1 Introduction

The following question, though vague, is one of the central questions in extremal graph theory.

Question 1.1. *If G is a triangle-free graph, why?*

There is an obvious answer: *because G contains no triangles.*

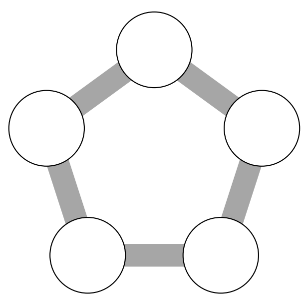
However, there is arguably a better answer, which gets at some of the central results in extremal graph theory: *because G "looks like" a smaller triangle-free graph.*

More precisely, there are a number of results that say that under certain assumptions, there is a *constant-sized* graph Γ which is triangle-free, and whose structure is related to that of G. In particular, the triangle-freeness can be fully explained by the structural similarity to Γ, and thus we obtain a "constant-sized certificate" for the triangle-freeness of G.

This general phenomenon is best explained by some examples, so let's see a few.

Theorem 1.2 (Andrásfai 1964)**.** *If G is an n-vertex triangle-free graph with minimum degree strictly greater than $\frac{2}{5}n$, then G is bipartite.*

It is perhaps not obvious why this is an instance of the general phenomenon I discussed above, and we'll return to this shortly. But before doing so let me remark that the value $\frac{2}{5}n$ appearing in Theorem 1.2 is best possible, as shown by the following example:

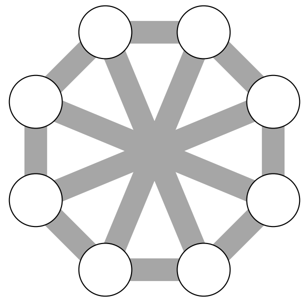


This graph, which is called a (balanced) *blowup* of C_5, is obtained by partitioning the vertex set into five parts of size $n/5$, putting complete bipartite graphs between cyclically consecutive pairs of parts, and no other edges. The concept of blowups will recur throughout the talk, and it will always mean this: a *blowup* of a graph Γ is obtained by replacing each vertex of Γ by an independent set, replacing each edge of Γ by a complete bipartite graph, and putting in no other edges.

It is easy to check that the blowup of C_5 is triangle-free and has minimum degree exactly $\frac{2}{5}n$, but it is not bipartite. As such, Theorem 1.2 is best possible. However, if we slightly weaken the conclusion, we can obtain the following generalization.

Theorem 1.3 (Häggkvist 1982)**.** *If G is an n-vertex triangle-free graph with minimum degree strictly greater than $\frac{3}{8}n$, then G is bipartite or a subgraph of a blowup of C_5.*

In other words, while Theorem 1.2 is tight—in the sense that at minimum degree $\frac{2}{5}n$ a new structure, the blowup of C_5, becomes possible—this is the *only* new structure for another interval of minimum degrees, until $\frac{3}{8}n$. Again, Theorem 1.3 is best possible, as shown by another blowup, this time of an 8-vertex graph:



One can check that this blowup has minimum degree $\frac{3}{8}n$, is triangle-free, and is neither bipartite nor a subgraph of a blowup of C_5.

As we will be discussing subgraphs of blowups a lot throughout this talk, it is useful to give this concept a name.

Definition 1.4. Let G, Γ be graphs. We say that G is *homomorphic* to Γ, and write $G \to \Gamma$, if there is a map $V(G) \to V(\Gamma)$ that maps edges of G to edges of Γ.

It is not hard to check that $G \to \Gamma$ if and only if G is a subgraph of a blowup of Γ. Indeed, if G is a subgraph of a blowup of Γ, then we may define a homomorphism $G \to \Gamma$ by sending each vertex v of G to the vertex of Γ corresponding to the blob of the blowup that v lies in. Conversely, a homomorphism $G \to \Gamma$ describes an embedding of G into any sufficiently large blowup of Γ, by sending vertices of G to distinct elements of the blowup blob corresponding to their image under the homomorphism.

Note that a graph is bipartite if and only if it's a subgraph of a blowup of a single edge, i.e. homomorphic to K_2. Thus, with this terminology, we can restate Theorem 1.2 as saying that if G is a triangle-free graph with minimum degree greater than $\frac{2}{5}n$, then $G \to K_2$. Similarly, Theorem 1.3 says that if G is a triangle-free graph with minimum degree greater than $\frac{3}{8}n$, then G is homomorphic to one of K_2 and C_5.

