal{F}\text{-free } r\text{-graphs on } n \text{ labeled vertices}\} \le \sum_{m=0}^{\mathrm{ex}(n,\mathcal{F})} \binom{\binom{n}{r}}{m} \le 2^{\mathrm{ex}(n,\mathcal{F}) r \log n}.$$

In many instances it is known that the lower bound is closer to the truth. Early invesigations into this and related problems led to the first seeds of the important container method, as developed by Kleitman and Winston [27] and by Sapozhenkho [43, 44, 45]. More recently, influential works of Balogh, Morris, and Samotij [3] and of Saxton and Thomason [46] pushed these ideas considerably further and showed their broad applicability.

Using the container method, Palmer, Tait, Timmons, and Wagner [36] proved that the number of r-graphs on n vertices with girth greater than 4 is at most $2^{c_r n^{3/2}}$ for an appropriate constant c_r. We improve this bound when the girth is greater than 5, strengthening Corollary 1.10. We refer the reader to Appendix C for the proof.

Theorem 1.11. *For every fixed $r \ge 3$, the number of r-graphs on n vertices with girth greater than 5 is $2^{o(n^{3/2})}$.*

1.3. **Number-theoretic applications.** It was already noted by Ruzsa and Szemerédi that their results imply Roth's theorem [39], the statement that every subset of $[n] := \{1, 2, \dots, n\}$ without 3-term arithmetic progressions has size $o(n)$. Here we discuss some number-theoretic applications of our sparse removal results along similar lines. We first illustrate our results with two specific applications, beginning with the following theorem.[3]

Theorem 1.12. *Every subset of $[n]$ without a nontrivial solution to the equation*

$$x_1 + x_2 + 2x_3 = x_4 + 3x_5 \tag{1}$$

has size $o(\sqrt{n})$. Here a trivial solution is one of the form $(x_1, \dots, x_5) = (x, y, y, x, y)$ or (y, x, y, x, y) for some $x, y \in \mathbb{Z}$.

Any set of integers without a nontrivial solution to (1) must be a *Sidon set*, with no nontrivial solution to the equation $x_1 + x_2 = x_4 + x_5$, since any nontrivial solution automatically extends to a nontrivial solution of (1) by setting $x_3 = x_5$. In particular, the upper bound for the size of Sidon sets, $(1+o(1))\sqrt{n}$, is also an upper bound for the size of a subset of $[n]$ without a nontrivial solution to (1). Our Theorem 1.12 improves on this simple bound, though it remains an open problem to determine whether the bound can be improved further to $n^{1/2-\epsilon}$ for some $\epsilon > 0$.

We now give a second number-theoretic application, this time restricting to Sidon sets.

Theorem 1.13. *The maximum size of a Sidon subset of $[n]$ without a solution in distinct variables to the equation*

$$x_1 + x_2 + x_3 + x_4 = 4x_5$$

is at most $o(\sqrt{n})$ and at least $n^{1/2-o(1)}$.

In other words, we are simultaneously avoiding

(a) nontrivial solutions to the Sidon equation $x_1 + x_2 = x_3 + x_4$ and
(b) distinct variable solutions to the linear equation $x_1 + x_2 + x_3 + x_4 = 4x_5$.

There exist Sidon sets of size $(1+o(1))\sqrt{n}$, as well as sets of size $n^{1-o(1)}$ avoiding (b) (by a standard modification of Behrend's construction [4] of large sets without 3-term arithmetic progressions). However, Theorem 1.13 shows that by simultaneously avoiding nontrivial solutions to both equations, the maximum size is substantially reduced.

[3]Though we focus here on applications to linear equations with five variables, all of our results extend to linear equations with more than five variables by using the counting lemma for longer cycles mentioned in an earlier footnote.

This is the first example showing a lack of "compactness" for linear equations. In extremal graph theory, the *Erdős–Simonovits compactness conjecture* [17] is a well-known conjecture saying that, for every finite set $\mathcal{F}$ of graphs, $\mathrm{ex}(n,\mathcal{F}) \geq c_{\mathcal{F}} \min_{F\in\mathcal{F}} \mathrm{ex}(n,F)$ for some constant $c_{\mathcal{F}} > 0$. The analogous statement is false for r-graphs with $r \geq 3$ (by the Ruzsa–Szemerédi theorem and a simple generalisation to r-graphs noted in [15, Theorem 1.9]), but remains open for graphs. Our Theorem 1.13 shows that it also fails for linear equations.

Theorem 1.13 also sheds some light on the fascinating open problem (see, for example, Gowers' blog post [25]) of understanding the structure of Sidon sets with near-maximum size, say within a constant factor of $\sqrt{n}$, showing that any such set must contain five distinct elements with one of them being the average of the others. More generally, we have the following result, showing that a large Sidon set must contain solutions to a wide family of translation-invariant linear equations in five variables. We note that the lower bound of $n^{1/2-o(1)}$ simply comes from intersecting a Sidon set of set $(1+o(1))n^{1/2}$ with a random translate of a subset of $[n]$ of size $n^{1-o(1)}$ that avoids nontrivial solutions to (2) (which again exists by a standard modification of Behrend's construction [4]).

Theorem 1.14. *Fix positive integers $a_1,\dots,a_4$. The maximum size of a Sidon subset of $[n]$ without a solution in distinct variables to the equation*

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = (a_1+a_2+a_3+a_4)x_5 \tag{2}$$

is at most $o(\sqrt{n})$ and at least $n^{1/2-o(1)}$.

Similarly, Theorem 1.12 is a special case of the following statement.

Theorem 1.15. *Fix positive integers a and b. Every subset of $[n]$ without a nontrivial solution to the equation*

$$ax_1 + ax_2 + bx_3 = ax_4 + (a+b)x_5 \tag{3}$$

has size $o(\sqrt{n})$. Here a trivial solution is one of the form $(x_1,\dots,x_5) = (x,y,y,x,y)$ or (y,x,y,x,y) (or (y,y,x,x,y) if $a=b$) for some $x,y\in\mathbb{Z}$.

Both Theorem 1.15 and the upper bound in Theorem 1.14 are special cases of the following more robust theorem (applied with $X_1 = \cdots = X_5$), whose proof can be found in Section 6. Indeed, to prove the upper bound in Theorem 1.14, we apply Lemma 1.17 below to check that every Sidon subset of $[n]$ contains $O(n)$ solutions to (2) where not all variables are distinct, thereby verifying hypothesis (b) of Theorem 1.16 (with a $O(n)$ bound instead of $o(n^{3/2})$). To prove Theorem 1.15, we note, by setting $x_3 = x_5$ in (3), that the subset satisfying the hypothesis of Theorem 1.15 must be a Sidon set. Finally, when $X_1 = \cdots = X_5$, the removal statement in the conclusion of Theorem 1.16 implies that $|X_1| = o(\sqrt{n})$ since $x_1 = \cdots = x_5$ is always a solution due to $a_1 + \cdots + a_5 = 0$.

Theorem 1.16. *Fix nonzero integers $a_1,\dots,a_5$ with $a_1+\cdots+a_5 = 0$. Suppose that $X_1,\dots,X_5$ are subsets of $[n]$ satisfying*

(a) *each X_i has $o(n)$ nontrivial solutions to $x_1+x_2 = x_3+x_4$*
(a trivial solution here is one with $(x_1,x_2) = (x_3,x_4)$ or (x_4,x_3)) and
(b) *$o(n^{3/2})$ solutions to $a_1x_1+\cdots+a_5x_5 = 0$ with $x_1\in X_1, \dots, x_5\in X_5$.*

Then one can remove $o(\sqrt{n})$ elements from each X_i to remove all solutions to $a_1x_1+\cdots+a_5x_5 = 0$ with $x_1\in X_1, \dots, x_5\in X_5$.

Lemma 1.17. *Let $a_1,\dots,a_4$ be nonzero integers and X be a Sidon subset of $[n]$. Then X contains $O(n)$ solutions to the equation $a_1x_1+a_2x_2+a_3x_3+a_4x_4 = 0$.*

Since our proofs rely on the graph removal lemma, they give poor quantitative bounds, the best bound on that lemma [21] having tower-type dependencies. However, for our number-theoretic applications, it is possible to use the best bounds for the relevant Roth-type theorem, together with a weak arithmetic regularity lemma and our C_5-counting lemma to obtain reasonable bounds.

This is similar in spirit to the arithmetic transference proof of the relative Szemerédi theorem given in [53], though we omit the details. A follow-up work of Prendiville [37] giving a Fourier-analytic proof of our number-theoretic results also yields comparable bounds.

As a final remark, we note that the results of this subsection carry over essentially verbatim to arbitrary abelian groups. Following Král'–Serra–Vena [31] (see also [32, 49]), one may also use our sparse graph removal lemma to derive a sparse removal lemma that is meaningful in arbitrary groups. We again omit the details, but refer the interested reader to [12, Theorem 1.2] for a result which is similar in flavor.

2. A weak sparse regularity lemma

In this section, we develop a sparse version of the Frieze–Kannan weak regularity lemma [22]. A sparse version of Szemerédi's regularity lemma was originally developed by Kohayakawa [28] and Rödl (see [24]) under an additional "no dense spots" hypothesis, but Scott [48] showed that, with a slight variation in the statement, this additional hypothesis is not needed. The approach we use here for proving a sparse version of the weak regularity lemma will be similar to that of Scott. With this result (and an appropriate counting lemma) in hand, we will then be able to "transfer" the removal lemma from the dense setting to the sparse setting.

In order to give an analytic formulation of the weak regularity lemma, we first make some definitions. Given a pair of probability spaces V_1 and V_2, which are usually vertex sets with the uniform measure (or, if the vertices carry weights, then with the probability measure that is proportional to the vertex weights), we define the *cut norm* for a measurable function $f\colon V_1 \times V_2 \to \mathbb{R}$ (we will sometimes omit mentioning the measurability requirement when it is clear from context) by

$$\|f\|_\square := \sup_{\substack{A\subset V_1\\ B\subset V_2}} \left|\mathbb{E}_{x\in V_1, y\in V_2} f(x,y)1_A(x)1_B(y)\right|, \tag{4}$$

where A and B range over all measurable subsets and x and y are chosen independently according to the corresponding probability measures.

Given a partition $\mathcal{P}$ of some probability space V and a function $f\colon V\times V\to\mathbb{R}$, we write $f_\mathcal{P}\colon V\times V\to\mathbb{R}$ for the function obtained from f by "averaging" over blocks $A\times B$ where A and B are parts of $\mathcal{P}$, i.e., $f_\mathcal{P}(x,y) = \frac{1}{\mu(A)\mu(B)}\int_{A\times B} f$ for all $(x,y)\in A\times B$, where $\mu(\cdot)$ is the probability measure on V. We may ignore zero-measure parts.

The weak regularity lemma of Frieze and Kannan may be rephrased in the following way (for example, see [34, Corollary 9.13]), saying that all bounded functions can be approximated in terms of the cut norm by a step function with a bounded number of blocks. Furthermore, the step function can be obtained by averaging the original function over steps. This analytic perspective on the weak regularity lemma has been popularized by the development of graph limits [35].

