The Turán density of tight cycles in three-uniform hypergraphs

Nina Kamčev* Shoham Letzter† Alexey Pokrovskiy‡

Abstract

The *Turán density* of an r-uniform hypergraph $\mathcal{H}$, denoted $\pi(\mathcal{H})$, is the limit of the maximum density of an n-vertex r-uniform hypergraph not containing a copy of $\mathcal{H}$, as $n \to \infty$.

Denote by $\mathcal{C}_\ell$ the 3-uniform tight cycle on ℓ vertices. Mubayi and Rödl gave an "iterated blow-up" construction showing that the Turán density of $\mathcal{C}_5$ is at least $2\sqrt{3} - 3 \approx 0.464$, and this bound is conjectured to be tight. Their construction also does not contain $\mathcal{C}_\ell$ for larger ℓ not divisible by 3, which suggests that it might be the extremal construction for these hypergraphs as well. Here, we determine the Turán density of $\mathcal{C}_\ell$ for all large ℓ not divisible by 3, showing that indeed $\pi(\mathcal{C}_\ell) = 2\sqrt{3} - 3$. To our knowledge, this is the first example of a Turán density being determined where the extremal construction is an iterated blow-up construction.

A key component in our proof, which may be of independent interest, is a 3-uniform analogue of the statement "a graph is bipartite if and only if it does not contain an odd cycle".

1 Introduction

For an r-uniform hypergraph $\mathcal{H}$, the *Turán number* of $\mathcal{H}$, denoted $\text{ex}(n, \mathcal{H})$, is defined as the maximum number of edges an n-vertex r-uniform hypergraph can have without containing a copy of $\mathcal{H}$ as a subgraph. For (2-uniform) graphs, we have a fairly good understanding of Turán numbers. The first theorem proved about them is Mantel's theorem [25], which says that, for the triangle, we have $\text{ex}(n, K_3) = \lfloor n^2/4 \rfloor$. This was generalised by Turán [34] who showed that $\text{ex}(n, K_r) \approx (1 - \frac{1}{r-1})\binom{n}{2}$. For non-complete graphs we know less, and usually only know what the Turán number of a graph is asymptotically, up to $o(n^2)$ terms. Because of this, we study the *Turán density* of an r-uniform hypergraph $\mathcal{H}$, denoted $\pi(\mathcal{H})$, and defined as $\pi(\mathcal{H}) = \lim_{n\to\infty} \frac{\text{ex}(n,\mathcal{H})}{\binom{n}{r}}$. This limit is known to exist, and, moreover, it is clear that $\pi(\mathcal{H}) \in [0, 1]$ for every H. The Turán densities

*Department of Mathematics, Faculty of Science, University of Zagreb, Croatia. Email: `nina.kamcev@math.hr`.
Research supported by the European Union's Horizon 2020 research and innovation programme [MSCA GA No 101038085].

†Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK. Email: `s.letzter@ucl.ac.uk`. Research supported by the Royal Society.

‡Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK. Email: `dralexeypokrovskiy@gmail.com`.

of (2-uniform) graphs were completely determined by Erdős and Stone [9], who showed that every graph H satisfies $\pi(H) = 1 - \frac{1}{\chi(H)-1}$.

The special case of the Erdős–Stone theorem for bipartite graphs can be generalised to higher uniformities, as follows (see [8]): every r-partite r-uniform hypergraph $\mathcal{H}$ satisfies $\pi(\mathcal{H}) = 0$. (An r-uniform hypergraph $\mathcal{H}$ is said to be *r-partite* if its vertices can be r-coloured so that every edge has one vertex of each colour.) Nevertheless, in general, our understanding of Turán numbers in higher uniformities is very limited, and there are only a small number of hypergraphs whose Turán densities are known; see Keevash [19] for a comprehensive survey of the topic listing a number of such hypergraphs. A notable, relatively early example is a result of de Caen and Füredi [7] showing that the Turán density of the Fano plane is 3/4 (see also [14, 20]). More recently, the impactful computer-assisted "flag-algebra" technique has been used to obtain a number of sharpest known upper bounds on Turán densities (see [1, 19, 31] and the references therein).

Given the sporadicity of hypergraphs whose Turán densities are known, it is unsurprising that there are many conjectures about Turán densities of specific hypergraphs. The most famous of these is Turán's conjecture [35], that the Turán density of the *tetrahedron* $K_4^{(3)}$ is 5/9. Frankl and Füredi [13] conjectured that the Turán density of the 3-edge subgraph of $K_4^{(3)}$ (usually denoted K_4^-) is 2/7. A particularly relevant conjecture for us concerns tight cycles. The *r-uniform tight cycle* of length ℓ, denoted $\mathcal{C}_\ell^r$, is defined to be the hypergraph with vertex set $\{1, \ldots, \ell\}$ and hyperedges all sets of the form $\{x, x+1, \ldots, x+r-1 \pmod{\ell}\}$. The following conjecture, usually attributed to Mubayi and Rödl, appears for instance in [11, 26].

Conjecture 1.1. *$\pi(\mathcal{C}_5^3) = 2\sqrt{3} - 3$.*

The lower bound $\pi(\mathcal{C}_5^3) \geq 2\sqrt{3} - 3 \approx 0.464$ was found by Mubayi and Rödl (see Example 1.2 below for a description of their example) and the best upper bound is due to Razborov [29], who showed $\pi(\mathcal{C}_5^3) \leq 0.468$.

One basic reason why hypergraphs are more difficult than graphs is that the extremal $\mathcal{H}$-free hypergraphs can be much more complicated than the extremal graphs. In the 2-uniform case, the Erdős–Stone theorem shows that all optimal graphs are close to being complete multipartite. For higher-uniformity hypergraphs there have been numerous papers discovering more complicated possible extremal hypergraphs, for instance [3, 4, 13], as well as Conjectures 1.1 and 7.1. For some hypergraphs, such as $K_4^{(3)}$, the conjectured extremal constructions are even non-unique and very different from each other [4, 12, 21, 30].

One class of extremal examples, which does not occur for graphs, is an "iterated blow-up construction". The conjectured extremal example for Conjecture 1.1 is an instance of such a construction.

Example 1.2 (Iterated blow-up construction with no copies of $\mathcal{C}_5^3$)**.** *Consider nested vertex sets $V_1 \supseteq \ldots \supseteq V_t$ with $|V_i| - |V_{i+1}| = x_i$ for $i \in [t]$, with the convention $|V_{t+1}| = \emptyset$. Let $\mathcal{H}(x_1, \ldots, x_t)$*

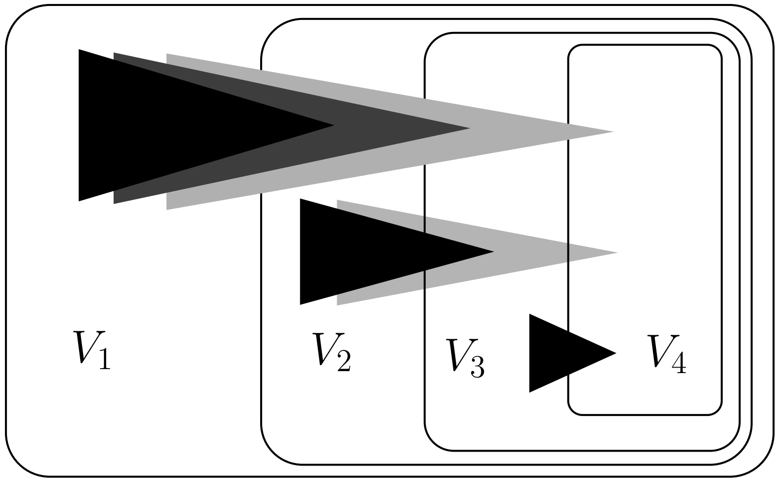

Figure 1: An illustration of the hypergraph $\mathcal{H}(x_1, x_2, x_3, x_4)$.

be a 3-uniform hypergraph on the vertex set V_1, where xyz is an edge whenever $x, y \in V_i \setminus V_{i+1}$ and $z \in V_{i+1}$ for some i (see Figure 1).

We claim that there is no copy of $\mathcal{C}_5^3$. To see this, say that an edge with two vertices in $V_i \setminus V_{i+1}$ and one vertex in V_{i+1} has type i, *and observe that if two edges e and f intersect in two vertices, they are of the same type. Thus, if $C = (u_1 \dots u_5)$ is a cycle, then its edges all have the same type, say i. Without loss of generality, $u_1, u_2 \in V_i \setminus V_{i+1}$ and $u_3 \in V_{i+1}$. It follows that $u_4 \in V_i \setminus V_{i+1}$, and thus $u_5 \in V_i \setminus V_{i+1}$. But then $u_4u_5u_1$ is not an edge of $\mathcal{H}(x_1, \dots, x_t)$, a contradiction.*

Thus $\pi(\mathcal{C}_5^3) \geq e(\mathcal{H}(x_1, \dots, x_t))/\binom{n}{3}$ for all choices of $x_1, \dots, x_t$ with $x_1 + \dots + x_t = n$. Let $f(n)$ denote the maximum number of edges that such a hypergraph on n vertices can have i.e $f(n) :=$ $\max(e(\mathcal{H}(x_1, \dots, x_t) : x_1, \dots, x_t \geq 1,\ x_1 + \dots + x_t = n)$. It is possible to show $\lim_{n\to\infty} f(n)/\binom{n}{3} =$ $2\sqrt{3} - 3$ (see Section 4 for details), which gives $\pi(\mathcal{C}_5^3) \geq 2\sqrt{3} - 3$.