To bring this discussion back to Question 1.1, we note the following simple observation: if $G \to \Gamma$ and Γ is triangle-free, then G is triangle-free as well. Indeed, if Γ is triangle-free then so is every blowup of it, and thus so is every subgraph of a blowup. But the point is that if G is homomorphic to a triangle-free graph Γ, which we think of as "small", then we obtain a "small" explanation of the triangle-freeness of G. The structure of G can be described as (contained within) the structure of Γ, and this information is sufficient to conclude that G is triangle-free. In particular, if Γ is of constant size, we obtain a constant-sized certificate of the triangle-freeneess of Γ.

With this perspective, we see that results like Theorems 1.2 and 1.3 fit in the framework of Question 1.1. They tell us that if G has sufficiently high minimum degree, then the "reason" why G is triangle-free is that it looks like a very small triangle-free graph Γ (namely,

$\Gamma \in \{K_2, C_5\}$). Of course, Theorem 1.3 is not the end of the story of such results. As you might expect, there is another interval of minimum degrees (until $\frac{4}{11}n$) where the three structures we've already seen are the only ones that can appear, at which point a fourth structure appears. These four remain the only ones for another interval, and then a fifth appears. However, the sequence of minimum degree conditions has a limit point at $\frac{1}{3}$, at which point the pattern breaks; the precise result is given in the following pair of theorems.

Theorem 1.5 (Łuczak 2006)**.** *For every $\alpha > \frac{1}{3}$, there exists a finite set $\{\Gamma_1, \ldots, \Gamma_m\}$ of triangle-free graphs such that the following holds. If G is an n-vertex triangle-free graph with minimum degree strictly greater than αn, then $G \to \Gamma_i$ for some i.*

Theorem 1.6 (Hajnal 1973)**.** *For every $\alpha < \frac{1}{3}$, every integer K, and all sufficiently large n, there exists an n-vertex triangle-free graph G with minimum degree strictly greater than αn which is not homomorphic to any graph Γ on at most K vertices.*

Note that in Theorem 1.6, we obtain that G is not homomorphic to *any* constant-sized graph, regardless of whether it's triangle-free or not. Moreover, let me remark that later work of Brandt and Thomassé has completely characterized the graphs $\Gamma_1, \ldots, \Gamma_m$ appearing in Theorem 1.5, and hence gives the most general extension of Theorems 1.2 and 1.3 possible.

Note that all of the results we've discussed so far have some serious limitations. First, by Theorem 1.6, they can only apply for rather dense graphs, namely those of minimum degree greater than $n/3$. Second, even when they do apply, the assumption is rather stringent: minimum degree conditions are quite rigid, and theorems like Theorems 1.2 and 1.3 become false if we allow even a single vertex of lower degree. Of course, the conclusion given by these results is itself rather strong, namely that the triangle-freeness of G is *fully* explained by the constant-sized certificate Γ. In what follows, we will mostly focus on theorems that have a weaker hypothesis—and thus apply much more generally—at the expense of a weaker, more *approximate*, conclusion.

2 Removal and approximate homomorphisms

Actually, before continuing our discussion of graphs with *no* triangles, let us briefly turn our attention to graphs with *few* triangles. In this direction, the most important result is the following fundamental theorem, now called the *triangle removal lemma.*

Theorem 2.1 (Ruzsa–Szemerédi 1978)**.** *For every $\varepsilon > 0$, there exists some $\delta = \delta(\varepsilon) > 0$ such that the following holds. If G is an n-vertex graph with at most δn^3 triangles, then G can be obtained from a triangle-free graph G_0 by adding at most εn^2 edges.*

In other words, the *only* way to find a graph with few triangles is to start with a graph G_0 with *no* triangles, and then add to it a small number of edges. Despite its simple statement, this is an extraordinarily deep result, with a large number of applications in disparate fields such as number theory, computer science, discrete geometry, and graph theory.

There is a great deal we still do not understand about the triangle removal lemma. Perhaps most importantly, we really do not understand the optimal quantitative dependence of δ on ε. At present, the best known bounds are

$$\left(\frac{1}{\varepsilon}\right)^{\Omega(\log \frac{1}{\varepsilon})} \leqslant \frac{1}{\delta(\varepsilon)} \leqslant \left.2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}\right\} O(\log \tfrac{1}{\varepsilon}),$$

due to Ruzsa–Szemerédi and Fox, respectively. The precise details are not so important: the thing I want to stress is that the two bounds are extremely far apart (it is a major open problem to narrow the gap), and that the lower bound is *super-polynomial*: we cannot take δ to depend polynomially on ε in Theorem 2.1.

