Abstract

The square G^2 of a graph G is the graph on $V(G)$ with a pair of vertices uv an edge whenever u and v have distance 1 or 2 in G. Given graphs G and H, the Ramsey number $R(G, H)$ is the minimum N such that whenever the edges of the complete graph K_N are coloured with red and blue, there exists either a red copy of G or a blue copy of H.

We prove that for all sufficiently large n we have

$$R(P_{3n}^2, P_{3n}^2) = R(P_{3n+1}^2, P_{3n+1}^2) = R(C_{3n}^2, C_{3n}^2) = 9n - 3$$
$$\text{and } R(P_{3n+2}^2, P_{3n+2}^2) = 9n + 1.$$

We also show that for every $\gamma > 0$ and Δ there exists $\beta > 0$ such that the following holds: If G can be coloured with three colours such that all colour classes have size at most n, the maximum degree of G is at most Δ, and G has bandwidth at most βn, then $R(G, G) \leqslant (3 + \gamma)n$.

Mathematics Subject Classifications: 05D10

The Ramsey numbers of squares of paths and cycles

Peter Allen   Domenico Mergoni Cecchelli   Barnaby Roberts
Jozef Skokan

1 Introduction

In this paper we want to study Ramsey numbers of squares of graphs. For k a positive integer, the power G^k of a graph G is the graph on $V(G)$ with a pair of vertices uv an edge whenever u and v have distance at most k in G. The case with $k = 2$ is of interest here, and we define the *square* of G as G^2.

Given graphs G and H, the Ramsey number $R(G, H)$ is the minimum N such that whenever the edges of the complete graph on N vertices K_N are coloured with red and blue, there exists either a red copy of G or a blue copy of H.

The study of Ramsey numbers has a long history, and in general it is hard to find even good upper and lower bounds on $R(G, H)$. In this paper, we are interested in the case that G and H are sparse graphs. In this case, if G is connected and $v(G) \geqslant \sigma(H)$, one has the lower bound

$$R(G, H) \geqslant \big(\chi(H) - 1\big)\big(v(G) - 1\big) + \sigma(H). \tag{1}$$

Here $v(G)$ denotes the number of vertices of G, $\chi(H)$ is the chromatic number of H, and the chromatic surplus $\sigma(H)$ is the minimum, over all $\chi(H)$-colourings of H, of the smallest colour class size. This lower bound is due to Burr [5], with the corresponding construction being $\chi(H) - 1$ vertex-disjoint red cliques each on $v(G) - 1$ vertices, plus one further red clique on $\sigma(H) - 1$ vertices, and all other edges blue. When this construction gives the Ramsey number (i.e. when we have an equality in (1)), we say that G is H-*good*.

For fixed graphs H, the class of graphs G which are H-good is quite well understood; see Allen, Brightwell and Skokan [3] and Nikiforov and Rousseau [10]. However much less is known about the case when H grows with $v(G)$, or when $H = G$. Burr [5] conjectured that for fixed Δ, every connected graph G with $\Delta(G) \leqslant \Delta$ and $v(G)$ large enough is G-good. This statement holds for $G = P_n$ [7] and $G = C_n$ [4, 11]. However it was disproved by Graham, Rödl and Ruciński [8], who showed that it fails badly for expander graphs,

---

London School of Economics, London, WC2A 2AE, UK. (`p.d.allen@lse.ac.uk`, `d.mergoni@lse.ac.uk`, `j.skokan@lse.ac.uk`, `roberts.barnaby@gmail.com`).

and again in [3], where a lower bound on $R(P_n^k, P_n^k)$ better than (1) is shown for each $k \geqslant 2$. In the latter paper, however, it is shown that Burr's conjecture is off by at most a factor (roughly) 2 when G has bounded maximum degree and sublinear bandwidth. Here the bandwidth of G is the smallest k such that G is a subgraph of $P_{v(G)}^k$.

In [3], a value for the Ramsey numbers of squares of paths, and squares of cycles on a number of vertices divisible by 3, is conjectured. We observe that the conjectured value is wrong by one, and prove the modified conjecture.

Theorem 1. *There exists n_0 such that for all $n \geqslant n_0$ we have:*

$$R(P_{3n}^2, P_{3n}^2) = R(P_{3n+1}^2, P_{3n+1}^2) = R(C_{3n}^2, C_{3n}^2) = 9n - 3 \text{ and } R(P_{3n+2}^2, P_{3n+2}^2) = 9n + 1.$$

The lower bound part of this theorem is the following construction from [3]. We take disjoint vertex sets X_1, X_2, Y_1, Y_2 each with $2n - 1$ vertices, plus Z with $n - 1$ vertices. We colour edges within each X_i blue and within each Y_i red. We colour edges in the bipartite graphs (X_1, X_2) and (X_i, Z) red, and in (Y_1, Y_2) and (Y_i, Z) blue. We colour (X_1, Y_1) and (X_2, Y_2) blue, and (X_1, Y_2) and (X_2, Y_1) red. Finally, we add a single vertex z, which sends blue edges to $X_1 \cup X_2$ and red to $Y_1 \cup Y_2$. The edges within $Z \cup \{z\}$ may be coloured arbitrarily, as illustrated in Figure 1. A quick case analysis, part of our proof of Theorem 1, demonstrates that this construction does not contain a monochromatic P_{3n}^2. Furthermore, we can add one extra vertex to each of X_1, X_2, Y_1, Y_2 and still have no P_{3n+2}^2.
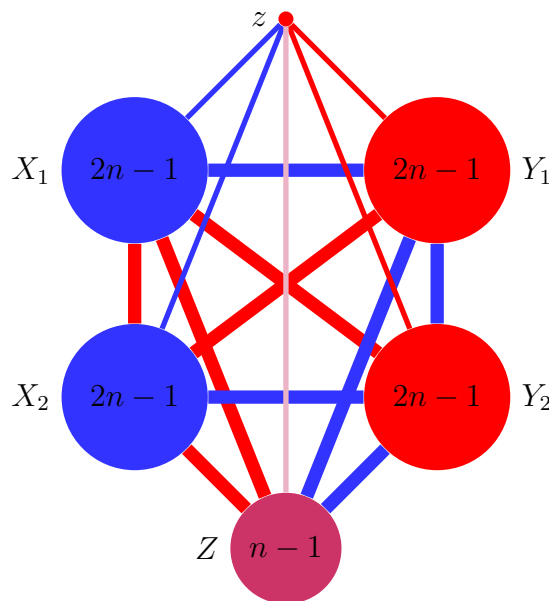

Figure 1: Lower bound construction

It is natural to ask for the Ramsey numbers $R(C_{3n+1}^2, C_{3n+1}^2)$ and $R(C_{3n+2}^2, C_{3n+2}^2)$. Observe that (for large n) both these graphs have chromatic number 4 and chromatic

surplus respectively 1 and 2, so Burr's construction shows $R(C_{3n+1}^2, C_{3n+1}^2) \geqslant 3(3n) + 1$ and $R(C_{3n+2}^2, C_{3n+2}^2) \geqslant 3(3n+1) + 2$ respectively. These numbers are matched by the construction of Figure 1 exactly (with X_1, X_2, Y_1, Y_2 having respectively $2n$ and $2n+1$ vertices); it seems reasonable to believe that this is the correct value for the Ramsey number, but we do not prove it; our approach requires 3-colourability of H.

In addition, we give a general upper bound on Ramsey numbers for 3-colourable graphs with bounded maximum degree and sublinear bandwidth, which P_{3n}^2 demonstrates is asymptotically tight.

Theorem 2. *Given $\gamma > 0$ and Δ, there exist $\beta > 0$ and n_0 such that for all $n \geqslant n_0$ the following holds. Suppose that H is a graph with $\Delta(H) \leqslant \Delta$, with bandwidth at most βn, and with a proper vertex 3-colouring all of whose colour classes have at most n vertices. Then $R(H, H) \leqslant (9 + \gamma)n$.*

We recall from [3] that the bandwidth restriction in this theorem is necessary. Graham, Rödl and Ruciński [8] constructed n-vertex graphs H such that $R(H, H) \geqslant 2^{c\Delta}n$ where $\Delta = \Delta(H)$; from this it can be shown that for each given $\beta > 0$, if Δ is large enough there are n-vertex graphs H with bandwidth at most βn and maximum degree at most Δ for which the theorem statement is false.

Our proof method uses the stability-extremal paradigm. Using the Szemerédi Regularity Lemma and the Blow-up Lemma, we will argue that to find a monochromatic square of a path (or cycle, or 3-colourable sparse graph as in Theorem 2) it is enough to find in the cluster graph (which has roughly $9t$ vertices) a monochromatic triangle factor which is 'triangle-connected' (which we will define later). This standard reduction leaves us looking, in a nearly complete edge-coloured graph, for a large monochromatic triangle-connected triangle factor (TCTF). The main technical work of the paper (Lemma 3) is then to prove that a 2-edge-coloured near complete graph on nearly $9t$ vertices will either contain a monochromatic TCTF on a little more than $3t$ vertices, or alternatively the graph must be close to the extremal example.

To prove the main lemma, we use a second partitioning method, as in [3]: by an iterative use of Ramsey's theorem, we partition most of the $9t$ vertices into a collection of bounded size (but quite large) monochromatic cliques. Obviously, it is easy to find a large red triangle factor in a collection of red cliques: in addition, we will see that two triangles in (or even using one edge of) the same red clique are 'triangle-connected' in red, and that if two red cliques are *not* red triangle-connected, then almost all the edges between them have to be blue. These observations were previously made in [3]. Where we improve compared to that paper is that we are able to deal with the interaction between cliques of different colours (whereas in [3] the minority colour cliques are thrown away).

2 Notation, main lemmas and organisation

Our graph notation is mainly standard. We will often write $|G|$ for the number of vertices in a graph G, and similarly $|M|$ for the number of vertices covered by a matching M (i.e.

twice the number of edges of M); we also write $G \setminus M$ for the graph $G[V(G) \setminus V(M)]$ and similarly for sets. We will often want to refer to edges (of a given colour) between two or three vertex sets. We write (A, B) or (A, B, C) for respectively $\{ab : a \in A, b \in B\}$ and $(A, B) \cup (A, C) \cup (B, C)$, the graph we refer to will always be clear from the context. We will work with 2-edge-coloured graphs, and refer to the two colours as 'red' and 'blue'.

Given a graph G, we say that edges uv and uw of G are *triangle-connected* if vw is an edge of G, we extend this to an equivalence relation on edges by transitive closure. We refer to the equivalence classes of this relation as *triangle components*. We will generally want to talk about monochromatic triangle connection. Thus, if the edges of G are 2-coloured, we say that two red edges are red triangle-connected if they are triangle-connected in the subgraph of G consisting only of red edges, we define red triangle component similarly. We also, slightly abusing notation, will say two red cliques (each with at least two vertices) are red triangle-connected if an edge (and so all edges) in one is red triangle-connected to an edge (so all edges) of the other. When the colour is clear from the context (as with *red* cliques) we will often just say that the two cliques are triangle-connected.

A *triangle factor* in a graph G is a collection of vertex-disjoint triangles of G. It is a *triangle-connected triangle factor* (TCTF) if all its edges lie in a single triangle component. Again, we will usually want to talk about monochromatic TCTFs in a 2-edge-coloured graph G, and as above a red TCTF means a TCTF in the subgraph of red edges of G.

At this point, we are in a position to give the case analysis proving the lower bound part of Theorem 1.

Proof of Theorem 1, lower bounds. We begin by describing the red triangle components of the lower bound construction for P_{3n}^2, P_{3n+1}^2 and C_{3n}^2. The edges in Y_1 and in $(Y_1, X_2 \cup \{z\})$, form a red triangle component. Similarly, the edges in Y_2 and $(Y_2, X_1 \cup \{z\})$ form a red triangle component. The edges (X_1, X_2, Z), together with all red edges in Z and all red edges from z to Z which lie in a red triangle, form a red triangle component. Finally, each red edge from z to Z which is not in a red triangle forms a triangle component. The blue components are analogous.

If the lower bound construction contains a red P_{3n}^2, then in particular it has a red triangle component which contains a red triangle factor with n triangles. Checking each entry in the list above, observe that removing Y_1 from the first leaves an independent set: $X_2 \cup \{z\}$ contains no red edges. But Y_1 contains only $2n - 1$ vertices, so there cannot be a $3n$-vertex triangle factor in this component. The symmetric argument deals with the symmetric second red triangle component. For the third case, removing Z leaves a bipartite graph: the only red edges are those in (X_1, X_2). But Z contains only $n - 1$ vertices, so this component too contains no $3n$-vertex red triangle factor. Finally, trivially the single-edge components contain no red triangle factor. The argument to exclude a blue P_{3n}^2 is symmetric.

For the modification for P_{3n+2}^2, adding one vertex to each of X_1, X_2, Y_1, Y_2, the description of triangle components above, and the explanation that the red triangle component containing (X_1, X_2, Z) does not contain P_{3n}^2 continues to work. Observe that P_{3n+2}^2 has independence number $n + 1$, so removing any $2n$ vertices leaves at least one edge. This

observation shows that the red component consisting of edges in Y_1 and $(Y_1, X_2 \cup \{z\})$ does not contain a red P^2_{3n+2}, and the other cases are symmetric. $\qquad\square$

The main work of this paper is to prove the following stability lemma, which states that a 2-edge-coloured nearly complete graph G on almost $9t$ vertices either contains a monochromatic TCTF on a little more than $3t$ vertices, or is close to the extremal example. To state it, we need one further definition.

Given an edge-coloured graph G, let $A \subseteq V(G)$ and v a vertex of G not in A. For $r \in \mathbb{R}$, we say that v is r-*blue to* A if va is a blue edge of G for all but at most r vertices $a \in A$. Similarly, given $A, B \subseteq V(G)$ disjoint, we say that (A, B) *is* r-*blue* if all but at most r vertices in A are r-blue to B and vice versa. We define similarly r-*red*.

We will generally use this notation with r much smaller than the sets A and B, so the reader can think of r-blue as meaning 'almost all blue'. Our main lemma is then the following.

Lemma 3. *There exists $\delta_0 > 0$ such that for every $0 < h, \lambda < \delta_0$ there exist $\varepsilon_0 > 0$ and $t_0 \in \mathbb{N}$ such that for every $t \geqslant t_0$ and $0 < \varepsilon < \varepsilon_0$ the following holds. Let G be a 2-edge-coloured graph on $(9 - \varepsilon)t$ vertices with minimum degree at least $(9 - 2\varepsilon)t$. Then either G contains a monochromatic TCTF on at least $3(1 + \varepsilon)t$ vertices or $V(G)$ can be partitioned in sets B_1, B_2, R_1, R_2, Z, T such that the following hold.*

(a) $(2 - h)t \leqslant |B_1|, |B_2|, |R_1|, |R_2| \leqslant (2 + h)t,$

(b) $(1 - h)t \leqslant |Z| \leqslant (1 + h)t,$

(c) all the edges in $G[B_1]$ and $G[B_2]$ are blue, and all the edges in $G[R_1]$ and $G[R_2]$ are red,

(d) all the edges between the pairs $(B_1, R_1), (B_2, R_2), (R_1, Z)$ and (R_2, Z) are blue, and those between the pairs $(B_1, R_2), (B_2, R_1), (B_1, Z)$ and (B_2, Z) are red,

(e) the pair (B_1, B_2) is λt-red, and the pair (R_1, R_2) is λt-blue, and

(f) $|T| \leqslant ht.$

We will prove this lemma in Sections 4–7.

By applying the Regularity Method in a standard way, we are able to upgrade Lemma 3 to the following superficially similar statement, in which we replace TCTF with the square of a path and cycle. We could generalise the following lemma to nearly-complete graphs easily (as in Lemma 3), but we do not need it for the proof.

Lemma 4. *For every $\alpha > 0$ there exists $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for every $n > n_0$ the following holds. Let $N \geqslant (9 - \delta)n$, and let G be a 2-edge-colouring of K_N. Then either G contains both a monochromatic copy of P^2_{3n+2} and of C^2_{3n}, or we can partition $V(G)$ into sets X_1, X_2, Y_1, Y_2, Z and R such that the following hold.*

(a) $(2 - \alpha)n \leqslant |X_1|, |X_2|, |Y_1|, |Y_2| \leqslant (2 + \alpha)n,$

(b) $(1 - \alpha)n \leqslant |Z| \leqslant (1 + \alpha)n$,

(c) $|R| \leqslant \alpha n$,

(d) Vertices in the following pairs have at most αn red neighbours in the opposite part: $(X_1, Y_1), (X_2, Y_2), (Y_1, Y_2), (Y_1, Z)$ and (Y_2, Z),

(e) Vertices in the following pairs have at most αn blue neighbours in the opposite part: $(X_1, X_2), (X_2, Y_1), (X_1, Y_2), (X_1, Z)$ and (X_2, Z),

(f) Vertices in X_1 and X_2 have at most αn red neighbours in their own part,

(g) Vertices in Y_1 and Y_2 have at most αn blue neighbours in their own part.

We deduce this lemma from Lemma 3 in Section 8.

To complete the proof of Theorem 1, we need to show that a complete graph which can be partitioned as in the above Lemma 4 and which has $9n - 3$ vertices necessarily contains both a monochromatic P_{3n+1}^2 and C_{3n}^2; and $9n + 1$ vertices suffices for P_{3n+2}^2. We do this in Section 9.

Finally, to prove Theorem 2 it suffices to observe that if G satisfies the conditions of Lemma 3 and can be partitioned as in that lemma, then it contains a monochromatic TCTF on nearly $3t$ vertices. Together with a standard application of the Regularity Method, which we sketch in Section 8, this completes the proof of Theorem 2.

3 Preliminary lemmas

In this section we prove some basic Ramsey-theoretic results which we will need to prove Lemma 3, but for which we do *not* assume the conditions of Lemma 3.

Lemma 5. *There exist $\varepsilon_0 > 0$ and $t \in \mathbb{R}$ such that the following holds for every $0 < \varepsilon < \varepsilon_0$ and $t > t_0$. Let G be a graph on at least $2(1 + 3\varepsilon)t$ vertices with minimum degree at least $|G| - \varepsilon t$. Any 2-edge-colouring of the edges of G contains a red matching on $2(1 + \varepsilon)t$ vertices or a blue connected matching on $\min \{|G| - (1 + 2\varepsilon)t, 2|G| - 4(1 + 2\varepsilon)t\}$ vertices.*

Proof. Let M be the largest red matching in G and let $Y = V(G) \setminus M$. We may assume that M has at most $2(1 + \varepsilon)t$ vertices. Since M is maximal, every edge in M has one endpoint with at most one red neighbour in Y. Indeed, if $xy \in M$ and both x and y have at least two neighbours in Y we can take x' in Y adjacent to x and y' distinct from x' adjacent to y in Y, and obtain a red matching which is larger than M by substituting xy with $x'x$ and $y'y$.

Let S be the set of vertices in M with at most one red neighbour in Y. We can now form a blue matching P (that we are going to show is connected) by greedily matching vertices in S with blue neighbours in Y. We claim that P has at least $\min \{|S|, |G| - |M| - 2\varepsilon t\}$ edges. Indeed, since the process is greedy we stop only by finishing all the vertices of S or when $S \setminus P$ is not empty, but no vertex in $S \setminus P$ has a blue neighbour in $Y \setminus P$, and this means that there are less than $2\varepsilon t$ vertices not yet covered by P in Y.

If we stopped for the first reason (if $|S| < |G| - |M| - 2\varepsilon t$) we can extend P to a larger blue matching P': the induced graph over Y contains only blue edges by maximality of M and there are some edges left in $Y \setminus P$. This extension of P can continue at least until all but εt vertices in Y are covered: we stop only when all edges in Y have one vertex covered by P'. Therefore we have

$$
|V(P')| \geqslant \overbrace{2\,|S|}^{\text{in } P} + \overbrace{|Y| - |S| - \varepsilon t}^{\text{in } Y}
$$
$$
\overset{2|S| \geqslant |M|}{\geqslant} |G| - \frac{|M|}{2} - \varepsilon t
$$
$$
\geqslant |G| - (1 + 2\varepsilon)t \,,
$$

as desired.

If on the other hand we stopped because no vertex in $S \setminus P$ has a blue neighbour in $Y \setminus P$ (but $S \setminus P$ is not empty). In particular, by definition of S this means that every vertex in $S \setminus P$ has at most one neighbour in $Y \setminus P$. This can only happen if $|Y \setminus P| < 2\varepsilon t$ and hence all but at most $2\varepsilon t$ vertices of Y are covered by P. This means that the size of P is at least

$$
|P| \geqslant 2(|Y| - 2\varepsilon t)
$$
$$
\geqslant 2(|G| - |M| - 2\varepsilon t)
$$
$$
\geqslant 2(|G| - 2(1 + \varepsilon)t - 2\varepsilon t)
$$
$$
= 2\,|G| - 4(1 + 2\varepsilon)t \,,
$$

as desired.