Let $\mathcal{G}_n = \mathcal{H}(x_1, \dots, x_t)$ for a choice of $x_1 \geq \dots \geq x_t$ such that $n = x_1 + \dots + x_t$ and $e(\mathcal{G}_n) = f(n)$.

Note that in the above construction, $\mathcal{G}_n$ has no tight cycles of lengths $\ell \equiv 1$ or $2 \pmod 3$ either. So it is plausible that Conjecture 1.1 could be strengthened to say that $\pi(\mathcal{C}_\ell^3) = 2\sqrt{3} - 3$ for all $\ell \geq 5$ with $\ell \equiv 1$ or $2 \pmod 3$ (notice that $\mathcal{C}_4^3 = K_4^{(3)}$, and there are known examples of $K_4^{(3)}$-free 3-uniform graphs with density at least $5/9 > 2\sqrt{3} - 3$). The main result of our paper is to show that this is true for sufficiently large ℓ.

Theorem 1.3. *Let ℓ be sufficiently large with $\ell \equiv 1$ or $2 \pmod 3$. Then $\pi(\mathcal{C}_\ell^3) = 2\sqrt{3} - 3$.*

To our knowledge, this is the first example of a Turán density being determined where the extremal construction is an iterated blow-up construction, and could be a step towards Conjecture 1.1. This is also one of the few examples of hypergraphs with irrational Turán densities. Such hypergraphs were recently found by Yan and Peng [37], as well as Wu [36], motivated by the work of Chung and

Graham [6], Baber and Talbot [1], and Pikhurko [28]. We remark that Conjecture 1.1 would imply Theorem 1.3 via Theorem 2.1 below (using the same argument as in the proof of Theorem 1.3 in Section 6).

One of our main tools, which may be of independent interest, is a 3-uniform analogue of the statement "a graph is bipartite if and only if it does not contain an odd cycle"; see Theorem 2.4. Thus, we characterise 3-uniform hypergraphs $\mathcal{H}$ which do not contain *homomorphic images* of cycles $\mathcal{C}_\ell^3$ with $3 \nmid \ell$, in terms of certain colourings of $V(\mathcal{H})^2$, as explained in the proof overview.

Throughout the paper we will informally refer to 3-uniform cycles of length $\ell \equiv 1$ or $2 \pmod 3$ as *odd cycles*, and we will often refer to 3-uniform hypergraphs as 3-graphs.

Related results

As we mentioned, there are very few hypergraphs with a known Turán density, but let us state some recent results on Turán-type problems for tight cycles. A well-studied hypergraph parameter is the so-called *uniform Turán density*, the infimum over all d for which any sufficiently large hypergraph with the property that all its linear-size subhypergraphs have density at least d contains $\mathcal{H}$. This line of research was initiated by Erdős and Sós [10] and, parallel to the classical Turán densities, the motivating questions in the area are determining the uniform Turán densities of the tetrahedron $K_4^{(3)}$ and its 3-edge subgraph K_4^-. The latter was found to be 1/4 by Glebov, Kráľ, and Volec [15] and later by Reiher, Rödl, and Schacht [32] with a different proof. In 2022, Bucić, Cooper, Kráľ, Mohr, and Munha Correia showed that for $\ell \geq 5$ and not divisible by 3, the uniform Turán density of $\mathcal{C}_\ell^3$ is $\frac{4}{27}$ [5].

Another question that has attracted a lot of interest in the last few years is, what is the extremal number of tight cycles (the maximum number of edges in an n-vertex r-uniform hypergraph containing no tight cycles)? For $r = 2$, the answer is of course $n-1$, but it turns out that the behaviour is rather different for $r \geq 3$. More specifically, after a series of results [17, 18, 22, 33], we know that the extremal number of tight r-uniform cycles lies between $\Omega\left(n^{r-1}\log n/\log\log n\right)$ and $O\left(n^{r-1}\log^5 n\right)$.

2 Proof overview

For an r-uniform hypergraph $\mathcal{H}$, the *t-blow-up* of $\mathcal{H}$, denoted $\mathcal{H}[t]$, is defined to be the r-uniform hypergraph with vertex set $V(\mathcal{H}) \times [t]$ and edges all r-tuples $\{(x_1, i_1), \ldots, (x_r, i_r)\}$ with $\{x_1, \ldots, x_r\} \in E(\mathcal{H})$. The starting point of our proof is the following theorem, which asserts that the blow-up of a hypergraph $\mathcal{H}$ has the same Turán density as $\mathcal{H}$.

Theorem 2.1 ([19], Theorem 2.2)**.** *Let t be an integer and let $\mathcal{H}$ be an r-uniform hypergraph. Then $\pi(\mathcal{H}[t]) = \pi(\mathcal{H})$.*

It shows that, rather than focusing on the Turán density $\pi(\mathcal{C}_k^3)$ for an odd cycle C_k^3, we can instead work out the Turán density of $\pi(\mathcal{H})$ for any hypergraph $\mathcal{H}$ whose blow-up $\mathcal{H}[t]$ contains $\mathcal{C}_k^3$ for some

t. We refer to such hypergraphs $\mathcal{H}$ as pseudocycles, and they can be equivalently defined as follows.

Definition 2.2. A *pseudocycle* of length ℓ in a 3-uniform hypergraph $\mathcal{H}$ is a sequence of (not necessarily distinct) vertices $v_1, \ldots, v_\ell$, such that for each $i \in [\ell]$, we have that $\{v_i, v_{i+1 \pmod{\ell}}, v_{i+2 \pmod{\ell}}\}$ is an edge of $\mathcal{H}$. A *pseudopath* of order ℓ is defined analogously.

It is easy to show that for a hypergraph $\mathcal{H}$, the properties "$\mathcal{H}[t]$ contains a $\mathcal{C}_k^3$ for some t" and "$\mathcal{H}$ contains a length k pseudocycle" are equivalent.

Thus, the starting point of our approach is, what is the maximum number of edges that a 3-uniform hypergraph can have without containing an odd pseudocycle? Later (after Corollary 2.6), we will discuss how to forbid only *short* pseudocycles. To understand our approach to this question, consider the analogous question about graphs — what is the maximum number of edges in a (2-uniform) graph with no odd circuits? By Kotzig's Lemma, a graph has no odd circuit if, and only if, it is bipartite. Thus, the maximum number of edges in an n-vertex bipartite graph is $\left\lfloor \frac{n^2}{4} \right\rfloor$.

Our approach to the 3-uniform case is analogous to this. We first find the relevant generalisation of bipartite graphs, and then maximise the number of edges over this class of graphs. To define this generalisation, recall that a graph is bipartite if, and only if, it has a proper 2-vertex-colouring. In our context, we will be colouring the shadow of a 3-uniform hypergraph. The *shadow* of a hypergraph $\mathcal{H}$, denoted $\partial\mathcal{H}$, is the graph on vertices $V(\mathcal{H})$ whose edges are pairs xy that are contained in an edge in $\mathcal{H}$.

Definition 2.3. A *good colouring* of a 3-uniform hypergraph $\mathcal{H}$ is a colouring of its shadow, such that each edge xy in the shadow is either coloured blue or coloured red and given an orientation, and every edge e in $\mathcal{H}$ can be written as xyz where xy and xz are red and directed from x and yz is blue.

The key first step of our proof is to show that the notion of "good colouring" is exactly equivalent to $\mathcal{H}$ not containing an odd pseudocycle.

Theorem 2.4. *A 3-uniform hypergraph $\mathcal{H}$ has a good colouring if, and only if, $\mathcal{H}$ has no pseudocycle of length ℓ with $3 \nmid \ell$.*

This theorem is proved in Section 3. Having established the above theorem, we next wish to maximise the number of edges in a hypergraph with a good colouring. To this end, we define a *coloured graph* to be a complete graph whose edges are either coloured blue or coloured red and oriented. A *cherry* in a coloured graph G is a triple xyz such that xy and xz are red and directed from x and yz is blue. Denote by $c(G)$ the number of cherries in G. Notice that if we have a good colouring of the shadow of $\mathcal{H}$ (and the remaining vertex pairs can be coloured arbitrarily), then all edges of $\mathcal{H}$ will be cherries in the resulting coloured graph. Thus, let $m_{\text{cherry}}(n)$ be the maximum number of cherries in an n-vertex coloured graph. The quantity $m_{\text{cherry}}(n)$ has been studied before by Falgas-Ravry and Vaughan [11], who used flag algebras to show that $\lim_{n\to\infty} m_{\text{cherry}}/\binom{n}{3} = 2\sqrt{3}-3$. Huang

[16] worked on the area further and determined the maximum number of "induced out-stars" of size t in an n-vertex coloured graph. We remark that the afore-mentioned authors used an equivalent reformulation – they studied the maximum number of "induced out-stars" in an uncoloured directed graph, and there is a clear correspondence between a coloured graph (in our sense) and a directed graph. The following special case of their results is relevant for us.