The triangle removal lemma allows us, in some sense, to reduce the study of graphs with few triangles to the study of graphs with no triangles. Moreover, there is a simple strengthening of Theorem 2.1 which allows us to obtain substantially more control on the mysterious triangle-free graph G_0. This strengthening appears to have first been noted by Tao.

Theorem 2.2 (Bounded-complexity triangle removal lemma)**.** *For every $\varepsilon > 0$, there exists some $M = M(\varepsilon) \in \mathbb{N}$ such that the following holds. In Theorem 2.1, we may take G_0 to be homomorphic to some triangle-free graph Γ on at most M vertices.*

In other words, not only can we decompose G as a triangle-free graph G_0 plus some small amount of noise, but moreover, we can get a constant-sized certificate for the triangle-freeness of G_0, in the same sense as previously.

It is helpful to give a name to the kind of information given by Theorem 2.2.

Definition 2.3. Let G, Γ be two graphs, where G has n vertices. We say that G is ε-*approximately homomorphic* to Γ, and write $G \xrightarrow{\varepsilon} \Gamma$ if there is a map $V(G) \to V(\Gamma)$ which maps all but at most εn^2 edges of G to edges of Γ.

Thus, $G \xrightarrow{\varepsilon} \Gamma$ if and only if G can be obtained from Γ by blowing up, passing to a subgraph, and then adding up to εn^2 "noise" edges. Let me stress that this notion is only meaningful for dense graphs G: if G has at most εn^2 edges already, then it is trivially ε-approximately homomorphic to any graph.

With this terminology, Theorem 2.2 states that if G contains at most δn^3 triangles, then $G \xrightarrow{\varepsilon} \Gamma$ for some triangle-free graph Γ on at most M vertices. Moreover, note that this statement is meaningful and interesting even for graphs with *no* triangles; as this is our main focus today, let me explicitly record this.

Theorem 2.4 (Approximate homomorphism theorem/triangle-free lemma)**.** *For every $\varepsilon > 0$, there exists some $M = M(\varepsilon) \in \mathbb{N}$ such that the following holds. If G is a triangle-free graph, then $G \xrightarrow{\varepsilon} \Gamma$ for some triangle-free graph Γ on at most M vertices.*

Note that this gives us yet another version of the answer to Question 1.1: *any* triangle-free graph can have its triangle-freeness explained—up to a small amount of noise—by a constant-sized certificate, where the constant depends only on the amount of noise we are willing to tolerate. This is an extremely general and, in my opinion, quite astonishing theorem.

Unfortunately, it turns out that the quantitative aspects of Theorem 2.4 are closely related to—but even worse than—those of the triangle removal lemma. Concretely, we have the following result, which I will state a bit imprecisely.

"Theorem" 2.5 (Hoppen–Kohayakawa–Lang–Lefmann–Stagni '20, Fox–Zhao '22)**.** *There are upper and lower bounds on $M(\varepsilon)$ of the form $\exp(1/\delta(\varepsilon))$, where $\delta(\varepsilon)$ is the optimal constant in Theorem 2.1. As a consequence,*

$$2^{(1/\varepsilon)^{\Omega(\log \frac{1}{\varepsilon})}} \leqslant M(\varepsilon) \leqslant \left. 2^{2^{\cdot^{\cdot^{\cdot^{2}}}}} \right\} O(\log \tfrac{1}{\varepsilon}). \tag{1}$$

In particular, this result shows us that $M(\varepsilon)$ is *super-exponential* in $1/\varepsilon$, i.e. grows faster than 2 to any fixed power of $1/\varepsilon$. Also, it shows that narrowing the gap in (1) is essentially the same problem as narrowing the gap for the best bounds in the triangle removal lemma, which is a notorious open problem.

3 New results

Our new main result shows that the value of $M(\varepsilon)$ can change rather radically if one imposes mild extra conditions on G. Namely, we will assume that G is both triangle-free and C_ℓ-free, for some odd ℓ. Let me remark here that it is not interesting to require G to be C_ℓ-free for any even ℓ, since it is known that if G is C_ℓ-free for any fixed even ℓ, then G has $o(n^2)$ edges, and hence is ε-approximately homomorphic to any graph (if n is sufficiently large). Our theorem shows that there is a surprising "double phase transition" in the value of $M(\varepsilon)$, depending on the value of ℓ.