In order to conclude, we must now argue that the matching P (or P') we obtained is blue connected. But this is the case, indeed, every edge of P (or P') has at least one vertex in Y. Indeed $|Y| = |G| - |M| \geqslant 4\varepsilon t$ and all edges in Y are blue. By the minimum degree of G each vertex of Y is non-adjacent to at most εt vertices of Y, so any pair of vertices of Y has a common neighbour in Y, and therefore Y is blue-connected. $\square$

Lemma 6. *Let G be a graph with minimum degree strictly greater than $\frac{2}{3}\,|G|$. Then all the edges of G are triangle-connected. Moreover, there exists a TCTF on all but at most 2 vertices of G.*

Proof. We may notice that every three vertices of G share a common neighbour by the minimum degree condition and the pigeonhole principle. As any pair of adjacent edges spans three vertices, and these three vertices would have a neighbour in common outside of themselves by our previous claim, we get that any pair of adjacent edges is triangle-connected. This observation implies that connected components and triangle components coincide in G (because of the minimum degree condition we have that G is connected and therefore every pair of edges is triangle-connected). The existence of the TCTF is given by a theorem of Corradi and Hajnal [6]. $\square$

Lemma 7. *There exist $\varepsilon_0 > 0$ and $t \in \mathbb{R}$ such that the following holds for every $0 < \varepsilon < \varepsilon_0$, and every $t > 0$. Let G be a graph on at least $(5 + 100\varepsilon)t$ vertices with minimum degree at least $|G| - \varepsilon t$. Any 2-edge-colouring of the edges of G contains a red connected matching over $2(1 + \varepsilon)t$ vertices or a blue TCTF on $3(1 + \varepsilon)t$ vertices.*

Proof. Without loss of generality, we may assume G has $(5 + 100\varepsilon)t$ vertices. We separate cases.

<u>Case 1</u>: G has a maximal red connected component A that spans at least $(4 + 5\varepsilon)t$ vertices.

Let M be the largest red matching in A. Since A is a red connected component, we may assume $|M| < 2(1 + \varepsilon)t$ (recall that we use $|M|$ for the number of vertices in M). Since M is a maximal red matching in A, we know that every edge in $A \setminus M$ is blue.

Because of our assumption on the size of A, we have that $|A \setminus M| > (2 + 3\varepsilon)t$. We construct a matching P of size $2(1 + \varepsilon)t$ in $A \setminus M$ greedily, which is possible by the minimum degree of G. By Lemma 6, every pair of edges in $A \setminus M$ is blue triangle-connected. In particular P is blue triangle-connected.

We now greedily extend the edges of P to blue triangles by taking vertices in $X = V(G) \setminus (P \cup M)$. We have no red edges from vertices of X to vertices of P: if $x \in X$ is not in A, this is since A is a red component, while if $x \in X \cap A$ then it is by maximality of M. We have $|X| \geqslant (1 + 96\varepsilon)t$, and by the minimum degree of G any edge of P makes a triangle with all but at most $2\varepsilon t$ vertices of X, so the greedy extension succeeds.

<u>Case 2</u>: G has a maximal red connected component A that spans at least $3(1 + 2\varepsilon)t$ but less than $(4 + 5\varepsilon)t$ vertices.

If G has a red connected matching over $2(1 + \varepsilon)t$ vertices we are done, so we assume it does not. By Lemma 5 applied to A, we obtain a blue connected matching P in A of size at least $2(1 + \varepsilon)t$. Now as in the previous case, we can greedily extend all the edges of P to a blue triangle factor using vertices of $V(G) \setminus A$. Observe that every two blue adjacent edges in A share a neighbour in $V(G) \setminus A$, therefore every blue connected component in A is also blue triangle-connected. In particular, P, and hence the blue triangle factor containing it, are triangle-connected.

<u>Case 3</u>: G has two maximal red connected components A_1 and A_2 covering at least $(5 + 12\varepsilon)t$ vertices in total, and we are not in Cases 1 or 2.

Because we are not in Cases 1 or 2, A_1 and A_2 both span less than $3(1 + 2\varepsilon)t$ vertices and hence at least $(2 + 6\varepsilon)t$ vertices. In addition, neither component contains a red matching on $2(1 + \varepsilon)t$ vertices, because otherwise we would be done. Therefore, each A_i contains a blue connected matching P_i on precisely $\min\left(2|A_i| - 4(1 + 2\varepsilon)t, 2t\right)$ vertices by Lemma 5. Indeed, for the possible values of $|A_i|$, we have $2|A_i| - 4(1 + 2\varepsilon)t < |A_i| - (1 + 2\varepsilon)t$. Observe that every edge between A_1 and A_2 is blue and therefore $P_1 \cup P_2$ is a blue connected matching. We have $|P_1|, |P_2| \geqslant 4\varepsilon t$ and hence if $|P_1| = 2t$ we see that $P_1 \cup P_2$ has at least $(1 + 2\varepsilon)t$ edges. Similarly if $|P_2| = 2t$. If $|P_1|, |P_2| < 2t$ then we have at least $|A_1| + |A_2| - 4(1 + 2\varepsilon)t \geqslant (1 + 4\varepsilon)t$ edges, in any case we have in $P_1 \cup P_2$ at least $(1 + 2\varepsilon)t$ edges. Let $Y_i = A_i \setminus P_i$. We extend greedily the edges of P_1 to a set of disjoint blue triangles T_1 using vertices of Y_2, and in the same way we greedily extend the edges of P_2 to a set of

disjoint blue triangles T_2 using vertices of Y_1. Note that $|Y_i| = 4(1+2\varepsilon)t - |A_i| > (1+2\varepsilon)t$, and therefore we are able to extend the edges of $P_1 \cup P_2$, so we obtain a blue triangle factor with at least $(1 + \varepsilon)t$ triangles.

It now suffices to show that the triangle factor $T_1 \cup T_2$ is triangle-connected. Because every two blue incident edges in A_1 share a neighbour in A_2 and vice versa, we have that both T_1 and T_2 are TCTFs. Without loss of generality we assume that $|P_1| \leqslant |P_2|$. We know that $|P_1| = 2|A_1| - 4(1+2\varepsilon)t > 4\varepsilon$. Let xy be an edge in P_2, because every edge between A_1 and A_2 is blue, and because of the minimum degree condition we have that x and y share at least $|P_1| - 2\varepsilon t$ blue neighbours in P_1. Because P_1 has a blue matching, every set in P_1 of size strictly bigger than $\frac{|P_1|}{2}$ has an edge from P_1. Therefore we have that there exists zw in P_1 such that $G[\{x, y, z, w\}]$ is a blue clique with xy in P_2 and zw in P_1. Because both P_1 and P_2 are triangle-connected, we are done.

<u>Case 4</u>: G is not in any of cases 1–3, i.e. there is no red component of size $3(1 + 2\varepsilon)t$ or bigger, and no two red components cover $(5 + 12\varepsilon)t$ or more vertices.

Let $A_1, A_2, \ldots$ be the maximal red connected components, ordered by decreasing cardinality. We have $|A_1| < 3(1+2\varepsilon)t$ and $|A_1| + |A_2| < (5+12\varepsilon)t$, and we can assume that G does not have a red connected matching over $2(1 + \varepsilon)t$ vertices since otherwise we are done.

Claim 8. *The set of blue edges of G is triangle-connected.*

Proof. Every blue edge in a component A_i is in a blue triangle with some vertex in a different component A_j, so it suffices to prove that the edges between distinct components all lie in the same triangle component. In particular, it is enough to show that for any $j, k \geqslant 2$ distinct, any $a_1 a_j$ an edge between A_1 and A_j, and any $b_j b_k$ an edge between A_j and A_k, then $a_1 a_j$ and $b_j b_k$ are triangle-connected. This last equivalence is due to the fact that there are at least three red components (indeed, $|V(G)| - |A_1 \cup A_2| > 88\varepsilon t$).

Given a_1, a_j, b_j, b_k as above, let c be a common blue neighbour of a_1, a_j, b_j not in $A_1 \cup A_j$. This exists by minimum degree condition and by considering that a_1, a_j, b_j are all in $A_1 \cup A_j$ and there are at least $88\varepsilon t$ vertices not in $A_1 \cup A_j$. Now let us take d a common blue neighbour of c, a_j, b_j, b_k in A_1: this exists since c, a_j, b_j, b_k are not in A_1, and using the minimum degree condition. We can now conclude since $(a_1 a_j c, a_j c d, c d b_j, d b_j b_k)$ is a sequence of blue triangles that proves that $a_1 a_j$ and $b_j b_k$ are triangle-connected. □

Because we showed that every blue edge is triangle-connected, it is sufficient to find $(1 + \varepsilon)t$ disjoint blue triangles. We are in one of the following cases.

<u>Case A</u>: Both A_1 and A_2 are larger than $2(1 + 20\varepsilon)t$.

By Lemma 5 we can find blue matchings $M_i \subseteq A_i$ on $2|A_i| - 4(1+2\varepsilon)t$ vertices for $i = 1, 2$. Indeed, because $A_i < 3(1+2\varepsilon)t$ we have $2|A_i| - 4(1+2\varepsilon)t \leqslant |A_i| - (1+2\varepsilon)t$. We can greedily extend the matching M_1 to a blue triangle factor using vertices in $A_2 \setminus M_2$: because $|A_2 \setminus M_2| = (4 + 8\varepsilon)t - |A_2| > |A_1| - (2 + 4\varepsilon)t + 2\varepsilon t = \frac{|M_1|}{2} + 2\varepsilon t$ we are able to extend every edge in M_1 to a blue triangle. Similarly we can extend all the matching M_2 to a blue triangle factor using vertices in $A_1 \setminus M_1$. This two triangle factors are disjoint

and therefore they form a unique triangle factor that we denote with T. We can observe that $|T| = \frac{3}{2}|M_1| + |M_2| = 3(|A_1| + |A_2|) - 12(1 + 2\varepsilon)t$.

Let us now denote $U_1 = A_1 \setminus T$, $U_2 = A_2 \setminus T$ and $W = V(G) \setminus (A_1 \cup A_2)$. We have

$$
\begin{aligned}
|U_1| &= |A_1| - |M_1| - \frac{|M_2|}{2} \\
&= |A_1| - 2|A_1| + 4(1 + 2\varepsilon)t + 2(1 + 2\varepsilon)t - |A_2| \\
&= 6(1 + 2\varepsilon)t - (|A_1| + |A_2|) \geqslant t.
\end{aligned}
$$

Similarly we have $|U_2| \geqslant t$. We can also notice that $|W| = (5 + 100\varepsilon)t - (|A_1| + |A_2|) \geqslant 88\varepsilon t$. Finally, let us observe that $|U_1|, |U_2| > |W| + 4\varepsilon t$ by our assumption on $|G|$. Therefore we can find a blue triangle factor on (U_1, U_2, W) covering $3|W|$ vertices. Adding this triangle factor to T we get a TCTF on

$$
3(5 + 100\varepsilon)t - 3(|A_1| + |A_2|) + 3(|A_1| + |A_2|) - 12(1 + 2\varepsilon)t = (3 + 276\varepsilon)t
$$

vertices.

<u>Case B</u>: A_1 is larger than $2(1 + 3\varepsilon)t$ but all the other red components are smaller than $2(1 + 3\varepsilon)t$.

Let M_1 be a blue matching in A_1 on $2|A_1| - 4(1 + 2\varepsilon)t$ vertices. Let $U_1 = A_1 \setminus M_1$ and notice $|U_1| \geqslant 4(1 + 2\varepsilon)t - |A_1|$. Because all the other red components are smaller than $2(1 + 3\varepsilon)t$, we claim there exists j such that $(1 + 3\varepsilon)t < \left| \bigcup_{i=2}^{j} A_i \right| \leqslant 2(1 + 3\varepsilon)t$, and write $U_2 = \bigcup_{i=2}^{j} A_i$. Indeed, if $|A_2| > (1 + 3\varepsilon)t$ we can take $j = 2$, while if not then we can increase j sequentially until the lower bound is satisfied. Since in the latter situation we have $|A_j| \leqslant |A_2| \leqslant (1 + 3\varepsilon)t$ the upper bound is not exceeded. Finally, let $W = V(G) \setminus (A_1 \cup U_2)$ and note that $|W| \geqslant (3 + 94\varepsilon)t - |A_1|$.

Because of the size of U_2, we can extend edges of the blue matching M_1 to form a triangle factor T in $M_1 \cup U_2$ over $3|A_1| - 6(1 + 2\varepsilon)t$ vertices. We have that $|U_2 \setminus T| \geqslant 3(1 + 2\varepsilon)t - |A_1|$. Because $|W|, |U_1| > |U_2 \setminus T| + 4\varepsilon$, we can find a blue triangle factor on $(U_1, U_2 \setminus T, W)$ covering at least $3|U_2 \setminus T|$ vertices. Therefore combining this triangle factor with the one previously obtained over $M_1 \cup T$ we have a TCTF over at least $3|A_1| - 6(1 + 2\varepsilon)t + 3(3(1 + 2\varepsilon)t - |A_1|) = 3(1 + 2\varepsilon)t$ vertices.

<u>Case C</u>: Assume all connected components are smaller than $2(1 + 2\varepsilon)t$. This means that we can partition $V(G)$ in three sets U_1, U_2 and W such that $(1 + 3\varepsilon)t < |U_1|, |U_2| \leqslant 2(1 + 3\varepsilon)t$ by choosing unions of components as in the previous case to get U_1 and U_2, and let W be the union of the remaining components. Thus there are no red edges between any two of U_1, U_2 and W. Because $|W| = 5(1 + 100\varepsilon)t - |U_1| + |U_2|$ we have that all three sets U_1, U_2 and W have size at least $(1 + 3\varepsilon)t$ and that the largest of the three has at least $(1 + 6\varepsilon)t$ vertices. We can find a blue matching between the smallest two of U_1, U_2, W greedily of size $(1 + 2\varepsilon)t$, and extend this to a blue TCTF of size $3(1 + 2\varepsilon)t$ vertices greedily, using the largest component. $\qquad \square$

Lemma 9. *For $n \in \mathbb{N}$ sufficiently large, let G be a tripartite graph over $3n$ vertices with partition sets of the same size. Assume that every vertex has at least $\frac{3n}{4}$ neighbours in*

each of the two partition sets of which it is not part of. There exists a TCTF that covers every vertex of G.

Also, every pair of edges in G is triangle-connected.

Proof. Let $m = \frac{3n}{4}$ and X, Y and Z denote the sets which partition G. We first use Hall's theorem to prove that there exists a perfect matching M between X and Y. Indeed, let S be a subset of X. If $|S| \leqslant m$, because every vertex in S has at least m neighbours in Y we have that the neighbourhood of S in Y has size not smaller than the size of S itself. If $|S| > m$ observe that by inclusion-exclusion principle we have that every vertex in Y has a neighbour in S. We shall now define a bipartite support graph H over the sets M, Z. We add an edge between xy and z if the vertices xyz form a triangle in G. We can observe that the existence of a perfect matching in H gives us a triangle factor that covers all vertices of G. Let xy be in M, we can notice that since both x and y have at least m neighbours in Z we have that at least $\frac{n}{2}$ of the vertices of Z are neighbours of both x and y. Therefore every edge of M has minimum degree at least $\frac{n}{2}$ in H. Also, every vertex in Z has minimum degree at least $\frac{n}{2}$ in H, since in G it has minimum degree at least m in both X and Y. We can then repeat the above piece of proof and use Hall's theorem to prove that we can find a perfect matching in H and therefore a perfect triangle factor in G.

Let us now show that every pair of edges in G is triangle-connected. Let us first observe that if xy and xy' are both edges with $x \in X$ and $y, y' \in Y$ then we have that x, y, y' share a neighbour in Z and therefore they are triangle-connected. This implies that the set of edges between X and Y is in the same triangle component. We can easily conclude noticing that every triangle has one edge in each of the components $(X, Y), (Y, Z)$ and (Z, X) which are therefore all the same triangle component. $\qquad\square$

Corollary 10. *For $n \in \mathbb{N}$ sufficiently large let $k, r \in \mathbb{N}$ such that $6r + 4k < n$, let also G be a tripartite graph over $3n$ vertices with partition sets X, Y and Z of the same size n. Moreover, assume every vertex in G is adjacent to all but at most k of the vertices in each of the two partition sets it is not a part of. Let us fix a 2-edge-colouring of G such that (X, Y), (Y, Z) and (X, Z) are r-red. We can find a red TCTF formed by at least $n - 2r$ red triangles.*

Also, all but at most $3r^2$ red edges of G are in the same red triangle component.

Proof. Let $X' \subseteq X, Y' \subseteq Y$ and $Z' \subseteq Z$ of size exactly $n' = n - 2r$ such that every vertex in $X' \cup Y' \cup Z'$ has at most r blue vertices in each of the other two parts. We can apply Lemma 9 to $G' = G^{Red}[X' \cup Y' \cup Z']$ considering that each vertex in G' is adjacent to all but at most $r + k < \frac{3}{4}n'$ vertices in each of the two partitioning sets. $\qquad\square$

Lemma 11. *There exists $\varepsilon_0 \in \mathbb{R}$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ we have the following. Let G be a graph of minimum degree at least $|G| - \varepsilon t$ whose edges are 2-edge-coloured. If there exist in G two disjoint sets X and Y of size respectively $(1 + 5\varepsilon)t$ and $(5 + 200\varepsilon)t$ such that (X, Y) is εt-red, then G contains a monochromatic TCTF on at least $3(1 + \varepsilon)t$ vertices.*

Proof. Let Y' be the set of vertices in Y that have at least $|X| - \varepsilon t$ red neighbours in X. We have that $|Y'| \geqslant (5 + 100\varepsilon)t$ and also $G[Y']$ has minimum degree at least $|Y'| - \varepsilon t$. By Lemma 7 applied to Y', we find either a blue TCTF of size $3(1 + \varepsilon)t$ or a red connected matching on $2(1 + \varepsilon)t$ vertices. In the first case we are done, so we can assume we have a red connected matching on $2(1 + \varepsilon)t$ vertices, let us denote it by M. By Lemma 9 we can extend M to a triangle factor T of size at least $3(1 + \varepsilon)t$. We claim that this triangle factor is triangle-connected. Indeed, every adjacent pair of red edges in Y' is triangle-connected since any three vertices in Y' share a red neighbour in X. Since being triangle-connected is a transitive property and because M is red connected, we can conclude that T is triangle-connected. $\qquad\square$

4   General setting

To prove Lemma 3, we will use a decomposition of $V(G)$ into red and blue cliques, and some associated notation. In this section, we describe the decomposition, define the notation, and prove that the decomposition exists under the assumptions of Lemma 3.

Setting 1. *Given $\varepsilon, t > 0$, let $m = \frac{1}{4} |\log \varepsilon|$.*

Given a graph G with $(9 - \varepsilon)t$ vertices and minimum degree at least $(9 - 2\varepsilon)t$, suppose that $E(G)$ is 2-edge-coloured and that there is no monochromatic TCTF with at least $3(1 + \varepsilon)t$ vertices.

We fix a partition of $V(G)$ into a set V_{bin} of size at most $\varepsilon^{1/2}t + \frac{40t}{\sqrt{m}}$ and a collection of at most $\frac{9t}{m}$ monochromatic cliques each of size between 2 and m such that the following holds.

For each vertex u which is in a blue clique C of the partition, we assume that at most $\frac{20t}{\sqrt{m}}$ blue edges go from u to vertices in blue cliques of the partition which are not blue triangle-connected to C. We assume a similar statement replacing red with blue. Moreover, for every positive integer k, the number of cliques of size less than $(1 - \frac{1}{k})m$ is at most $\frac{400k}{|\log \varepsilon|^{3/2}}t$.

We write B_1 for a blue triangle component of blue cliques of the partition covering the largest number of vertices, B_2 for the next largest, and so on. We break ties arbitrarily, and define similarly R_1 for the largest red triangle component of red cliques of the partition and so on. We write $B_{\geqslant 3} := B_3 \cup B_4 \cup \dots$, and $R_{\geqslant 3} := R_3 \cup R_4 \cup \dots$.

It is important to note that while we will care about which vertices contain the triangles of a TCTF, we will not care which vertices are used for the triangle connections between these triangles: when we ask whether two (say red) edges are red triangle-connected, we will always mean red triangle-connected in the entire graph G. Thus 'there is a red TCTF in X of size $3s$' means that there is a set of s vertex-disjoint red triangles contained in the set X, which are all in the same red triangle component of G. In particular, the set B_1 is a collection of blue cliques which are blue triangle-connected in G, the connections might well use vertices outside B_1.

In the following sections, we will often state lemmas referring to a 'decomposition as in Setting 1'. When we do this, we intend to fix a specific decomposition which will remain

unchanged in the proof, and statements we make refer only to this decomposition. Thus 'there is no red TCTF of size $3s$ contained in the red cliques' should be understood as meaning that the union of the red cliques *of the fixed partition* do not contain such a TCTF. It might be that there is a different partition which does contain such a TCTF.