Theorem 2.1 (Weak regularity lemma, dense setting)**.** *Let $\epsilon > 0$. Let V be a probability space and $f\colon V\times V\to[0,1]$ be a measurable symmetric function (i.e., $f(x,y) = f(y,x)$ for all $x,y\in V$). Then there exists a partition $\mathcal{P}$ of V into at most $2^{O(\epsilon^{-2})}$ parts such that*

$$\|f - f_\mathcal{P}\|_\square \le \epsilon.$$

For sparse graphs, one would like to have control on the error term that is commensurate with the overall edge density of the graph. More explicitly, for an n-vertex graph with on the order of pn^2 edges for some $p = p_n \to 0$, one would like to have error terms of the form ϵp in the above inequality. To capture this scaling, we renormalize f by dividing the edge-indicator function of the graph by the edge density p. Thus, the sparse setting corresponds to unbounded functions f with L^1 norm $O(1)$.

Our sparse weak regularity lemma is stated below. The proof follows an energy increment strategy, as is usual with proofs of regularity lemmas. Since this is now fairly standard, we refer the reader to Appendix A for the details.

Theorem 2.2. *Let $\epsilon > 0$. Let V be a probability space and $f\colon V \times V \to [0,\infty)$ be a measurable symmetric function. Then there exists a partition $\mathcal{P}$ of V into at most $2^{32\mathbb{E}f/\epsilon^2}$ parts such that*

$$\|(f - f_{\mathcal{P}})1_{f_{\mathcal{P}}\le 1}\|_{\square} \le \epsilon.$$

Here $(f - f_{\mathcal{P}})1_{f_{\mathcal{P}}\le 1}$ denotes the function with value $f(x,y) - f_{\mathcal{P}}(x,y)$ if $f_{\mathcal{P}}(x,y) \le 1$ and 0 otherwise. The cutoff term $1_{f_{\mathcal{P}}\le 1}$ means that we neglect those parts of the graph that are too dense. Indeed, it would be too much to ask for $\|f - f_{\mathcal{P}}\|_{\square}$ to be small without this cutoff. For example, if one had $|V| = n$, $f(x,x) = n$, and $f(x,y) = 0$ for all $x \ne y$, then one would not be able to partition V into $O_\epsilon(1)$ parts so that $\|f - f_{\mathcal{P}}\|_{\square} \le \epsilon$.

For our applications, it will be necessary to show that the cutoff has little effect on graphs with few C_4's. In the next two lemmas, we show that such graphs have a negligible number of edges lying between pairs of parts whose edge density greatly exceeds the average. A similar argument was also presented in [1], though we include the complete proof here for the convenience of the reader.

Lemma 2.3. *Let G be a bipartite graph with nonempty vertex sets A and B and m edges. If $m \ge 4|A||B|^{1/2} + 4|B|$, then the number of 4-cycles in G is at least $m^4|A|^{-2}|B|^{-2}/64$.*

Proof. Writing $\binom{x}{2} = x(x-1)/2$ for all real x, we have $\binom{x}{2} \ge x^2/4$ for all $x \ge 2$ and so, by the convexity of $x \mapsto \binom{x}{2}$,

$$\sum_{\{x,y\}\in\binom{A}{2}} \operatorname{codeg}(x,y) = \sum_{v\in B}\binom{\deg(v)}{2} \ge |B|\binom{m/|B|}{2} \ge \frac{m^2}{4|B|},$$

where $\operatorname{codeg}(x,y)$ is the number of common neighbors of x and y. Thus, the number of 4-cycles in G is

$$\sum_{\{x,y\}\in\binom{A}{2}}\binom{\operatorname{codeg}(x,y)}{2} \ge \binom{|A|}{2}\binom{\binom{|A|}{2}^{-1}\frac{m^2}{4|B|}}{2} \ge \frac{m^4}{64|A|^2|B|^2}. \qquad \square$$

We write $e(U,W)$ for the number of pairs $(u,w) \in U \times W$ forming an edge in the given graph G.

Lemma 2.4 (Dense pairs)**.** *Let G be an n-vertex graph and $V(G) = V_1 \cup V_2 \cup \cdots \cup V_M$ be a partition of the vertex set of G. Let $q > 0$ be a real number. Let T denote the number of 4-cycles in G. Then the number of edges of G lying between parts V_i and V_j with $e(V_i,V_j) \ge q|V_i||V_j|$ (here $i = j$ is allowed) is $O((1/q + M^2 + M^{1/4}T^{1/4})n)$.*

Proof. For each $i \in [M]$, let W_i denote the union of all V_j such that $e(V_i,V_j) \ge q|V_i||V_j|$. Then $e(V_i,W_i) \ge q|V_i||W_i|$. Let T_i denote the number of 4-cycles with the first and third vertices in V_i and second and fourth vertices in W_i. We claim that

$$e(V_i,W_i) \lesssim \frac{n}{Mq} + \frac{|V_i|}{q} + Mn + n^{1/2}|V_i|^{1/2}T_i^{1/4}. \tag{5}$$

Indeed, if $|V_i| \le 1/Mq$, then $e(V_i,W_i) \le |V_i||W_i| \le n/Mq$. So assume $|V_i| > 1/Mq$. If $e(V_i,W_i) < 4|V_i||W_i|^{1/2} + 4|W_i|$, then $e(V_i,W_i)^2 \le 32|V_i|^2|W_i| + 32|W_i|^2$ and

$$e(V_i,W_i) \le \frac{e(V_i,W_i)^2}{q|V_i||W_i|} \lesssim \frac{|V_i|}{q} + \frac{|W_i|}{q|V_i|} \le \frac{|V_i|}{q} + Mn,$$

where in the final inequality we use $|W_i| \le n$ and $|V_i| > 1/Mq$. Otherwise, by Lemma 2.3, we have $T_i \gtrsim e(V_i,W_i)^4|V_i|^{-2}|W_i|^{-2}$, so

$$e(V_i,W_i) \lesssim |V_i|^{1/2}|W_i|^{1/2}T_i^{1/4} \le n^{1/2}|V_i|^{1/2}T_i^{1/4}.$$

This proves (5). Summing over all i, we obtain that the number of edges of G lying between parts V_i and V_j with $e(V_i, V_j) \geq q\,|V_i|\,|V_j|$ is at most

$$\begin{aligned}\sum_{i=1}^{M} e(V_i, W_i) &\lesssim \sum_{i=1}^{M} \left(\frac{n}{Mq} + \frac{|V_i|}{q} + Mn + n^{1/2}\,|V_i|^{1/2}\,T_i^{1/4} \right) \\ &\leq \frac{2n}{q} + M^2 n + n^{1/2} \left(\sum_{i=1}^{M} 1 \right)^{1/4} \left(\sum_{i=1}^{M} |V_i| \right)^{1/2} \left(\sum_{i=1}^{M} T_i \right)^{1/4} \\ &\leq \frac{2n}{q} + M^2 n + 2M^{1/4} T^{1/4} n,\end{aligned}$$

where we applied Hölder's inequality in the second step and also that $\sum_i T_i \leq 2T$. □

Remark. We will apply this lemma with $q = K/\sqrt{n}$ for some large constant K and $T = o(n^2)$.

3. A sparse counting lemma for 5-cycles

This section contains a novel counting lemma for 5-cycles in sparse graphs. As discussed in the introduction, counting lemmas are often the key obstacles to the development of the sparse regularity method. As such, our counting result, which says that if a graph does not have too many copies of C_4, then the number of copies of C_5 is approximated by the C_5 count in the weak regularity approximation, may be seen as our main contribution.

Combinatorially, the intuition is that in a graph with few C_4's and edge density on the order of $n^{-1/2}$, the second neighborhood of a typical vertex has linear size, so that regularity ensures enough edges within this second neighborhood to generate many 5-cycles. This observation was already used to considerable effect in [1].

Let V be a probability space. Given a symmetric function $f\colon V \times V \to [0, \infty)$ (again *symmetric* means $f(x, y) = f(y, x)$), the *homomorphism density* of H in f is given by

$$t(H, f) = \int_{V^{V(H)}} \prod_{uv \in E(H)} f(x_u, x_v) \prod_{v \in V(H)} dx_v.$$

As in our formulation of the regularity lemma in Section 2, one should think of the function f as a normalized edge-indicator function of the form $p^{-1} 1_{E(G)}$. Our counting lemma is now as follows.

Theorem 3.1. *Let $0 < \epsilon < 1$ and $C \geq 1$. Let V be a probability space and let $f\colon V \times V \to [0, \infty)$ and $g\colon V \times V \to [0, 1]$ be measurable symmetric functions satisfying $t(C_4, f) \leq C$ and $\|f - g\|_\square \leq \epsilon^4$. Then*

$$t(C_5, f) \geq t(C_5, g) - 11C\epsilon.$$

In practice, we will prove a multipartite version of this counting lemma which will be useful for applications. The two versions are essentially equivalent.

Notation and conventions. Let $V_1, \ldots, V_5$ be probability spaces with indices considered mod 5. Vertices in V_i are denoted by x_i and, when appearing in the subscript of an expectation symbol (e.g., $\mathbb{E}_{x_i}$), it is assumed that x_i varies independently over V_i according to its probability distribution (e.g., a uniformly random vertex if V_i comes from an unweighted graph). For all $i \neq j$ differing by one in $\mathbb{Z}/5\mathbb{Z}$, we write $f_{ij} = f_{i,j}$ for a measurable function on $V_i \times V_j$ and assume that $f_{ij}(x_i, x_j) = f_{ji}(x_j, x_i)$. Given functions $f_{12}\colon V_1 \times V_2 \to \mathbb{R}$ and $f_{23}\colon V_2 \times V_3 \to \mathbb{R}$, we write $f_{12} \circ f_{23}\colon V_1 \times V_3 \to \mathbb{R}$ for the function

$$(f_{12} \circ f_{23})(x_1, x_3) = \mathbb{E}_{x_2 \in V_2} f_{12}(x_1, x_2) f_{23}(x_2, x_3),$$

which can be thought of as the number of 2-edge paths between the vertices x_1 and x_3, appropriately normalized. The notation $\circ$ is used since it can be viewed as a composition of linear operators.