Theorem 3.1 (Gishboliner–Hurley–W. 2025+)**.** *Fix an odd integer ℓ and some $\varepsilon > 0$. Let G be an n-vertex graph that is both triangle-free and C_ℓ-free. Then $G \xrightarrow{\varepsilon} \Gamma$, for some triangle-free graph Γ with at most $M_\ell(\varepsilon)$ vertices, where*

$$M_\ell(\varepsilon) = \begin{cases} \textit{super-exponential} & \textit{if } \ell = 3, \\ 2^{\mathrm{poly}(1/\varepsilon)} & \textit{if } \ell \in \{5,7\}, \\ O(1/\varepsilon) & \textit{if } \ell \geqslant 9. \end{cases}$$

Some remarks are in order. First, if $\ell = 3$, then we are back in the setting of Theorem 2.4, so the first stated bound is simply what is given in (1). Second, in case $\ell \in \{5,7\}$, then we can prove *both* lower and upper bounds of the form $2^{\mathrm{poly}(1/\varepsilon)}$. That is, assuming that G is C_5-free or C_7-free allows us to improve the super-exponential bound from (1) to an exponential bound, but no further. Furthermore, if $\ell \geqslant 9$, then this bound can be further improved all the way to linear (we do not know if this linear bound is optimal, but it is easy to prove a polynomial lower bound). Finally, let me stress that there is an *asymmetry* in the statement of Theorem 3.1: we assume that G is triangle-free and C_ℓ-free, but the graph Γ we obtain is only triangle-free. This asymmetry is essential: if we demand identical conditions for both G and Γ, then it is known that super-exponential bounds are again required.

I don't know of any other result in graph theory that looks like this! It is fairly common to find problems in which triangles behave in one way, and all other odd cycles behave in some other way. But I know of no other problem that features such a double phase transition, where C_5 and C_7 behave differently both from triangles and from all longer odd cycles. In the rest of the talk, my goal is to sketch the proofs of the three new bounds, and in particular to try to indicate what changes between length 7 and length 9.

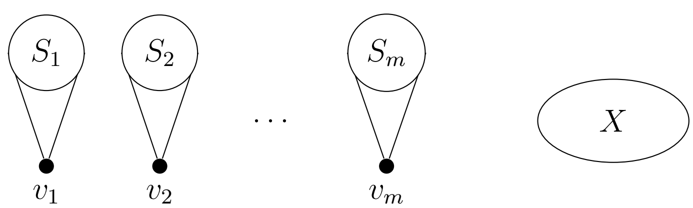
4 The upper bound for C_9

We begin by proving the upper bound for $\ell = 9$ in Theorem 3.1. Essentially the same proof works for all odd $\ell \geqslant 9$.

So let G be a triangle-free and C_9-free n-vertex graph, and fix some $\varepsilon > 0$. Our goal is to find a triangle-free graph Γ on $O(1/\varepsilon)$ vertices, as well as a $O(\varepsilon)$-approximate homomorphism $G \overset{O(\varepsilon)}{\to} \Gamma$. Equivalently, we want to delete at most $O(\varepsilon n^2)$ edges from G to end up with a graph that is homomorphic to Γ.

We can greedily find vertices $v_1, \ldots, v_m \in V(G)$, as well as disjoint sets $S_1, \ldots, S_m$, where $m = O(1/\varepsilon)$, such that:

- v_i is adjacent to all vertices in S_i, and
- if we let $X = V(G) \setminus (S_1 \cup \cdots \cup S_m \cup \{v_1, \ldots, v_m\})$, then the vertices in X are incident to at most εn^2 edges in total.



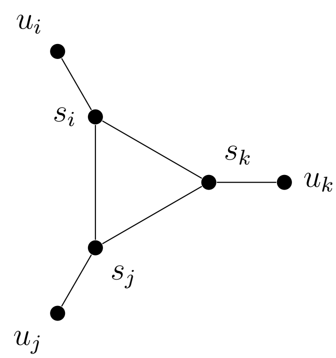
Indeed, we can find these by greedily picking out vertices v_i of highest degree, letting S_i be their neighbors that we have not yet taken, and repeating until we are left with few edges. Alternately, if we are OK with obtaining the weaker bound $m = O(\frac{1}{\varepsilon} \log \frac{1}{\varepsilon})$, then we can simply randomly sample $v_1, \ldots, v_m$.