Theorem 2.5 (Falgas-Ravry–Vaughan [11]; Huang [16])**.** *Every coloured graph on n vertices contains at most $f(n)$ cherries.*

Combining Theorem 2.4 and Theorem 2.5 already yields the following weakening of Theorem 1.3.

Corollary 2.6. *If $\mathcal{H}$ is a 3-uniform hypergraph on n vertices which does not contain a pseudocycle of length ℓ for any ℓ with $3 \nmid \ell$, then $e(\mathcal{H}) \leq f(n)$.*

Notice that in Theorem 1.3 we forbid odd pseudocycles of a single length, whereas in Corollary 2.6 odd pseudocycles of *all* lengths are forbidden. Thus, the next goal is to prove a version of the above corollary which holds when forbidding *short* odd pseudocycles, yielding a finite family of forbidden hypergraphs. This is done by controlling the diameter of the hypergraph $\mathcal{H}$.

Definition 2.7. The *diameter* of a hypergraph $\mathcal{H}$ is the minimum ℓ such that the following holds: for every $x, y, z, w \in V(\mathcal{H})$ (where x, y are distinct and z, w are distinct) whenever there is a pseudopath from xy to zw, there is such a pseudopath of order at most ℓ.

In Section 6, we show that, for $\ell \gg \epsilon^{-1}$, every 3-uniform hypergraph $\mathcal{H}$ contains a subhypergraph $\mathcal{H}'$ with $e(\mathcal{H}') \geq e(\mathcal{H}) - \epsilon n^3$ such that $\mathcal{H}'$ has diameter at most ℓ (see Proposition 6.4). Then we show that in every 3-uniform hypergraph of diameter ℓ, if there is some odd pseudocycle, then there is also an odd pseudocycle of length at most 4ℓ (see Proposition 6.3). Combining these with Corollary 2.6 shows the following

Corollary 2.8. *Let $1/n \ll 1/\ell \ll \epsilon \ll 1$, and let $\mathcal{H}$ be an n-vertex hypergraph with no odd pseudocycles of length at most ℓ. Then $e(\mathcal{H}) \leq f(n) + \epsilon n^3$.*

Note that this is still not strong enough to combine with Theorem 2.1 to yield Theorem 1.3. The issue is that the length of the cycle ℓ depends on ϵ — therefore, when combined with Theorem 2.1, we would only get that $\lim_{m\to\infty} \pi(\mathcal{C}_m^3) = 2\sqrt{3} - 3$. To go further, we prove a "stability version" of Theorem 2.5. We show that if a coloured graph D on n vertices contains more than $f(n) - \epsilon n^3$ cherries, then D must have a very constrained structure similar to the iterated blow-up construction (see Theorem 5.2 for the precise statement). Once we have this, we can obtain the following strengthening of Corollary 2.8

Theorem 2.9. *There exists $L > 0$ such that the following holds. If $\mathcal{H}$ is a 3-uniform hypergraph on n vertices which does not contain a pseudocycle of length ℓ for any $\ell \leq L$ with $3 \nmid \ell$, then $e(\mathcal{H}) \leq f(n) + O(1)$.*

This theorem easily combines with Theorem 2.1 in order to give our main result, Theorem 1.3 (see Section 6).

3 Finding a good colouring

Recall that a *pseudocycle* of order m (or m-*pseudocycle*) is a sequence $v_1 \dots v_m$ of not necessarily distinct vertices such that $v_i v_{i+1} v_{i+2}$ is an edge for $i \in [m]$ (indices taken mod 3). A *pseudopath* of order m is defined analogously. A hypergraph is called *tightly connected* if there is a pseudopath between any two edges. Given vertices x, y, z, w (not necessarily distinct), a pseudopath from xy to zw (where xy and zw are ordered pairs) is a pseudopath whose first two vertices are x and y (in this order) and the last two vertices are z and w. The *shadow* of a hypergraph $\mathcal{H}$, denoted $\partial\mathcal{H}$, is the graph on vertices $V(\mathcal{H})$ whose edges are pairs xy that are contained in an edge in $\mathcal{H}$.

Recall that a *good colouring* of a hypergraph $\mathcal{H}$ is a colouring of its shadow, such that each edge xy in the shadow is either coloured blue or coloured red and given an orientation, and every edge e in $\mathcal{H}$ can be written as xyz where xy and xz are red and directed from x and yz is blue. Such an edge is called a *cherry*, and the vertex x is called its *apex*.

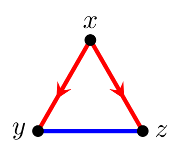

Figure 2: A cherry xyz with apex x

In this section, we will prove Theorem 2.4, restated here.

Theorem 2.4. *A 3-uniform hypergraph $\mathcal{H}$ has a good colouring if, and only if, $\mathcal{H}$ has no pseudocycle of length ℓ with $3 \nmid \ell$.*

It is easy to see that a hypergraph with a good colouring has no pseudocycles of length ℓ with $3 \nmid \ell$, so the main effort will be put into proving the "if" direction. Namely, we need to show that every hypergraph with no odd pseudocycles has a good colouring. Before specifying such a colouring, let us give some intuition. Any proper path (that is, with no repetitions) $v_1 \dots v_k$ has a good colouring, and this colouring is unique given the colour of $v_1 v_2$ (see Figure 3 for the three good colourings of a path of order 9, and notice that each such colouring colours each edge in the shadow differently).

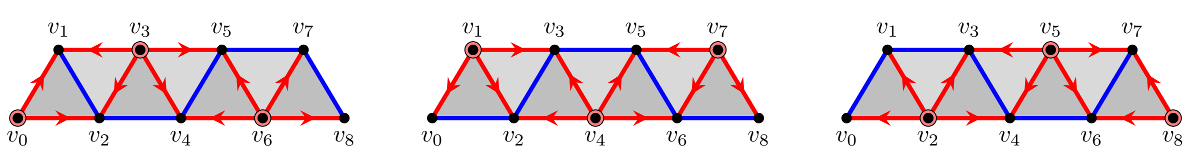

Figure 3: The three good colourings of a path of order 9

A proper cycle has a good colouring if and only if it is tripartite (i.e. the number of vertices is divisible by 3). See Figure 4 for a good colouring of a cycle of length 18 and notice that every third vertex is an apex.

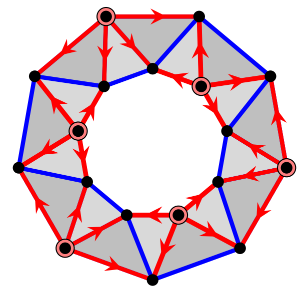

Figure 4: A good colouring of a cycle whose length is divisible by 3

Moreover, if there is a path $P = xy \dots yx$, then the order of P uniquely determines the colour of xy. This fact will be used to construct a good colouring of our hypergraph $\mathcal{H}$ – we will start from a specific pair xy and extend the colouring uniquely along pseudopaths. The difficulty is to show that this colouring is well defined, so the actual colouring definition will involve some more formalism. For a pseudopath $P = v_1 \dots v_k$, define $\tilde{P}$ by

$$\tilde{P} := v_{k-1}v_kv_{k-2}v_{k-1}v_{k-3} \dots v_4v_2v_3v_1v_2; \tag{1}$$

note that $\tilde{P}$ is a pseudopath from $v_{k-1}v_k$ to v_1v_2 of order $2k-2$ (because every vertex but v_1 and v_k appears twice).

Proof of Theorem 2.4. Whenever we talk about a path or cycle in this proof, we mean a pseudopath or pseudocycle.

As we said above, it is easy to show that a hypergraph with a good colouring has no odd cycles, so it suffices to show that if $\mathcal{H}$ has no odd cycles then $\mathcal{H}$ has a good colouring. Note that we may assume that $\mathcal{H}$ is tightly connected, by adding edges if necessary.

Let P_0 be a shortest path with the property that its first two vertices are the same as the last two but in reversed order, if such a path exists. Write $P_0 := v_0 \dots v_k$ and denote $x := v_0 = v_k$ and $y := v_1 = v_{k-1}$.

Define $\sigma \in \{0, 1, 2\}$ by

$$\sigma \equiv 2k \pmod 3. \tag{2}$$

Intuitively, σ is defined so that if P_0 has a good colouring, then the apexes in this colouring are at index $\sigma \pmod 3$. If P_0 does not exist, define $\sigma = 2$.