Theorem 3.2. *Let $0 < \epsilon < 1$ and $C \geq 1$. Let $V_1, \dots, V_5$ be probability spaces. For each i (taken mod 5), let $f_{i,i+1}, g_{i,i+1} \colon V_i \times V_{i+1} \to [0,\infty)$ with $0 \leq g_{i,i+1} \leq 1$ pointwise, $\|f_{i,i+1} - g_{i,i+1}\|_\square \leq \epsilon^4$, and*

$$\|f_{i-1,i} \circ f_{i,i+1}\|_2^2 \leq C. \tag{6}$$

Then

$$\mathbb{E}_{x_1,\dots,x_5} \prod_{i=1}^5 f_{i,i+1}(x_i, x_{i+1}) \geq \mathbb{E}_{x_1,\dots,x_5} \prod_{i=1}^5 g_{i,i+1}(x_i, x_{i+1}) - 11C\epsilon. \tag{7}$$

Remark. For an n-vertex graph with edge density on the order of p, the hypothesis (6) translates into a $O(n^4p^4)$ upper bound on the number of 4-cycles with vertices in $V_{i-1}, V_i, V_{i+1}, V_i$ in turn, since

$$\|f_{01} \circ f_{12}\|_2^2 = \mathbb{E}_{x_0,x_1,x_1',x_2} f_{01}(x_0,x_1) f_{01}(x_0,x_1') f_{12}(x_1,x_2) f_{12}(x_1',x_2).$$

Applying the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\|f_{01} \circ f_{12}\|_2^2 &= \mathbb{E}_{x_1,x_1'}[(\mathbb{E}_{x_0} f_{01}(x_0,x_1) f_{01}(x_0,x_1'))(\mathbb{E}_{x_2} f_{12}(x_1,x_2) f_{12}(x_1',x_2))] \\
&\leq \left(\mathbb{E}_{x_1,x_1'}[(\mathbb{E}_{x_0} f_{01}(x_0,x_1) f_{01}(x_0,x_1'))^2]\right)^{1/2} \left(\mathbb{E}_{x_1,x_1'}[(\mathbb{E}_{x_2} f_{12}(x_1,x_2) f_{12}(x_1',x_2))^2]\right)^{1/2} \\
&= \left(\mathbb{E}_{x_0,x_0',x_1,x_1'}[f_{01}(x_0,x_1) f_{01}(x_0',x_1) f_{01}(x_0,x_1') f_{01}(x_0',x_1')]\right)^{1/2} \\
&\quad \cdot \left(\mathbb{E}_{x_1,x_1',x_2,x_2'}[f_{12}(x_1,x_2) f_{12}(x_1',x_2) f_{12}(x_1,x_2') f_{12}(x_1',x_2')]\right)^{1/2} \\
&= t(K_{2,2}, f_{01})^{1/2} t(K_{2,2}, f_{12})^{1/2}.
\end{aligned}$$

So we could replace (6) by the stronger condition that $t(K_{2,2}, f_{i,i+1}) \leq C$ for each i, i.e., a $O(n^4p^4)$ bound on the number of 4-cycles in the bipartite graph between V_i and V_{i+1} for every i.

Let us collect some basic facts about the cut norm. It is a standard fact that the definition of the cut norm (4) is equivalent to

$$\|f\|_\square = \sup_{\substack{a \colon V_1 \to [0,1] \\ b \colon V_2 \to [0,1]}} \left|\mathbb{E}_{x \in V_1, y \in V_2} f(x,y) a(x) b(y)\right|, \tag{8}$$

where now a and b are measurable functions taking values in $[0,1]$. The equivalence can be seen since the expectation expression in (8) is bilinear in a and b and thus its extrema must occur when a and b take values from the set $\{0,1\}$. But then a and b may be thought of as indicator functions of sets A and B, returning us to the earlier definition (4) of the cut norm.

Lemma 3.3. *Let $f_{12} \colon V_1 \times V_2 \to \mathbb{R}$ and $f_{23} \colon V_2 \times V_3 \to [0,\infty)$ with $\mathbb{E}_{x_3 \in V_3} f_{23}(x_2,x_3) \leq M$ for all $x_2 \in V_2$. Then*

$$\|f_{12} \circ f_{23}\|_\square \leq M \|f_{12}\|_\square.$$

Proof. For any $u_1 \colon V_1 \to [0,1]$ and $u_3 \colon V_3 \to [0,1]$, one has

$$\begin{aligned}
\mathbb{E}_{x_1,x_3} u_1(x_1)(f_{12} \circ f_{23})(x_1,x_3) u_3(x_3) &= \mathbb{E}_{x_1,x_2,x_3} u_1(x_1) f_{12}(x_1,x_2) f_{23}(x_2,x_3) u_3(x_3) \\
&= M \cdot \mathbb{E}_{x_1,x_2} u_1(x_1) f_{12}(x_1,x_2) u_2(x_2),
\end{aligned}$$

where

$$u_2(x_2) = \frac{1}{M} \mathbb{E}_{x_3} f_{23}(x_2,x_3) u_3(x_3) \leq \frac{1}{M} \mathbb{E}_{x_3} f_{23}(x_2,x_3) \leq 1.$$

The claim then follows from (8) since $u_1(x_1) \in [0,1]$ and $u_2(x_2) \in [0,1]$. □

Lemma 3.4 (Triangle counting lemma for dense graphs)**.** *Let $f_{ij} \colon V_i \times V_j \to \mathbb{R}$ for $ij \in \{12,13,23\}$ and assume that f_{12} and f_{13} take only nonnegative values. Then*

$$\mathbb{E}_{x_1 \in V_1, x_2 \in V_2, x_3 \in V_3} f_{12}(x_1,x_2) f_{13}(x_1,x_3) f_{23}(x_2,x_3) \leq \|f_{12}\|_\infty \|f_{13}\|_\infty \|f_{23}\|_\square.$$

Proof. The above inequality is true if we fix any choice of x_1 on the LHS, due to the definition of $\|f_{23}\|_\square$ in (8), and hence it remains true if we take the expectation over x_1. □

Proof of Theorem 3.2. For fixed $i, j \in \mathbb{Z}/5\mathbb{Z}$ with $i - j = \pm 1$ and $f_{ij} \colon V_i \times V_j \to [0, \infty)$, write

$$f'_{ij}(x_i, x_j) = f_{ij}(x_i, x_j) 1_{S_{ij}}(x_i),$$

where

$$S_{ij} = \{x_i \in V_i : \mathbb{E}_{x_j \in V_j} f_{ij}(x_i, x_j) \le \epsilon^{-2}\}.$$

Since $f_{ij}, g_{ij} \ge 0$, note that $\|f_{ij}\|_1 - \|g_{ij}\|_1 = \mathbb{E}_{x_i, x_j}(f_{ij} - g_{ij})$, which, by definition, is at most $\|f_{ij} - g_{ij}\|_\square$. Hence, writing $\mu(\cdot)$ for the measure of a subset of V_i (where $\mu(V_i) = 1$), we have

$$\mu(V_i \setminus S_{ij})\epsilon^{-2} \le \|f_{ij}\|_1 \le \|g_{ij}\|_1 + \|f_{ij} - g_{ij}\|_\square \le 2.$$

Therefore, $\mu(V_i \setminus S_{ij}) \le 2\epsilon^2$ and

$$\begin{aligned} \|f'_{ij} - g_{ij}\|_\square &\le \|(f_{ij} - g_{ij}) 1_{S_{ij} \times V_j}\|_\square + \|g_{ij} 1_{(V_i \setminus S_{ij}) \times V_j}\|_\square \\ &\le \|f_{ij} - g_{ij}\|_\square + \mu(V_i \setminus S_{ij}) \\ &\le 3\epsilon^2. \end{aligned}$$

Let i, j, k be three consecutive elements of $\mathbb{Z}/5\mathbb{Z}$ (in ascending or descending order). By Lemma 3.3,

$$\left\|(f_{ij} - g_{ij}) \circ f'_{jk}\right\|_\square \le \epsilon^{-2} \|f_{ij} - g_{ij}\|_\square \le \epsilon^2$$

and

$$\left\|g_{ij} \circ (f'_{jk} - g_{jk})\right\|_\square \le \left\|f'_{jk} - g_{jk}\right\|_\square \le 3\epsilon^2.$$

Putting the above two inequalities together, we obtain

$$\left\|f_{ij} \circ f'_{jk} - g_{ij} \circ g_{jk}\right\|_\square \le \left\|(f_{ij} - g_{ij}) \circ f'_{jk}\right\|_\square + \left\|g_{ij} \circ (f'_{jk} - g_{jk})\right\|_\square \le 4\epsilon^2.$$

For every $A > 0$ and any function $h(z)$, define $h_{\le A}(z) = h(z)$ if $h(z) \le A$ and 0 otherwise. Then we have

$$\begin{aligned} \left\|(f_{ij} \circ f'_{jk})_{\le A} - g_{ij} \circ g_{jk}\right\|_\square &\le \left\|f_{ij} \circ f'_{jk} - g_{ij} \circ g_{jk}\right\|_\square + \|(f_{ij} \circ f_{jk})_{> A}\|_1 \\ &\le 4\epsilon^2 + A^{-1} \|f_{ij} \circ f_{jk}\|_2^2 \\ &\le 4\epsilon^2 + CA^{-1}. \end{aligned}$$

In particular, setting $A = \epsilon^{-1}$, we obtain

$$\left\|(f_{34} \circ f'_{45})_{\le \epsilon^{-1}} - g_{34} \circ g_{45}\right\|_\square \le 5C\epsilon \tag{9}$$

and setting $A = \epsilon^{-2}$, we obtain

$$\left\|(f_{32} \circ f'_{21})_{\le \epsilon^{-2}} - g_{32} \circ g_{21}\right\|_\square \le 5C\epsilon^2. \tag{10}$$

Repeatedly applying Lemma 3.4 in the 3rd, 4th and 5th inequalities, we have

$$\begin{aligned} \text{LHS of (7)} &\ge \mathbb{E}_{x_1, x_3, x_5}(f_{34} \circ f'_{45})(x_3, x_5)(f_{32} \circ f'_{21})(x_3, x_1) f_{51}(x_5, x_1) && \text{[pointwise bound]} \\ &\ge \mathbb{E}_{x_1, x_3, x_5}(f_{34} \circ f'_{45})_{\le \epsilon^{-1}}(x_3, x_5)(f_{32} \circ f'_{21})_{\le \epsilon^{-2}}(x_3, x_1) f_{51}(x_5, x_1) && \text{[pointwise bound]} \\ &\ge \mathbb{E}_{x_1, x_3, x_5}(f_{34} \circ f'_{45})_{\le \epsilon^{-1}}(x_3, x_5)(f_{32} \circ f'_{21})_{\le \epsilon^{-2}}(x_3, x_1) g_{51}(x_5, x_1) - \epsilon && \text{[since } \|f_{51} - g_{51}\|_\square \le \epsilon^4\text{]} \\ &\ge \mathbb{E}_{x_1, x_3, x_5}(f_{34} \circ f'_{45})_{\le \epsilon^{-1}}(x_3, x_5)(g_{32} \circ g_{21})(x_3, x_1) g_{51}(x_5, x_1) - 6C\epsilon && \text{[by (10)]} \\ &\ge \mathbb{E}_{x_1, x_3, x_5}(g_{34} \circ g_{45})(x_3, x_5)(g_{32} \circ g_{21})(x_3, x_1) g_{51}(x_5, x_1) - 11C\epsilon && \text{[by (9)]} \\ &= \text{RHS of (7)}. \end{aligned}$$

□

4. Removal lemmas in sparse graphs

Our proof of the sparse removal lemma uses, as a black box, the usual graph removal lemma (for dense graphs). We use the following formulation of the graph removal lemma allowing both vertex and edge weights. A standard sampling argument shows that this formulation (at least with a finite V) is equivalent to the more usual version without weights. Alternatively, the standard proof of the graph removal lemma using Szemerédi's regularity lemma can easily be amended to give the weighted version.