The idea of our proof of Lemma 3 is now roughly as follows. We suppose that G contains no large monochromatic TCTF. Our initial aim is then to show that each of B_1, B_2, R_1, R_2 has roughly $2t$ vertices, while $B_3 \cup R_3$ has roughly t vertices, these give us the five large sets of the partition of Lemma 3. We will see that once the size bounds are obtained, it is not too hard to show that the edge colours are as claimed. Our proof for the claimed size bounds will go over several steps of finding increasingly strong upper and lower bounds on these sizes.

We obtain Setting 1 by iterative application of Ramsey's theorem followed by removing a few vertices to V_{bin}. The following Lemma 13 states that this is always possible, provided ε is small enough and t large enough.

Claim 12. *For n sufficiently large, let G be a graph over $2n$ vertices, and let A, B be disjoint cliques of size n in G. If there are more than $2(n-1)$ edges between A and B, the graph is triangle-connected.*

Proof. Equivalently, we can show that if H is subgraph of $K_{n,n}$ without a path of length three, then H has at most $2(n-1)$ edges. Assume H is a subgraph of $K_{n,n}$ without paths of length three. In particular this means that every edge has one endpoint with degree exactly one. Therefore the number of edges in H is at most equal to the number of vertices in H with degree one. If we have less than $2n-2$ vertices of degree one we are done. If we have $2n$ vertices with degree exactly one we know that H is a perfect matching. It cannot be the case that $2n-1$ vertices have degree exactly one. Therefore we covered all cases and we can conclude that the number of edges in H is at most $2(n-1)$. □

Lemma 13. *There exists $\varepsilon_0 \in \mathbb{R}$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ the following holds. Given a graph G with at least $(9-\varepsilon)t$ vertices, with minimum degree at least $|G| - \varepsilon t$, whose edges are 2-edge-coloured, there exist sets $R_1, \ldots$ and $B_1, \ldots$ of monochromatic red and blue cliques respectively satisfying the properties of Setting 1.*

Proof. Let us start by proving that we can find disjoint monochromatic copies of K_m covering all but at most $\varepsilon^{\frac{1}{2}}t$ vertices of G.

First, notice that we do not want all cliques to be of the same colour, we just want monochromatic cliques. Let us start by selecting greedily as many monochromatic copies of K_m as possible in G, this means that we start by selecting an arbitrary monochromatic K_m, then we remove its vertices and we repeat the process over the remaining vertices of G.

Let us assume by contradiction that when this process stops more than $\varepsilon^{\frac{1}{2}}t$ vertices of G remain. Let W be a set of size $\varepsilon^{\frac{1}{2}}t$ not containing any monochromatic clique of size m. Because of the minimum degree condition over G, we have that each vertex

of $G[W]$ has degree at least $(\varepsilon^{\frac{1}{2}} - \varepsilon)t$ and therefore $G[W]$ contains at least $\varepsilon^{\frac{1}{2}}(\varepsilon^{\frac{1}{2}} - \varepsilon)t^2 = \left(1 - \frac{1}{\varepsilon^{-\frac{1}{2}}}\right)(\varepsilon^{\frac{1}{2}}t)^2$ edges. By Turán's theorem, we have that $G[W]$ contains a (not necessarily monochromatic) clique K of size $\varepsilon^{-\frac{1}{2}}$. By a Ramsey's upper bound on diagonal Ramsey numbers we have that $R(m, m) \leqslant 4^m$, this value is smaller than $\varepsilon^{-\frac{1}{2}}$ for ε small enough. Indeed, for $\varepsilon < 1$ we have $\varepsilon = e^{-4m}$ and hence we can rewrite the inequality as $R(m, m) \leqslant 4^m \leqslant e^{2m} = \varepsilon^{-\frac{1}{2}}$ which holds for m large enough. Therefore we can find a monochromatic clique K' of size m in W. This contradicts the stopping of our greedy algorithm.

We can now focus on the number of vertices in blue cliques that witness more than $\frac{20t}{\sqrt{m}}$ blue edges that have endpoints in distinct triangle components of blue K_m.

- There are at most $\frac{9t}{m}$ disjoint copies of K_m in G. This, combined with Claim 12 gives us that at most $\frac{(10t)^2}{m}$ blue edges have endpoints in distinct triangle components of blue K_m.

- At most $\frac{(20t)^2}{m}$ vertices in blue cliques of G witness a blue edge with its two extremities in two distinct triangle components of blue K_m. Therefore at most $\frac{20t}{\sqrt{m}}$ vertices in blue cliques witness more than $\frac{20t}{\sqrt{m}}$ such edges.

- We can do the same for red and obtain again at most $\frac{20t}{\sqrt{m}}$ vertices in red cliques that witness more than $\frac{20t}{\sqrt{m}}$ edges with their two extremities in two distinct triangle components of red cliques.

We denote with V_{bin} the set of vertices that were not in the original partition of cliques, together with the at most $\frac{40t}{\sqrt{m}}$ vertices that we selected in the previous point.

For each positive integer k, we want to count how many monochromatic cliques in $V(G) \setminus V_{\mathrm{bin}}$ can have less than $(1 - \frac{1}{k})m$ vertices. In other words, we want to bound the number of cliques of G with more than $\frac{m}{k}$ vertices in V_{bin}. It is not difficult to see that this number is less than $\frac{40t}{\sqrt{m}} \cdot \frac{k}{m} \leqslant \frac{400k}{|\log \varepsilon|^{\frac{3}{2}}}t$. Therefore at most $\frac{100k}{|\log \varepsilon|^{\frac{1}{2}}}t$ vertices are in cliques of size at most $(1 - \frac{1}{k})m$. $\qquad\square$

5 First upper bounds on the component size

In this section, we prove that $|B_i|, |R_i|$ cannot be much bigger than $\frac{7}{3}t$ (Lemma 14) and that we cannot have both B_1 and B_2 (or R_1 and R_2) much bigger than $2t$ (Lemma 15).

Lemma 14. *There exists $h_0 > 0$ such that for every $0 < h < h_0$ there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ we have the following. Let G be a 2-edge coloured graph with $(9 - \varepsilon)t$ vertices and minimum degree at least $(9 - 2\varepsilon)t$. Fix a collection of red and blue cliques as in Setting 1 with parameters ε and t. If G has a set of blue triangle-connected cliques covering more than $(\frac{7}{3} + h)t$ vertices, then G contains a monochromatic TCTF with $(1 + \varepsilon)t$ triangles. The same holds replacing blue with red.*

Proof. Let A be a triangle-connected set of blue cliques that covers more than $(\frac{7}{3} + h)t$ vertices. If $|A| \geqslant 3(1 + \frac{50}{|\log \varepsilon|})t$ then we greedily construct a blue TCTF within A that leaves out at most two vertices from each clique and obtain a blue TCTF covering at least $3(1 + \varepsilon)t$ vertices as desired, so we may now assume $|A| < 3(1 + \frac{50}{|\log \varepsilon|})t$.

Because of this bound on the size of A, and the condition of Setting 1 there are at most $\frac{40000}{|\log \varepsilon|^{\frac{3}{2}}}t$ cliques with less than $\frac{99}{100}m$ vertices, we have that there are at most

$$\frac{3(1 + 10\varepsilon)t}{\frac{99}{100}m} + \frac{40000}{|\log \varepsilon|^{\frac{3}{2}}}t \leqslant \frac{16t}{|\log \varepsilon|}$$

blue cliques in A and at least $|V(G)| - 3(1 + \frac{50}{|\log \varepsilon|})t$ vertices in $V(G) \setminus A$. In succession for each blue clique in A, we greedily construct a blue triangle factor T using one edge in the selected clique and one vertex outside of A. There are two possible cases.

<u>Case A</u>: The greedy construction provides us with a set T of $\frac{2}{3}(1 + \varepsilon)t$ triangles.

We can extend T to a triangle factor T' by adding triangles from within the cliques in A. When we stop, at most two vertices for each cliques are being unused and hence we obtained a blue TCTF covering at least

$$3 \cdot \frac{2}{3}(1 + \varepsilon)t + \left(\left(\tfrac{7}{3} + h \right) - 2 \cdot \frac{2}{3}(1 + \varepsilon) - 2 \cdot \frac{16}{|\log \varepsilon|} \right) t$$

vertices. Note that this means that T' covers at least $3(1 + \varepsilon)t$ vertices.

<u>Case B</u>: The greedy construction stops before we get $\frac{2}{3}(1 + \varepsilon)t$ triangles.

Let $Y = V(G) \setminus (A \cup T)$. We have that

$$|Y| \geqslant (9 - \varepsilon)t - 3(1 + \frac{50}{|\log \varepsilon|})t - \frac{2}{3}(1 + \varepsilon)t$$
$$\geqslant (5 + h)t \geqslant \left(5 + \frac{20000}{\sqrt{|\log \varepsilon|}} \right) t \,.$$

Let us denote by X the set of all the vertices in $A \setminus T$ which are in cliques that have at least three vertices in $A \setminus T$. At most $\frac{4}{3}(1 + \varepsilon)t + 2 \cdot \frac{16}{|\log \varepsilon|}t$ vertices are in A but not in X. Therefore we have that

$$|X| \geqslant \left(1 + \frac{100}{\sqrt{|\log \varepsilon|}} \right) t \,.$$

Because we stopped the greedy procedure, we cannot extend T using an edge in a clique of X and a vertex in Y, therefore each vertex in Y has at most one blue neighbour in each clique of X. This means that there are at most $\frac{16t}{|\log \varepsilon|} \cdot |Y| < \frac{16t}{|\log \varepsilon|} \cdot \left(9 - \frac{7}{3} \right) t \leqslant \frac{20^2 t^2}{|\log \varepsilon|}$ blue edges between X and Y. Hence we have that (X, Y) is $\frac{20}{\sqrt{|\log \varepsilon|}}t$-red. We can now apply Lemma 11 with input $\frac{20}{\sqrt{|\log \varepsilon|}}$. We conclude that G contains a monochromatic TCTF on at least

$$3 \left(1 + \frac{20}{\sqrt{|\log \varepsilon|}} \right) t > 3(1 + \varepsilon)t$$

vertices. $\qquad\square$

Lemma 15. *There exists $h_0 > 0$ such that for every $0 < h < h_0$ there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ we have the following. Let G be a 2-edge coloured graph with $(9 - \varepsilon)t$ vertices and minimum degree at least $(9 - 2\varepsilon)t$. Fix a collection of red and blue cliques as in Setting 1 with parameters ε and t. If G contains two disjoint sets of blue triangle-connected cliques, with each set of cliques covering more than $(2+h)t$ vertices, then G contains a monochromatic TCTF with $(1+\varepsilon)t$ triangles. The same holds replacing blue with red.*

Proof. Let A and B be disjoint sets of triangle-connected blue cliques, each covering at least $(2 + h)t$ vertices. We may suppose $h_0 \leqslant \frac{1}{30}$. Let C denote the collection of all the remaining vertices in blue cliques, if any exist. By Lemma 14, either we have the desired monochromatic TCTF or both A and B are smaller than $\left(\frac{7}{3} + h\right)t$. Therefore by Setting 1 with $k = 100$ they both contain at most the following number of blue cliques:

$$\frac{\frac{71}{30}t}{\frac{99}{100} \cdot \frac{1}{4} \left|\log \varepsilon\right|} + \frac{40000t}{\left|\log \varepsilon\right|^{\frac{3}{2}}} \leqslant \frac{10t}{\left|\log \varepsilon\right|}.$$

Moreover, by Claim 12 there are less than $2m$ blue edges between any blue clique in A and any clique in B. Therefore, between A and B there are less than $2m \cdot \frac{10t}{\left|\log \varepsilon\right|} \cdot \frac{10t}{\left|\log \varepsilon\right|} \leqslant \frac{50}{\left|\log \varepsilon\right|}t^2$ blue edges. Hence, (A, B) is $\frac{8}{\sqrt{\left|\log \varepsilon\right|}}t$-red. Let us set $\lambda = \frac{8}{\sqrt{\left|\log \varepsilon\right|}}$.

Let us greedily build a blue triangle factor T_A by extending blue edges in blue cliques of A to blue triangles using vertices outside of A. Let Y_A be the set of vertices in $V(G) \setminus A$ used in this way and A' the set of remaining vertices in A. We can independently do the same construction with B and obtain a triangle factor T_B and some similar sets Y_B and B'. Finally, let us denote $Z = V(G) \setminus (A \cup Y_A \cup B \cup Y_B)$.

Because we can extend T_A to a blue TCTF that covers all but at most two vertices for each clique of A (and similarly for B), we have that $\left|A \cup Y_A\right|, \left|B \cup Y_B\right| \leqslant (3 + 3\varepsilon + \frac{8}{m})t$. This implies

$$\left|Z\right| \geqslant \left|V\right| - (\left|A \cup Y_A\right| + \left|B \cup Y_B\right|) \geqslant (9 - h)t - 2(3 + h)t = 3(1 - h)t.$$

We also have that $\left|Y_A\right|, \left|Y_B\right| \leqslant (1 + \varepsilon)t$, which implies that $\left|A'\right|, \left|B'\right| \geqslant (h - 2\varepsilon)t$.

Each vertex of Z has at most one blue neighbour per clique in each of A' and B', since we cannot further extend T_A or T_B. Since $\left|Z\right| \leqslant 5t$ and there are at most $\frac{10t}{\left|\log \varepsilon\right|}$ cliques in each of A' and B' we have that both (A', Z) and (B', Z) have at most $\frac{50}{\left|\log \varepsilon\right|}t^2$ blue edges and hence they are both λt-red.

Claim 16. *We claim that all red edges in Z are triangle-connected. Moreover, if $\left|C \cap Z\right| \geqslant \frac{1}{3}t$ then we can find a TCTF in $(A, B, C \cap Z)$ on $\left|C \cap Z\right| - ht$ triangles that is triangle-connected to the red triangle component of Z.*

Proof. Let xy and uv be two red edges in Z, let N_A be the set of vertices in A' red adjacent to all vertices x, y, u and v and let N_B be defined similarly. To prove that xy and uv are

triangle-connected it suffices to show that there exists a red edge between N_A and N_B. Because of the lower bound on the size of A' and B', because of the minimum degree condition and because every vertex in Z is adjacent in red to all but at most $\frac{10t}{|\log \varepsilon|}$ of its neighbours in A' and B', we have that $|N_A|, |N_B| \geqslant (h - 2\varepsilon)t - 4 \cdot \varepsilon t - 4 \cdot \frac{10t}{|\log \varepsilon|} \geqslant \frac{3h}{4}t$. Since (A, B) is λt-red, there is a red edge between N_A and N_B. Therefore all the red edges in Z are in the same triangle component.

Let us now create a red TCTF Δ in $(A, B, C \cap Z)$ as follows. We first find a largest TCTF Δ' in $(A', B', C \cap Z)$. By Corollary 10, we have that Δ' has at least $\frac{h}{2}t$ vertices, since we have a lower bound on both $|A'|$ and $|B'|$.

We can now use Corollary 10 to find a red TCTF in $(A \setminus \Delta', B \setminus \Delta', (C \cap Z) \setminus \Delta')$ that covers almost all $(C \cap Z) \setminus \Delta'$. Let us call Δ the union of the two triangle factors. By Lemma 9 that Δ is triangle-connected.

It now suffices to show that Δ' is triangle-connected to the red triangle component of Z. Let xy be a red edge in Z, let N_A be the set of vertices in $A' \cap \Delta'$ red adjacent to both x and y, and let N_B be defined similarly in $B' \cap \Delta'$. To prove that xy and Δ' are triangle-connected it suffices to show that there exists an edge of Δ' between N_A and N_B. Because every vertex in Z is adjacent in red to all but at most $\frac{10t}{|\log \varepsilon|}$ of its neighbours in A' and B', we have that $|N_A|, |N_B| \geqslant \frac{99h}{100}t$. Since Δ' is a matching in (A', B') of large size, some of its edges are between N_A and N_B. $\qquad \square$

$Z \setminus C$ can be extended to a set of triangle-connected red cliques of G, possibly adding vertices from Y_A and Y_B. Therefore, we have $|Z \setminus C| \leqslant \left(\frac{7}{3} + h\right)t$ and this in particular implies that $|C \cap Z| \geqslant (\frac{2}{3} - 4h)t$. We form a red TCTF as follows. We start by using our last claim to construct a TCTF T_C over at least $|C \cap Z| - ht \geqslant (\frac{2}{3} - 5h)t$ triangles between A, B and $C \cap Z$ that is also triangle-connected to the red triangle component of Z. We then extend this TCTF by taking triangles in cliques of $Z \setminus C$. This is enough to conclude. $\qquad \square$

6 Colours and connection, and the sharp upper bound

In this section we begin by proving two lemmas which show that certain patterns of edges between triangle components imply triangle connections, which we need in both this section and the next. We then establish several inequalities about sizes of the components (Lemma 21), most of which imply that various components cannot be too small. In particular, we establish the useful inequality $|B_2| \geqslant |B_{\geqslant 3}|$, and similarly for red. Building on this, we finally prove the sharp upper bound we want: none of the components can contain much more than $2t$ vertices (Lemma 22). These are the two statements we need to complete the proof of Lemma 3 in the next section.

6.1 Colours and connection

Claim 17. *For every $h > 0$ there exists $\varepsilon > 0$ such that if we use that ε for Setting 1 we have the following. Let A, B be two disjoint sets of vertices in blue cliques such that there*

are no blue triangle components with some vertices in A and some vertices in B. Then the pair (A, B) is ht-red. The same works for red.

Proof. By Setting 1, in G there are at most $\frac{9t}{m} \leqslant \frac{40t}{|\log \varepsilon|}$ cliques. Therefore, by Claim 12 we can have at most $2m \cdot \frac{20t}{|\log \varepsilon|} \cdot \frac{20t}{|\log \varepsilon|} \leqslant \frac{200t^2}{|\log \varepsilon|}$ blue edges between A and B. In particular this means that the pair (A, B) is $\sqrt{\frac{200}{|\log \varepsilon|}}t$-red. For ε small enough we have the result we wanted. $\square$

Lemma 18. *There exists $h_0 > 0$ such that for every $0 < h < h_0$ there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ we have the following. Let G be a 2-edge coloured graph with $(9 - \varepsilon)t$ vertices and minimum degree at least $(9 - 2\varepsilon)t$. Fix a collection of red and blue cliques as in Setting 1 with parameters ε and t. Let Y_1, Y_2, Y_3 be subsets of the red cliques of size at least $10ht$ in distinct red triangle components, and let X be a set of size at least ht of vertices in blue cliques which all have more than $2ht$ blue neighbours in each of two of the Y_is. Then at least one of the blue edges in a clique of X is triangle-connected to the large blue TCTF in (Y_1, Y_2, Y_3). Everything still works if we invert red and blue.*

Proof. First, note that for ε small enough and by Claim 17 we have that each pair in Y_1, Y_2, Y_3 is $\frac{h^3}{2}t$-blue. Let R_i be the set of vertices in Y_i with more than h^3t-red edges in one of the other Y_j. Without loss of generality let us assume that the set S of vertices in X with more than $2ht$ blue neighbours in both Y_1 and Y_2 has size at least $\frac{ht}{3}$. Then each vertex in S has at least $(2h - h^3)t$ blue neighbours in both $Y_1 \setminus R_1$ and $Y_2 \setminus R_2$. Then we have a vertex y_1 in $Y_1 \setminus R_1$ which is incident in blue to at least $(2h - h^3)t \cdot \frac{ht}{3} \cdot \frac{1}{9t} \geqslant \frac{1}{15}h^2t$ vertices in S. So for t large enough y_1 is incident in blue to at least two vertices of S that lie in the same clique, let us call two such vertices x_1 and x_2. Since y_1 has at least $|Y_2| - (\varepsilon + h + h^3)t \geqslant |Y_2| - (2h + h^3)t$ blue neighbours in $Y_2 \setminus R_2$, we have that y_1 and x_1 have a common blue neighbour y_2. This implies that $x_1 x_2$ is blue-triangle-connected to $y_1 y_2$ and this by minimum degree condition means that $x_1 x_2$ is triangle-connected to the large blue TCTF over (Y_1, Y_2, Y_3) given by Lemma 9. $\square$

Lemma 19. *There exists $h_0 > 0$ such that for every $0 < h < h_0$ there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ we have the following. Let G be a 2-edge coloured graph with $(9 - \varepsilon)t$ vertices and minimum degree at least $(9 - 2\varepsilon)t$. Fix a collection of red and blue cliques as in Setting 1 with parameters ε and t, and let Y_1, Y_2 be subsets of size at least $10ht$ of vertices in red cliques in distinct red triangle components, and let X_1, X_2 be subsets of size at least $10ht$ of vertices in blue cliques in distinct blue triangle components. Finally, assume that X_1 is ht-red to each of Y_1 and Y_2. Then at most $2ht$ vertices in X_2 have more than $2ht$ red neighbours in both Y_1 and Y_2. Everything still works if we invert red and blue.*

Proof. First, note that for ε small enough and by Claim 17 we have that (X_1, X_2) is ht-red. Let S be the set of vertices in X_2 which have more than $2ht$ red neighbours in both Y_1 and Y_2. Assume by contradiction $|S| \geqslant 2ht$. Note that there is a vertex x_1 in X_1 which

has at most ht blue neighbours in each of X_2, Y_1 and Y_2, so x_1 is red-adjacent to some vertex $x_2 \in S$. Now x_1 and x_2 have at least $\frac{h}{4}t$ common red neighbours in each Y_i and therefore they have at least two common red neighbours from the same clique in each of the Y_i. But this is absurd because it would mean that a clique in Y_1 is triangle-connected to a clique in Y_2. $\qquad\square$

6.2 Some lower bounds

Claim 20. *Let k be a positive integer and let $b_1 \geqslant \ldots \geqslant b_k > 0$ be positive reals such that $\sum_{i>1} b_i > b_1$. Then we can partition $\{1, \ldots, k\}$ into two sets A, B such that if $\alpha := \sum_{i \in A} b_i$ and $\beta := \sum_{i \in B} b_i$ we have $2\alpha \geqslant \beta \geqslant \alpha$.*

Proof. We can construct such a partition greedily in two steps.