Note that since G is triangle-free, each set S_i is an independent set. We now define a graph Γ on vertex set $\{u_1, \ldots, u_m, s_1, \ldots, s_m, x\}$ as follows: x is an isolated vertex, each u_i is adjacent to s_i and to no other vertex, and $s_i s_j \in E(\Gamma)$ if and only if there is at least one edge between S_i and S_j.

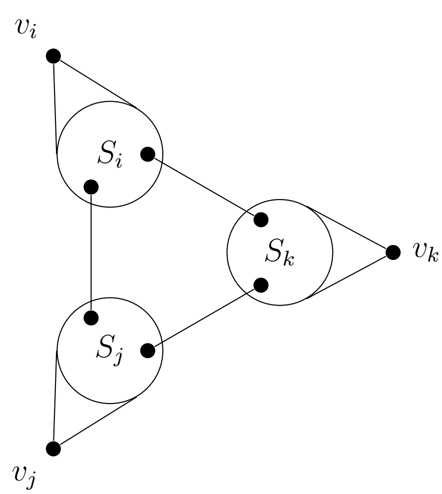
There is an obvious map $V(G) \to V(\Gamma)$, namely the one sending $v_i \mapsto u_i, S_i \mapsto s_i, X \mapsto x$. Which edges of G are mapped to non-edges of Γ? The only ones are those incident to X

and those between v_i and S_j for $j \neq i$, because S_i is an independent set, and every edge between S_i and S_j is mapped to an edge of Γ by the way we defined $E(\Gamma)$. By definition there are at most εn^2 edges incident to X, and the edges incident to some u_i contribute at most $O(mn) = O(n/\varepsilon) = O(\varepsilon n^2)$ edges, for n sufficiently large. Hence $G \xrightarrow{O(\varepsilon)} \Gamma$, and Γ has $2m + 1 = O(1/\varepsilon)$ vertices.

Therefore, to complete the proof, we just need to show that Γ is triangle-free. So suppose for contradiction that there is a triangle in Γ. By definition, it must go between three distinct vertices s_i, s_j, s_k. Thus, in Γ, the picture must look like this:



The fact that $s_i s_j \in E(\Gamma)$ implies that there is an edge in G between S_i and S_j, and similarly for the other two pairs. Therefore, in G, this corresponds to the following picture:



And this gives a C_9 in G! This is a contradiction, hence Γ is triangle-free, as claimed.

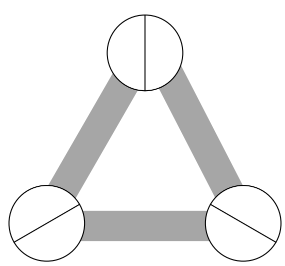
Note that this proof basically works because C_9 is the double subdivision of K_3, hence finding a triangle in Γ allows us to find a C_9 in G. Because of this, the same proof technique allows us to prove a version of the final bound in Theorem 3.1 whenever we assume that G is a graph containing no copies both of H and of the double subdivision of H; in that case, we can prove that $G \xrightarrow{\varepsilon} \Gamma$, where $|\Gamma| = O(1/\varepsilon)$ and Γ is H-free.

5 The lower bound for C_5

Let us now turn to the proof of the exponential lower bound in Theorem 3.1 in the case of $\ell = 5$. That is, we would like to exhibit a graph G which is both triangle-free and C_5-free, such that whenever $G \xrightarrow{\varepsilon} \Gamma$ for a triangle-free graph Γ, we have that $|\Gamma| \geqslant 2^{\mathrm{poly}(1/\varepsilon)}$. This proof is a variant of the proof given by Fox–Zhao to prove the lower bound on $M(\varepsilon)$ in Theorem 2.5.

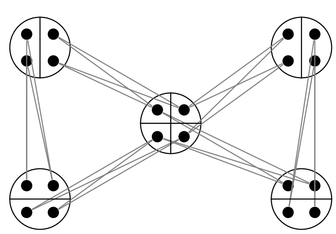
Our input is going to be a graph F on k vertices, for some $k = \mathrm{poly}(1/\varepsilon)$, with the following properties. F is a C_5-free graph, and moreover $E(F)$ is a union of edge-disjoint triangles, such that each vertex of F lies in $t = k^c$ such triangles, for some small $c > 0$. To construct such a graph, one can start with a high-girth linear 3-uniform hypergraph of polynomial degree (these can be easily constructed by a random deletion argument) and then replacing each of its hyperedges by a (graph) triangle.