Theorem 4.1 (Weighted graph removal lemma, dense setting)**.** *For every graph F and $\epsilon > 0$, there exists some $\delta > 0$ such that for a probability space V and measurable symmetric $g\colon V \times V \to [0,1]$ with $t(F,g) \le \delta$, there exists a measurable symmetric subset $A \subseteq V \times V$ such that $t(F, g1_A) = 0$ and $\|g - g1_A\|_1 \le \epsilon$.*

Now we are ready to prove our main sparse removal lemma. We will first state and prove a version which highlights the hypotheses involved and then show that it implies both Theorems 1.2 and 1.6.

Proposition 4.2 (Removal)**.** *For every $\epsilon, K, C > 0$, there exist $n_0, M, \delta > 0$ such that if $n > n_0$, $p \in (C^{-1}n^{-1/2}, 1]$, and G is an n-vertex graph satisfying*

(a) *(Dense pairs condition) for every partition of $V(G)$ into M parts $V_1 \cup \cdots \cup V_M$, at most a total of $\epsilon pn^2/4$ edges of G lie between pairs (V_i, V_j) with $e(V_i, V_j) \ge Kp\,|V_i|\,|V_j|$,*
(b) *(Not too many 4-cycles) G has at most $(Cpn)^4$ copies of C_4, and*
(c) *(Few 5-cycles) G has at most $\delta p^5 n^5$ copies of C_5,*

then G can be made C_3-free and C_5-free by removing at most ϵpn^2 edges.

Proof. The value of $\delta = \delta(\epsilon, K, C)$ will be given later in the proof using Theorem 4.1. For now, we simply assume that it has a fixed value.

Write $V = V(G)$ and $f = p^{-1}1_G\colon V \times V \to [0, p^{-1}]$ for the normalized edge-indicator function of G. Hypothesis (a) implies that G has at most $(K+\epsilon)pn^2/2$ edges. Hence, $\mathbb{E}f \le K + \epsilon$.

Apply Theorem 2.2, the sparse weak regularity lemma, to the function f/K to obtain a partition $\mathcal{P}$ of V into at most $M = M(K, \epsilon, \delta)$ parts such that

$$\|(f - f_{\mathcal{P}})1_{f_{\mathcal{P}} \le K}\|_{\square} \le K\delta^4,$$

which can be rewritten as

$$\|\widetilde{f} - \widetilde{g}\|_{\square} \le K\delta^4,$$

where

$$\widetilde{f} = f1_{f_{\mathcal{P}} \le K} \quad \text{and} \quad \widetilde{g} = \widetilde{f}_{\mathcal{P}} = f_{\mathcal{P}}1_{f_{\mathcal{P}} \le K}.$$

Note that $\widetilde{g}$ can be obtained from $\widetilde{f}$ by averaging over pairs of parts of $\mathcal{P}$.

By hypothesis (a), G has at most $\epsilon pn^2/4$ edges in $\{f_{\mathcal{P}} > K\} \subseteq V \times V$, since the latter is precisely the union of pairs $V_i \times V_j$ of $\mathcal{P}$ with $e(V_i, V_j) > Kp\,|V_i|\,|V_j|$. The graph obtained after removing these edges from G is represented by $\widetilde{f}$.

Now we would like to use the 5-cycle counting lemma (Theorem 3.2) to deduce that $t(C_5, \widetilde{g})$ must be small from the fact that G has few 5-cycles. This is basically true, but one has to be a bit careful in the application of the C_5-counting lemma. The reason is that while we know that G has few 5-cycles, we have not ruled out the possibility that the triangles of G give rise to many homomorphic copies of C_5. We address this somewhat technical issue by splitting each part of $\mathcal{P}$ arbitrarily into five nearly equal parts labeled by elements of $\mathbb{Z}/5\mathbb{Z}$ and only considering 5-cycles where the i-th vertex is embedded into a part labeled by i, so that all five vertices of the cycle are forced to be distinct. It is in this step that we need the $n > n_0$ hypothesis in the statement of the theorem, in order to guarantee that all five parts are nonempty. Indeed, the theorem as stated is false without the $n > n_0$ hypothesis, with $G = K_3$ being a counterexample. On the other hand,

if we only wish to obtain a C_5-free graph, then the $n > n_0$ hypothesis can be trivially removed by taking δ small enough that for $n \le n_0$ the hypothesis (c) would already imply that G is C_5-free.

For the details, we begin by removing some more edges. Indeed, if some part V_i of the partition $\mathcal{P}$ has at most 100 vertices, then we remove all edges from G with at least one vertex in V_i. We delete at most $100MKpn \le \epsilon pn^2/10$ edges this way, provided that $n > n_0(\epsilon, \delta, K)$ is sufficiently large. From now on, we may therefore assume that all parts of $\mathcal{P}$ have more than 100 vertices.

Partition each $V_i \in \mathcal{P}$ arbitrarily into five parts of nearly equal size (differing by at most 1), $V_i = V_i^{(1)} \cup \cdots \cup V_i^{(5)}$. Let $V^{(r)} = \bigcup_{i\in[M]} V_i^{(r)}$ for each $r \in [5]$. For each $r, s \in [5]$, write

$$f_{rs} = K^{-1}\widetilde{f}1_{V^{(r)}\times V^{(s)}} \quad \text{and} \quad g_{rs} = K^{-1}\widetilde{g}1_{V^{(r)}\times V^{(s)}} \tag{11}$$

(both functions $V \times V \to [0, \infty)$ with the latter taking values in $[0, 1]$). Then

$$\|f_{rs} - g_{rs}\|_\square = K^{-1}\|(\widetilde{f} - \widetilde{g})1_{V^{(r)}\times V^{(s)}}\|_\square \le K^{-1}\|\widetilde{f} - \widetilde{g}\|_\square \le \delta^4.$$

By (b), G has at most $(Cpn)^4$ copies of C_4. We would actually like to upper bound the number of homomorphic copies of C_4 (i.e., closed walks of length 4). Let $r(x, y)$ denote the number of walks of length 2 from x to y in G. Then the number of homomorphic copies of C_4 in G is at most (here we use the inequality $t^2 \le 4\binom{t}{2} + 1$ for all nonnegative integers t)

$$n^2 + 2\sum_{\substack{x,y\in V(G)\\ x\ne y}} r(x, y)^2 \le n^2 + 2\sum_{\substack{x,y\in V(G)\\ x\ne y}} \left(4\binom{r(x, y)}{2} + 1\right) = 3n^2 + O\big((Cpn)^4\big) = O\big((Cpn)^4\big),$$

where the final step uses the hypothesis $p \ge C^{-1}n^{-1/2}$. After normalization, we obtain $t(C_4, f) = O(C^4)$. Hence,

$$\|f_{r-1,r} \circ f_{r,r+1}\|_2^2 \le K^{-4}t(C_4, f) = O(C^4K^{-4}).$$

Applying the 5-cycle counting lemma (Theorem 3.2) to the functions $f_{r,r+1}$ and $g_{r,r+1}$, with $r \in \mathbb{Z}/5\mathbb{Z}$, we obtain

$$\mathbb{E}_{x_1,\dots,x_5\in V}\prod_{r=1}^{5} f_{r,r+1}(x_r, x_{r+1}) \ge \mathbb{E}_{x_1,\dots,x_5\in V}\prod_{r=1}^{5} g_{r,r+1}(x_r, x_{r+1}) - O(C^4K^{-4}\delta),$$

which, by (11), can be rewritten as

$$\mathbb{E}_{x_1,\dots,x_5\in V}\prod_{r=1}^{5} \widetilde{f}(x_r, x_{r+1})1_{V^{(r)}}(x_r) \ge \mathbb{E}_{x_1,\dots,x_5\in V}\prod_{r=1}^{5} \widetilde{g}(x_r, x_{r+1})1_{V^{(r)}}(x_r) - O(C^4K\delta). \tag{12}$$

Each V_i has size at least 100, so each of its 5 parts $V_i^{(r)}$, $r \in [5]$, has at least a 1/6-fraction of the vertices. Thus, $|V^{(r)}| \ge |V|/6$ for each $r \in [5]$. Note that $\widetilde{g}$ is constant on each $V_i \times V_j$. It follows that the RHS of (12) is at least $6^{-5}t(C_5, \widetilde{g}) - O(C^4K\delta)$. On the other hand, the LHS of (12) is $p^{-5}n^{-5}$ times the number of 5-cycles in G with the r-th vertex in $V^{(r)}$ for each $r \in [5]$ and so the LHS of (12) is at most δ by hypothesis (c) that G has at most δn^5p^5 copies of C_5. Putting these two bounds together, we obtain

$$\delta \ge 6^{-5}t(C_5, \widetilde{g}) - O(C^4K\delta).$$

Thus, $t(C_5, \widetilde{g}) \lesssim (1 + C^4K)\delta$. Now apply Theorem 4.1, the graph removal lemma, for C_5 and choose $\delta = \delta(\epsilon, K, C)$ small enough so that the above upper bound on $t(C_5, \widetilde{g})$ guarantees that there exists some symmetric $A \subseteq V \times V$, which is a union of pairs of parts of the partition $\mathcal{P}$, such that $\|\widetilde{g} - \widetilde{g}1_A\|_1 \le \epsilon/3$ and $t(C_5, \widetilde{g}1_A) = 0$. Note that we need to apply the dense removal lemma to the weighted graph with vertices being the parts of $\mathcal{P}$ and vertex weights proportional to the sizes of the parts so that the dense removal lemma outputs an A of the desired form.

So $t(C_5, \widetilde{g}1_A) = 0$ and, consequently, $t(C_3, \widetilde{g}1_A) = 0$ as well. Thus, $t(C_5, \widetilde{f}1_A) = 0$ and $t(C_3, \widetilde{f}1_A) = 0$, since if f takes some positive values on $V_i \times V_j$, then so must g. Since $\widetilde{g}$ is obtained from $\widetilde{f}$ by averaging over each pair of parts of $\mathcal{P}$,

$$\|\widetilde{f} - \widetilde{f}1_A\|_1 = \|\widetilde{g} - \widetilde{g}1_A\|_1 \le \epsilon/3.$$

To conclude, the graph G can be made C_3-free and C_5-free by removing at most ϵpn^2 edges: at most $\epsilon pn^2/4$ edges between pairs of parts (V_i, V_j) with $e(V_i, V_j) \ge Kp\,|V_i|\,|V_j|$, at most $\epsilon pn^2/10$ edges with at least one endpoint in some part with at most 100 vertices, and at most $\epsilon pn^2/3$ edges outside the set A in the last step above. □

Recall the following equivalent way of stating Theorem 1.2:

> For every $\epsilon > 0$, there exist n_0 and $\delta > 0$ such that, for every $n \ge n_0$, every n-vertex graph with at most $\delta n^{5/2}$ copies of C_5 and at most δn^2 copies of C_4 can be made C_3-free and C_5-free by deleting at most $\epsilon n^{3/2}$ edges.

Proof of Theorem 1.2. Let G be an n-vertex graph with at most $\delta n^{5/2}$ copies of C_5 and at most δn^2 copies of C_4. Applying Lemma 2.4, we see that for every partition of $V(G)$ into M parts $V_1 \cup \cdots \cup V_M$, the number of edges of G that lie between "dense pairs" (V_i, V_j) with $e(V_i, V_j) \ge Kn^{-1/2}\,|V_i|\,|V_j|$ is $O(n^{3/2}/K + M^2 n + M^{1/4}\delta^{1/4} n^{3/2})$. For $p = n^{-1/2}$, this is at most $\epsilon pn^2/4$ provided that K is a sufficiently large constant times $1/\epsilon$, δ is a sufficiently small constant times ϵ^4/M, and n is sufficiently large (depending on ϵ and M). Thus, condition (a) of Proposition 4.2 is satisfied.