If $b_1 + b_3 \leqslant 2(b_2 + b_4)$ we set $1, 3 \in B$ and $2, 4 \in A$. Otherwise we set $b_1 \in B$ and $2, 3, \ldots, \ell \in A$ with an ℓ such that $b_1 > \sum_{i=2}^{\ell} b_i > \frac{b_1}{2}$ (such an ℓ exists because of the hypotheses and because $b_1 > b_2 + b_3$).

We now proceed by induction. Assume we already partitioned $1, \ldots, i-1$ such that the requests of the lemma are satisfied and let α and β be as in the statement of the lemma. If $2\alpha \geqslant \beta + b_i$ we can add $i \in B$. Otherwise, we have $\beta > 2\alpha - b_i \geqslant \alpha + b_i$, where the last inequality is given by the fact that the b_i are ordered in decreasing order and $|A| \geqslant 2$. In this second case we can add i to the set A. $\qquad\square$

Lemma 21. *There exists $h_0 > 0$ such that for every $0 < h < h_0$ there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ we have the following. Let G be a 2-edge coloured graph with $(9 - \varepsilon)t$ vertices and minimum degree at least $(9 - 2\varepsilon)t$. Fix a collection of red and blue cliques as in Setting 1 with parameters ε and t, and define $B_1, B_2, \ldots$ and $R_1, R_2, \ldots$ as in Setting 1.*

(i) If $|B_1| \leqslant \frac{7}{6}t$ then $\left|\bigcup_i B_i\right| \leqslant \left(\frac{7}{2} + h\right)t$.

(ii) If $|B_1| \geqslant \frac{7}{6}t$ and $|B_2| \leqslant \frac{7}{6}t$ then $\left|\bigcup_{i \neq 1} B_i\right| \leqslant \left(\frac{7}{3} + h\right)t$.

(iii) If $|B_1|, |B_2| \geqslant \frac{7}{6}t$ then $|\cup_i B_i| \leqslant (\frac{16}{3} + h)t$.

(iv) We have $\frac{43}{12}t \leqslant |\cup_i B_i|, |\cup_i R_i| \leqslant (\frac{16}{3} + h)t$. We also have $|B_1| > \frac{7}{6}t$.

(v) If $|B_2| < |\cup_{i \geqslant 3} B_i|$ then we can find a red TCTF in $\cup_i B_i$ of size at least $\frac{3}{2}|\cup_{i \geqslant 2} B_i| - ht$.

(vi) If $|B_2| \leqslant \frac{8}{7}t$ then $|\cup_i B_i| < (\frac{9}{2} - h)t$.

(vii) We have $|B_2| \geqslant |\cup_{i \geqslant 3} B_i|$.

The corresponding results also hold for red and $R_1, R_2, R_3, \ldots$. Moreover, by (vi) we have that at most one of B_2 or R_2 can be smaller than $\frac{8}{7}t$.

Proof. We are going to prove these results in order, and we are sometimes going to use previous points already proved.

<u>Proof of 21(i)</u>: Suppose for a contradiction that $|B_1| \leqslant \frac{7}{6}t$ and $|\bigcup_i B_i| > \left(\frac{7}{2} + h\right)t$. Observe that by Setting 1 with $k = 100$, all but at most $\frac{40000}{|\log \varepsilon|^{\frac{1}{2}}}t$ vertices of G are in cliques fixed in Setting 1 with at least $\frac{99}{100}m$ vertices. We let for each i the set B_i' consist of all vertices in blue cliques of B_i with at least $\frac{99}{100}m$ vertices.

We want to study how many edges have endpoints in two distinct B_i'. For each fixed i, the maximum number of blue edges that have one endpoint in B_i' and the other in some B_j' with $j \neq i$, is less than

$$2m \cdot \frac{|B_i'|}{\frac{99}{100}m} \cdot \frac{\left|\bigcup_{j \neq i} B_j'\right|}{\frac{99}{100}m} \leqslant 3\frac{|B_i'| \cdot \left|\bigcup_{j \neq i} B_j'\right|}{m} \leqslant \frac{27}{m}t\, |B_i'|\ .$$

Let us now observe that the number of vertices in B_i' that have more than $\frac{h}{100}t$ blue neighbours outside of B_i' is at most $\frac{27}{m}t\, |B_i'| \cdot \frac{100}{ht} \leqslant \frac{10^4}{mh}|B_i'|$.

Let us remove from each B_i' all the vertices with more than $\frac{h}{100}t$ blue neighbours in $\bigcup_{j \neq i} B_j'$, let us call the result B_i''. By the last observation, we have that

$$\left|\bigcup_i B_i''\right| \geqslant \left(1 - \frac{10^4}{mh}\right)\left|\bigcup_i B_i'\right|$$
$$\geqslant \left(1 - \frac{10^4}{mh}\right)\left(\left|\bigcup_i B_i\right| - \frac{40000}{|\log \varepsilon|^{\frac{1}{2}}}t\right)$$
$$\geqslant \left(1 - \frac{10^4}{mh}\right) \cdot \left(\frac{7}{2} + \frac{3h}{4}\right)t$$
$$\geqslant \left(\frac{7}{2} + \frac{3h}{4} - \frac{4 \cdot 10^4}{mh} - \frac{10^4}{m}\right)t$$
$$\geqslant \left(\frac{7}{2} + \frac{h}{2}\right)t\ .$$

In $G^{Red}\left[\bigcup_i B_i''\right]$ every vertex has red degree at least $\left|\bigcup_i B_i'\right| - (\frac{7}{6} + \varepsilon + \frac{h}{100})t$ which is more than $\frac{2}{3}\left|\bigcup_i B_i''\right|$. So by Lemma 6, $G^{Red}[\cup_i B_i'']$ contains a red TCTF of size $\frac{7}{2}t$.

<u>Proof of 21(ii)</u>: Let B_1^* be a set of the fixed blue cliques in B_1 covering between $\frac{7}{6}t - m$ and $\frac{7}{6}t$ vertices. We may assume $|B_2| \leqslant |B_1^*|$, by swapping these two sets of cliques if necessary.

Repeating what we did in Lemma 21(i) to the sets $B_1^*, B_2, B_3, \ldots$, we obtain $\left|B_1^* \cup \bigcup_{i \geqslant 2} B_i\right| \leqslant \left(\frac{7}{2} + h\right)t$. Since $|B_1^*| \leqslant \frac{7}{6}t$, we have $\left|\bigcup_{i \geqslant 2} B_i\right| \leqslant \left(\frac{7}{3} + h\right)t$ as desired.

<u>Proof of 21(iii)</u>: By Corollary 10 we have that $|\cup_{i \geqslant 3} B_i| \leqslant (1 + \frac{h}{3})t$ because otherwise we can find a red TCTF over more than $3(1 + \varepsilon)t$ vertices. By Lemmas 14 and 15 we have that $|B_1| \leqslant (\frac{7+h}{3})t$ and $|B_2| \leqslant (\frac{6+h}{3})t$. Summing these bounds completes the proof.

<u>Proof of 21(iv)</u>: By Lemmas 21(i), (ii), (iii) we have that for any possible size of B_1 and R_1 we always have $|\cup_i B_i|, |\cup_i R_i| \leqslant (\frac{16}{3} + h)t$. Because $|\cup_i B_i| + |\cup_i R_i| \geqslant (9 - h)t$ we therefore must have $\frac{43}{12}t \leqslant |\cup_i B_i|, |\cup_i R_i|$. By Lemma 21(i) this implies that $|B_1|, |R_1| > \frac{7}{6}t$.

<u>Proof of 21(v)</u>: Let us take a set of vertices $B_1' \subseteq B_1$ such that $|B_1'| = \frac{1}{2}|\cup_{i \geqslant 2} B_i| - \frac{1}{100}ht$ (we know that B_1 is large enough, indeed we know $|B_1| \geqslant \frac{7}{6}t$ and it cannot be the case that $|B_2| \geqslant \frac{7}{6}t$ because otherwise we would find a large red TCTF over $(B_1, B_2, \cup_{i \geqslant 3} B_i)$). By Claim 17 all but at most $\frac{1}{100}ht$ vertices of B_1' have red degree in $G[B_1' \cup \bigcup_{i \geqslant 2} B_i]$ at least $|\cup_{i \geqslant 2} B_i| - \frac{1}{100}ht$. Let B_1'' be a subset of size $\frac{1}{2}|\cup_{i \geqslant 2} B_i| - \frac{2}{100}ht$ such that every vertex in B_1'' has red degree in $G[B_1'' \cup \bigcup_{i \geqslant 2} B_i]$ at least $|\cup_{i \geqslant 2} B_i| - \frac{1}{100}ht \geqslant \frac{2}{3}|B_1'' \cup \bigcup_{i \geqslant 2} B_i|$. Because every vertex in $\cup_{i \geqslant 2} B_i$ is in a triangle component of size significantly smaller than $\frac{2}{3}|B_1'' \cup \bigcup_{i \geqslant 2} B_i|$ we can conclude by Lemma 6 that we can find a red TCTF over all but at most two vertices of $B_1'' \cup \bigcup_{i \geqslant 2} B_i$. Which is, we can find a red TCTF over at least $\frac{3}{2}|\cup_{i \geqslant 2} B_i| - ht$ vertices.

<u>Proof of 21(vi)</u>: Fix some $h > 0$ arbitrarily small, depending on which we can choose our ε. By Lemma 21(i) we can assume $|B_1| \geqslant \frac{7}{6}t$. Remember also that we have by Lemma 14, $(\frac{7}{3} + h)t \geqslant |B_1|$. Assume by contradiction $|B_2| \leqslant \frac{8}{7}t$ and $|\cup_i B_i| \geqslant (\frac{9}{2} - h)t$. Then we would have $|\cup_{i \geqslant 3} B_i| \geqslant (\frac{9}{2} - h)t - (\frac{7}{3} + h)t - \frac{8}{7}t = (\frac{43}{42} - 2h)t$. By Corollary 10 and Claim 17 it cannot be the case that $|B_2| \geqslant (\frac{43}{42} - 2h)t$ because otherwise we would find a large red TCTF over $(B_1, B_2, \cup_{i \geqslant 3} B_i)$. Therefore we must have $|B_2| < |B_3|$, and therefore by Lemma 21(v) we must have that $\frac{3}{2}|\cup_{i \geqslant 2} B_i| - ht < (3 + h)t$ which is to say that $|\cup_{i \geqslant 2} B_i| < \frac{25}{12}t$. We can conclude that $|\cup_i B_i| < (\frac{7}{3} + h)t + \frac{25}{12}t < (\frac{9}{2} - h)t$.

<u>Proof of 21(vii)</u>: First, let us note that we cannot have both $|B_2| < |\cup_{i \geqslant 3} B_i|$ and $|R_2| < |\cup_{i \geqslant 3} R_i|$. Indeed, by 21(vi) at least one between B_2 and R_2 has cardinality at least $\frac{8}{7}t$. Let us say without loss of generality that $|R_2| \geqslant \frac{8}{7}t$, then it cannot be $|\cup_{i \geqslant 3} R_i| > |R_2|$ because of Corollary 10.

Let us now assume by contradiction that $|\cup_{i \geqslant 3} B_i| > |B_2|$. By Lemmas 14 and 15 we have that $|R_1| \leqslant (\frac{7}{3} + h)t$ and $|R_2| \leqslant (2 + h)t$. Moreover, by Corollary 10 we have $|R_3| \leqslant (1 + h)t$. Therefore we have $|B_2 \bigcup \cup_{i \geqslant 3} B_i| \geqslant (\frac{4}{3} - 5h)t$.

By Claim 20, since both B_3 and B_4 are non-trivial (by our contradiction hypothesis), we can partition the sets B_i into collections B_1', B_2' and B_3' such that $B_1' = B_1$ and $|B_2'| \geqslant |B_3'|$ and also $|B_2'| \leqslant \frac{2}{3}|\cup_{i \geqslant 2} B_i|$. In particular this means $2|B_3'| \geqslant |B_2'| \geqslant |B_3'|$ and $|B_2'| \geqslant (\frac{2}{3} - 5h)t$ and $|B_3'| \geqslant (\frac{4}{9} - 5h)t$.

Notice that by Lemma 21(v) we have $|B_2 \bigcup \cup_{i \geqslant 3} B_i| \leqslant (2 - 2h)t$. We claim that no blue clique in B_1' is triangle-connected to the blue TCTF in (R_1, R_2, R_3). Indeed we have that this would create a blue TCTF of size at least $3|R_3| + |B_1|$ and we have $|R_3| \geqslant 9t - |B_1| - |B_2 \bigcup \cup_{i \geqslant 3} B_i| - |R_1| - |R_2| \geqslant (\frac{1}{3} - 5h)t$ and $|R_3| + |B_1| \geqslant (\frac{8}{3} - 4h)t$. Which implies that $3|R_3| + |B_1| > (3 + h)t$.

In particular, by Lemma 18 this implies that all but at most ht vertices in B_1' have less than $2ht$ blue neighbours in two of the R_1, R_2 or $R_{\geqslant 3}$. This means that there is a set $T \subset B_1'$ of size at least $\frac{1}{3}(|B_1'| - ht)$ such that $(T, R_i), (T, R_j)$ are ht-red and $i, j \in \{1, 2, \geqslant 3\}$. Let us assume that (T, R_2) is ht-red (if not, then we have (T, R_1) is ht-red and this is strictly better in the following computations). We claim that (R_2, B_2') and (R_2, B_3') are ht-blue. Indeed by Lemma 21(v) and by the lower bound $|B_2 \bigcup \cup_{i \geqslant 3} B_i| \geqslant (\frac{4}{3} - 5h)t$ we got earlier, we have a red TCTF in $B_1' \cup B_2' \cup B_3'$ of size at least $(2 - 10h)t$, since $|R_2| \geqslant \frac{8}{7}t$ we must have that each clique in R_2 is not triangle-connected to the large TCTF between

the B_i components. By Lemma 18 and since (T, R_2) is red, we get that (R_2, B_2') and (R_2, B_3') are ht-blue.

We now claim that there is a B_i in B_2' such that $(B_i, R_{\geqslant 3})$ is ht-red, in particular, this means that each red clique in $R_{\geqslant 3}$ is in the same triangle component of (B_1', B_2', B_3'). There exists such a B_i because B_2' is formed by at least two distinct blue triangle components, which cannot therefore be triangle-connected among themselves. But we also know that (B_2', R_2) is ht-blue, so if we had that more than one blue component in B_2' has blue neighbours in R_3 we would get that these blue components are triangle-connected.

Now we claim that we must have $|R_{\geqslant 3}| \geqslant (1 - 20h)t$. As observed above, there is a red TCTF in $B_1' \cup B_2' \cup B_3'$ of size at least $\frac{3}{2} |B_2' \cup B_3'|$, and its triangles are triangle-connected in red to $R_{\geqslant 3}$, we have a red TCTF of size $\frac{3}{2} |B_2' \cup B_3'| + |R_3| - ht$, moreover, we have $|R_3 \cup B_3' \cup B_2'| \geqslant 9 - |B_1' \cup R_1 \cup R_2| \geqslant (\frac{7}{3} - 10h)t$ which gives us a red TCTF over more than $(3 + h)t$ vertices, unless $|R_3| \geqslant (1 - 20h)t$.

In particular, we can say that we can find a blue TCTF in $(R_1, R_2, R_{\geqslant 3})$ of size at least $3(1 - 20h)t$. Since we have already that (R_2, B_2') and (R_2, B_3') are ht-blue, and since we cannot extend the blue TCTF in $(R_1, R_2, R_{\geqslant 3})$ at all, this means that (R_1, B_2') and (R_1, B_3') must be ht-red, but this is absurd since it would create a red TCTF in $(B_1', B_2', B_3') \cup R_1$ of size at least $\frac{3}{2} |B_3' \cup B_2'| + |R_1| - ht > 3(1 + h)t$. $\qquad\square$

6.3 The sharp upper bound

Lemma 22. *There exists $h_0 > 0$ such that for every $0 < h < h_0$ there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ we have the following. Let G be a 2-edge coloured graph with $(9 - \varepsilon)t$ vertices and minimum degree at least $(9 - 2\varepsilon)t$. Fix a collection of red and blue cliques as in Setting 1 with parameters ε and t, and define $B_1, B_2, \ldots$ as in Setting 1. We have that $|B_1|, |R_1| \leqslant (2 + h)t$.*

Proof. Let us denote with $B_{\geqslant 3}$ the set $\cup_{i \geqslant 3} B_i$ and similarly for red. By Lemmas 14 and 15 we can assume $|B_1|, |R_1| \leqslant (\frac{7}{3} + h)t$ and $|B_2|, |R_2| \leqslant (2 + h)t$. Let us assume by contradiction that $|B_1| \geqslant (2 + h)t$. We construct greedily a blue TCTF T_B as follows. Select an edge in a blue clique of B_1, and extend it (if possible) to a blue triangle using a vertex outside of B_1 not used yet in the process. We can repeat greedily until there are no blue edges in cliques of B_1 that can be extended outside of T_B. Let us denote with Y_B the set of vertices $T_B \setminus B_1$ used to extend the edges in B_1, and let us denote with B_1' the set $B_1 \setminus T_B$ of remaining vertices.

Because T_B is triangle-connected, we have that the size of T_B is smaller than $3(1 + \varepsilon)t$ and therefore in particular $|B_1'| = |B_1| - |B_1 \cap T_B| > \frac{h}{2}t$. Let $h' := \min\left\{\frac{|B_1'|}{200t}, h\right\} \geqslant h^{\frac{3}{2}}$.

Because we stopped the greedy construction of T_B only when we could not extend T_B anymore, we have that every vertex in $V \setminus (B_1 \cup Y_B)$ has at most as many blue neighbours in B_1' as the number of cliques with at least two vertices that are in B_1'. This means that the number of blue edges in $(B_1', V \setminus (B_1 \cup Y_B))$ is at most $7t \cdot (\frac{7}{3} \frac{1}{\frac{99}{100} m} + \frac{k}{|\log \varepsilon|^{\frac{3}{2}}})t \leqslant (\frac{5}{\sqrt{m}} t)^2$.

Therefore we have that the pair $(B_1', V \setminus (B_1 \cup Y_B))$ is λt-red for $\lambda = \frac{5}{\sqrt{m}}$.

We now separate four cases.

<u>Case A</u>: We already have $|B_1| \geqslant (2+h)t$, and $|B_1|, |R_1| \leqslant (\frac{7}{3}+h)t$, and $|B_2|, |R_2| \leqslant (2+h)t$. Assume now $|R_{\geqslant 3}| \leqslant ht$.

It follows that $|\cup_i R_i| \leqslant (\frac{13}{3}+3h)t$, and from this it follows that $|\cup B_i| \geqslant (\frac{14}{3}-4h)t$ and therefore $|B_2 \cup B_{\geqslant 3}| \geqslant (\frac{7}{3}-5h)t$. Since $|B_2| \geqslant |B_{\geqslant 3}|$ by 21(vii), we have that $|B_2| > (1+h)t$ and therefore by 9 we have $|B_{\geqslant 3}| < (1+h)t$. By what stated above, we also get:

we get the following:

$$|B_{\geqslant 3}| \geqslant \begin{cases} (\frac{8}{3}-5h)t - |R_1| \\ (\frac{7}{3}-5h)t - |R_2| \geqslant (\frac{1}{3}-6h)t \\ (\frac{14}{3}-4h)t - |B_1| - |R_2| \end{cases}$$

Since $|B_{\geqslant 3}| < (1+h)t$, we must have $|R_2| > (\frac{4}{3}-6h)t$. Let us call C_B the red triangle component in $(B_1, B_2, B_{\geqslant 3})$ that by Corollary 9 contains almost all red edges of $(B_1, B_{\geqslant 3})$ and $(B_2, B_{\geqslant 3})$.