To obtain G from F, we first blow up each vertex of F to an independent set of size 2^t. However, unlike in the usual blowup of F, we only keep some of the edges. Namely, for each triangle in F, we split each of the corresponding vertex blobs in half, and only place complete bipartite graphs between opposite halves. Thus, a triangle in F yields the following picture in G:



Thus, each triangle in F yields, in G, a disjoint union of three complete bipartite graphs, each with parts of size $2^t/2$.

Because each vertex of F lies in t triangles, we do this t times per blob in G. Since each blob has 2^t vertices, we can pick the partitions for each triangle to be "orthogonal". For example, if F is the bowtie graph consisting of two triangles sharing a vertex, then G will look as follows:



Now, the key properties of G are as follows. First, G is both triangle-free and C_5-free. The reason is that any potential triangle in G must arise from a triangle in F, and as we discussed, each triangle in F actually yields a disconnected bipartite graph in G. Similarly, G is C_5-free, because any C_5 in G must come from a closed walk of length 5 in F. As F is C_5-free, each such closed walk consists of a triangle plus an edge traversed twice, and it is fairly easy to convince yourself that this also cannot form a C_5 in F. Crucially, however, even if F is C_9-free, there *will be* copies of C_9 in G. The reason is essentially the same reason as the key step in the proof from Section 4, namely that a C_9 can be built by traversing a triangle in F, except using three extra vertices to "hop between" the two halves of the partition. In this way, we find that the current proof of the lower bound is in some sense "dual" to the upper bound proof from Section 4, and hence both proofs give the same "reason" for the extra phase transition in Theorem 3.1 between $\ell = 7$ and $\ell = 9$.

Continuing with the properties of G, we also have that

$$n := |G| = 2^t \cdot |F| = 2^{k^c} \cdot k = 2^{\mathrm{poly}(1/\varepsilon)},$$

as $k = \mathrm{poly}(1/\varepsilon)$ and $c > 0$ is a constant.

Finally, the crucial property of G is that it has no ε-approximate homomorphism to any triangle-free graph Γ on "substantially fewer" than n vertices. I won't prove this rigorously, but here is the basic idea. Fix some small graph Γ, and fix a map $V(G) \to V(\Gamma)$. As $|\Gamma| \ll |G|$, we must map many vertices of G to the same vertex of Γ. If we collapse many vertices from distinct blobs, then we must map many edges of G to non-edges of Γ, hence this map cannot be an ε-approximate homomorphism. Alternately, we may mostly collapse vertices within the same blob. But in this case, we must identify many vertices that come from opposite sides of the partitions used in constructing G, which in turn must create a triangle in Γ. Thus, if $G \xrightarrow{\varepsilon} \Gamma$ for a small Γ, then Γ cannot be triangle-free.

While this basic idea captures the main point of the proof, the details are substantially more involved. We use a variant of an elegant entropy argument developed by Fox and Zhao, which very conveniently handles many of the technicalities that arise in the proof.

We note that the same proof also works for the lower bound in Theorem 3.1 in the case $\ell = 7$. However, as discussed above, it cannot possibly work for $\ell = 9$, because the graph G we construct definitely has copies of C_9. This is, of course, not surprising, as we know that such a lower bound cannot hold when $\ell = 9$, by the result proved in Section 4.

6 The upper bound for C_5

Finally, we prove the upper bound for C_5 in Theorem 3.1. The same proof actually gives a similar upper bound for all odd $\ell \geqslant 5$, but of course, once $\ell \geqslant 9$ we have a much stronger upper bound from the argument in Section 4. We actually found two completely different proofs of this upper bound, one using "approximate" techniques like those we've seen so far, and one using "exact" techniques like those used to prove the results in Section 1.

The "approximate" technique is much more robust, and can be used to prove analogous bounds in much greater generality than the case of triangle-free and C_5-free graphs. Nevertheless, in order to see some ideas of a rather different flavor, let's begin with the other proof.