Condition (b) of Proposition 4.2 is also automatically satisfied with $C = 1$ provided that $\delta < 1$, as is Condition (c) for δ sufficiently small. It thus follows from Proposition 4.2 that one can remove at most $\epsilon n^{3/2}$ edges to make G both C_3-free and C_5-free. □

Theorem 1.6 is the 5-partite version of the sparse C_5-removal lemma, where we instead assume that there are at most δn^2 copies of C_4 between any two consecutive vertex sets.

Proof of Theorem 1.6. The proof is nearly identical to the proof of Theorem 1.2. Set $p = n^{-1/2}$ as earlier. As before, we apply Lemma 2.4 to each bipartite graph $V_i \times V_{i+1}$ to yield that at most $\epsilon pn^2/4$ edges lie between dense pairs for any partition into at most M parts. This verifies condition (a) of Proposition 4.2.

To verify condition (b) of Proposition 4.2, note that the number of 4-cycles spanning three parts V_{i-1}, V_i, V_{i+1} is given by (the sums are taken over all unordered pairs of distinct vertices $x, y \in V_i$ and $\mathrm{codeg}_j(x, y)$ is the number of common neighbors of x and y in V_j)

$$\sum_{x \ne y \in V_i} \mathrm{codeg}_{i-1}(x,y)\,\mathrm{codeg}_{i+1}(x,y) \le \left(\sum_{x \ne y \in V_i} \mathrm{codeg}_{i-1}(x,y)^2\right)^{1/2} \left(\sum_{x \ne y \in V_i} \mathrm{codeg}_{i+1}(x,y)\right)^{1/2},$$

where the inequality follows from Cauchy–Schwarz. Using that $t^2 \le 1 + 4\binom{t}{2}$ for all nonnegative integers t and provided that $\delta < 1/8$, we have

$$\sum_{x \ne y \in V_i} \mathrm{codeg}_{i-1}(x,y)^2 \le \binom{n}{2} + 4 \sum_{x \ne y \in V_i} \binom{\mathrm{codeg}_{i-1}(x,y)}{2} \le \binom{n}{2} + 4\delta n^2 \le n^2,$$

where we apply hypothesis (a) of Theorem 1.6 that there are at most δn^2 C_4's between V_{i-1} and V_i. Likewise,

$$\sum_{x \ne y \in V_i} \mathrm{codeg}_{i+1}(x,y)^2 \le n^2.$$

Hence, the number of 4-cycles spanning the vertex sets V_{i-1}, V_i, V_{i+1} is at most n^2 for each i. Thus, condition (b) of Proposition 4.2 is satisfied with $C = 1$.

Finally, condition (c) of Proposition 4.2 is satisfied due to hypothesis (b) of Theorem 1.6.

The C_5-removal claim thus follows. Note that the "sufficiently large n" condition is superfluous here, since we can always make δ smaller to take care of the finite number of potentially exceptional values of n. □

5. Removal lemma corollaries

Here we prove Corollary 1.3 of the sparse removal lemma Theorem 1.2, saying that an n-vertex graph with $o(n^2)$ C_5's can be made triangle-free by removing $o(n^{3/2})$ edges.

In each of the following proofs, we let G be the graph in the statement and G' be a subgraph of G whose edge set is the union of a maximal collection of edge-disjoint triangles in G. In order to show that G can be made triangle-free by deleting $o(n^{3/2})$ edges, it is sufficient to show that G' has $o(n^{3/2})$ edge-disjoint triangles, which is in turn equivalent to showing that G' can be made triangle-free by deleting $o(n^{3/2})$ edges. This last statement is what we will show.

Let us first prove a corollary of Theorem 1.2 that may be of independent interest. The *house* graph is depicted below.

Corollary 5.1. *An n-vertex graph with $o(n^2)$ C_4's that extend to houses and $o(n^{5/2})$ C_5's can be made triangle-free by deleting $o(n^{3/2})$ edges.*

Proof. It is easy to check that in an edge-disjoint union of triangles, every C_4 extends to a house. Following the notation above, since each C_4 in G' extends to a house in G', G' has $o(n^2)$ C_4's. Moreover, G' has $o(n^{5/2})$ C_5's, since the same is true in G. Applying Theorem 1.2 then yields that G' can be made triangle-free by removing $o(n^{3/2})$ edges. □

Proof of Corollary 1.3. Since G' is an edge-disjoint union of triangles, every C_4 in G' extends to a house, which contains a C_5. Moreover, each C_5 in G' can arise from at most five different C_4's in this way. Since G' has $o(n^2)$ C_5's, we see that G' has $o(n^2)$ C_4's. Thus, Theorem 1.2 implies that G' can be made triangle-free by removing $o(n^{3/2})$ edges. □

6. Number-theoretic applications

Suppose that $a_1, \dots, a_5$ are fixed nonzero integers summing to zero and $X_1, \dots, X_5$ are subsets of $[n]$ such that each X_i has $o(n)$ nontrivial solutions to $x_1+x_2 = x_3+x_4$ and there are $o(n^{3/2})$ solutions to $a_1x_1 + \cdots + a_5x_5 = 0$ with $x_1 \in X_1, \dots, x_5 \in X_5$. Then Theorem 1.16, which we now prove, says that we can remove $o(\sqrt{n})$ elements from each X_i to remove all solutions to $a_1x_1 + \cdots + a_5x_5 = 0$ with $x_1 \in X_1, \dots, x_5 \in X_5$.

Proof of Theorem 1.16. Embed $X_1, \dots, X_5$ into $\mathbb{Z}/N\mathbb{Z}$, where N is the smallest integer greater than $(|a_1| + \cdots + |a_5|)n$ (to avoid wraparound issues) which is coprime to each of $a_1, \dots, a_5$. We consider the 5-partite graph G with vertex sets $V_1, \dots, V_5$, each with elements indexed by $\mathbb{Z}/N\mathbb{Z}$, and edges $(s, s + a_ix_i) \in V_i \times V_{i+1}$ for all $i \pmod 5$, $s \in \mathbb{Z}/N\mathbb{Z}$, and $x_i \in X_i$.

Now we verify that G satisfies the hypotheses of the sparse 5-cycle removal lemma, Theorem 1.6:

(a) All 4-cycles lying between V_i and V_{i+1} are of the form

$$s, \quad s + a_ix_1, \quad s + a_i(x_1 - x_2), \quad s + a_i(x_1 - x_2 + x_3)$$

and we must have $s = s + a_i(x_1 - x_2 + x_3 - x_4)$ for some x_4 to close off the cycle. In order for the 4-cycle to have distinct vertices, $(x_1, x_2, x_3, x_4) \in X_i^4$ must be a nontrivial solution to $x_1+x_3 = x_2+x_4$. But there are $o(N)$ such 4-tuples, while the choice of $s \in \mathbb{Z}/N\mathbb{Z}$ is arbitrary, so there are $o(N^2)$ 4-cycles between each pair (V_i, V_{i+1}).

(b) The number of 5-cycles in G equals N times the number of solutions to $a_1x_1 + \cdots + a_5x_5 = 0$ with $x_1 \in X_1$, …, $x_5 \in X_5$, so there are $o(N^{5/2})$ C_5's.

Thus, by Theorem 1.6, G can be made C_5-free by removing $o(N^{3/2})$ edges. In each X_i, we now remove the element x_i if at least $N/5$ edges of the form $(s, s + a_i x_i)$ have been removed. Since we removed $o(N^{3/2})$ edges from G, we remove $o(\sqrt{N}) = o(\sqrt{n})$ elements from each X_i. Let X_i' denote the remaining elements of X_i. For any solution to $a_1x_1 + \cdots + a_5x_5 = 0$ with $x_1 \in X_1, \dots, x_5 \in X_5$, consider the N edge-disjoint 5-cycles in the graph G that arise from this solution. We must have removed at least one edge from each of these 5-cycles and so we must have removed at least $N/5$ edges of the form $(s, s + a_i x_i)$ for some i, which implies that $x_i \notin X_i'$. There must therefore be no solution to $a_1x_1 + \cdots + a_5x_5 = 0$ with $x_1 \in X_1', \dots, x_5 \in X_5'$, as required. □

Recall that the following lemma, saying that a Sidon set has $O(n)$ solutions to $a_1x_1+a_2x_2+a_3x_3+a_4x_4 = 0$ for any nonzero integers $a_1, \dots, a_4$, was needed to derive Theorem 1.14 from Theorem 1.16.

Proof of Lemma 1.17. Writing

$$\widehat{1}_X(t) = \sum_{x \in X} e^{2\pi i x t},$$

we have, by a standard Fourier identity (easy to see by expansion), that

$$\begin{aligned}
&\left|\left\{(x_1, x_2, x_3, x_4) \in X^4 : a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = 0\right\}\right| \\
&\quad = \int_0^1 \widehat{1}_X(a_1t)\widehat{1}_X(a_2t)\widehat{1}_X(a_3t)\widehat{1}_X(a_4t)\, dt \\
&\quad \le \prod_{j=1}^4 \left(\int_0^1 |\widehat{1}_X(a_jt)|^4\, dt\right)^{1/4}.
\end{aligned}$$

Moreover,

$$\int_0^1 |\widehat{1}_X(a_jt)|^4\, dt = \left|\left\{(x_1, x_2, x_3, x_4) \in X^4 : x_1 + x_2 = x_3 + x_4\right\}\right| = 2|X|^2 - |X| = O(n),$$

since $X \subset [n]$ is a Sidon set. □

7. Some constructions

In the previous section, we deduced our number-theoretic results by starting with a set of integers avoiding solutions to certain equations and building an associated graph to which we could apply our removal lemma. We now use this same idea to prove Proposition 1.7, which asserts the existence of an n-vertex C_4-free graph with $n^{3/2-o(1)}$ edges where every edge lies in exactly one 5-cycle.

Proof of Proposition 1.7 . First, we note that there exists a set $X \subseteq [n]$ with $|X| \ge e^{-O(\sqrt{\log n})}n^{1/2}$ with

(1) no nontrivial solutions to the equation $a(x_1 - x_2) = b(x_3 - x_4)$ for all $a, b \in \{1, 2, 3, 4, 10\}$ (here a solution is called trivial if $x_1 = x_2$ and $x_3 = x_4$ or $a = b$, $x_1 = x_3$ and $x_2 = x_4$) and
(2) no nontrivial solutions to the equation $x_1 + 2x_2 + 3x_3 + 4x_4 = 10x_5$ (here a solution is called trivial if $x_1 = \cdots = x_5$).