Claim 23. *No red edge in a clique of R_1 or R_2 is red triangle-connected to C_B.*

Proof. If R_1 were red triangle-connected to C_B we could extend a large red TCTF of size $3|B_{\geqslant 3}| - ht$ (which is given us by Corollary 9) using vertices of R_1 and obtain a TCTF over

$$3|B_{\geqslant 3}| + |R_1| - 2ht \geqslant 3((\frac{8}{3}-5h)t - |R_1|) + |R_1| - 2ht$$
$$= (8-17h)t - 2|R_1| > 3(1+\varepsilon)t$$

vertices. It is also absurd that R_2 is red triangle-connected to C_B. Indeed we would have:

$\underline{Case\ 1}$: If $|R_2| \leqslant \frac{17}{9}t$ then we have a red TCTF over $3|B_{\geqslant 3}| + |R_2| - ht \geqslant (7-16h)t - 2|R_2| > 3(1+\varepsilon)t$ vertices.

$\underline{Case\ 2}$: If $|R_2| \geqslant \frac{17}{9}t$, we can greedily construct a TCTF T as follows. We select edges in red cliques of $R_2 \setminus Y_B$ and we extend them to disjoint triangles using vertices of B_1'. Because $(R_2 \setminus Y_B, B_1')$ is λt-red we have that we can continue this process until we almost finish red edges in red cliques of $R_2 \setminus Y_B$ (we can have at most ht vertices remaining in $R_2 \setminus Y_B$) or vertices in B_1' with enough red neighbours in $R_2 \setminus Y_B$ (if we stopped because of this we have that at most $h't$ vertices in B_1' are not used, because otherwise we would have vertices with high red degree in $R_2 \setminus Y_B$ and because $|R_2 \setminus Y_B|$ would be at least $h't$ we could find an edge in a clique between two neighbours of the same vertex of B_1'). At this point we can extend T with triangles from cliques of R_2 and obtain a TCTF over at least $\min\left\{|R_2| + |B_1'|, \frac{3}{2}|R_2 \setminus Y_B| + |R_2 \cap Y_B|\right\} - 3ht$ vertices. T intersects B_1 in at most t vertices, therefore we can again extend T using the tripartition $(B_1 \setminus T, B_2, B_{\geqslant 3})$, in this way we are adding at least $3|B_{\geqslant 3}| - ht$ vertices since $|B_3| \leqslant (1+h)t$, $|B_2| > \frac{7}{6}$ and $|B_1| \geqslant (2+h)t$. Therefore we end up with a red TCTF over $\min\left\{|R_2| + |B_1'|, \frac{1}{2}|R_2 \setminus Y_B| + |R_2|\right\} + 3|B_{\geqslant 3}| - 4ht = 3|B_{\geqslant 3}| + |R_2| +$

$\min\left\{\left|B'_1\right|, \frac{1}{2}\left|R_2 \setminus Y_B\right|\right\} - 4ht$ vertices. We can notice at this point that

$$\begin{aligned}
\left|B'_1\right| + \left|B_{\geqslant 3}\right| &\geqslant \left|B_1\right| + \left|B_{\geqslant 3}\right| - (2+h)t \\
&\geqslant \frac{14}{3}t - 4ht - \left|B_2\right| - (2+h)t \\
&\geqslant \left(\frac{2}{3} - 6h\right)t\,.
\end{aligned}$$

Since $\left|B_{\geqslant 3}\right| \geqslant \left(\frac{1}{3}6h\right)t$ and $\left|R_2 \setminus Y_B\right| \geqslant \frac{5}{6}t$, we are done. Indeed we have $3\left|B_{\geqslant 3}\right| + \left|R_2\right| + \min\left\{\left|B'_1\right|, \frac{1}{2}\left|R_2 \setminus Y_B\right|\right\} - 4ht \geqslant 3\left(\frac{1}{3}6h\right)t + \frac{17}{9}t + \frac{1}{3}t > 3(1+h)t$. $\qquad\square$

Now we know that neither R_1 nor R_2 are triangle-connected to the large triangle component of the tripartition $(B_1, B_2, B_{\geqslant 3})$. In order to use Lemma 18 efficiently, we first need to remember that B'_1, $R_1 \setminus Y_B$ and $R_2 \setminus Y_B$ are all non-trivial and that $(B'_1, R_1 \setminus Y_B)$ and $(B'_1, R_2 \setminus Y_B)$ are both λt-red. Now we can use Lemma 18 to conclude that at most $2\lambda t$ vertices in $(R_1 \cup R_2) \setminus Y_B$ can have more than $2\lambda t$ red neighbours in each of B_2 and $B_{\geqslant 3}$. But this is absurd because of Lemma 19.

<u>Case B</u>: We already have $\left|B_1\right| \geqslant (2+h)t$, and $\left|B_1\right|, \left|R_1\right| \leqslant \left(\frac{7}{3}+h\right)t$, and $\left|B_2\right|, \left|R_2\right| \leqslant (2+h)t$. Assume now that $\left|B_{\geqslant 3}\right| \leqslant ht$.

We can also assume that $\left|R_1\right| \leqslant (2+h)t$ because otherwise we would be in the same situation as case A under switching colours. By Corollary 21(iv) we have $\left|\cup_i B_i\right| \geqslant \frac{43}{12}t$ and this implies $\left|B_2\right| \geqslant \frac{6}{5}t$. We can consider that $\left|R_2 \cup R_{\geqslant 3}\right| \geqslant (9-h)t - \left|R_1\right| - \left|\cup_i B_i\right| \geqslant \left(\frac{8}{3} - 5h\right)t$, which gives us $\left|R_{\geqslant 3}\right| \geqslant \left(\frac{2}{3} - 6h\right)t$. By Lemma 21(vii) we have $\left|R_2\right| \geqslant \left(\frac{4}{3} - 3h\right)t$. By Corollary 10 this also implies that there is a red TCTF on $(R_1, R_2, R_{\geqslant 3})$ covering at least $3\left|R_{\geqslant 3}\right| - ht \geqslant (2-19h)t$ vertices. This gives us the upper bound $\left|R_{\geqslant 3}\right| \leqslant \frac{1+h}{t}$. This also implies that $\left|R_2\right| \geqslant \left(\frac{5}{3} - 6h\right)t$.

Since both B_1 and B_2 are bigger than $\frac{8}{7}t$ we have that neither B_1 nor B_2 can be blue triangle-connected to the large TCTF over $(R_1, R_2, R_{\geqslant 3})$.

By Lemma 18 this means that at most ht vertices from each of B_1 and B_2 can be blue adjacent to more than $2ht$ vertices in any two of R_1, R_2 or $R_{\geqslant 3}$. But we know also that B'_1, $R_1 \setminus Y_B$ and $R_2 \setminus Y_B$ are non trivial, and therefore $(B'_1, R_1 \setminus Y_B)$ and $(B'_1, R_2 \setminus Y_B)$ are λt-red. Hence, by Lemma 19 it can not not be the case that there are more than $2ht$ vertices of B_2 with more than $2ht$ red neighbours in both $R_1 \setminus Y_B$ and $R_2 \setminus Y_B$. Therefore by Lemma 18 there are at most $3ht$ vertices in B_2 which have more than $2ht$ blue neighbours in $R_{\geqslant 3}$.

This means that we can find a set S_1 of at least $\frac{1}{2}\left|B_2\right| - 10ht$ vertices in B_2 such that every vertex in S_1 has at most $2ht$ blue neighbours both in $R_{\geqslant 3}$ and one of $R_1 \setminus Y_B$ or $R_2 \setminus Y_B$ (say R_2, it is the same if it was R_1). Therefore by applying Lemma 19 with S_1 and B'_1 on one side and R_2 and $R_{\geqslant 3}$ on the other side, we get that there are at most $6h't$ vertices in B'_1 which have more than $3h't$ red neighbours in $R_{\geqslant 3}$, and this means that $(B'_1, R_{\geqslant 3})$ is $6h't$-blue. By Lemma 18 we know that (B'_1, R_1) and (B'_1, R_2) are $9h't$-red, and in the same way we know that almost all the vertices of B_2 are $2h't$-red to one of R_1 or R_2. As an example, we are going to assume that we have a subset S_2 of B_2 of size at least $\frac{\left|B_2\right| - 20h't}{2}$ such that every vertex in S_2 has at most $2h't$ blue neighbours to R_2.

Therefore $(S_2, R_{\geqslant 3})$ and (S_2, R_2) are $2h't$-red. Because (B_1', R_1) and (B_1', R_2) are both $9h't$-red, by Lemma 19 we have that (B_1, R_1) is $9h't$-red. By Lemma 18 as above, at most $6h't$ vertices in B_1 can have more than $2h't$ blue neighbours in any two of R_1, R_2 and $R_{\geqslant 3}$. We can find $S' \subseteq B_1$ of size at least $\frac{|B_1| - 20h't}{2}$ that is either $10h't$-red to $R_{\geqslant 3}$ or to R_2. In the first case, we find a large red TCTF using triangles in $(S', S_2, R_{\geqslant 3})$ and then triangles in B_1. In the latter case, we can find a red TCTF on (S_2, S', R_2) over at least

$$2 \cdot \min \{|S_2|, |S'|, |R_2|\} + |R_2| - 20h't$$

vertices. We claim that this is enough, indeed we have $|R_{\geqslant 3}| \leqslant (1 + h')t$, and therefore we get the lower bound $(\frac{5}{3} - 10h')t$ for $|R_2|$ and $t - 10h't$ for $|S'|$.

<u>Case C</u>: We already have $|B_1| \geqslant (2 + h)t$, and $|B_1|, |R_1| \leqslant (\frac{7}{3} + h)t$, and $|B_2|, |R_2| \leqslant (2 + h)t$. Assume now $|R_1| \leqslant (2 + h)t$ and $|B_{\geqslant 3}|, |R_{\geqslant 3}| \geqslant ht$.

We have two cases.

<u>*Case C.1*</u>: Let us assume $|R_2| \leqslant \frac{8}{7}t$.

Claim 24. *Neither B_1 nor B_2 is blue connected to the TCTF over $(R_1, R_2, R_{\geqslant 3})$. Also, R_1 is not triangle-connected to $(B_1, B_2, B_{\geqslant 3})$.*

Proof. By Corollary 21(iv) we have that $|R_1 \cup R_2 \cup R_{\geqslant 3}| \geqslant \frac{43}{12}t$ and hence $|R_2 \cup R_{\geqslant 3}| \geqslant (\frac{19}{12} - h)t$.

By Lemma 21(vii) we have $|R_2| \geqslant |R_{\geqslant 3}|$ and by Lemma 21(vi) we have $|B_2| \geqslant \frac{8}{7}t$ and since $R_1, \ldots, B_{\geqslant 3}$ form a partition of G, we have $|R_{\geqslant 3}| \geqslant (9 - \frac{7}{3} - 1 - 2 - \frac{8}{7} - 3h)t - |B_2| > (\frac{5}{2} + h)t - |B_2|$. By Corollary 9 we can find a blue TCTF over $(R_1, R_2, R_{\geqslant 3})$ of size at least $3|R_{\geqslant 3}| \geqslant \frac{15}{2}t - 3|B_2|$. In particular this implies that both B_1 and B_2 are not triangle-connected to the blue TCTF over $(R_1, R_2, R_{\geqslant 3})$.

We now prove that R_1 is not triangle-connected to $(B_1, B_2, B_{\geqslant 3})$.

If $|R_2 \cup R_{\geqslant 3}| > (\frac{8}{7} + 1 + h)t$, then by Lemma 21(vii) we have $|R_2| > |R_{\geqslant 3}|$, since $|R_2| \leqslant \frac{8}{7}t$ we have $|R_{\geqslant 3}| \geqslant (1 + h)t$ and by Corollary 10 we again obtain a blue TCTF of size $(3 + h)t$.

If on the other hand we have $|R_2 \cup R_{\geqslant 3}| \leqslant (\frac{8}{7} + 1 + h)t$, it follows that $|B_{\geqslant 3}| \geqslant (9 - \frac{7}{3} - 2 - \frac{8}{7} - 1 - 4h)t - |R_1|$ which means $3|B_{\geqslant 3}| + |R_1| \geqslant \frac{24}{7}t$. Therefore it cannot be that R_1 is red triangle-connected to the large TCTF over $(B_1, B_2, B_{\geqslant 3})$. $\square$

Since R_1 is not connected to $(B_1, B_2, B_{\geqslant 3})$ we have by Lemma 18 that at most $h^5 t$ vertices in R_1 have more than $2h^5 t$ red neighbours in two of the B_i. Since $R_1 \setminus Y_B$ is nontrivial we have that $(B_1', R_1 \setminus Y_B)$ is λt-red. Therefore we must have that $(R_1 \setminus Y_B, B_2), (R_1 \setminus Y_B, B_{\geqslant 3})$ are $h^2 t$-blue. We can now apply Lemma 18 again knowing that B_2 is not blue triangle-connected to the blue triangle component over $(R_1, R_2, R_{\geqslant 3})$ and therefore at most $h^5 t$ vertices of B_2 have more than $2h^5 t$ blue neighbours in two of the R_i. Hence, (B_2, R_2) and $(B_2, R_{\geqslant 3})$ are $h^2 t$-red.

Since they are not red triangle-connected among themselves, we have that either R_2 or $R_{\geqslant 3}$ is not red triangle-connected to the red triangle component over $(B_1, B_2, B_{\geqslant 3})$. Let R_2 the one not red triangle-connected, and $R_{\geqslant 3}$ the other one (if the situation is reversed we get better bounds). Then by Lemma 18 we have that R_2 is $h^2 t$-blue to B_1 and $B_{\geqslant 3}$, and therefore by the same Lemma we have that $(B_1, R_{\geqslant 3})$ is $h^2 t$-red. Then

$(B_1, B_2, B_{\geqslant 3} \cup R_{\geqslant 3})$ is a dense red tripartition with $|B_1|, |B_2| \geqslant \frac{8}{7}t$. We have $|B_{\geqslant 3} \cup R_{\geqslant 3}| \geqslant (9 - \frac{7}{3} - 2 - 2 - \frac{8}{7} - 3h)t \geqslant \frac{3}{2}t$ which is enough to conclude by Corollary 9.

Case C.2: Let us now assume $|R_2| \geqslant \frac{8}{7}t$.

Then both $R_1 \setminus Y_B$ and $R_2 \setminus Y_B$ are non trivial and λt-red to B_1'. We cannot have that both R_1 and R_2 are red triangle-connected to $(B_1, B_2, B_{\geqslant 3})$ (because otherwise they would be red triangle-connected among themselves). By Lemma 18 this means that one between $R_1 \setminus Y_B$ and $R_2 \setminus Y_B$ must be $h^2 t$-blue to both B_2 and $B_{\geqslant 3}$, we are going to work with the example in which $R_1 \setminus Y_B$ is $h^2 t$-blue to both B_2 and $B_{\geqslant 3}$ (it would be the same if we had R_2).

We cannot have both B_2 and $B_{\geqslant 3}$ to be blue triangle-connected to $(R_1, R_2, R_{\geqslant 3})$ (otherwise they would be in the same connected component) and therefore we split our case depending on whether or not B_2 is blue triangle-connected to $(R_1, R_2, R_{\geqslant 3})$.

Let us assume that it is so. Then $B_{\geqslant 3}$ is not blue triangle-connected to $(R_1, R_2, R_{\geqslant 3})$ and so $(B_{\geqslant 3}, R_2)$ and $(B_{\geqslant 3}, R_{\geqslant 3})$ are $h^2 t$-red. By Lemma 18 this implies that R_2 is red triangle-connected to $(B_1, B_2, B_{\geqslant 3})$ and therefore $R_{\geqslant 3}$ is not. Therefore $(R_{\geqslant 3}, B_1)$ and $(R_{\geqslant 3}, B_2)$ are $h^2 t$-blue. Therefore (B_1, R_1) and (B_1, R_2) must be $h^2 t$-red. Therefore we can find a blue TCTF over $3|R_{\geqslant 3}| + |B_2|$ vertices by taking triangles from $(R_1, R_2, R_{\geqslant 3})$ and B_2. We can also find a red TCTF over $3|B_{\geqslant 3}| + \frac{3}{2}|R_2|$ vertices by taking triangles from $(B_1, B_2, B_{\geqslant 3})$ and by taking edges from R_2 and extending them with vertices from B_1. We conclude by taking the average of the size of these two TCTFs.

Let us now assume that B_2 is not blue triangle-connected to $(R_1, R_2, R_{\geqslant 3})$. Then (B_2, R_2) and $(B_2, R_{\geqslant 3})$ are $h^2 t$-red, since $(R_1 \setminus Y_B, B_2)$ is ht-blue, and this implies that R_2 is red triangle-connected to $(B_1, B_2, B_{\geqslant 3})$. Therefore $R_{\geqslant 3}$ is $h^2 t$-blue to B_1 and $B_{\geqslant 3}$ and so $B_{\geqslant 3}$ is blue triangle-connected to $(R_1, R_2, R_{\geqslant 3})$. This also means that B_1 must be $h^2 t$-red to both R_1 and R_2 in order not to be blue triangle-connected to $(R_1, R_2, R_{\geqslant 3})$ but this leaves us with a dense red $(B_1, B_2, B_{\geqslant 3} \cup R_2)$.

<u>Case D</u>: We already have $|B_1| \geqslant (2 + h)t$, and $|B_1|, |R_1| \leqslant (\frac{7}{3} + h)t$, and $|B_2|, |R_2| \leqslant (2 + h)t$. Let us assume that $|R_1| \geqslant (2 + h)t$ and both $B_{\geqslant 3}$ and $R_{\geqslant 3}$ contain more than ht vertices (otherwise without loss of generality we are in case B). We can also assume without loss of generality that $|B_2| \geqslant |R_2|$ and therefore by Lemma 21(vi) we also have $|B_2| \geqslant \frac{8}{7}t$.

We can greedily extend blue edges in cliques of B_1 to a blue TCTF T_B by using vertices outside of B_1. Since $|B_1| \geqslant (2 + h)t$ we can either create a TCTF over more than $3(1 + \varepsilon)t$ vertices or we have to stop at some point. Since $|B_1| \geqslant (2 + h)t$, this means that $B_1 \setminus T_B$ is non trivial. We can do the same with a red TCTF T_R extending red edges in R_1 (since we are assuming $|R_1| \geqslant (2 + h)t$). Let us call $B_1' := B_1 \setminus T_B$ and $R_1' := R_1 \setminus T_R$. Since the TCTFs T_R and T_B are maximal, we have that $(B_1', V(G) \setminus T_B)$ is ht-red, while $(R_1, V(G) \setminus T_R)$ has to be ht-blue. In particular, there are non-trivial subsets $S_{B_1} \subseteq B_1$ of size at least $(1 + \frac{h}{2})t$, $S_{B_2} \subseteq B_2$ of size at least $(\frac{1}{7} + \frac{h}{2})t$ and $S_{R_1} \subseteq R_1$ of size at least $(1 + \frac{h}{2})t$ such that (S_{B_1}, R_1') and (S_{B_2}, R_1') are ht-blue and (S_{R_1}, B_1') is ht-red.

There are two cases:

<u>Case D.1</u>: B_1 is blue triangle-connected to the large TCTF in $(R_1, R_2, R_{\geqslant 3})$. Then we know that B_2 and $B_{\geqslant 3}$ are not triangle-connected to the same TCTF. In particular, since

(S_{B_2}, R'_1) is ht-blue, we must have that both (S_{B_2}, R_2) and $(S_{B_2}, R_{\geqslant 3})$ are ht-red. Now, either R_1 is red triangle-connected to the large TCTF in $(B_1, B_2, B_{\geqslant 3})$ or not.

In the first case, we have that both R_2 and $R_{\geqslant 3}$ are not triangle-connected to the large TCTF in $(B_1, B_2, B_{\geqslant 3})$. Because (S_{B_2}, R_2) and $(S_{B_2}, R_{\geqslant 3})$ are ht-red, this means that $(R_{\geqslant 3}, B_{\geqslant 3}), (R_{\geqslant 3}, B_1)$ and $(R_2, B_{\geqslant 3}), (R_2, B_1)$ are ht-blue, which is absurd because it would mean that B_1 and $B_{\geqslant 3}$ are in the same blue-connected component.

If R_1 is not red triangle-connected to the large TCTF in $(B_1, B_2, B_{\geqslant 3})$, then $(S_{R_1}, B_{\geqslant 3})$ and (S_{R_1}, B_2) have to be ht-blue. But now we get a contradiction since $(B_{\geqslant 3}, R_{\geqslant 3})$ and $(B_{\geqslant 3}, R_2)$ need to be ht-red or otherwise $B_{\geqslant 3}$ is going to be triangle-connected to the blue TCTF in $(R_1, R_2, R_{\geqslant 3})$, and also $(B_2, R_{\geqslant 3})$ and (B_2, R_2) need to be ht-red or otherwise B_2 is going to be triangle-connected to the blue TCTF in $(R_1, R_2, R_{\geqslant 3})$. This is enough to say that R_2 and $R_{\geqslant 3}$ are in the same red-connected component.