Let $\{z, w\}$ be an edge in the shadow of $\mathcal{H}$, and let $P = xyv_2v_3 \dots v_k$ be a path from xy whose last three vertices contain z and w. Let $i_w \in \{k-2, k-1, k\}$ be the index of w (namely, $v_{i_w} = w$), and define i_z analogously, and note that $i_w \neq i_z$. Define the index

$$\eta(P, \{z, w\}) = \begin{cases} w, & i_w \equiv \sigma \pmod 3, \\ z, & i_z \equiv \sigma \pmod 3, \\ *, & \text{otherwise.} \end{cases}$$

In particular, this defines $\eta(xy, \{x, y\})$. We claim that $\eta(P, \{z, w\})$ is independent of the choice of the path P.

Claim 3.1. *Let $z, w \in V(\mathcal{H})$ be distinct. Let $P = v_0 \dots v_p$ and $Q = u_0 \dots u_q$ be two paths starting at xy such that z and w are among their last three vertices. Then*

$$\eta(P, \{z, w\}) = \eta(Q, \{z, w\}).$$

Proof. Let $i_z \in \{p-2, p-1, p\}$ such that $v_{i_z} = z$, and define j_z similarly with respect to Q. It suffices to prove that $i_z \equiv \sigma \pmod 3$ if and only if $j_z \equiv \sigma \pmod 3$. Indeed, this implies the same equivalence for w, thus showing that $\eta(P, \{z, w\}) = *$ if and only if $\eta(Q, \{z, w\}) = *$.

First, we modify P so as to assume that P ends with zw or wz. If this is not the case, then up to swapping z and w we have that P ends with either $zw*$ or $z * w$. In the former case remove the last vertex of P, and in the latter case append z to P. It is easy to see that the statement of the claim holds for the original P if and only if it holds for the modified path. Similarly, we may assume that Q ends with zw or wz.

Assume first that P and Q both end with zw. Then $\tilde{Q}$ (defined as in (1)) is a path from zw to xy of order $2(q+1) - 2 = 2q$. Hence $v_2 v_3 \dots v_{p-2} \tilde{Q}$ is a cycle, and by assumption its order is divisible by 3. That is, $p - 3 + 2q \equiv 0 \pmod 3$, and thus $p \equiv q \pmod 3$. Since $i_z = p - 1$ and $j_z = q - 1$, this proves Claim 3.1. The same argument holds when P and Q both end with wz.

Secondly, assume that P is a path from xy to zw and Q is a path from xy to wz. Note that this case only arises if P_0 is defined, as $v_0 v_1 \dots v_{p-2} zw u_{q-2} \dots u_0$ is a path from xy to yx. Then consider the cycle $v_2 \dots v_{p-2} zw u_{q-2} \dots u_2 \tilde{P}_0$. This is indeed a cycle because $u_1 u_0 = yx$, $v_0 v_1 = xy$ and $\tilde{P}_0$ is a path from yx to xy. The order of this cycle is $p - 3 + 2 + q - 3 + 2k \equiv p + q + 2 + \sigma \pmod 3$, using (2). Now substitute $p = i_z + 1$ and $q = j_z$. We have $i_z + j_z + \sigma \equiv 0 \pmod 3$, so $i_z \equiv \sigma \pmod 3$ if and only if $j_z \equiv \sigma \pmod 3$. □

Note that for every edge $\{z, w\}$ in the shadow of $\mathcal{H}$ there is a path P starting at xy whose last three vertices contain z and w. Indeed, as $\mathcal{H}$ is tightly connected, there is a path Q such that x and y are among its first three vertices and z and w among its last three vertices. Using a modification as in the proof of Claim 3.1 we may assume that Q starts with xy or yx. If it starts with xy we are done, and otherwise the reverse of the path $\tilde{Q}$ satisfies the requirements.

Given an edge $\{z, w\}$ in the shadow of $\mathcal{H}$, define $\eta(zw) = \eta(P, \{zw\})$, where P is any path from xy whose last three vertices contain z and w (which exists by the previous paragraph). This parameter is well defined by Claim 3.1. Now define χ as follows: let zw be blue if $\eta(zw) = *$, and let it be red and oriented away from $\eta(zw)$ otherwise.

Finally, we show that χ is a good colouring. To see this, consider an edge uvw of G. Let P be a path from xy whose last three vertices are u, v, w (in some order); such a path exists by the paragraph above. Write $P := xyv_2v_3 \dots v_{p-2}v_{p-1}v_p$, let $i \in \{p-2, p-1, p\}$ with $i \equiv \sigma \pmod 3$, and we may assume that $v_i = u$. Then $\eta(uv) = \eta(uw) = u$ and $\eta(vw) = *$, which implies that uvw is a cherry with apex u. $\square$

Remark. Our proof actually shows that if G does not contain a path P_0 starting and ending with xy and yx respectively, then the graph is tripartite.

Notice that a good colouring of $\mathcal{H}$ can be extended from the shadow of $\mathcal{H}$ to K_n with no restrictions. Thus, in what follows it will suffice to analyse colourings of complete graphs by blue edges and red oriented edges (we will call such graphs *coloured graphs*).

4 Maximising the number of cherries

The results of the previous section establish a connection between maximising the number of edges in an odd-pseudocycle-free hypergraph and maximising the number of *cherries* in colourings of K_n (formally defined below). It will turn out that both problems have the same extremal construction which yields the maximum $f(n)$. Recall that we have defined $f(n)$ as the maximum number of edges in a hypergraph $\mathcal{H}(x_1, \dots, x_k)$ with $\sum_i x_i = n$. An explicit expression for f is

$$f(n) = \max_{k \geq 1} \max_{\substack{x_1, \dots, x_k \geq 1: \\ x_1 + \dots + x_k = n}} \left\{ \sum_{1 \leq i < j \leq k} \binom{x_i}{2} \cdot x_j \right\}. \tag{3}$$

Equivalently, we have the recursive characterisation

$$\begin{aligned} &f(1) = 0, \\ &f(n) = \max_{k \in [n-1]} \binom{k}{2}(n-k) + f(n-k) \quad \text{for } n \geq 2. \end{aligned} \tag{4}$$

Write

$$\beta = \frac{3 - \sqrt{3}}{2} \approx 0.634 \qquad \text{and} \qquad \alpha = \frac{\beta(1-\beta)}{2(3 - 3\beta + 3\beta^2)} = \frac{\sqrt{3}}{3} - \frac{1}{2} \approx 0.077. \tag{5}$$

The following proposition will be proved in Section 5.2.

Proposition 4.1. *$f(n) = \alpha n^3 + o(n^3)$.*

We remark that the density of the corresponding hypergraph $\mathcal{H}_n$ is $6\alpha = 2\sqrt{3} - 3$, as already noted by Mubayi and Rödl [27].

As in Section 2, we call a graph G *coloured* if it is a complete graph whose edges are either coloured blue or coloured red and oriented. A *cherry* in a coloured graph G is a triple xyz such that xy and xz are red and directed from x and yz is blue. Denote by $c(G)$ the number of cherries in G. Theorem 2.5 states that $c(G) \le f(n)$ for any n-vertex coloured graph G. Recall that this was originally proved by Falgas-Ravry and Vaughan [11] (who used flag algebras and also proved a similar result for out-directed stars on four vertices) and by Huang [16] (who used a symmetrisation argument, and proved a similar result for out-directed stars on k vertices, for every $k \ge 3$). Nevertheless, we provide a proof, both for completeness and because we need most of the groundwork to prove a stability version of Theorem 2.5.

As mentioned in the proof overview (Section 2), Corollary 2.6, which is a weak version of our main result and is restated here, follows directly from Theorem 2.5 (proved in the next section) and Theorem 2.4 (proved in the previous section).

Corollary 2.6. *If $\mathcal{H}$ is a 3-uniform hypergraph on n vertices which does not contain a pseudocycle of length ℓ for any ℓ with $3 \nmid \ell$, then $e(\mathcal{H}) \le f(n)$.*

5 Stability with symmetrisation

Most of the work in this section will go into proving the following lemma, providing a stability version of Theorem 2.5. It will then be iterated to prove a stability result about cherries in coloured graphs; recall that $\beta = (3 - \sqrt{3})/2$ (see (5)).

We point out that this stability result is somewhat similar to a general result due to Liu–Pikhurko–Sharifzadeh–Staden [23] which allows one to obtain stability versions of a class of extremal results that can be proved using a symmetrisation argument. However, while we indeed prove the extremal result in Theorem 2.5 using a symmetrisation argument, the result in [23] does not apply to automatically convert it into a stability result.