The existence of such a set follows from two ingredients, both essentially noted by Ruzsa [41]. Indeed, sets of size $e^{-O(\sqrt{\log n})}n^{1/2}$ satisfying the first property can be constructed through a minor modification of [41, Theorem 7.3], as noted in [8]. Moreover, a set of size $e^{-O(\sqrt{\log n})}n$ satisfying the second property exists by a standard adaptation [41, Theorem 2.3] of Behrend's construction [4]. Since the second property is translation invariant, taking the intersection of a random translation of the second set with the first set gives a set with the claimed size satisfying both properties.

Let $N = 60n + 1$. Let $V_1, \dots, V_5$ be vertex sets each with vertices indexed by $\mathbb{Z}/N\mathbb{Z}$. Add a 5-cycle $(s, s + x, s + 3x, s + 6x, s + 10x) \in V_1 \times \cdots \times V_5$ to the graph for each $s \in \mathbb{Z}/N\mathbb{Z}$ and $x \in X$. This construction is similar to the construction used in the proof of Theorem 1.16.

We now show that this graph has all of the required properties. First note that the cycles are edge-disjoint, since if, for instance, $s+3x = s'+3x'$ and $s+6x = s'+6x'$, then $s = s'$ and $x = x'$, where we used that N is coprime to $\{1,2,3,4,10\}$. Moreover, there are no further C_5's in this graph, since any other such 5-cycle would give a nontrivial solution to the equation $x_1+2x_2+3x_3+4x_4 = 10x_5$ in X.

To see that there are no C_4's, note that there are no C_4's between any pair of vertex sets since X is Sidon. Moreover, there are no C_4's across three vertex sets, since, for instance, any 4-cycle spanning V_2, V_3, V_4 would induce a nontrivial solution to the equation $2(x_1-x_2) = 3(x_3-x_4)$. □

Remark. We do not know how to show that the exponent in Corollary 1.4 is best possible. Recall the statement, that every n-vertex C_5-free graph can be made triangle-free by deleting $o(n^{3/2})$ edges. To match this with a lower bound, we would like to construct a tripartite graph between sets of order n which is the union of $n^{3/2-o(1)}$ edge-disjoint triangles but containing no C_5. Following the strategy above, this would require a set $X \subseteq [n]$ of size $n^{1/2-o(1)}$ which satisfies properties similar to (1) and (2), one particular case being that there should be no nontrivial solutions to the equation $x+2y+2z = 2u+3w$. At present, we do not know how to construct such a set satisfying even this latter property on its own. In fact, it is entirely plausible that no such set exists.

Finally, we prove Proposition 1.5, which says that there exist n-vertex graphs with $o(n^{2.442})$ C_5's that cannot be made triangle-free by deleting $o(n^{3/2})$ edges.

Proof of Proposition 1.5. Let G be the m-th tensor power of a triangle. In other words, its vertices can be labeled by $\mathbb{F}_3^m$ and two vertices are adjacent iff they differ in every coordinate. Note that every edge of G lies in a unique triangle and this graph has 6^{m-1} such triangles.

Write $\hom(H,G)$ for the number of homomorphisms from H to G. The number of closed walks of length k in C_3 is equal to the k-th moment of the eigenvalues of the adjacency matrix of C_3 and so $\hom(C_k, C_3) = 2^k + 2(-1)^k$. Thus,

$$\hom(C_3, G) = \hom(C_3, C_3)^m = 6^m,$$
$$\hom(C_4, G) = \hom(C_4, C_3)^m = 18^m,$$

and

$$\hom(C_5, G) = \hom(C_5, C_3)^m = 30^m.$$

Let G' be the graph obtained from G by keeping every triangle independently with probability $p = (\sqrt{3}/2)^m$, so that in expectation $p6^{m-1} = \Theta(3^{3m/2})$ triangles remain in G'.

To estimate the expected number of C_5 in G', note that the number of C_5 that intersect exactly four triangles of G is $O(18^m)$, since every such C_5 extends to a house, which contains a C_4. Furthermore, it is impossible for a C_5 in G to intersect at most three triangles, as every edge of G is contained in exactly one triangle. Thus, the expected number of C_5 in G' is $O(p^5 30^m + p^4 18^m) = O((3^{7/2} \cdot 5/16)^m) = o(3^{2.442m})$.

Therefore, by Markov's inequality, we see that with positive probability G' is a graph on $n = 3^m$ vertices with $o(n^{2.442})$ C_5's which is an edge-disjoint union of $\Theta(n^{3/2})$ triangles. □

Acknowledgments

We would like to thank Jacques Verstraëte and József Solymosi for helpful comments.

References

[1] Peter Allen, Peter Keevash, Benny Sudakov, and Jacques Verstraëte, *Turán numbers of bipartite graphs plus an odd cycle*, J. Combin. Theory Ser. B **106** (2014), 134–162.

[2] Noga Alon and Clara Shikhelman, *Many T copies in H-free graphs*, J. Combin. Theory Ser. B **121** (2016), 146–172.

[3] József Balogh, Robert Morris, and Wojciech Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), 669–709.

[4] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331–332.

[5] Béla Bollobás and Ervin Győri, *Pentagons vs. triangles*, Discrete Math. **308** (2008), 4332–4336.

[6] W. G. Brown, *On graphs that do not contain a Thomsen graph*, Canad. Math. Bull. **9** (1966), 281–285.

[7] W. G. Brown, P. Erdős, and V. T. Sós, *Some extremal problems on r-graphs*, New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich, 1971), 1973, pp. 53–63.

[8] Javier Cilleruelo and Craig Timmons, *k-fold Sidon sets*, Electron. J. Combin. **21** (2014), Paper 4.12, 9.

[9] D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math. (2) **184** (2016), 367–454.

[10] D. Conlon, W. T. Gowers, W. Samotij, and M. Schacht, *On the KŁR conjecture in random graphs*, Israel J. Math. **203** (2014), 535–580.

[11] David Conlon and Jacob Fox, *Graph removal lemmas*, Surveys in combinatorics 2013, London Math. Soc. Lecture Note Ser., vol. 409, Cambridge Univ. Press, Cambridge, 2013, pp. 1–49.

[12] David Conlon, Jacob Fox, and Yufei Zhao, *Extremal results in sparse pseudorandom graphs*, Adv. Math. **256** (2014), 206–290.

[13] David Conlon, Jacob Fox, and Yufei Zhao, *The Green-Tao theorem: an exposition*, EMS Surv. Math. Sci. **1** (2014), 249–282.

[14] David Conlon, Jacob Fox, and Yufei Zhao, *A relative Szemerédi theorem*, Geom. Funct. Anal. **25** (2015), 733–762.

[15] P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), 113–121.

[16] P. Erdős, A. Rényi, and V. T. Sós, *On a problem of graph theory*, Studia Sci. Math. Hungar. **1** (1966), 215–235.

[17] P. Erdős and M. Simonovits, *Compactness results in extremal graph theory*, Combinatorica **2** (1982), 275–288.

[18] Paul Erdős, *Problems and results in combinatorial number theory*, Journées Arithmétiques de Bordeaux (Conf., Univ. Bordeaux, Bordeaux, 1974), Astérisque, vol. 24-25, Soc. Math. France, Paris, 1975, pp. 295–310.

[19] Beka Ergemlidze and Abhishek Methuku, *Triangles in C_5-free graphs and hypergraphs of girth six*, arXiv:1811.11873.

[20] Beka Ergemlidze, Abhishek Methuku, Nika Salia, and Ervin Győri, *A note on the maximum number of triangles in a C_5-free graph*, J. Graph Theory **90** (2019), 227–230.

[21] Jacob Fox, *A new proof of the graph removal lemma*, Ann. of Math. (2) **174** (2011), 561–579.

[22] Alan Frieze and Ravi Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), 175–220.

[23] Zoltán Füredi and Lale Özkahya, *On 3-uniform hypergraphs without a cycle of a given length*, Discrete Appl. Math. **216** (2017), 582–588.

[24] Stefanie Gerke and Angelika Steger, *The sparse regularity lemma and its applications*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 227–258.

[25] W. T. Gowers, *What are dense Sidon subsets of $\{1, 2, \dots, n\}$ like?*, blog post `https://gowers.wordpress.com/2012/07/13/what-are-dense-sidon-subsets-of-12-n-like/`.

[26] Ben Green and Terence Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), 481–547.

[27] Daniel J. Kleitman and Kenneth J. Winston, *On the number of graphs without 4-cycles*, Discrete Math. **41** (1982), 167–172.

[28] Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230.

[29] Y. Kohayakawa, T. Łuczak, and V. Rödl, *On K^4-free subgraphs of random graphs*, Combinatorica **17** (1997), 173–213.

[30] J. Komlós and M. Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993), Bolyai Soc. Math. Stud., vol. 2, János Bolyai Math. Soc., Budapest, 1996, pp. 295–352.

[31] Daniel Král, Oriol Serra, and Lluís Vena, *A combinatorial proof of the removal lemma for groups*, J. Combin. Theory Ser. A **116** (2009), 971–978.

[32] Daniel Kráľ, Oriol Serra, and Lluís Vena, *A removal lemma for systems of linear equations over finite fields*, Israel J. Math. **187** (2012), 193–207.

[33] Felix Lazebnik and Jacques Verstraëte, *On hypergraphs of girth five*, Electron. J. Combin. **10** (2003), Research Paper 25, 15 pp.

[34] László Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012.

[35] László Lovász and Balázs Szegedy, *Szemerédi's lemma for the analyst*, Geom. Funct. Anal. **17** (2007), 252–270.

[36] Cory Palmer, Michael Tait, Craig Timmons, and Adam Zsolt Wagner, *Turán numbers for Berge-hypergraphs and related extremal problems*, Discrete Math. **342** (2019), 1553–1563.

[37] Sean Prendiville, *Solving equations in dense Sidon sets*, Math. Proc. Cambridge. Phil. Soc., to appear.
[38] Vojtěch Rödl and Mathias Schacht, *Regularity lemmas for graphs*, Fete of combinatorics and computer science, Bolyai Soc. Math. Stud., vol. 20, János Bolyai Math. Soc., Budapest, 2010, pp. 287–325.
[39] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.
[40] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam-New York, 1978, pp. 939–945.
[41] Imre Z. Ruzsa, *Solving a linear equation in a set of integers. I*, Acta Arith. **65** (1993), 259–282.
[42] Wojciech Samotij, *Counting independent sets in graphs*, European J. Combin. **48** (2015), 5–18.
[43] A. A. Sapozhenko, *On the number of independent sets in extenders*, Diskret. Mat. **13** (2001), 56–62.
[44] A. A. Sapozhenko, *Asymptotics of the number of sum-free sets in abelian groups of even order*, Dokl. Akad. Nauk **383** (2002), 454–457.
[45] A. A. Sapozhenko, *The Cameron-Erdős conjecture*, Dokl. Akad. Nauk **393** (2003), 749–752.
[46] David Saxton and Andrew Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), 925–992.
[47] Mathias Schacht, *Extremal results for random discrete structures*, Ann. of Math. (2) **184** (2016), 333–365.
[48] Alexander Scott, *Szemerédi's regularity lemma for matrices and sparse graphs*, Combin. Probab. Comput. **20** (2011), 455–466.
[49] Asaf Shapira, *A proof of Green's conjecture regarding the removal properties of sets of linear equations*, J. Lond. Math. Soc. (2) **81** (2010), 355–373.
[50] V. T. Sós, P. Erdős, and W. G. Brown, *On the existence of triangulated spheres in 3-graphs, and related problems*, Period. Math. Hungar. **3** (1973), 221–228.
[51] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.
[52] Endre Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401.
[53] Yufei Zhao, *An arithmetic transference proof of a relative Szemerédi theorem*, Math. Proc. Cambridge Philos. Soc. **156** (2014), 255–261.