<u>Case D.2</u>: B_1 is not blue triangle-connected to the large TCTF in $(R_1, R_2, R_{\geqslant 3})$ but R_1 is red triangle-connected to the large TCTF in $(B_1, B_2, B_{\geqslant 3})$. Since B_1 is not blue triangle-connected to the large TCTF in $(R_1, R_2, R_{\geqslant 3})$ and because (S_{B_1}, R'_1) is ht-blue, we have that (S_{B_1}, R_2) and $(S_{B_1}, R_{\geqslant 3})$ are ht-red. Now, since R_1 is red triangle-connected to the large TCTF in $(B_1, B_2, B_{\geqslant 3})$ we have that R_2 and $R_{\geqslant 3}$ are not, because (S_{B_1}, R_2) and $(S_{B_1}, R_{\geqslant 3})$ are ht-red this implies that $(B_2, R_2), (B_{\geqslant 3}, R_2)$ and $(B_2, R_{\geqslant 3}), (B_{\geqslant 3}, R_{\geqslant 3})$ are ht-blue, which is absurd because it implies that both B_2 and $B_{\geqslant 3}$ are connected to the large TCTF in $(R_1, R_2, R_{\geqslant 3})$.

<u>Case D.3</u>: B_1 is not blue triangle-connected to the large TCTF in $(R_1, R_2, R_{\geqslant 3})$ and R_1 is not blue triangle-connected to the large TCTF in $(B_1, B_2, B_{\geqslant 3})$. In which case we notice that the blue cliques in B_1 are not triangle-connected to the large blue TCTF in $(R_1, R_2, R_{\geqslant 3})$ and similarly the red cliques in R_1 are not triangle-connected to the large red TCTF in $(B_1, B_2, B_{\geqslant 3})$. In particular, this implies that $(S_{B_1}, R_{\geqslant 3})$ and (S_{B_1}, R_2) are ht-red, because we have that (S_{B_1}, R'_1) is ht-blue and $(R'_1, \cup_{i \geqslant 2} R_i)$ is ht-blue. Likewise, we have that $(S_{R_1}, B_{\geqslant 3})$ and (S_{R_1}, B_2) are ht-blue. But this leaves us in a contradiction, indeed, neither B_2 nor $B_{\geqslant 3}$ can be triangle-connected to (R_1, R_2, R_3). Since $(S_{R_1}, B_{\geqslant 3})$ is ht-blue this means that $(R_2, B_{\geqslant 3})$ and $(R_{\geqslant 3}, B_{\geqslant 3})$ are ht-red. This is enough to get a contradiction, since we have $(R_2, B_{\geqslant 3})$ and $(R_{\geqslant 3}, B_{\geqslant 3})$ are ht-red but also $(S_{B_1}, R_{\geqslant 3})$ and (S_{B_1}, R_2) are ht-red. $\qquad\square$

7 The colours of edges

In this section we complete the proof of Lemma 3. We first deduce an approximate version, proving that $B_{\geqslant 3} \cup R_{\geqslant 3}$ cannot have much more than t vertices (which implies all components have roughly the correct size) and that most edges in various pairs have the 'correct' colour. We then prove Lemma 3 by arguing that any edges with the 'wrong' colour lead to triangle components which are much larger than they should be. The following is our approximate version.

Lemma 25. *There exists $h_0 > 0$ such that for every $0 < h < h_0$ there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ there exists t_0 such that for every $t > t_0$ we have the following. Let*

G be a 2-edge coloured graph with $(9-\varepsilon)t$ vertices and minimum degree at least $(9-2\varepsilon)t$. Fix a collection of red and blue cliques as in Setting 1 with parameters ε and t, and define $B_1, B_2, \ldots$ as in Setting 1. Then it holds:

- $(2-h)t \leqslant |B_1|, |B_2|, |R_1|, |R_2| \leqslant (2+h)t$,

- $(1-h)t \leqslant |B_{\geqslant 3} \cup R_{\geqslant 3}| \leqslant (1+h)t$,

- $|G \setminus \cup_i (B_i \cup R_i)| \leqslant ht$,

- *The following pairs are h^2t-blue: (B_1, R_1), (B_2, R_2), (R_1, R_2), $(R_1, B_{\geqslant 3} \cup R_{\geqslant 3})$ and $(R_2, B_{\geqslant 3} \cup R_{\geqslant 3})$,*

- *The following pairs are h^2t-red: (B_1, B_2), $(B_1, B_{\geqslant 3} \cup R_{\geqslant 3})$, $(B_2, B_{\geqslant 3} \cup R_{\geqslant 3})$, (B_1, R_2) and (B_2, R_1).*

Proof. By Lemma 22 we know that for $\varepsilon > 0$ small enough we have $|B_1|, |B_2|, |R_1|, |R_2| \leqslant (2+h^{\frac{3}{2}})t$ and therefore we have $|B_{\geqslant 3} \cup R_{\geqslant 3}| \geqslant (1-5h^{\frac{3}{2}})t$. Without loss of generality, let us assume $|B_{\geqslant 3}| \geqslant |R_{\geqslant 3}|$.

Claim 26. *We have that R_1 and R_2 are not red triangle-connected to $(B_1, B_2, B_{\geqslant 3})$. Moreover, without loss of generality, we have $(R_1, B_1), (R_1, B_{\geqslant 3})$ and $(R_2, B_2), (R_2, B_{\geqslant 3})$ are h^2t-blue.*

Proof. Notice that $|B_1| \geqslant \frac{(1-5h^{\frac{3}{2}})t}{2}$. Let us consider first that R_1 and R_2 are not red triangle-connected to $(B_1, B_2, B_{\geqslant 3})$. Indeed, assume this is not the case and we have $3|B_{\geqslant 3}| + |R_i| < (3+h)t$ for some $i \in \{1, 2\}$. Then we have $|R_1| + |R_2| + |B_{\geqslant 3}| < (3+h+2+h^{\frac{3}{2}})t - 2|B_{\geqslant 3}| < (4-3h)t$ which is clearly absurd because it implies $|B_1| + |B_2| + |R_{\geqslant 3}| \geqslant (5+2h)t$.

We now claim that there is an ordering i, j, k of $\{1, 2, \geqslant 3\}$ such that $(R_1, B_i), (R_1, B_j)$ and $(R_2, B_k), (R_2, B_j)$ are h^2t-blue. Indeed, by Lemma 18 we know that up to removing at most h^5t vertices from each of R_1 and R_2, every vertex in $R_1 \cup R_2$ has many blue edges in at least two among $\{B_1, B_2, B_{\geqslant 3}\}$. This means that we can partition (not in a unique way) almost all the vertices of R_1 among the sets $S_{B_h}^{R_1}$, where the vertices in $S_{B_h}^{R_1}$ have their red neighbour in $\cup_\ell B_\ell$ contained in B_h. We define similarly $S_{B_h}^{R_2}$. We claim that just one of the $S_{B_h}^{R_1}$ is not trivial.

Assume by contradiction that $S_{B_i}^{R_1}$ and $S_{B_j}^{R_1}$ have size at least ht. We cannot have that $S_{B_i}^{R_2}$ or $S_{B_j}^{R_2}$ have size at least ht, because otherwise we would have that B_j and B_k or B_i and B_k are connected respectively. Therefore we must have that $S_{B_k}^{R_2}$ contains almost all the vertices of R_2 and in particular is not trivial. Therefore we have that $S_{B_i}^{R_1}, S_{B_j}^{R_1}$ and $S_{B_k}^{R_2}$ are not trivial, which gives us that both B_i and B_j are in the same triangle component. This implies that just one of the $S_{B_h}^{R_1}$ is nontrivial, and by symmetry the same is true for R_2. Moreover, we have that $S_{B_{\geqslant 3}}^{R_i}$ is trivial, because otherwise we would find a large blue TCTF in $(B_1, B_2, R_i \cup B_{\geqslant 3})$.

Finally, since by Lemma 19 we cannot have that R_1 and R_2 are $h^2 t$-blue to the same pair, we know that each of R_1 and R_2 is $h^2 t$-blue to $B_{\geqslant 3}$ and one between B_1 and B_2. We are going to assume without loss of generality that (R_1, B_1) and $(R_1, B_{\geqslant 3})$ are $h^2 t$-blue, and that (R_2, B_2) and $(R_2, B_{\geqslant 3})$ are $h^2 t$-blue, as we wanted.

$\square$

By the claim, we have that (R_1, B_1), $(R_1, B_{\geqslant 3})$ and (R_2, B_2), $(R_2, B_{\geqslant 3})$ are $h^2 t$-blue. In particular, this means that we can find a blue TCTF in $(R_1, R_2, B_{\geqslant 3} \cup R_{\geqslant 3})$. This gives us immediately that $|B_{\geqslant 3} \cup R_{\geqslant 3}| \leqslant (1 + h^{\frac{3}{2}})t$ and in particular $|B_1|, |B_2|, |R_1|, |R_2| \geqslant (2-h)t$.

Also, we get that (B_1, R_2) and (B_2, R_1) are $h^2 t$-red. This holds because otherwise we would have both $B_{\geqslant 3}$ and B_2 are in the same connected component, indeed, $(B_{\geqslant 3}, R_1)$, $(B_{\geqslant 3}, R_2)$ and (B_2, R_2) are $h^2 t$ blue.

Assume now that $|R_{\geqslant 3}| \geqslant h^{\frac{3}{2}} t$. We have that $(R_1, B_1), (R_1, B_{\geqslant 3})$ and $(R_2, B_2), (R_2, B_{\geqslant 3})$ are $h^2 t$-blue, this gives us that $B_{\geqslant 3}$ is blue triangle-connected to $(R_1, R_2, R_{\geqslant 3})$ (which is a non-trivial TCTF) which in turn gives us that B_1 and B_2 are not. From this last fact we can conclude that $(B_1, R_{\geqslant 3})$ and $(B_2, R_{\geqslant 3})$ are all $h^2 t$-red.

So we have the construction that we wanted up to change the indices between B_1, B_2 and R_1, R_2 respectively. $\square$

Let us now prove Lemma 3 that we restate for convenience.

Lemma 27. *There exists $\delta_0 > 0$ such that for every $0 < h, \lambda < \delta_0$ there exist $\varepsilon_0 > 0$ and $t_0 \in \mathbb{N}$ such that for every $t \geqslant t_0$ and $0 < \varepsilon < \varepsilon_0$ the following holds. Let G be a 2-edge-coloured graph on $(9 - \varepsilon)t$ vertices with minimum degree at least $(9 - 2\varepsilon)t$. Then either G contains a monochromatic TCTF on at least $3(1 + \varepsilon)t$ vertices or $V(G)$ can be partitioned in sets B_1, B_2, R_1, R_2, Z, T such that the following hold.*

(a) $(2 - h)t \leqslant |B_1|, |B_2|, |R_1|, |R_2| \leqslant (2 + h)t$,

(b) $(1 - h)t \leqslant |Z| \leqslant (1 + h)t$,

(c) all the edges in $G[B_1]$ and $G[B_2]$ are blue, and all the edges in $G[R_1]$ and $G[R_2]$ are red,

(d) all the edges between the pairs (B_1, R_1), (B_2, R_2), (R_1, Z) and (R_2, Z) are blue, and those between the pairs (B_1, R_2), (B_2, R_1), (B_1, Z) and (B_2, Z) are red,

(e) the pair (B_1, B_2) is λt-red, and the pair (R_1, R_2) is λt-blue, and

(f) $|T| \leqslant ht$.

Proof of Lemma 3. We are going to refine Lemma 25 to obtain an exact result.

By Lemma 25 we have that there exists $\delta_0 > 0$ such that for $\delta_0 > h, \lambda > 0$ there exists $\varepsilon_0 > 0$ and $t_0 \in \mathbb{N}$ such that for every $t > t_0$ and $\varepsilon_0 > \varepsilon > 0$ if G is a 2-edge-coloured graph over $(9 - \varepsilon)t$ vertices with minimum degree at least $(9 - 2\varepsilon)$ and without a monochromatic TCTF on at least $3(1 + \varepsilon)t$ vertices, then we can partition $V(G)$ in

the sets B_1, B_2, R_1, R_2, Z, T (where the B_i and R_i are as in Setting 1 and where where $Z = B_{\geqslant 3} \cup R_{\geqslant 3}$ and T is the set of vertices which are not already counted) such that the following holds:

- $(2-h)t \leqslant |B_1|, |B_2|, |R_1|, |R_2| \leqslant (2+h)t$,

- $(1-h)t \leqslant |Z| \leqslant (1+h)t$,

- $|T| \leqslant ht$,

- The following pairs are λt-blue: (B_1, R_1), (B_2, R_2), (R_1, R_2), (R_1, Z) and (R_2, Z),

- The following pairs are λt-red: (B_1, B_2), (B_1, Z), (B_2, Z), (B_1, R_2) and (B_2, R_1).

We first need to slightly prune our sets. We start by removing from B_1 (and putting in T) the vertices with more than $\frac{1}{8}\lambda$ red neighbours to R_1 and the vertices with more than λ blue neighbours to either B_2, R_2 or Z. We do the same to B_2, R_1 and R_2 accordingly to the colour of the pairs we are considering.

Up to reducing ε_0, we are still respecting all the bounds on the sizes that we need for Lemma 3, but we have a slightly better result on the state of the "problematic" edges. Indeed, we know that there are no vertices outside T that witness more than λ "problematic" edges.

We now just need to prove that $G[B_1]$, $G[B_2]$, $G[R_1]$, $G[R_2]$ and (B_1, R_1), (B_2, R_2), (R_1, Z), (R_2, Z), (B_1, Z), (B_2, Z), (B_1, R_2), (B_2, R_1) are entirely monochromatic.

The proofs to show that $G[B_1]$, $G[B_2]$, $G[R_1]$, $G[R_2]$ are monochromatic have the same structure. Therefore without loss of generality we show that $G[B_1]$ is entirely blue. Assume by contradiction that we can find u, v in B_1 such that uv is red. By our earlier pruning we know that both u and v have at most $\frac{1}{8}\lambda$ blue neighbours in R_2. Therefore uv is triangle-connected to one of the red cliques of R_2 (and therefore to all red cliques of R_2). Let us now prove that uv is also triangle-connected to the large red TCTF in (B_1, B_2, Z) (which is enough to conclude since we would then be able to find a large triangle-connected triangle component). Almost all the red edges in $(\{u\}, B_2)$ are triangle-connected to uv, indeed, all but at most $\frac{1}{8}\lambda$ of them are in a red triangle with uv, the same holds for the red edges in $(\{u\}, Z)$. This means that there are at most λ vertices in either B_2 or Z that witness a red edge in (B_2, Z) which is not triangle-connected to uv. But this is absurd, as we mentioned before, since it implies that a large red TCTF in (B_1, B_2, Z) is triangle-connected to uv.

Let us now prove that (B_1, R_1), (B_2, R_2), (R_1, Z), (R_2, Z), (B_1, Z), (B_2, Z), (B_1, R_2), (B_2, R_1) are entirely monochromatic. The structure of these proofs is always the same, so without loss of generality we prove that (B_1, R_1) is monochromatic. Assume it is not, and let uv be a red edge between B_1 and R_1 (with $u \in B_1$). We prove that uv is triangle-connected both to one clique of R_1 (and therefore all cliques of R_1) and to the large red triangle component in (B_1, B_2, Z), which is absurd since this would give a large red TCTF.

We first show that uv is triangle-connected to R_2, let $w_1 \in R_2$ such that vw_1 is an edge (which has to be red by our previous proof that $G[R_1]$ is entirely red). Then by our

pruning we know that u, v and w_1 share a red neighbour in B_2. We now observe that if $w_2 w_3$ is a red edge between B_2 and Z (with $w_2 \in B_2$) such that w_2 is a red neighbour of both u and v and w_3 is a red neighbour of w_2 and u, then $w_2 w_3$ is triangle-connected to uv. By the pruning we did earlier, we can say that most of the red edges between B_2 and Z are triangle-connected to uv, which is what we wanted.

Up to changing the roles of the clusters, the other proofs have the same structure. $\square$

8 Regularity Method: proofs of Lemma 4 and Theorem 2

In this section we state the Regularity Lemma and Blow-up Lemma, and use them to deduce Lemma 4 and Theorem 2 from Lemma 3.

Definition 28 (density, ε-regular). Let G be a graph and let X, Y be disjoint subsets in $V(G)$. We define the *density* $d(X, Y)$ between X and Y to be:

$$d(X, Y) := \frac{e(X, Y)}{|X| \, |Y|}.$$

Given $\varepsilon > 0$, we say that (X, Y) is ε-regular if for every $X' \subseteq X$, $Y' \subseteq Y$ such that $|X'| > \varepsilon |X|$ and $|Y'| > \varepsilon |Y|$ we have $|d(X', Y') - d(X, Y)| < \varepsilon$.

We use the following version of the Regularity Lemma. We will apply this to the graph of red edges within K_n, and observe that if (X, Y) is ε-regular in red edges then, since the blue edges are the complement of the red edges, it is also ε-regular in blue.

Lemma 29 (Regularity Lemma). *For every $\varepsilon \in (0, 1)$ there are $M, N_0 \in \mathbb{N}$ such that the following holds. Let G be a graph on $n \geqslant N_0$ vertices, then there is a partition $\{V_0, \ldots, V_m\}$ of $V(G)$ such that the following conditions hold. We have $|V_0| \leqslant \varepsilon^{-1}$ and $\varepsilon^{-1} \leqslant m \leqslant M$. We have $|V_1| = \cdots = |V_m|$. Finally, for any given $i \in [m]$, for all but at most εm choices of $j \in [m]$ the pair (V_i, V_j) is ε-regular in G.*

This version follows from the original version of Szemerédi [12] (which is similar but bounds the total number of irregular pairs by εm^2 rather than the number of irregular pairs meeting a part) applied with parameter $\frac{1}{8}\varepsilon^2$, followed by removing parts incident to more than $\frac{1}{2}\varepsilon m$ irregular pairs (of which there are at most $\frac{1}{2}\varepsilon m$) to V_0; we leave the details to the reader.

Given $\varepsilon, d > 0$, a 2-edge-coloured complete graph G and a partition obtained by applying Lemma 29 with parameter ε to the subgraph of red edges, we define the (ε, d)-*reduced graph* of G (with respect to the partition) to be the graph H on vertex set $[m]$, in which an edge ij is present if it is ε-regular, and assigned the colour red if its density in red is at least $1 - d$, blue if its density in blue is at least $1 - d$, and otherwise purple.

We will see that for the purposes of embedding a graph into G, we can treat purple edges as being either red or blue as we desire, so that a large TCTF in (red $\cup$ purple) edges, or in (blue $\cup$ purple) edges in the reduced graph implies the existence of the square paths and cycles in G we need. In order to apply Lemma 3 in this setting, we deduce the following consequence, which roughly says that either we are done or we get essentially the same partition as in Lemma 3, in particular there are very few purple edges.

Lemma 30. *For every $\delta > 0$ there exists $\varepsilon > 0$ such that for all $t \geqslant \frac{1}{\varepsilon}$, if G is a {red, blue, purple}-edge-coloured graph on $(9 - \varepsilon)t$ vertices with minimum degree at least $(9 - 2\varepsilon)t$, then either there is a choice of a colour between blue and red such that if we colour all the purple edges of that colour we can find a monochromatic TCTF on at least $3(1 + \varepsilon)t$ vertices in G or $V(G)$ can be partitioned in sets $\{B_1, B_2, R_1, R_2, Z, T\}$ such that the following hold.*

(a) $(2 - \delta)t \leqslant |B_1|, |B_2|, |R_1|, |R_2| \leqslant (2 + \delta)t$,

(b) $(1 - \delta)t \leqslant |Z| \leqslant (1 + \delta)t$,

(c) all the edges in $G[B_1]$ and $G[B_2]$ are blue, and all the edges in $G[R_1]$ and $G[R_2]$ are red,

(d) the pairs (B_1, R_1), (B_2, R_2), (R_1, Z) and (R_2, Z) are entirely blue. Moreover, the pairs (B_1, R_2), (B_2, R_1), (B_1, Z) and (B_2, Z) are entirely red,

(e) the pair (B_1, B_2) is δt-red, while the pair (R_1, R_2) is δt-blue, and

(f) $|T| \leqslant \delta t$.

Proof. Let ε be given by Lemma 3 for input $h = \lambda = \frac{1}{1000}\delta$; without loss of generality we may suppose δ is sufficiently small for this application.

Let G be a coloured graph satisfying the conditions of the lemma, and suppose there is neither a red-purple TCTF over $3(1 + \varepsilon)t$ vertices nor a blue-purple TCTF over $3(1 + \varepsilon)t$ vertices.