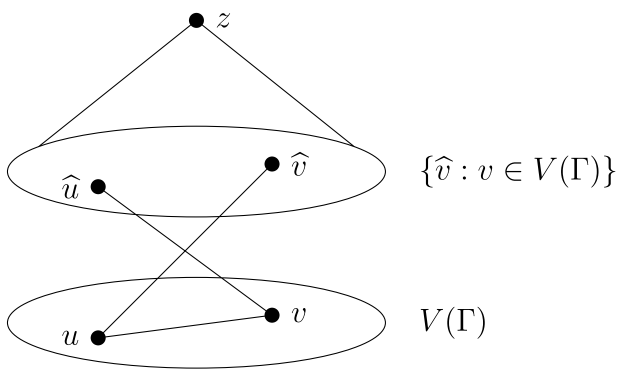
6.1 Proof via exact homomorphisms

The key input of the proof is the following simple lemma, which deals with exact, and not approximate, homomorphisms. It allows us to extend a homomorphism defined on a certain induced subgraph of G to a homomorphism defined on all of G, while preserving the triangle-freeness of the image and not enlarging it too much.

Lemma 6.1. *Let G be a graph with vertex set $A \sqcup B \sqcup C$. Suppose that A and B are independent sets, and there are no edges between A and C.*

Let Γ be a graph, and suppose that $G[B \cup C] \to \Gamma$. Then there exists a graph $\widehat{\Gamma}$ with $|\widehat{\Gamma}| = 2|\Gamma| + 1$, as well as a homomorphism $G \to \widehat{\Gamma}$. Moreover, if Γ is triangle-free, then so is $\widehat{\Gamma}$.

Proof. We let $\widehat{\Gamma}$ be the *Mycielskian* of Γ, defined as follows. For every vertex $v \in V(\Gamma)$, we make a copy $\widehat{v}$. Then $V(\widehat{\Gamma}) = V(\Gamma) \cup \{\widehat{v} : v \in V(\Gamma)\} \cup \{z\}$, where z is a special new vertex. The edges of $\widehat{\Gamma}$ are defined as follows: for all $uv \in E(\Gamma)$, we make $uv, \widehat{u}v, u\widehat{v}$ all edges of $\widehat{\Gamma}$ (but *not* $\widehat{u}\widehat{v}$). Finally, we make z adjacent to all vertices $\widehat{v}$.



It is easy to see that $\widehat{\Gamma}$ is triangle-free if Γ is. Indeed, $\widehat{\Gamma} \setminus \{z\}$ is homomorphic to Γ, so any triangle in $\widehat{\Gamma}$ must involve z, but its neighborhood is an independent set.

Now, let $\phi : G[B \cup C] \to \Gamma$ be a homomorphism. We define a map $\psi : V(G) \to V(\Gamma)$ as follows:

- For all $c \in C$, we set $\psi(c) = \phi(c)$.
- For all $b \in B$, we set $\psi(b) = \widehat{\phi(b)}$.
- For all $a \in A$, we set $\psi(a) = z$.

The fact that A and B are both independent sets, and that there are no edges from A to C, implies that the only edges in G are inside C, between A and B, or between B and C. From the construction, we immediately see that edges between A and B are mapped to edges of $\widehat{\Gamma}$, as are edges inside C (since ϕ is a homomorphism). Finally, edges between B and C are also mapped to edges, since the neighborhood of $\widehat{v}$ in $V(\Gamma)$ is the same as that of v. □

Given Lemma 6.1, it is not hard to conclude the proof. Let G be an n-vertex graph that is triangle-free and C_5-free. As in Section 4, we may find vertices $v_1, \ldots, v_m$ and sets $S_1, \ldots, S_m, X$ with $m = O(1/\varepsilon)$ such that each v_i is adjacent to all of S_i, and such that X is incident to at most εn^2 edges. We delete the edges incident to X to obtain a new graph G'. It suffices to prove that $G' \to \Gamma$, for some triangle-free graph Γ with $|\Gamma| \leqslant 2^{O(1/\varepsilon)}$.

Let $T_i = N(S_i)$ be the neighborhood of S_i. Note that both S_i and T_i are independent sets, since G is triangle-free and C_5-free. This means that we can iteratively apply Lemma 6.1 to build up the whole of G' from the empty graph. We start with $B = C = \varnothing$ and $A = S_1$. At the ith step, having constructed some graph G_{i-1}, we set $B_i = T_i \cap V(G_{i-1})$ and $C_i = V(G_{i-1}) \setminus T_i$, and define $A_i = S_i$. By continuing in this way, we can construct all of G' except potentially $\{v_1, \ldots, v_m\}$, which can be included by at most m additional steps of this process. We thus do at most $2m = O(1/\varepsilon)$ steps, and each step doubles the size of the target graph Γ. So at the end of the process, we construct a triangle-free graph Γ with $|\Gamma| \leqslant 2^{O(1/\varepsilon)}$, as well as a homomorphism $G' \to \Gamma$. Recalling that we deleted at most εn^2 edges to pass from G to G', we conclude that $G \xrightarrow{\varepsilon} \Gamma$.