Appendix A. Proof of the sparse weak regularity lemma

Here we prove Theorem 2.2, our sparse weak regularity lemma. As with nearly all proofs of regularity lemmas, we keep track of some "energy" function and show, by using a certain defect inequality, that the energy must increase significantly at every iteration of a partition refinement process. As in Scott [48], the energy of a partition will be defined using the convex function

$$\Phi(x) := \begin{cases} x^2 & \text{if } 0 \le x \le 2 \\ 4x - 4 & \text{if } x > 2. \end{cases}$$

The defect inequality is now captured by the following lemma.

Lemma A.1. *Every real-valued random variable $X \ge 0$ with expectation $\mu \le 1$ satisfies*

$$\frac{1}{4}(\mathbb{E}|X-\mu|)^2 \le \mathbb{E}\Phi(X) - \Phi(\mathbb{E}X).$$

Proof. Let $p_1 = \mathbb{P}(X < \mu)$ and $p_2 = \mathbb{P}(X \ge \mu)$. Let $\mu_1 = \mathbb{E}[X|X < \mu]$ and $\mu_2 = \mathbb{E}[X|X \ge \mu]$ (setting $\mu_1 = \mu$ if $p_1 = 0$). Note that $\mu = p_1\mu_1 + p_2\mu_2$. We consider two cases.

Case 1: $\mu_2 \le 2$. By the convexity of Φ,

$$\begin{aligned}\mathbb{E}\Phi(X) - \Phi(\mathbb{E}X) &\ge p_1\Phi(\mu_1) + p_2\Phi(\mu_2) - \Phi(\mu) = p_1\mu_1^2 + p_2\mu_2^2 - \mu^2 \\ &= p_1(\mu-\mu_1)^2 + p_2(\mu_2-\mu)^2 \ge (p_1(\mu-\mu_1) + p_2(\mu_2-\mu))^2 = (\mathbb{E}|X-\mu|)^2.\end{aligned}$$

Case 2: $\mu_2 \ge 2$. Let $\Phi_\mu(x) := \Phi(x) - 2\mu x + \mu^2$, which is convex. Indeed, Φ_μ is a quadratic function on $[0,2]$ with minimum at μ and is linear on $[2,\infty)$. Since $0 \le \mu_1 \le \mu \le 1$ and $\mu_2 \ge 2$, we have $\Phi_\mu(\mu_1) \le \Phi_\mu(0) = \mu^2 \le 1 \le (2-\mu)^2 = \Phi_\mu(2) \le \Phi_\mu(\mu_2)$. Thus,

$$\begin{aligned}\mathbb{E}\Phi(X) - \Phi(\mathbb{E}X) &= \mathbb{E}\Phi_\mu(X) \ge p_1\Phi_\mu(\mu_1) + p_2\Phi_\mu(\mu_2) \ge \Phi_\mu(\mu_1) = (\mu_1-\mu)^2 \\ &\ge \frac{1}{4}(2p_1(\mu-\mu_1))^2 = \frac{1}{4}(p_1(\mu-\mu_1) + p_2(\mu_2-\mu))^2 = \frac{1}{4}(\mathbb{E}|X-\mu|)^2. \qquad \square\end{aligned}$$

Lemma A.2. *Let V be a probability space and $f\colon V \times V \to [0,\infty)$ be a measurable symmetric function. Let $\mathcal{P}$ and $\mathcal{Q}$ be measurable partitions of V such that $\mathcal{Q}$ refines $\mathcal{P}$. Then*

$$\mathbb{E}[\Phi \circ f_{\mathcal{Q}}] - \mathbb{E}[\Phi \circ f_{\mathcal{P}}] \geq \frac{1}{4}(\mathbb{E}[|f_{\mathcal{Q}} - f_{\mathcal{P}}|\, 1_{f_{\mathcal{P}} \leq 1}])^2.$$

Proof. For every pair (U,W) of parts of $\mathcal{P}$, the function $f_{\mathcal{P}}$ is constant on $U \times W$ with value the average of f on $U \times W$, which is the same as the average of $f_{\mathcal{Q}}$ on $U \times W$ since $\mathcal{Q}$ is a refinement of $\mathcal{P}$. So, by the convexity of Φ, we always have $\mathbb{E}_{U\times W}[\Phi \circ f_{\mathcal{Q}}] \geq \mathbb{E}_{U\times W}[\Phi \circ f_{\mathcal{P}}]$. Furthermore, by applying Lemma A.1 on those $U \times W$ where $f_{\mathcal{P}} \leq 1$, we deduce that

$$\mathbb{E}_{U\times W}[\Phi \circ f_{\mathcal{Q}}] - \mathbb{E}_{U\times W}[\Phi \circ f_{\mathcal{P}}] \geq \frac{1}{4}(\mathbb{E}_{U\times W}[|f_{\mathcal{Q}} - f_{\mathcal{P}}|\, 1_{f_{\mathcal{P}} \leq 1}])^2$$

for all $(U,W) \in \mathcal{P} \times \mathcal{P}$. Finally, summing the above inequality over all pairs (U,W) and applying the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\mathbb{E}[\Phi \circ f_{\mathcal{Q}}] - \mathbb{E}[\Phi \circ f_{\mathcal{P}}] &= \sum_{U,W\in\mathcal{P}} \mu(U)\mu(W)\left(\mathbb{E}_{U\times W}[\Phi \circ f_{\mathcal{Q}}] - \mathbb{E}_{U\times W}[\Phi \circ f_{\mathcal{P}}]\right) \\
&\geq \frac{1}{4} \sum_{U,W\in\mathcal{P}} \mu(U)\mu(W)(\mathbb{E}_{U\times W}[|f_{\mathcal{Q}} - f_{\mathcal{P}}|\, 1_{f_{\mathcal{P}} \leq 1}])^2 \\
&\geq \frac{1}{4}\left(\sum_{U,W\in\mathcal{P}} \mu(U)\mu(W)\mathbb{E}_{U\times W}[|f_{\mathcal{Q}} - f_{\mathcal{P}}|\, 1_{f_{\mathcal{P}} \leq 1}]\right)^2 \\
&= \frac{1}{4}(\mathbb{E}[|f_{\mathcal{Q}} - f_{\mathcal{P}}|\, 1_{f_{\mathcal{P}} \leq 1}])^2. \qquad \square
\end{aligned}$$

Now we prove the sparse weak regularity lemma. Recall the statement, that, given $\epsilon > 0$ and $f\colon V \times V \to [0,\infty)$, there exists a partition $\mathcal{P}$ of V into at most $2^{32\mathbb{E}f/\epsilon^2}$ parts such that

$$\|(f - f_{\mathcal{P}})1_{f_{\mathcal{P}} \leq 1}\|_{\square} \leq \epsilon. \tag{13}$$

Proof of Theorem 2.2. Starting with the trivial partition $\mathcal{P}_0$ of V (i.e., consisting of a single part), consider the following iterative process: for each $i \geq 0$, if (13) is satisfied for $\mathcal{P} = \mathcal{P}_i$, then we stop the iteration; otherwise, by the definition of the cut norm, there exist measurable subsets $A, B \subset V$ such that $|\mathbb{E}[(f - f_{\mathcal{P}})1_{f_{\mathcal{P}}\leq 1}1_{A\times B}]| > \epsilon$ and we set $\mathcal{P}_{i+1}$ to be the common refinement of the partition $\mathcal{P}_i$ and the four-part partition determined by A and B. Write

$$\Phi(\mathcal{P}) := \mathbb{E}[\Phi \circ f_{\mathcal{P}}] = \mathbb{E}_{x,y\in V}[\Phi(f_{\mathcal{P}})]$$

for the "energy" of the partition $\mathcal{P}$.

With $\mathcal{P} = \mathcal{P}_i$, $\mathcal{Q} = \mathcal{P}_{i+1}$, and $A, B \subset V$ as above, we have, by Lemma A.2, that

$$\Phi(\mathcal{Q}) - \Phi(\mathcal{P}) \geq \tfrac{1}{4}(\mathbb{E}[|f_{\mathcal{Q}} - f_{\mathcal{P}}|\, 1_{f_{\mathcal{P}} \leq 1}])^2.$$

Since A and B are unions of parts of $\mathcal{Q}$, we have

$$\mathbb{E}[|f_{\mathcal{Q}} - f_{\mathcal{P}}|\, 1_{f_{\mathcal{P}} \leq 1}] \geq |\mathbb{E}[(f_{\mathcal{Q}} - f_{\mathcal{P}})1_{f_{\mathcal{P}}\leq 1}1_{A\times B}]| = |\mathbb{E}[(f - f_{\mathcal{P}})1_{f_{\mathcal{P}}\leq 1}1_{A\times B}]| > \epsilon.$$

Thus, for every $i \geq 0$, one has

$$\Phi(\mathcal{P}_{i+1}) - \Phi(\mathcal{P}_i) \geq \tfrac{1}{4}\epsilon^2.$$

On the other hand, since $0 \leq \Phi(x) \leq 4x$ for all $x \geq 0$, for every partition $\mathcal{P}$ of V, we have

$$\Phi(\mathcal{P}) \leq 4\mathbb{E}f_{\mathcal{P}} = 4\mathbb{E}f.$$

Thus, the iteration must terminate after at most $16\mathbb{E}f/\epsilon^2$ steps, at which point (13) is satisfied. The number of parts increases by a factor of at most 4 in each iteration, so the final number of parts is at most $2^{32\mathbb{E}f/\epsilon^2}$. $\square$

Appendix B. Connecting Brown–Erdős–Sós to the extremal problem for girth

Here we prove Proposition 1.9, which says that $f_r(n,(r-1)e,e) = h_r(n,e)$ for sufficiently large $n \geq n_0(r,e)$. Recall that $f_r(n,v,e)$ is the maximum number of edges in an n-vertex r-graph without a (v,e)-configuration, a subgraph with e edges and at most v vertices. Moreover, $h_r(n,g)$ is the maximum number of edges in an n-vertex r-graph of girth greater than g.

Lemma B.1. *An r-graph H on n vertices without an $((r-1)e,e)$-configuration and with a cycle of length at most e has at most $\max\{e-1, n/(r-1)\}$ edges.*

Proof. By assumption, H contains a cycle C of length $\ell \leq e$. The vertices of this cycle span at most $(r-1)\ell$ vertices. Growing the connected component containing C one edge at a time, we see that each additional edge adds at most $r-1$ new vertices. Let P_0 be the connected component containing C. In particular, if P_0 has at least e edges, then it would contain e edges spanning at most $(r-1)e$ vertices, contradicting that H has no $((r-1)e,e)$-configuration. Hence, P_0 has fewer than e edges. Let $P_1,\dots,P_j$ denote the remaining connected components of H. For $0 \leq i \leq j$, let n_i and e_i denote the number of vertices and edges, respectively, of P_i. Let $p_i = (r-1)e_i - n_i$ and assume, by reordering if necessary, that $p_1 \geq \cdots \geq p_j$. By construction, $p_0 \geq 0$ and $e_0 < e$.