Lemma 5.1. *Let $1/n \ll \varepsilon \ll 1$ and let G be a coloured graph on n vertices satisfying $c(G) \ge f(n) - \varepsilon^2 n^3$. Then there is a coloured graph G' on $V(G)$ satisfying: $c(G') \ge c(G)$; the graphs G and G' differ on at most $800\varepsilon^{1/2}n^2$ edges; moreover, there is a set $Q \subseteq V(G)$ satisfying $\big||Q| - \beta n\big| \le 100\varepsilon n$; Q is a blue clique in G'; and all other edges in G' that are incident with Q are red and oriented towards Q.*

The proof consists of two main parts: first we show that G has a blue almost-clique on a vertex set Q' of size roughly βn. Then we show that most $(V \setminus Q', Q')$ edges are red and point towards Q'. In both parts, we make use of a "symmetrisation procedure" which builds blue cliques without decreasing the number of cherries.

A *blue clone-clique* in a coloured graph G is a set of vertices Q such that Q is a blue clique in G, and for any $v \notin Q$, either all edges between v and Q are blue, or they are all red and have the same orientation (namely, they all point towards v or all point away from v). A *full blue clone-clique* is a blue clone-clique Q such that all $(V \setminus Q, Q)$ edges are red.

The symmetrisation procedure, which will be described in detail in the next section, receives as input a vertex x in a graph G, and produces a graph G' on the same vertex set, which has at least as many cherries as G and has a full blue clone-clique Q in G' that contains x.

The symmetrisation procedure can be applied repeatedly to a coloured graph G to find a coloured graph G' with at least as many cherries as G, and whose vertices can be partitioned into full blue clone-cliques. Some calculus (detailed in Section 5.2) will show that such a G' contains a full blue clone-clique Q' of size approximately βn.

To proceed we need two lemmas (Lemmas 5.6 and 5.7; see Section 5.3) that together tell us the following. Suppose that a symmetrisation procedure on G resulted in a full blue clone-clique Q, of size approximately βn. Then (even before performing symmetrisation) almost all edges in $G[Q, V \setminus Q]$ are red and point towards Q, and almost all edges in $G[Q]$ are blue.

Applying these lemmas to the previously found blue clone-clique Q', we conclude first that $G[Q']$ is almost fully blue. We then show that there is a particular instance of the symmetrisation procedure that results in a graph G' and full blue clone-clique Q such that Q and Q' differ on only few vertices. Lemma 5.6 implies that almost all $G[Q', V \setminus Q']$ edges are red and point towards Q'. This essentially completes the proof. This part is detailed in Section 5.4.

In Section 5.5, we iterate Lemma 5.1 to prove the following result.

Theorem 5.2. *Let $1/n \ll \varepsilon_1 \ll \varepsilon_2 \ll 1$. Let G be a coloured graph on n vertices satisfying $c(G) \geq f(n) - \varepsilon_1 n^3$. Then there exists a coloured graph G' on the same vertex set, satisfying:*

(a) $c(G') \geq c(G)$,

(b) *G and G' differ on at most $\varepsilon_2 n^2$ edges,*

(c) *the vertices of G' can be partitioned into $Q_1, \ldots, Q_t$ such that:*

  (i) $|Q_1| \geq \ldots \geq |Q_t|$,

  (ii) *all edges in Q_i are blue, for $i \in [t]$,*

  (iii) *all edges in (Q_i, Q_j) are red and directed towards Q_i, for $1 \leq i < j \leq t$,*

  (iv) $\big||Q_i| - \beta \cdot |Q_i \cup \ldots \cup Q_t|\big| \leq \varepsilon_2 n$ *for $i \in [t]$.*

In a coloured graph G, let $N_G^-(x)$ be the red in-neighbourhood of x and let $N_G^+(x)$ be the red out-neighbourhood of x (we sometimes omit the subscript G).

The symmetrisation procedure $S_G(x)$

Input: a coloured graph G on vertex set V and a vertex $x \in V$.

Output: a graph G' and a full blue clone-clique Q in G' containing x.

The process: the algorithm builds sequences $x_1, \ldots, x_t$ and $y_1, \ldots, y_t$ of vertices in V and $G_1, \ldots, G_t$ of graphs on V, for some t, as follows.

1. Set $x_1, y_1 := x$ and $G_1 := G$.
2. Suppose $x_1, \ldots, x_k$ and $G_1, \ldots, G_k$ are given and $\{x_1, \ldots, x_k\}$ is a blue clone-clique in G_k.
3. If there are no vertices in $V \setminus \{x_1, \ldots, x_k\}$ whose edges to $\{x_1, \ldots, x_k\}$ are all blue, put $Q := \{x_1, \ldots, x_k\}$ and $G' := G_k$, and return G' and Q.
4. Otherwise, let x_{k+1} be a vertex in $V \setminus \{x_1, \ldots, x_k\}$ which sends blue edges to $\{x_1, \ldots, x_k\}$ (x_{k+1} can be chosen arbitrarily or judiciously).
5. For $y \in V$, let $G_{k+1}(y)$ be the graph obtained by replacing $\{x_1, \ldots, x_{k+1}\}$ by $k+1$ copies of y and letting the new vertices form a blue clique.

   If $c(G_{k+1}(x_{k+1})) - c(G_k) \geq k \cdot (c(G_{k+1}(x_1)) - c(G_k))$, set $y_{k+1} = x_{k+1}$, and otherwise let $y_{k+1} = x_1$. Define $G_{k+1} := G_{k+1}(y_{k+1})$, and return to step 2, considering the sequences $x_1, \ldots, x_{k+1}$ and $G_1, \ldots, G_{k+1}$.

Figure 5: Description of the symmetrisation process $S_G(x)$

5.1 The symmetrisation procedure

Given $x \in V(G)$, the *symmetrisation procedure* $S_G(x)$ (or $S(x)$ in short) builds a blue clone-clique containing x; see Figure 5 for a detailed description. The result of the procedure depends on the choice of x_{k+1} in step 4, but we suppress this dependence in the notation $S_G(x)$.

We now show that the procedure $S_G(x)$ does not decrease the number of cherries. In fact, we prove a stronger quantitative claim.

Claim 5.3. *Let $x_1, \ldots, x_t$, $y_1, \ldots, y_t$ and $G_1, \ldots, G_t$ be sequences produced by $S_G(x)$, let $k \in [t-1]$, and use $N^-(u)$ as a shorthand for $N^-_{G_k}(u)$. Then, one of the following holds.*

(i) $y_{k+1} = x_1$ *and* $c(G_{k+1}) - c(G_k) \geq \frac{k+1}{4} \cdot \left|N^-(x_1) \triangle N^-(x_{k+1})\right|$,

(ii) $y_{k+1} = x_{k+1}$ *and* $c(G_{k+1}) - c(G_k) \geq \frac{k(k+1)}{4} \cdot \left|N^-(x_1) \triangle N^-(x_{k+1})\right|$.

In particular, $c(G_{k+1}) \geq c(G_k)$.

Proof. Let $c(x_i)$ denote the number of cherries in G_k containing x_i and no other vertices in $x_1, \ldots, x_k$. Recall that for $y \in \{x_1, \ldots, x_{k+1}\}$ the graph $G_{k+1}(y)$ is obtained from G_k by replacing $x_1, \ldots, x_{k+1}$

by copies of y that form a blue clique. Write $\Delta_j := c(G_{k+1}(x_j)) - c(G_k)$. We claim that

$$\Delta_j = (k+1)c(x_j) - \sum_{i\in[k+1]} c(x_i) + \binom{k+1}{2}\big|N^-(x_j)\big| - \frac{1}{2}\sum_{i_1\neq i_2}\big|N^-(x_{i_1})\cap N^-(x_{i_2})\big|.$$

Indeed, the first two terms account for the triples with only one vertex in $\{x_1,\ldots,x_k\}$. For the second two terms, notice that a triple (x_{i_1}, x_{i_2}, v) with $1 \le i_1 < i_2 \le k+1$ is a cherry in $G_{k+1}(x_j)$ if and only if $v \in N^-(x_j)$, and it is a cherry in G_k if and only if $v \in N^-(x_{i_1}) \cap N^-(x_{i_2})$. Summing over $j \in [k+1]$, we obtain

$$\begin{aligned}\sum_j \Delta_j &= \binom{k+1}{2}\sum_j \big|N^-(x_j)\big| - \frac{k+1}{2}\sum_{i\neq j}\big|N^-(x_i)\cap N^-(x_j)\big| \\ &= \frac{k+1}{2}\sum_{i\neq j}\big|N^-(x_j)\setminus N^-(x_i)\big|.\end{aligned}$$

In particular, since $\{x_1,\ldots,x_k\}$ is a blue clone-clique,

$$\begin{aligned}k\Delta_1 + \Delta_{k+1} &= \frac{k(k+1)}{2}\cdot\left(\big|N^-(x_{k+1})\setminus N^-(x_1)\big| + \big|N^-(x_1)\setminus N^-(x_{k+1})\big|\right) \\ &= \frac{k(k+1)}{2}\cdot\big|N^-(x_1)\triangle N^-(x_{k+1})\big|.\end{aligned}$$

Now, $\max(k\Delta_1, \Delta_{k+1})$ is at least one half of the RHS. Thus, if $y_{k+1} = x_{k+1}$ then $\Delta_{k+1} \ge k\Delta_1$ and so $\Delta_{k+1} = \max(k\Delta_1, \Delta_{k+1}) \ge \frac{k(k+1)}{4}\cdot\big|N^-(x_1)\triangle N^-(x_{k+1})\big|$, and if $y_1 = x_1$ then $k\Delta_1 > \Delta_{k+1}$ and so $\Delta_1 = \max(k\Delta_1, \Delta_{k+1}) \ge \frac{k+1}{4}\cdot|N^-(x_1)\triangle N^-(x_{k+1})\big|$. □

Theorem 2.5 follows easily from the above claim.