Let G^r be the graph obtained from G by recolouring the purple edges red, and similarly G^b by recolouring them blue. Let $R_1^r, R_2^r, B_1^r, B_2^r, X^r, T^r$ be the partition obtained by applying Lemma 3 to G^r, and define similarly the partition for G^b replacing r with b. Observe that if we swap R_r^1 and R_2^r, and also B_r^1 and B_2^r, we still have a partition satisfying the conclusion of Lemma 3. If $\left| R_2^r \cap R_1^b \right| > \left| R_1^r \cap R_b^1 \right|$, we perform this swap (and in an abuse of notation continue to use the same letters for the swapped classes).

We define $R_i := R_i^r \cap R_i^b$ and $B_i := B_i^r \cap B_i^b$ for each $i = 1, 2$, and $Z := Z_r \cap Z_b$ and finally $T := V(G) \setminus (B_1 \cup B_2 \cup R_1 \cup R_2 \cup Z)$. We will now prove this partition satisfies the conclusions of the lemma. Observe that the statements in (c), (d) and (e) about sets or pairs being entirely red, or δt-red, follow directly from the same statements for the partition of G^b, and the corresponding ones about being blue from the partition of G^r; what remains is to prove these sets have the correct sizes.

To begin with, observe that all edges in R_1^b are red in G^b and so also in G. It follows that R_1^b intersects B_i^r in at most one vertex for each $i = 1, 2$, since otherwise B_i^r would contain a red edge. Thus R_1^b has at least $(2 - \frac{1}{100}\delta)t - 2$ vertices which are not in $B_1^r \cap B_2^r$. These vertices cannot all be in $T^r \cup Z^r$, which is too small, so R_1^b has a vertex in at least one of R_1^r and R_2^r. Now R_b^1 cannot have vertices in Z^r, since all edges from Z^r to $R_1^r \cup R_2^r$ are not red. It follows that all but at most $\frac{1}{1000}\delta t + 2$ vertices of R_b^1 are in $R_1^r \cup R_2^r$, and by the observation above there are at least as many vertices in R_1^r as R_2^r. Since (R_1^r, R_2^r)

is $\frac{1}{1000}\delta t$-blue, and all edges in R_b^1 are red, we see R_b^1 has at most $\frac{1}{1000}\delta t$ vertices in R_2^r. Finally, we conclude $|R_1| \geqslant (2 - \frac{1}{100}\delta)t$. We also have $|R_1| \leqslant |R_1^r| \leqslant (2 + \frac{1}{1000}\delta)t$. By a similar argument (noting that R_1^b and R_2^b are disjoint) we obtain

$$(2 - \tfrac{1}{100}\delta)t \leqslant |R_i| \leqslant (2 + \tfrac{1}{1000}\delta)t$$

for each $i = 1, 2$.

We make a similar argument for B_1^r. As above, we can conclude that all but at most $\frac{1}{1000}\delta)t + 2$ vertices of B_1^r are in $B_1^b \cup B_2^b$. However we can now observe that all edges from B_1^r to $R_1 \subset R_1^r$ are blue, while the edges from B_2^b to $R_1 \subset R_1^b$ are red. It follows that B_1^r is disjoint from B_2^b, and we obtain

$$(2 - \tfrac{1}{100}\delta)t \leqslant |B_i| \leqslant (2 + \tfrac{1}{1000}\delta)t$$

for $i = 1$, and, by the similar argument, for $i = 2$.

Now Z^r and Z^b are two sets of size at least $(1 - \frac{1}{1000}\delta)t$ in $V(G) \setminus (R_1 \cup R_2 \cup B_1 \cup B_2)$, which has size at most $(9 - \varepsilon)t - 4(2 - \frac{1}{100}\delta)t \leqslant t + \frac{2}{50}\delta t$. It follows their intersection Z has size at least $(1 - \frac{1}{10}\delta)t$, and at most $|Z^r| \leqslant (1 + \frac{1}{1000}\delta)t$. Finally, putting these size bounds together we have (a), (b) and an upper bound on $|T|$ giving (f). □

To go with the above lemma, we state the following two embedding lemmas. The first is a corollary of [1, Lemma 7.1], though one could use the original Blow-up Lemma of Komlós, Sárközy and Szemerédi [9] with some extra technical work in the proof of Theorem 2. To deduce the following statement from [1, Lemma 7.1], we take R' to be the graph with zero edges and $\Delta_{R'} = 1$, we take $\kappa = 2$, and we add to H for each $i \in R$ a set of $|V_i| - |\phi^{-1}(i)|$ isolated vertices which (extending ϕ) we map to i and let be the buffer vertices $\tilde{X}_i$.

Theorem 31. *Given $d, \gamma > 0$ and $\Delta \in \mathbb{N}$, there exists $\varepsilon > 0$ such that for each given T, the following holds for all $m \geqslant m_0$. Let R be any graph on $[t]$, where $t \leqslant T$. Let $V_1, \ldots, V_t$ be disjoint vertex sets with $m \leqslant |V_i| \leqslant 2|V_j|$ for each $i, j \in [t]$, and suppose that G is a graph on $V_1 \cup \cdots \cup V_t$ such that (V_i, V_j) is an ε-regular pair of density at least d for each $ij \in E(R)$. Suppose that H is any graph with $\Delta(H) \leqslant \Delta$ such that there exists a graph homomorphism $\phi : H \to R$ satisfying $\left|\phi^{-1}(i)\right| \leqslant (1 - \gamma)|V_i|$. Then H is a subgraph of G.*

The next is a consequence of the (original) Blow-up Lemma derived in [2].

Lemma 32 (Embedding Lemma, Allen, Böttcher, Hladký [2])**.** *For all $d > 0$ there exists $\varepsilon_{EL} > 0$ with the following property. Given $0 < \varepsilon < \varepsilon_{EL}$, for every $m_{EL} \in \mathbb{N}$ there exists $n_{EL} \in \mathbb{N}$ such that the following holds for each graph G on $n > n_{EL}$ vertices with (ε, d)-reduced graph R on $m \leqslant m_{EL}$ vertices. Let $\xi(R)$ be the size of the largest TCTF in R, then for every $\ell \in \mathbb{N}$ with $3\ell \leqslant (1 - d)\xi(R)\frac{n}{m}$ we have $C_{3\ell}^2 \subset G$.*

We are now in a position to prove Lemma 4, which we restate for convenience.

Lemma 33. *For every $\alpha > 0$ there exists $\delta > 0$ and $n_0 \in \mathbb{N}$ such that for every $n > n_0$ the following holds. Let $N \geqslant (9 - \delta)n$, and let G be a 2-edge-colouring of K_N. Then either G contains both a monochromatic copy of P^2_{3n+2} and of C^2_{3n}, or we can partition $V(G)$ into sets X_1, X_2, Y_1, Y_2, Z and R such that the following hold.*

(a) $(2 - \alpha)n \leqslant |X_1|, |X_2|, |Y_1|, |Y_2| \leqslant (2 + \alpha)n$,

(b) $(1 - \alpha)n \leqslant |Z| \leqslant (1 + \alpha)n$,

(c) $|R| \leqslant \alpha n$,

(d) *Vertices in the following pairs have at most αn red neighbours in the opposite part:* $(X_1, Y_1), (X_2, Y_2), (Y_1, Y_2), (Y_1, Z)$ *and* (Y_2, Z),

(e) *Vertices in the following pairs have at most αn blue neighbours in the opposite part:* $(X_1, X_2), (X_2, Y_1), (X_1, Y_2), (X_1, Z)$ *and* (X_2, Z),

(f) *Vertices in X_1 and X_2 have at most αn red neighbours in their own part,*

(g) *Vertices in Y_1 and Y_2 have at most αn blue neighbours in their own part.*

Proof. Given $\alpha > 0$, let d be such that $10000\alpha^{-2}d$ is returned by Lemma 30 when we use as input $\frac{\alpha^2}{10000}$. Let ε_{EL} be returned by Lemma 32 for input d, and let $\varepsilon = \min(\frac{1}{10}d, \varepsilon_{EL}, \frac{1}{10000}\alpha^2)$. Let now N_0 and M be returned by Lemma 29 with input ε. We let n_{EL} be returned by Lemma 32 for input d, ε and $m_{EL} = M$. Finally, let $\delta = d$ and $n_0 = \max(100\varepsilon^{-1}, N_0, N_1)$ be the constants returned by the lemma..

Let us now fix some $n > n_0$ and, for $N > (9 - \delta)n$, a 2-edge-colouring G of K_N.

We apply Lemma 29 with parameter as above to the red subgraph of G to get a partition $V_0, \ldots, V_m$ of $V(G)$, with $\varepsilon^{-1} \leqslant m \leqslant M$, as in Lemma 29. Let H be the (ε, d)-reduced graph of G. Since each cluster V_i is in at most εm irregular pairs, we have $\delta(H) \geqslant (1 - \varepsilon)m - 1$. Let $t = \frac{m}{9 - 10\alpha^{-1}d}$, so that H has $(9 - 10\alpha^{-1}d)t$ vertices and, by choice of ε, minimum degree at least $(9 - 20\alpha^{-1}d)t$. By Lemma 30 with constants as above, one of the following occurs.

It could be that H contains a red-purple TCTF over $3(1 + 10d)t = \frac{3(1+10d)}{9-10\alpha^{-1}d}m \geqslant \frac{1}{3}(1 + 10d)m$ vertices. Applying Lemma 32 with constants as above, we conclude that G contains a red C^2_{3s} for each $s \leqslant (1 - d) \cdot \frac{1}{3}(1 + 10d) \cdot (9 - d)n \geqslant 3(1 + d)n$. But then in particular G contains a red copy of C^2_{3n} and P^2_{3n+2} and we are done. Similarly, if H contains a blue-purple TCTF over $3(1 + 10d)t$ vertices then G contains a blue copy of C^2_{3n} and P^2_{3n+2} and we are done.

Alternatively, by Lemma 30 we get a partition of $V(H)$ in sets B_1, B_2, R_1, R_2, Z'' and T. We obtain from this a partition of $V(G)$, setting $X'_j = \bigcup_{i \in B_j} V_i$ and $Y'_j = \bigcup_{i \in R_j} V_i$ for $j = 1, 2$, setting $Z' := \bigcup_{i \in Z''} V_i$, and letting R' be the remaining vertices. Since we applied Lemma 30 with input $\frac{\alpha^2}{10000}$ and by choice of ε, we have properties (a) and (b) of Lemma 25 with $\frac{1}{1000}\alpha^2$ instead of α.

Since (B_1, B_2) is $\frac{1}{10000}\alpha^2 t$-red, the number of blue edges in G between X'_1 and X'_2 is at most

$$d|X'_1||X'_2| + \tfrac{1}{1000}\alpha^2 n|X'_2| + \tfrac{1}{1000}\alpha^2 n|X'_1| \leqslant \tfrac{1}{200}\alpha^2 n^2,$$

where the inequality uses $d \leqslant \frac{1}{10000}\alpha^2$. In particular there are less than $\frac{1}{200}\alpha n$ vertices in X_1' which have more than αn blue neighbours in X_2', and similarly swapping X_1' and X_2'. By the identical calculation, an analogous statement holds for Y_1' and Y_2' in red.

We now claim that at most $\varepsilon|X_1'|$ vertices in X_1' send αn or more red edges to Y_1'. Suppose for a contradiction this statement is false. By averaging, there is a cluster V_i with $i \in B_1$ such that a set S of $\varepsilon|V_i|$ vertices of V_i all send αn or more red edges to Y_1'. Since V_i is in irregular pairs with at most εm other clusters, at most $2\varepsilon n$ red edges from each $s \in S$ go to clusters of R_1 that make irregular pairs with V_i. The remaining at least $\frac{1}{2}\alpha n|S|$ edges from S therefore go to the remaining less than $3m$ clusters V_j with $j \in R_1$, which all form ε-regular pairs with V_i that have density at most d in red. Again by averaging, there is a cluster V_j with $j \in R_1$ such that (V_i, V_j) is ε-regular and has red density at most d, but also receives at least $\frac{\alpha n|S|}{6m} > 2d|V_j||S|$ red edges from S. But this, since $\varepsilon < d$ and $|S| \geqslant \varepsilon|V_i|$, is a contradiction to regularity of (V_i, V_j).

By a similar argument, at most $\varepsilon|X_i'|$ vertices in X_i' send edges of the 'wrong' colour to each Y_j' or to Z' or vice versa. We can modify the argument slightly to show that at most $\varepsilon|X_1'|$ vertices of X_1' have more than αn red neighbours in X_1': again we can find a set S in a cluster V_i with $i \in B_1'$ whose vertices all have more than αn red neighbours in X_1', but we need to discard both red edges in irregular pairs at V_i and also edges within V_i. Since $|V_i| \leqslant \frac{m}{n} \leqslant \varepsilon n$, there are in total at most $2\varepsilon n$ such edges, which is the same bound we used above and from this point the proof above works as written.

We now let X_1 be obtained from X_1' by discarding all vertices which have more than αn edges of the 'wrong' colour to any of X_i' or Y_i' or Z'. By the above calculations, in total we discard at most $4\varepsilon|X_1'| + \frac{1}{200}\alpha n \leqslant \frac{1}{100}\alpha n$ vertices of X_1'. We define similarly X_2, Y_1, Y_2, Z, and similarly remove at most $\frac{1}{100}\alpha n$ vertices in each case. Finally, we let R denote the set of all vertices not in $X_1 \cup X_2 \cup Y_1 \cup Y_2 \cup Z$. By construction, each X_i, Y_i and Z has the claimed size; and $|R| \leqslant \alpha n$ follows since each vertex of R was either in V_0, or V_i for some $i \in T$, or removed from X_i' or Y_i' or Z'. There are at most $\varepsilon n + \frac{\alpha^2}{10000}n + 5 \cdot \frac{1}{100}\alpha n$ such.

Finally, by definition each vertex of X_1 has at most αn edges of the 'wrong' colour to any of X_i', Y_i' or Z', which are supersets of X_i, Y_i, Z respectively, giving the required bounds on 'wrong' coloured edges at X_1. The same holds for X_2, Y_1, Y_2, Z by the similar argument. $\qquad\square$

Finally, we prove Theorem 2. First, we deduce from Lemma 3 that if G satisfies the conditions of Theorem 2, then the reduced graph R of G is an m-vertex graph which contains a monochromatic TCTF on nearly $\frac{1}{3}m$ vertices. Suppose this is red. We then show how to construct a homomorphism from any given H satisfying the conditions of Theorem 2 to the red subgraph of R which does not overload any vertex i of R, i.e. map too many vertices to i, and finally apply Theorem 31 to find the desired monochromatic copy of H in G.

The only tricky step of this sketch is to construct the required homomorphism. We split up $V(H)$ into *chunks* and *fragments*, which are intervals in the bandwidth ordering, alternating between chunks and fragments. Each fragment is of equal length and their total size is tiny compared to the size of a cluster, and the chunks are of equal length and

much bigger than the fragments (but still much smaller than the size of a cluster). Given our TCTF in R, we put an order $T_1, \ldots, T_k$ (arbitrarily) on the triangles of the TCTF, and fix for each $1 \leqslant i \leqslant k - 1$ a tight walk from T_i to T_{i+1}. We assign each chunk of H to some T_i where i is chosen uniformly and independently from $[k]$. We claim that it is possible to now construct a homomorphism where each chunk will be mapped entirely to its assigned triangle, using the fragments to connect up along the fixed tight walks, and that this homomorphism will with positive probability not overload any vertex of R: the point here will be to analyse the assignment of chunks, since the total size of all fragments is tiny.

Proof of Theorem 2. Given $\gamma > 0$ and Δ, we fix $h \leqslant \frac{\gamma}{1000}$ and $\lambda > 0$ (which will play no further rôle in this proof) sufficiently small for Lemma 3, and let $2\varepsilon'$ and t_0 be the returned constants. We let $\varepsilon > 0$ be returned by Theorem 31 for input $d = \frac{1}{2}$, $\frac{\gamma}{100}$ and Δ. Without loss of generality, we may presume $\varepsilon < \frac{1}{10}\min(t_0^{-1}, \varepsilon', \gamma)$. We input ε and $d = \frac{1}{2}$ to Theorem 29 and let M, N_0 be the returned constants. We input $T = M$ to Theorem 31, with the other parameters as above, and choose N_1 such that the returned constant $m_0 \leqslant N_1/M$.

We set $\rho = \frac{1}{60000}M^{-3}\gamma^2$ and $\beta = \frac{1}{100}M^{-4}\gamma$.

Suppose now $n \geqslant \max(N_0, N_1)$.

Let $N = (9 + \gamma)n$. Given a 2-edge-coloured K_N, we apply Lemma 29 with constants as above, to the graph of red edges in K_N, to obtain a partition $V(K_N) = V_0 \cup V_1 \cup \cdots \cup V_m$, where $\varepsilon^{-1} \leqslant m \leqslant M$. By construction, the number of vertices in each part V_i with $1 \leqslant i \leqslant m$ is at least $\frac{(9+\gamma/2)n}{m}$.

Let R be the corresponding coloured reduced graph on $[m]$, in which we colour a pair ij red if (V_i, V_j) is ε-regular and has density in red at least $\frac{1}{2}$, blue if it is ε-regular and has density in blue strictly bigger than $\frac{1}{2}$, and otherwise (i.e. if the pair is irregular) we do not put an edge ij. By construction, we have $\delta(R) \geqslant (1 - \varepsilon)m$.

Let $t = m/(9 - \varepsilon')$, so that R has $(9 - \varepsilon')t$ vertices and minimum degree at least $(9 - 2\varepsilon')t$. By Lemma 3, either R contains a monochromatic $3(1 + \varepsilon')t$-vertex TCTF, or we obtain a partition of $V(R)$ as described in that lemma. In particular, there is a set B_1 of at least $(2 - h)t$ vertices and a disjoint set R_1 of at least $(2 - h)t$ vertices, such that any triangle with two vertices in B_1 and one in R_1 is monochromatic blue (and so all such triangles are in a blue triangle component). It follows that choosing $(1 - h)t$ disjoint such triangles greedily we obtain a monochromatic TCTF with $3(1 - h)t$ vertices. We see that in all cases R contains a monochromatic TCTF on at least $3(1 - h)t \geqslant \frac{1}{3}(1 - h)m =: 3k$ vertices.

Fix such a TCTF, let its triangles be $T_1, \ldots, T_k$ and suppose without loss of generality that it is red. By definition of red triangle-connectedness, for each $1 \leqslant i \leqslant k - 1$ there is a red triangle walk in R from T_i to T_{i+1}, and we fix for each i one such W_i chosen to be of minimum length. Thus W_i is a sequence of triangles, starting with T_i and ending with T_{i+1}, in which each pair of consecutive triangles shares two vertices. Finally, we assign labels $1, 2, 3$ to the vertices of all these triangles as follows: we label the vertices of T_1 in an arbitrary order, then assign labels to the successive triangles of $W_1, W_2, \ldots, W_{k-1}$ in order as follows: when we assign labels to the next triangle, we keep the labels of the

two vertices it shares with the previous triangle, and give the missing label to the third vertex. Note that a given vertex, or a given edge, might receive different labellings for different triangles, and indeed if a triangle appears in two different walks it might receive different labellings in the different walks.

Let H be a graph with maximum degree at most Δ, bandwidth at most βn, and a fixed 3-vertex colouring in which no colour class has more than n vertices. We split $V(H)$ into consecutive intervals $C_1, F_1, C_2, F_2, \ldots, F_{s-1}, C_s$ as follows: we let each C_i (except perhaps the last two, which can be of any size) consist of ρn vertices, and each F_i be of size $M^2 \beta n$. For each $1 \leqslant i \leqslant s$ we pick $\pi(i) \in [k]$ uniformly and independently at random. We now define a homomorphism $\psi : H \to R$ as follows. If x is a vertex of the chunk C_i for some i, and its colour in the fixed 3-colouring of H is $j \in [3]$, then we set $\psi(x)$ equal to the vertex of $T_{\pi(i)}$ with label j. We now describe how to construct ψ on the fragment F_1; the same procedure is used for each subsequent fragment with the obvious updates. We separate F_1 into intervals of length βn. If x is in the ith interval, and has colour j in the 3-colouring, then we set $\psi(x)$ equal to the vertex of the ith triangle after T_1 in W_1 with label j. If there is no such triangle (i.e. the walk has already reached T_2) then we set $\psi(x)$ equal to the vertex labelled j in T_2. We claim that this last event occurs for the final interval. Indeed, if two triangles of W_1 both contain a given edge e of R, then by minimality they are consecutive triangles in the walk, so W_1 has less than M^2 triangles.

We claim that this construction gives a homomorphism from H to the red subgraph of R. Indeed, suppose xy is an edge of H. Then x and y have different colours in the 3-colouring, and they are separated by at most βn in the bandwidth ordering. By construction, x is assigned to a vertex of some triangle T according to its colour. The vertex y is assigned to a triangle T' according to its colour; and either $T = T'$ or T and T' are consecutive triangles on one of the fixed walks, in particular they share two vertices and their labels are consistent on those two vertices. Either way, x and y are mapped to a red edge of R (the only non-edge is if $T \neq T'$ and it goes between the two vertices of the symmetric difference of T and T', which both have the same label: but x and y have different colours).