We may assume that H has at least e edges, as otherwise we are done. Let k be the largest index such that $e_0 + \cdots + e_k < e$ and so $e_0 + \cdots + e_k + e_{k+1} \geq e$, as H has at least e edges. Let $e' = e - (e_0 + \cdots + e_k)$. Let Q be a connected subset of P_{k+1} with e' edges formed by starting with a single edge in P_{k+1} and adding edges one at a time, keeping the resulting subset connected, until we get exactly e' edges. By construction, Q has at most $(r-1)e'+1$ vertices. Let U_0 consist of the union of the connected components $P_0,\dots,P_k$ together with Q, so that U_0 has e edges and at most $((r-1)e_0 - p_0) + \cdots + ((r-1)e_k - p_k) + (r-1)(e-(e_0+\cdots+e_k)) + 1 = (r-1)e + 1 - p_0 - \cdots - p_k$ vertices. As H has no $((r-1)e,e)$-configuration, we must have $p_0 + \cdots + p_k \leq 0$. As $p_1 \geq \cdots \geq p_j$, it follows that $p_0 + p_1 + \cdots + p_j \leq 0$. Equivalently, the number of edges in H is at most $n/(r-1)$. □

Lemma B.2. *For $r \geq 2$ and $g \geq 2$, there exists $n_0(r,g)$ such that $h_r(n,g) > n/(r-1)$ for all $n \geq n_0(r,g)$.*

Proof. Construct an r-graph by fixing two vertices u and v and adding edge-disjoint "paths" between u and v, where each "path" consists of $\lfloor g/2 \rfloor + 1$ edges, the first containing u, the last containing v, and where consecutive edges share exactly one vertex. Add as many paths as one can without exceeding n total vertices. Each new path uses $(\lfloor g/2 \rfloor + 1)(r-1) - 1$ new vertices (not counting u and v). The resulting r-graph has girth greater than g and

$$h_r(n,g) \geq \left\lfloor \frac{n-2}{(\lfloor g/2 \rfloor + 1)(r-1) - 1} \right\rfloor (\lfloor g/2 \rfloor + 1)$$

edges, which exceeds $n/(r-1)$ for sufficiently large n. □

Proof of Proposition 1.9. If an r-graph has girth greater than e, then every subset of e edges contains no cycle and hence spans more than $(r-1)e$ vertices. Thus, the r-graph has no $((r-1)e,e)$-configuration and $h_r(n,e) \leq f_r(n,(r-1)e,e)$ for all n.

Conversely, by the previous paragraph and Lemma B.2, we have that for sufficiently large n the largest n-vertex r-graph with no $((r-1)e,e)$-configuration has more than $n/(r-1)$ edges. Therefore, by Lemma B.1, it has no cycle of length at most e. Thus, $f_r(n,(r-1)e,e) \leq h_r(n,e)$. □

Appendix C. Counting 3-graphs with girth greater than 5

Here we prove Theorem 1.11, which says that for every fixed $r \geq 3$, the number of r-graphs on n vertices with girth greater than 5 is $2^{o(n^{3/2})}$.

Proof of Theorem 1.11. Let $F_r(n)$ denote the number of r-graphs on n labeled vertices with girth greater than 5.

First we show that, for every $r \geq 3$, one has $F_r(n) \leq F_3(n)^{\binom{r}{3}}$, thereby reducing the problem to $r = 3$. Indeed, given an r-graph H, color the triples contained in each edge of H arbitrarily from 1 to $\binom{r}{3}$, using one color for each triple. The color classes give a list $H_1, \dots, H_{\binom{r}{3}}$ of 3-graphs, each with girth greater than 5, and thus there are at most $F_3(n)^{\binom{r}{3}}$ possibilities for such a list. Furthermore, one can recover H from the list $H_1, \dots, H_{\binom{r}{3}}$ since two triples in $H_1 \cup \cdots \cup H_{\binom{r}{3}}$ share two vertices if and only if they are contained in the same edge in H (recall that H has no 2-cycles). Thus, there are at most $F_3(n)^{\binom{r}{3}}$ possibilities for H.

Now it remains to show that $F_3(n) = 2^{o(n^{3/2})}$. Let H be a 3-graph with girth greater than 5. By Corollary 1.10, H has $o(n^{3/2})$ edges. Let G be the underlying shadow graph (a pair of vertices form an edge of G if they are contained in a triple of H). Since H has girth greater than 5, every edge in G lies in a unique triangle, and G is C_4-free with $o(n^{3/2})$ edges. Then, by Proposition C.1 below, there are $2^{o(n^{3/2})}$ possibilities for G. The 3-graph H can be recovered uniquely from G and thus the number of such H is also $2^{o(n^{3/2})}$. □

Proposition C.1. *The number of C_4-free graphs on n vertices with $o(n^{3/2})$ edges is $2^{o(n^{3/2})}$.*

Proposition C.1 can be proved by modifying the proof of the following classic result of Kleitman and Winston [27].

Theorem C.2 (Kleitman–Winston)**.** *The number of C_4-free graphs on n vertices is $2^{O(n^{3/2})}$.*

We follow the exposition of Samotij [42, Theorem 8] in his survey on counting independent sets in graphs via graph containers. We begin with the following key lemma.

Lemma C.3 ([42, Lemma 1])**.** *Let G be an n-vertex graph. Suppose that an integer q and reals R and $\beta \in [0,1]$ satisfy $R \geq e^{-\beta q} n$. Suppose that every subset $U \subset V(G)$ with $|U| \geq R$ induces at least $\beta\binom{|U|}{2}$ edges in G. Then, for every integer $m \geq q$, the number of m-element independent sets in G is at most $\binom{n}{q}\binom{R}{m-q}$.*

As in [42], let $g_n(d)$ denote the maximum number of ways to attach a vertex of degree d to an n-vertex C_4-free graph with minimum degree at least $d-1$ in such a way that the resulting graph remains C_4-free. In [42], it was proved that $\max_{d \leq n} g_n(d) \leq e^{O(\sqrt{n})}$. We modify the proof of this statement to obtain the following bound.

Lemma C.4. *If $i \leq n$ and $d = o(\sqrt{n})$, then $g_i(d) = e^{o(\sqrt{n})}$.*

Proof. If $d \leq \sqrt{n}/(\log n)^2$, then $g_i(d) \leq \binom{i}{d} \leq n^d = e^{o(\sqrt{n})}$. So assume $d > \sqrt{n}/(\log n)^2$.

Let G be an i-vertex C_4-free graph with minimum degree at least $d-1$. Let H be the square of G, i.e., $V(H) = V(G)$ and two vertices are adjacent in H if and only if they are connected by a path of two edges in G. Note that attaching a new vertex to G will not create any 4-cycles if and only if the neighborhood of the new vertex is an independent set in H. It remains to upper bound the number of d-element independent sets in H using Lemma C.3.

Let $R = 2n/(d-1)$, $\beta = (d-1)^2/2n$, and $q = \lceil 3(\log n)^5 \rceil$. Since $d > \sqrt{n}/(\log n)^2$, for n sufficiently large we have $\beta q \geq \log n$ and thus $e^{-\beta q} i \leq e^{-\beta q} n \leq 1 \leq R$.

Since G has minimum degree at least $d-1$, every $B \subseteq V(H)$ satisfies $\sum_{z \in V(G)} \deg_G(z, B) \geq (d-1)\,|B|$. Thus, if $|B| \geq R = 2n/(d-1)$, then the number of edges that B induces in H is equal

to

$$\sum_{z\in V(G)}\binom{\deg_G(z,B)}{2} \geq i\binom{\sum_{z\in V(G)}\deg_G(z,B)/i}{2}$$
$$\geq i\cdot\frac{(d-1)\,|B|}{2i}\left(\frac{(d-1)\,|B|}{i}-1\right) \geq \frac{(d-1)^2}{2n}\binom{|B|}{2} = \beta\binom{|B|}{2},$$

where the first inequality uses the convexity of $x\mapsto\binom{x}{2}=\frac{x(x-1)}{2}$.

Applying Lemma C.3, the number of d-element independent sets in H is at most

$$\binom{i}{q}\binom{R}{d-q} \leq n^q\left(\frac{eR}{d-q}\right)^{d-q} \leq e^{O((\log n)^6)}\left(\frac{2ne}{(d-q)^2}\right)^{d-q}.$$

Applying $\lim_{x\to 0}\sqrt{x}\log(1/x)=0$ with $x=(d-q)^2/(2ne)=o(1)$, we see that the right-hand side above is $e^{o(\sqrt{n})}$. Thus, the number of d-element independent sets in H is $e^{o(\sqrt{n})}$. □

Lemma C.5. *A C_4-free graph with minimum degree d must have more than $d^3/2-d^2/2$ edges.*

Proof. Let G be a C_4-free graph with minimum degree d and v a vertex of degree d. As G is C_4-free, each vertex in $N(v)$ is adjacent to at most one vertex in $N(v)$. Thus, each vertex in $N(v)$ has at least $d-2$ neighbors not in $\{v\}\cup N(v)$. As G is C_4-free, each vertex not in $\{v\}\cup N(v)$ has at most one neighbor in $N(v)$, so G has at least $1+d+d(d-2)=d^2-d+1$ vertices. Hence, the number of edges in G is at least $|V(G)|d/2\geq(d^2-d+1)d/2>d^3/2-d^2/2$. □

Proof of Proposition C.1. By iteratively peeling off lowest-degree vertices, we see that every n-vertex graph has an ordering $v_1,\dots,v_n$ of vertices (in reverse order of peeling) such that, for each i, v_i is a minimum-degree vertex in the subgraph G_i induced by $\{v_1,\dots,v_i\}$. Letting d_i denote the degree of v_i in G_i, we see that the minimum degree of G_{i-1} is at least d_i-1.

By Lemma C.5, every induced subgraph of a C_4-free graph with m edges has minimum degree $O(m^{1/3})$. In particular, if the graph has n vertices and $m=o(n^{3/2})$, then $d_i=o(\sqrt{n})$ for all i.

For each fixed ordering of the vertices ($n!$ possibilities) and each fixed sequence $d_2,\dots,d_n$ of degrees (at most $n!$ possibilities), noting that there are at most $g_i(d_{i+1})$ ways to attach v_{i+1} to G_i, we see that the number of C_4-free graphs with these parameters is at most $g_1(d_2)g_2(d_3)\cdots g_{n-1}(d_n)$. By Lemma C.4, this count is $e^{o(n^{3/2})}$, even after summing over the at most $n!^2$ possibilities for the parameters. □

Conlon, Department of Mathematics, California Institute of Technology, Pasadena, CA, USA
Email address: dconlon@caltech.edu

Fox, Department of Mathematics, Stanford University, Stanford, CA, USA
Email address: jacobfox@stanford.edu

Sudakov, Department of Mathematics, ETH, Zurich, 8092, Switzerland
Email address: benjamin.sudakov@math.ethz.ch

Zhao, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA
Email address: yufeiz@mit.edu