Proof of Theorem 2.5. Let G be a coloured graph on n vertices. Run the following process: starting with $G' = G$, as long as there is a vertex x which is not in a full blue clone-clique in G', run $S_{G'}(x)$ and replace G' by the resulting graph. Let G_{final} be the graph G' at the end of the process (notice that the process will indeed end, because $S_{G'}(x)$ keeps full blue clone-cliques intact). Then $c(G_{\text{final}}) \ge c(G)$ by Claim 5.3, and the vertices of G_{final} can be partitioned into full blue clone-cliques $Q_1,\ldots,Q_t$; for convenience suppose that $|Q_1| \ge \ldots \ge |Q_t|$. Replace G_{final} by the graph G'_{final} obtained by directing the red edges between Q_i and Q_j towards Q_i, for $1 \le i < j \le t$. It is straightforward to verify that $c(G'_{\text{final}}) \ge c(G_{\text{final}})$, as the number of cherries in $Q_i \cup Q_j$ is larger when the arcs in (Q_i, Q_j) point towards the larger clique. Finally, denoting $q_i := |Q_i|$, observe that $c(G'_{\text{final}}) = \sum_{i\le j}\binom{q_i}{2}q_j \le f(n)$ (see (3)). Thus $c(G) \le f(n)$, as claimed. □

5.2 Optimising the clique size

Before proceeding to analyse the symmetrisation procedure, we prove the following lemma regarding the structure of a graph whose vertices are partitioned into full blue clone-cliques, mostly using calculus; recall that α and β are defined in (5).

Lemma 5.4. *Let $1/n \ll \varepsilon \ll 1$. Let G be a coloured graph on n vertices whose vertices can be partitioned into full blue clone-cliques, and suppose that $c(G) \geq f(n) - \varepsilon^2 n^3$. Then G has a full blue clone-clique Q satisfying $\big||Q| - \beta n\big| \leq 100\varepsilon n$.*

Define a function $g : [0,1] \to \mathbb{R}$ as follows.

$$g(x) = \frac{x(1-x)}{2(3-3x+x^2)}. \tag{6}$$

It will be convenient to note the following equation.

$$\big(1-(1-x)^3\big) \cdot g(x) = \frac{1}{2} \cdot x^2(1-x). \tag{7}$$

One can check that g' is decreasing and $g'(\beta) = 0$, showing that

$$g(x) \leq g(\beta) = \alpha \qquad \text{for } x \in [0,1]. \tag{8}$$

We first prove Proposition 4.1 regarding the value of $f(n)$.

Proof of Proposition 4.1. We show that $f(n) \leq \alpha n^3$ by induction on n. This is true for $n = 1$. Suppose that $f(m) \leq \alpha m^3$ for $m < n$.

Given $k \in [n-1]$ that maximises the LHS in (4), write $x = k/n$. The recursive definition of f implies that $\frac{f(n)}{n^3} \leq \frac{1}{2} \cdot x^2(1-x) + \alpha(1-x)^3$. Subtracting α and using (7), we obtain

$$\frac{f(n)}{n^3} - \alpha \leq \frac{1}{2} \cdot x^2(1-x) - \alpha(1-(1-x)^3) = (1-(1-x)^3)(g(x)-\alpha) \leq 0,$$

as required.

To verify that $f(n) \geq (\alpha + o(1))n^3$, set $x_i = \lfloor \beta(1-\beta)^i n \rfloor$ in (3). □

Proof of Lemma 5.4. Let $Q_1, \ldots, Q_t$ be the full blue clone-cliques in G, arranged in descending order according to their sizes. Let G' be obtained from G by orienting the (Q_i, Q_j) (red) edges towards Q_i, for $1 \leq i < j \leq t$. As explained before and by assumption on G, $c(G') \geq c(G) \geq f(n) - \varepsilon^2 n^3$.

Notice that $|Q_1| \geq 0.01n$, because otherwise $c(G) \leq n^2|Q_1| \leq 0.01n^3 < f(n) - \varepsilon^2 n^3$ (recall that $f(n) \approx 0.077n^3$, by Proposition 4.1).

Write $|Q_1| = \theta n$. Then, using $f(n) = \alpha n^3 + o(n^3) = g(\beta)n^3 + o(n^3)$ (which follows from Proposition 4.1 and the definition of α in (5)),

$$c(G') \leq \binom{|Q_1|}{2}(n - |Q_1|) + f(n - |Q_1|) \leq \frac{1}{2}\theta^2(1-\theta)n^3 + g(\beta)(1-\theta)^3 n^3 + o(n^3).$$

Thus, using (7),

$$\begin{aligned}\varepsilon^2 \geq \frac{f(n) - c(G')}{n^3} &\geq g(\beta) - g(\beta)(1-\theta)^3 - \frac{1}{2}\theta^2(1-\theta) + o(1)\\ &= \theta \cdot (3 - 3\theta + \theta^2) \cdot (g(\beta) - g(\theta)) + o(1)\\ &\geq 0.02 \cdot (g(\beta) - g(\theta)) + o(1).\end{aligned}$$

For the last inequality we used $\theta \geq 0.01$, which implies $\theta(3 - 3\theta + \theta^2) \geq 0.02$. By bounding the $o(1)$ term by $\varepsilon^2/2$ and using Claim 5.5 below, we get

$$100\varepsilon^2 \geq g(\beta) - g(\theta) \geq \min\{0.05(\beta - \theta)^2, 0.005\}.$$

Since ε is very small, we get $100\varepsilon^2 \geq 0.05(\beta - \theta)^2$, which implies $|\beta - \theta| \leq 100\varepsilon$. □

Claim 5.5. *For $x \in [0,1]$,*

$$g(\beta) - g(x) \geq \min\{0.05(\beta - x)^2, 0.005\}. \tag{9}$$

Proof. We use the following facts that can be checked easily.

- The function $g(x)$ is increasing on $[0, \beta]$ and decreasing on $[\beta, 1]$. In particular, its maximum is attained at β, and $g'(\beta) = 0$.
- $g(\beta) - g(0.5) \geq 0.005$.
- The second derivative $g''(x)$ (which is $\frac{-x(2x^2-9x+9)}{(x^2-3x+3)^3}$) is non-negative and decreasing on $[0, 1]$. In particular $g''(x) \leq g''(0.5) \leq -0.4$ for $x \in [0.5, 1]$.
- By Taylor's expansion: $g(x) = g(\beta) + \frac{1}{2}g'(\beta)(x - \beta) + \frac{1}{6}g''(c_x)(x - \beta)^2$ for every $x \in [0, 1]$ and some c_x between x and β.

By the first and second items (using $\beta > 0.5$), if $x \in [0, 0.5]$ then

$$g(\beta) - g(x) \geq g(\beta) - g(0.5) \geq 0.005.$$

By the first, third and fourth items, if $x \in [0.5, 1]$, then

$$g(\beta) - g(x) \geq \frac{0.4}{6}(x - \beta)^2 \geq 0.05(x - \beta)^2.$$

The two inequalities prove the claim. □

5.3 Blue clone-cliques before and after symmetrisation

The next two lemmas show that if a symmetrisation procedure on G produces a full blue clone-clique Q of size approximately βn, then almost all edges in $G[Q, V \setminus Q]$ are red and oriented towards Q

and almost all edges in $G[Q]$ are blue.

Lemma 5.6. *Let $1/n \ll \varepsilon \ll 1$. of a procedure $S_G(x)$, and suppose that $|Q| \geq 0.55n$. Then all but at most $10\varepsilon n^2$ edges in $G[Q, V \setminus Q]$ are red and directed towards Q.*

Proof. Set $U := V \setminus Q$, let V_{in} be the set of vertices u in U for which uq is a red arc in G' for every $q \in Q$, and let $V_{\text{out}} := U \setminus V_{\text{in}}$. We will show that V_{out} is small, and that not many pairs incident to V_{in} were recoloured during the symmetrisation procedure $S_G(x)$.