We still need to justify that with high probability ψ does not overload any vertex of R. To begin with, observe that the total number of vertices in the fragments is at most $M \cdot M^2 \beta n = M^3 \beta n \leqslant \frac{\gamma n}{100 m}$, which is much smaller than the size of any cluster V_i. In particular, if i is not in any triangle of the TCTF, then $|\psi^{-1}(i)| < \frac{1}{2}|V_i|$ as desired. Now consider the vertex u of T_i with label j. Apart from the at most $M^3 \beta n$ vertices of the fragments, the vertices of $\psi^{-1}(u)$ are vertices of chunks with colour j. There are at most n vertices in chunks of colour j in total, and each such chunk has probability $1/k$ of being assigned to T_i. We see that the expected number of chunk vertices in $\psi^{-1}(u)$ is at most n/k. The probability that the actual number of such vertices exceeds n/k by s is by Hoeffding's inequality at most

$$\exp\left(-\tfrac{s^2}{2 \cdot 3 \rho^{-1} \cdot (\rho n)^2}\right) = \exp\left(-\tfrac{s^2}{6 \rho n^2}\right),$$

where we used that there are at most $3\rho^{-1}$ chunks, and the maximum contribution of a given chunk to $|\psi^{-1}(u)|$ is at most ρn. Choosing $s = \frac{1}{100} \gamma M^{-1} n$, by choice of ρ the

probability that $|\psi^{-1}(u)| \geqslant n/k + s + M^3\beta n$ (the last term accounts for vertices in fragments) is at most $\exp(-M)$. In particular, with positive probability we have

$$|\psi^{-1}(u)| \leqslant \tfrac{n}{k} + \tfrac{1}{100}\gamma\tfrac{n}{M} + M^3\beta n$$

for every $u \in V(R)$. Suppose this event occurs. Substituting our values for β, k and finally h, we get

$$|\psi^{-1}(u)| \leqslant \tfrac{9n}{(1-h)m} + \tfrac{1}{100}\gamma\tfrac{n}{m} + \tfrac{1}{100}\gamma\tfrac{n}{m} \leqslant \tfrac{9n}{m}(1+2h) + \tfrac{1}{10}\gamma\tfrac{n}{m} \leqslant \frac{(9+\tfrac{\gamma}{5})n}{m}\,.$$

Since $|V_u| \geqslant \frac{(9+\gamma/2)n}{m}$, as observed at the start of this proof, we have $|\psi^{-1}(u)| \leqslant (1-\tfrac{\gamma}{100})|V_u|$ for every $u \in V(R)$. This is the required condition to apply Theorem 31.

Finally, by Theorem 31 we conclude that there is a red copy of H in the 2-coloured K_N. $\qquad\square$

9 Proof of Theorem 1

We are now ready to prove the main result of this paper, which we restate for convenience. Recall that we established the lower bound in Section 2, and what remains is to prove the corresponding upper bound. We give the full details for the square of a path, the square of a cycle case is similar.

Theorem 1. *There exists n_0 such that for all $n \geqslant n_0$ we have:*

$$R(P_{3n}^2, P_{3n}^2) = R(P_{3n+1}^2, P_{3n+1}^2) = R(C_{3n}^2, C_{3n}^2) = 9n-3 \text{ and } R(P_{3n+2}^2, P_{3n+2}^2) = 9n+1.$$

Proof of Theorem 1, upper bound for P_{3n+1}^2. Let n_0 and δ be given by Lemma 4 when we set $\alpha = \frac{1}{1000}$ (we are not trying to optimise this value) and then let us fix $n > \max(n_0, \frac{3}{\delta})$ and $N \geqslant 9n-3$. Let now G be any 2-edge-colouring of K_N. We suppose for a contradiction that G does not contain a monochromatic P_{3n+1}^2. By Lemma 4, since G does not contain a monochromatic P_{3n+1}^2, we have a partition X_1, X_2, Y_1, Y_2, Z, R of $V(G)$ with the conditions (a)-(g), which we fix.

We now want to refine these conditions by adapting repeatedly a greedy procedure. Since we are going to apply multiple times the same method, we will explain the greedy procedure and the arguments for existence only in the first instance.

Claim 34. X_1 and X_2 are entirely blue, while Y_1 and Y_2 are entirely red.

Proof. Assume by contradiction that there is a red edge $x_1 x_1'$ in X_1. Since x_1 and x_1' have at most αn blue neighbours in Z and $(1-\alpha)n \leqslant |Z|$, we have that x_1 and x_1' have a common red neighbour $z \in Z$. Similarly, by considering the common red neighbour of x_1 and x_1' in Y_2, we can find $y_2, y_2' \in Y_2$ such that $y_2 y_2', x_1 y_2, x_1' y_2', x_1' y_2, x_1 y_2'$ are all red.

We are now ready to extend the red path $P_0 = y_2, y_2', x_1, x_1', z$ (whose square is also monochromatic red) to a path P of length bigger than $3n + 2$ such that P^2 is also

monochromatic red. The idea is to greedily add to P_0 at least $\frac{3}{2}n$ vertices from Y_2 (using the fact that almost all the edges in Y_2 are red) and $2n$ vertices from (X_1, X_2, Z).

In order to do that, it suffices to show that we can find a path P_{Y_2} of length at least $\frac{3}{2}n$ in Y_2 that starts with $y_2 y_2'$ and such that $P_{Y_2}^2$ is monochromatic red. Assume we have built already a path $P_{Y_2} = y_2, y_2', \ldots, p_\ell$ with the aforementioned conditions, provided $\ell < \frac{3}{2}n$, we can extend P_{Y_2} simply by appending a vertex $p_{\ell+1}$ that is in the common red neighbour of p_ℓ and $p_{\ell-1}$ in $Y_2 \setminus P_{Y_2}$. But this is possible, indeed all but at most $\frac{2}{1000}n$ vertices in Y_2 have red edges to both p_ℓ and $p_{\ell-1}$.

We do a very similar procedure by greedily extending $P_{(X_1, X_2, Z)}$. Given a red path $P_{(X_1, X_2, Z)} = x_1', z, \ldots, p_\ell$ of length smaller than $2n$, we can extend it by taking a vertex in the common red neighbour of p_ℓ and $p_{\ell-1}$ and in the right component.

Since P_0^2 is monochromatic red, and since we showed how to extend the endpoints to form a long path whose square is also monochromatic, we are done.

The arguments for X_2, Y_1 and Y_2 are symmetric. $\qquad\square$

Claim 35. *The pairs (X_1, Z) and (X_2, Z) are entirely red, while the pairs (Y_1, Z) and (Y_2, Z) are entirely blue.*

Proof. Assume by contradiction that there is a blue edge $x_1 z$ between X_1 and Z. Let $y_1 \in Y_1$ be in the common blue neighbourhood of x_1 and z (which exists by arguments similar to the ones above).

Take $x_1' \in X_1 \setminus \{x_1\}$ in the common blue neighbourhood of y_1 and x_1 and let $P_0 = z, y_1, x_1, x_1'$. We have that P_0^2 is blue monochromatic. Also, we can greedily extend P_0 to a path P such that P^2 is also blue monochromatic and $|P| > 3n$ by extending x_1, x_1' to a path of length at least $\frac{3}{2}n$ in X_1 and extending the zy_1 end in (Y_1, Y_2, Z) by at least $2n$ vertices.

The argument for the other pairs is symmetric. $\qquad\square$

Claim 36. *The pairs (X_1, Y_1) and (X_2, Y_2) are entirely blue, while the pairs (X_1, Y_2) and (X_2, Y_1) are entirely red.*

Proof. Assume by contradiction that there is a red edge $x_1 y_1$ in (X_1, Y_1). The vertices x_1 and y_1 share a red neighbour x_2 in X_2. We can also find in $Y_1 \setminus \{y_1\}$ a common red neighbour y_1' of y_1 and x_2.

We can start with the path $P_0 = x_1, x_2, y_1, y_1'$, and then extend it using vertices in Y_1 on one side and vertices of (X_1, X_2, Z) on the other, until we get a path P such that P^2 is monochromatic red and covers at least $3n + 2$ vertices.

The argument for the other pairs is symmetric. $\qquad\square$

Claim 37. *The pair (X_1, X_2) has no blue P_4, while the pair (Y_1, Y_2) has no red P_4.*

Proof. Assume $x_1 x_2 x_1' x_2'$ formed a blue P_4 in (X_1, X_2). Since X_1 and X_2 are entirely blue, the edges $x_1 x_1'$ and $x_2 x_2'$ are blue. Each of these edges is the beginning of a square of a path covering the respective part. These join together to form a square of a path that is longer than allowed.

The argument for the other pair is symmetric. $\qquad\square$

From the claims above we can see that in the situation depicted by Lemma 4 we have $|X_1|, |X_2|, |Y_1|, |Y_2| \leqslant 2n - 1$. We can now partition the vertices of the remainder set R depending on their neighbourhoods as follows.

1) Let us denote with R_Z the set of vertices in R with more than $\frac{n}{4}$ red neighbours both in X_1 and X_2,

2) for $i = 1, 2$ let R_i be the vertices in R with more than $\frac{n}{4}$ blue neighbours in both X_i and Y_i,

3) let R_{12} denote the vertices in R with more than $\frac{n}{4}$ red neighbours in both X_1 and Y_2,

4) let R_{21} denote the vertices in R with more than $\frac{n}{4}$ red neighbours in both X_2 and Y_1,

5) let R^* denote any vertices in R that are not in any of the above sets.

Claim 38. *Vertices in R^* have at least $\frac{3}{2}n$ blue neighbours in each X_i and at least $\frac{3}{2}n$ red neighbours in each Y_i. Moreover, $|R^*| \leqslant 1$.*

Proof. The first part of the claim is true by construction. Let us now assume that there are two vertices u and v in R^*. Then u and v have more than $\frac{n}{2}$ common blue neighbours in each X_i and at least $\frac{n}{2}$ common red neighbours in each Y_i. Therefore if uv is blue it will create a blue square of a path with vertices from X_1 and X_2, while if it is red it will join long red squares of paths in Y_1 and Y_2. □

Claim 39. *We have the following bounds: $|X_1 \cup R_1|, |X_2 \cup R_2|, |Y_1 \cup R_{21}|, |Y_2 \cup R_{12}| \leqslant 2n - 1$.*

Proof. Assume by contradiction that $|X_1 \cup R_1| \geqslant 2n$. Recall that in previous claims we proved that all the edges in X_1 and (X_1, Y_1) are blue. Let us label the vertices in R_1 by $r_1, \ldots, r_\ell$. Recall that $\ell = |R_1| \leqslant |R| \leqslant \alpha n$.

Since every vertex in R_1 has at least $\frac{n}{4}$ blue neighbours in both X_1 and Y_1 we can find disjoint blue triangles $T_1, \ldots, T_\ell$ where triangle T_i contains the vertices r_i, x_i, y_i with $x_i \in X_1$ and $y_i \in Y_1$. We next find for each $i \in [\ell]$ vertices $a_i, b_i, c_i, a_i', b_i', c_i'$ as follows. We let c_i be a blue neighbour of r_i in Y_1, and $a_i, b_i \in X_1$, we let a_1' be a neighbour in X_1 of r_1, b_1' be in X_1, and c_1' be in Y_1. Observe that since $\ell \leqslant \alpha n$, we can ensure that all these vertices are different.

By construction, the vertex ordering $P_0 = (a_1, b_1, c_1, x_1, r_1, y_1, a_1', b_1', c_1', \ldots)$, where we repeat the same letter ordering for $i = 2$ and so on afterwards, is a blue square path. We extend P_0 further by choosing distinct vertices from X_1, X_1 and then Y_1 in this order, until no unused vertices remain in X_1. As $|X_1 \cup R_1| \geqslant 2n$, what we obtain is a blue square path with at least $3n$ vertices, if $|X_1 \cup R_1| \geqslant 2n + 1$ we obtain at least $3n + 1$ vertices. We can extend P_0 by one more vertex by adding a so far unused vertex of Y_1 at the start of the ordering. This gives the required $3n + 1$-vertex square path (and $3n + 2$ vertices if $|X_1 \cup R_1| \geqslant 2n + 1$). The arguments for the other pairs of sets are the same. □

Claim 40. *We have that $|Z \cup R_Z| \leqslant n - 1$.*

Proof. Let as assume that $|Z \cup R_Z| \geqslant n$ and let us label the vertices in R_Z by $r_1, \ldots, r_\ell$. Since (X_1, X_2) has no blue path on 4 vertices, there are at most 40 vertices in $X_1 \cup X_2$ with more than $\frac{n}{20}$ blue neighbours in the opposite part. Call the set of these vertices X_{bad}. Since each vertex in R_Z has more than $\frac{n}{4}$ red neighbours in each X_i, we can find disjoint red triangles $T_1, \ldots, T_\ell$ such that each T_i uses r_i, a vertex $x_i^1 \in X_1 \setminus X_{\text{bad}}$ and a vertex $x_i^2 \in X_2 \setminus X_{\text{bad}}$.

The idea is now to find for each $i \in [\ell]$ vertices $a_i, a_i' \in X_1$, $b_i, b_i' \in Z$, $c_i, c_i' \in X_2$ such that for every $i \in [\ell - 1]$ we have that $(x_i^1, r_i, x_i^2, a_i, b_i, c_i, a_i', b_i', c_i', x_{i+1}^1, r_{i+1}, x_{i+1}^2)$ is a red square of a path. But this can be done greedily since $\ell \leqslant \alpha n$. We now build the path $P_0 = (x_1^1, r_1, x_1^2, a_1, b_1, c_1, a_1', b_1', c_1', x_2^1, \ldots, x_\ell^2)$ which by construction has the property that P_0^2 is red.

We can extend P_0 by choosing distinct vertices from X_1, Z and then X_2 in this order, until no unused vertices remain in Z. As $|Z \cup R_Z| \geqslant n$, what we obtain is a red square path with at least $3n$ vertices. □

Putting the bounds from the last three claims together, we see $|G| \leqslant 1 + 4(2n - 1) + n - 1 = 9n - 4$, which contradiction completes the proof. □

The proof for P_{3n+2}^2 is almost verbatim as above (we actually worked with P_{3n+2}^2 in most of the claims), with the exception that in Claim 39 we obtain the upper bound $|X_1 \cup R_1| \leqslant 2n$, as explained in the proof of that claim, and consequently a final upper bound $|G| \leqslant 1 + 4(2n) + n - 1 = 9n$ for a contradiction.

Sketch proof of cycle case of Theorem 1. In order to prove that for n large enough we have $R(C_{3n}^2, C_{3n}^2) = 9n - 3$, is suffices to modify our previous proof. We start by constructing the same partition we built at the beginning of the proof of Theorem 1 to get the sets X_1, X_2, Y_1, Y_2, Z, R. Now, by using the same technique introduced in Claim 34 we can prove some weakened for of Claims 34, 35, 36, 37. Which is, we can prove that in X_1 we cannot find two disjoint red edges (the same holds for X_2), in Y_1 we cannot find two disjoint blue edges (the same holds for Y_2). Similarly, we cannot find two disjoint edges of the wrong colours in none of the following pairs: (X_1, Z), (X_2, Z), (Y_1, Z), (Y_2, Z), (X_1, Y_1), $(X_2, Y_2), (X_1, Y_2), (X_2, Y_1)$. Moreover, we cannot find two vertex-disjoint P_4 of the wrong colour in (X_1, X_2) nor in (Y_1, Y_2).

To From these results and the previously proved Lemma 4, we can see that also in this case we have $|X_1|, |X_2|, |Y_1|, |Y_2| \leqslant 2n - 1$. Now we can define the same partition of R in sets $R_Z, R_1, R_2, R_{12}, R_{21}$ and R^*. Let us point out that from this modified version of Claim 34 we have that there are two vertices $a, b \in X_1$ such that all edges in $G[X_1 \setminus \{a, b\}]$ are blue. In particular, from Claims 34, 35, 36, 37 we get that up to moving at most 10 vertices from X_1 to R_1 (and similarly from X_2 to R_2, from Y_1 to R_{21}, from Y_2 to R_{12} and from Z to R_Z) all the vertices in X_1 (and similarly in X_2, Y_1, Y_2, Z) are incident only to edges of the right colour in $G[X_1 \cup X_2 \cup Y_1 \cup Y_2 \cup Z]$, with the possible exception of edges in (X_1, X_2) and (Y_1, Y_2).

We now aim to explain how to modify Claim 38 to hold for cycles and how to modify

the proof of Claims 39, 40. The first part of Claim 38 holds by construction without any modifications. The second part of Claim 38 needs to be modified to state that we cannot find two parallel edges of the same colour in R^*. Indeed, otherwise we could find a long monochromatic blue cycle C such that C^2 is also blue by using vertices from X_1, X_2 and the two blue edges in R^*. In particular, this implies that $|R^*| \leqslant 4$. As a guide to show how to modify the proofs of Claim 39 and 40, we give a sketch of the modifications needed for Claim 39. If we assume by contradiction that $|X_1 \cup R_1| \geqslant 2n$ we can almost verbatim repeat the same proof, having care of extending our path P_0 in both directions and making sure that the two endpoints of P_0 and their neighbours are adjacent in blue to each other. This is possible because $G[X_1]$ is entirely blue as claimed above.

Claim 41. *If R^* contains a blue edge, then $|X_1 \cup R_1|, |X_2 \cup R_2| \leqslant 2n - 2$ (same holds for red, $Y_1 \cup R_{21}$ and $Y_2 \cup R_{12}$).*

Proof. Assume R^* contains a blue edge uv, then $|X_1 \cup R_1| \leqslant 2n - 2$ (the arguments for the other cases are the same). In order to prove this, it suffices to show that there exists a maximal matching T in X_1 such that we can build a blue cycle C that covers all the edges of X_1, the two vertices $u, v \in R^*$ and, for each edge in T, an extra vertex in Y_1. This can be done because by Claim 34 and Lemma 4 there is a vertex $w \in X_1$ such that the red neighbourhood of w in X_1 has size at most αn, but $G[X_1 \setminus w]$ has at most one red edge and because u and v have both at least $\frac{3}{2}n$ blue neighbours in X_1. Therefore, it is possible to build a cycle by replicating the construction in Claim 39 and by carefully adding the edge uv to the cycle. $\square$

This suffices to conclude. Indeed, if $|R^*| \leqslant 3$ then we still have

$$|X_1 \cup R_1 \cup R^* \cup X_2 \cup R_2| \leqslant 4n - 1,$$

while if $|R^*| = 4$ then we have both a red and a blue edge in R^* (since we cannot have two vertex-disjoint edges of the same colour). In this case we have the following inequalities: $|X_1 \cup R_1|, |X_2 \cup R_2|, |Y_1 \cup R_{21}|, |Y_2 \cup R_{12}| \leqslant 2n - 2$, which are enough to obtain the wanted bound. $\square$

References

[1] P. Allen, J. Böttcher, H. Hàn, Y. Kohayakawa, and Y. Person, *Blow-up lemmas for sparse graphs*, arXiv:1612.00622, 2016.

[2] P. Allen, J. Böttcher, and J. Hladký, *Filling the gap between Turán's theorem and Pósa's conjecture*, J. Lond. Math. Soc. (2) **84** (2011), no. 2, 269–302.

[3] P. Allen, G. Brightwell, and J. Skokan, *Ramsey-goodness—and otherwise*, Combinatorica **33** (2013), no. 2, 125–160.

[4] J. A. Bondy and P. Erdős, *Ramsey numbers for cycles in graphs*, J. Combinatorial Theory Ser. B **14** (1973), 46–54.

[5] S. A. Burr, *Ramsey numbers involving graphs with long suspended paths*, J. London Math. Soc. (2) **24** (1981), no. 3, 405–413.

[6] K. Corradi and A. Hajnal, *On the maximal number of independent circuits in a graph*, Acta Math. Acad. Sci. Hungar. **14** (1963), 423–439.

[7] L. Gerencsér and A. Gyárfás, *On Ramsey-type problems*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **10** (1967), 167–170.

[8] R. L. Graham, V. Rödl, and A. Ruciński, *On graphs with linear Ramsey numbers*, J. Graph Theory **35** (2000), no. 3, 176–192.

[9] J. Komlós, G. N. Sárközy, and E. Szemerédi, *Blow-up lemma*, Combinatorica **17** (1997), no. 1, 109–123.

[10] V. Nikiforov and C. C. Rousseau, *Ramsey goodness and beyond*, Combinatorica **29** (2009), no. 2, 227–262.

[11] V. Rosta, *On a Ramsey-type problem of J. A. Bondy and P. Erdős. I, II*, J. Combinatorial Theory Ser. B **15** (1973), 94–104; ibid. 15 (1973), 105–120.

[12] E. 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.