6.2 Proof via weak regularity

As before, we fix an n-vertex graph G which is both triangle-free and C_5-free.

The key input in this proof is the weak regularity lemma of Frieze and Kannan. Here is a (slightly informal) statement.

Lemma 6.2 (Frieze–Kannan 1999). *Let G be an n-vertex graph, and let $\varepsilon > 0$. There exists a partition of $V(G)$ into $M \leqslant 2^{\mathrm{poly}(1/\varepsilon)}$ equal-sized parts, as well as a "random model" $\widetilde{G}$ on this partition, such that for all graphs H we have*

$$|\#(\text{copies of } H \text{ in } G) - \#(\text{copies of } H \text{ in } \widetilde{G})| \leqslant C_H \varepsilon n^{|H|},$$

where C_H is a constant depending only on H.

Here, by a "random model", I mean a random graph defined as follows: for every pair of parts (V_i, V_j) of the partition, there is some density parameter $p_{ij} \in [0, 1]$, and $\widetilde{G}$ is obtained by including each edge in $V_i \times V_j$ independently with probability p_{ij} (this is often called the *stochastic block model*). Equivalently, we can think of $\widetilde{G}$ as a blowup of a *weighted* graph $\widetilde{\Gamma}$ on M vertices.

We apply Lemma 6.2 to G, with some parameter $\varepsilon' = \mathrm{poly}(\varepsilon)$ to be chosen later, and obtain the random model $\widetilde{G}$ and the weighted graph $\widetilde{\Gamma}$. Let Γ_0 be the (unweighted) graph obtained from $\widetilde{\Gamma}$ obtained by keeping each edge of weight at least ε. Note that there is an obvious map $V(G) \to V(\widetilde{\Gamma}) = V(\Gamma_0)$, namely the one that sends each vertex to the name of the part it lies in in the partition given by Lemma 6.2. Note too that since we only remove from $\widetilde{\Gamma}$ those edges that have weight less than ε, this map is actually an ε-approximate homomorphism.

If Γ_0 is triangle-free then we are done, since $|\Gamma_0| \leqslant 2^{\mathrm{poly}(1/\varepsilon')} = 2^{\mathrm{poly}(1/\varepsilon)}$, but of course there is no reason to expect it to be triangle-free. Nonetheless, we can split into two cases:

1. Suppose first that Γ_0 is *ε-close to triangle-free*, meaning that we can delete at most εM^2 edges from Γ_0 to obtain a triangle-free graph. This means that we can write Γ_0 as the union of a triangle-free graph Γ on the same vertex set, plus εM^2 additional edges. In this case we have $G \xrightarrow{2\varepsilon} \Gamma$, so we are done.

2. If Γ_0 is ε-far from triangle-free, then we are in the setting of removal lemmas. For example, Theorem 2.1 implies in this case that Γ_0 has δM^3 triangles, for some $\delta = \delta(\varepsilon) > 0$. Of course we don't want to use Theorem 2.1, since the quantitative aspects of it are terrible. Instead, we can use an *asymmetric* removal lemma, which in many cases has substantially better bounds. Concretely, we can use the following result.

   Theorem 6.3 (Gishboliner–Shapira–W. 2025)**.** *If an M-vertex graph Γ_0 is ε-far from triangle-free, then Γ_0 has at least $\operatorname{poly}(\varepsilon)M^5$ copies of C_5.*

   Thus, in this case, we learn that Γ_0 has at least $\operatorname{poly}(\varepsilon)M^5$ copies of C_5. This in turn implies that $\widetilde{\Gamma}$ has at least $\operatorname{poly}(\varepsilon)M^5$ copies of C_5 (potentially with a different polynomial). By the way we form $\widetilde{G}$ from $\widetilde{\Gamma}$, this in turn implies that $\widetilde{G}$ has $\operatorname{poly}(\varepsilon)n^5$ copies of C_5. Finally, so long as we picked ε' to be a sufficiently small power of ε, this actually contradicts the conclusion of Lemma 6.2, as we assumed that H is C_5-free. This contradiction completes the proof.