First, we show $|V_{\text{out}}| \leq 40\varepsilon^2 n$. Let G'' be obtained from G' by reorienting the edges in $G'[Q, V_{\text{out}}]$ to point towards Q. Notice that the cherries in G' that contain an edge in (Q, V_{out}) consist of one vertex in Q and two in V_{out}, and thus their number is at most $|Q|\binom{|V_{\text{out}}|}{2}$. Also, every set consisting of two vertices in Q and one in V_{out} is a cherry in G'' but not in G'. Thus, using $|Q| \geq 0.55n$ which implies $|Q| - |V_{\text{out}}| \geq 0.1n$,

$$c(G'') - c(G') \geq \binom{|Q|}{2}|V_{\text{out}}| - \binom{|V_{\text{out}}|}{2}|Q| = \frac{1}{2}|Q||V_{\text{out}}|(|Q| - |V_{\text{out}}|)$$
$$\geq \frac{1}{2} \cdot \frac{n}{2} \cdot \frac{n}{10} \cdot |V_{\text{out}}| = \frac{n^2}{40} \cdot |V_{\text{out}}|.$$

Recall that $c(G) \geq f(n) - \varepsilon^2 n^3$ by assumption, $c(G') \geq c(G)$ by Claim 5.3, and $c(G'') \leq f(n)$ by Theorem 2.5. Altogether, this implies $c(G'') - c(G') \leq \varepsilon^2 n^3$ and thus $|V_{\text{out}}| \leq 40\varepsilon^2 n$, as claimed.

Let R be the set of edges qv in $(Q, V \setminus Q)$ that are red and oriented towards Q in G' but not in G. We now upper-bound $|R|$. Notice that each such edge in R was recoloured to a red arc oriented towards Q at some point during $S_G(x)$ (possibly more than once). Let $G = G_1, \ldots, G_t = G'$ be the graphs obtained during the symmetrisation process on Q and let $x_1, \ldots, x_t$ be the corresponding sequence of vertices. For each $v \in V$ and $k \in [t]$, let $A_k(v)$ be the set of ordered pairs vq which changed to red arcs in step k (so they were recoloured from G_{k-1} to G_k).

We claim that $\sum_{k \geq \varepsilon n} \sum_{v \in V_{\text{in}}} |A_k(v)| \leq 4\varepsilon n^2$. To see this, fix $k \geq \varepsilon n$ and consider the k-th step. If $y_k = x_1$, then $A_k(v) = \{vx_k\}$ for $v \in N^-(x_1) \setminus N^-(x_k)$ and $A_k(v) = \emptyset$ otherwise, where $N^-(\cdot)$ refers to the in-neighbourhood with respect to G_{k-1}. Thus, using Claim 5.3 (i),

$$\sum_{v \in V_{\text{in}}} |A_k(v)| \leq \left|N^-(x_1) \setminus N^-(x_k)\right| \leq \frac{4}{k} \cdot \big(c(G_k) - c(G_{k-1})\big).$$

If $y_k = x_k$, then $A_k(v) = \{vx_1, \ldots, vx_{k-1}\}$ for $v \in N^-(x_k) \setminus N^-(x_1)$ and $A_k(v) = \emptyset$ otherwise. Thus, by Claim 5.3 (ii),

$$\sum_{v \in V_{\text{in}}} |A_k(v)| \leq (k-1) \cdot \left|N^-(x_1) \setminus N^-(x_k)\right| \leq \frac{4}{k} \cdot \big(c(G_k) - c(G_{k-1})\big),$$

In either case, we get that for $k \geq \varepsilon n$,

$$\sum_{v \in V_{\text{in}}} |A_k(v)| \leq \frac{4}{\varepsilon n}\big(c(G_k) - c(G_{k-1})\big).$$

Summing over $k \geq \varepsilon n$, we obtain the required inequality

$$\sum_{k \geq \varepsilon n} \sum_{v \in V_{\text{in}}} |A_k(v)| \leq \frac{4}{\varepsilon n}\big(c(G') - c(G_{\varepsilon n})\big) \leq 4\varepsilon n^2,$$

Where the last equality holds since $c(G') - c(G_{\varepsilon n}) \leq \varepsilon^2 n^2$.

Note that $|R| \leq \varepsilon n^2 + \sum_{k \geq \varepsilon n} \sum_{v \in V_{\text{in}}} |A_k(v)| \leq 5\varepsilon n^2$. In total, all but at most $(40\varepsilon^2 + 5\varepsilon)n^2 \leq 10\varepsilon n^2$ pairs in $(Q, V \setminus Q)$ are red and oriented towards Q. □

Lemma 5.7. *Let $1/n \ll \varepsilon \ll 1$. Let G be a coloured graph on n vertices with at least $f(n) - \varepsilon^2 n^3$ cherries. Suppose that G' and Q are the graph and full blue clone-clique produced by the procedure $S_G(x)$, and suppose that $0.55n \leq |Q| \leq 0.65n$. Then all but $1200\varepsilon n^2$ edges in $G[Q]$ are blue.*

Proof. Let F and F' be the graphs obtained from G and G' by colouring all $(Q, V \setminus Q)$ edges red and orienting them towards Q. Notice that F' can be obtained from F by colouring all edges in Q blue.

We will first derive an upper bound on $c(F') - c(F)$. By Lemma 5.6, the graphs G and F differ on at most $10\varepsilon n^2$ edges and thus $|c(G) - c(F)| \leq 10\varepsilon n^3$. Similarly, $|c(G') - c(F')| \leq 10\varepsilon n^3$ (the lemma is still applicable, as $S_{G'}(x)$ does not change the graph G'). By assumption on G we also have $c(G') - c(G) \leq \varepsilon^2 n^3$. Altogether,

$$c(F') - c(F) \leq c(G') - c(G) + 20\varepsilon n^3 \leq (\varepsilon^2 + 20\varepsilon)n^3 \leq 30\varepsilon n^3. \tag{10}$$

We now obtain a lower bound on the same quantity. Let e be the number of red edges in $G[Q]$. The number of cherries in F that are not cherries in F' is at most $\sum_{q \in Q} \binom{d^+(q)}{2}$, where $d^+(q)$ denotes the red out-degree of q in $F[Q]$. Notice that

$$\sum_{q \in Q} \binom{d^+(q)}{2} \leq \frac{1}{2} \sum_{q \in Q} (d^+(q))^2 \leq \frac{1}{2} e|Q|,$$

because $d^+(q) \leq |Q|$ and $e = \sum_q d^+(q)$.

On the other hand, the number of cherries in F' that are not cherries in F is exactly $e(n - |Q|)$. Thus,

$$c(F') - c(F) \geq e(n - |Q|) - \frac{1}{2} e|Q| = e \cdot \big(n - \frac{3}{2}|Q|\big) \geq \frac{en}{40}, \tag{11}$$

using $|Q| \leq 0.65n$.

By (10) and (11), we have $e \leq 1200\varepsilon n^2$, as claimed. □

5.4 Proof of Lemma 5.1

Finally, we start with the actual proof of Lemma 5.1. The first step is to find a set Q' of the right size almost all of whose edges in G are blue.

Lemma 5.8. *Let $1/n \ll \varepsilon \ll 1$. Let G be a coloured graph on n vertices, satisfying $c(G) \geq f(n) - \varepsilon^2 n^3$. Then there is a set $Q' \subseteq V(G)$ such that $\big||Q'| - \beta n\big| \leq 100\varepsilon n$ and all but at most $1200\varepsilon n^2$ edges in $G[Q']$ are blue.*

Proof. Similarly to the proof of Theorem 2.5, start with $G' = G$, and, as long as G' has a vertex x which is not in a full blue clone-clique, run the symmetrisation procedure $S_{G'}(x)$, and replace G' by the resulting graphs. Denote by G_{final} the graph at the end of the process (as before, the process is guaranteed to end). Then the vertices of G_{final} can be partitioned into full blue clone-cliques $Q_1, \ldots, Q_t$.

Let Q' be the vertex set of the largest clone-clique. By Lemma 5.4, we have $\big||Q'| - \beta n\big| \leq 100\varepsilon n$. In particular $|Q'| \in [0.55n, 0.65n]$.

Let F_1 be the graph created just before the symmetrisation procedure was started on an element of Q', and let F_2 be the graph just after Q' was built. Notice that $c(F_2) \geq c(F_1) \geq c(G) \geq f(n) - \varepsilon^2 n^3$. By Lemma 5.7, all but at most $1200\varepsilon n^2$ edges in $F_1[Q']$ are blue. Notice that during the above process, the edges in Q' remain untouched until right before a symmetrisation process is started on an element of Q'. It follows that all but at most $1200\varepsilon n^2$ edges in $G[Q']$ are blue. □

Now we can complete the proof by running a symmetrisation procedure in two phases. The first phase generates a blue clique Q which contains almost all the vertices of Q'. The second phase allows us to show that Q cannot be much larger than Q' and to control the remaining edges incident to Q.

Proof of Lemma 5.1. Apply Lemma 5.8 to find Q' such that $\big||Q'| - \beta n\big| \leq 100\varepsilon n$ and $G[Q']$ has at most $\delta^2 n^2$ red edges (with $\delta^2 = 1200\varepsilon$).

Claim 5.9. *We can run a symmetrisation procedure on G which results in a graph G' and a full blue clone-clique Q satisfying $|Q' \setminus Q| \leq 3\delta n$.*

Proof. Let A be the set of vertices in Q' with more than δn red (in- or out-) neighbours in $G[Q']$. The bound on the number of red edges in Q' gives $|A| < 2\delta n$. Define $Q'' := Q' \setminus A$.

We will run a symmetrisation procedure on G, but with a specific ordering of vertices. We start with $x_1 \in Q''$ (chosen arbitrarily). Assuming that $\{x_1, \ldots, x_k\}$ are defined and contained in Q'', if possible we pick x_{k+1} to also be in Q'' (we can do this as long as there is a vertex in $Q'' \setminus \{x_1, \ldots, x_k\}$ whose edges to $\{x_1, \ldots, x_k\}$ are blue). Once this is no longer possible, we continue with the symmetrisation procedure using an arbitrary order of vertices. Let Q be the full blue clone-clique built by this procedure.

Let k be largest such that $\{x_1, \ldots, x_k\} \subseteq Q''$. It is easy to see that throughout the procedure, until at least step k, every vertex in Q'' has at most δn non-blue neighbours in $Q'' \setminus \{x_1, \ldots, x_k\}$. Thus $k \geq |Q''| - \delta n \geq |Q| - 3\delta n$, as otherwise we could find a suitable x_{k+1} in Q'', contradicting the choice of k. It follows that $|Q' \setminus Q| \leq 3\delta n$. □

Let G' and Q be as in the above Claim. We claim that $|Q| \leq (\beta + 100\varepsilon)n$. Indeed, this follows from Lemma 5.4 by running symmetrisation procedures repeatedly, starting from G', until the vertices can be partitioned into full blue clone-cliques (one of which is Q). It follows that $|Q \setminus Q'| \leq 3\delta n + |Q| - |Q'| \leq (3\delta + 200\varepsilon)n \leq 5\delta n$. In particular, the number of red edges in $G[Q]$ is at most the number of red edges in $G[Q']$ plus the number of edges incident with $Q \setminus Q'$, which amounts to a total of at most $(\delta^2 + 5\delta)n^2 \leq 10\delta n^2$ red edges in $G[Q]$.

By Lemma 5.6, all but at most $10\varepsilon n^2$ edges in $G[Q, V \setminus Q]$ are red and oriented towards Q, and similarly for $G'[Q, V \setminus Q]$.

Since Q is a full blue clone-clique in G', the vertices in $V \setminus Q$ can be partitioned into V_{in} and V_{out}, where vq is a red arc for every $v \in V_{\text{in}}$ and $q \in Q$ and qv is a red arc for $v \in V_{\text{out}}$ and $q \in Q$. Thus, by the previous paragraph and because $|Q| \geq n/2$, $|V_{\text{out}}| \leq 20\varepsilon n$.

Let G'' be obtained from G' be reorienting all $(Q, V \setminus Q)$ edges towards Q. Then

$$c(G'') - c(G') \geq \binom{|Q|}{2}|V_{\text{out}}| - \binom{|V_{\text{out}}|}{2}|Q| = |Q||V_{\text{out}}| \cdot (|Q| - |V_{\text{out}}|) \geq 0.$$

It follows that $c(G'') \geq c(G') \geq c(G)$. Moreover, G'' and G' differ on at most $|V_{\text{out}}|n \leq 20\varepsilon n^2$ edges, and thus G and G'' differ on at most $(20\varepsilon + 10\varepsilon + 10\delta)n^2 \leq 20\delta n^2$ edges. Since G'' has the required structure, this proves Lemma 5.1. □

5.5 Full stability result

Proof of Theorem 5.2. Let $\varepsilon_1 \ll \eta \ll \varepsilon_2$. The idea is simply to iterate Lemma 5.1. We will find graphs $G_1, \ldots, G_s$ and sets $Q_1, \ldots, Q_s$, satisfying the following conditions, for $k \in [s]$ (for convenience set $G_0 := G$, $Q_0 := \emptyset$ and $V := V(G)$).

(1) G_k is a coloured graph on vertex set $V \setminus (Q_1 \cup \ldots \cup Q_{k-1})$.

(2) Q_k is a blue clique in G_k, all other edges incident with Q_k in G_k are red and point towards Q_k.

(3) $\big||Q_k| - \beta|G_k|\big| \leq \eta|G_k|$.

(4) G_k and $G_{k-1} \setminus Q_{k-1}$ differ on at most $\eta|G_k|^2$ edges.

(5) $c(G_k) \geq c(G_{k-1} \setminus Q_{k-1})$.

(6) $c(G_k \setminus Q_k) \geq f(|G_k \setminus Q_k|) - \varepsilon_1 n^3$.

To see how such a sequence can be built, suppose that $G_1, \ldots, G_{k-1}$ and $Q_1, \ldots, Q_{k-1}$ are defined and satisfy the above conditions. If $|G_{k-1} \setminus Q_{k-1}| \leq \eta n$, we stop the process and set $s := k-1$. Otherwise, we apply Lemma 5.1 to the graph $G_{k-1} \setminus Q_{k-1}$. Notice that by (6) and the assumption on $|G_{k-1} \setminus Q_{k-1}|$, we have $c(G_k \setminus Q_k) \geq f(|G_k \setminus Q_k) - \varepsilon_1 \eta^{-3} |G_k \setminus Q_k|^3$. Since $\varepsilon_1 \eta^{-3} \ll \eta$, the lemma is applicable. The lemma produces a graph G_k on vertex set $V(G_{k-1}) \setminus Q_{k-1} = V \setminus (Q_1 \cup \ldots \cup Q_{k-1})$ satisfying items (1) to (5). It remains to verify (6). Note that

$$c(G_k) = \binom{|Q_k|}{2} \cdot |G_k \setminus Q_k| + c(G_k \setminus Q_k).$$

Also

$$\begin{aligned} c(G_k) \geq c(G_{k-1} \setminus Q_{k-1}) &\geq f(|G_{k-1} \setminus Q_{k-1}|) - \varepsilon_1 n^3 \\ &= f(|G_k|) - \varepsilon_1 n^3 \\ &\geq \binom{|Q_k|}{2} |G_k \setminus Q_k| + f(|G_k \setminus Q_k|) - \varepsilon_1 n^3, \end{aligned}$$

where the last inequality follows from the definition of f. The two inequalities imply (6).

To finish, run a symmetrisation procedure on $G_s \setminus Q_s$ repeatedly, to obtain a graph H whose vertices are partitioned into full blue clone-cliques $Q_{s+1}, \ldots, Q_t$ (arranged in decreasing size); the edges between any two of them point towards the larger clique; and $c(H) \geq c(G_t \setminus Q_t)$. Let G' be the graph on vertex set V, such that $Q_1, \ldots, Q_t$ are blue cliques and the edges between any two of them are red and point towards the larger clique (note that $Q_1, \ldots, Q_t$ partition V).

To complete the proof of Theorem 5.2, we need to show that properties (a) to (c) hold. For (a), define G'_k to be the graph on vertex set V, obtained from G' by replacing $V \setminus (Q_1 \cup \ldots \cup Q_{k-1})$ by a copy of G_k (this makes sense due to (1)). It is easy to see that $c(G'_k) - c(G'_{k-1}) = c(G_k) - c(G_{k-1} \setminus Q_{k-1}) \geq 0$ for $k \in [s]$, using (5). Similarly, $c(G') \geq c(G'_s)$. Altogether, $c(G') \geq c(G'_1) = c(G)$, as required for (a).

Before continuing, we derive an upper bound on s. By (3) we have $|Q_k| \geq 0.55|G_k|$ for $k \in [s]$, so $|G_k| \leq 2^{-(k-1)} n$. Since $|G_t| \leq \eta n$, this implies that $s \leq 2 \log(1/\eta) \leq \eta^{-1/2}$, say.

By (4) we find that G' and G differ on at most $((s\eta + \eta) n^2 \leq 2\eta^{1/2} n^2 \leq \varepsilon_2 n^2$ edges. Property (b) follows.

Notice that the estimate $|Q_k| \geq 0.55|G_k|$, which follows from (3) implies $|Q_1| \geq \ldots \geq |Q_s|$. Thus c(i) to c(iii) clearly hold. Finally, c(iv) holds trivially for $k > s$ and, for $k \leq s$, it follows from (3) and $\eta \leq \varepsilon_2$. □

6 Hypergraphs with no short odd pseudocycles

In this section we leverage the stability result about cherries, Theorem 5.2, and the connection between hypergraphs with no odd pseudocycles to good colourings (Theorem 2.4) to prove the following result regarding the structure of a dense hypergraph with no short odd pseudocycles. In case of cycles and pseudocycles, the *length* (number of edges) and order (number of vertices) coincide, so, since there is no danger of confusion, we prefer the term *length*. Given vertex sets $X_1, X_2, X_3 \subset V(\mathcal{H})$, an $X_1X_2X_3$-*triple* in $\mathcal{H}$ is an (unordered) edge $x_1x_2x_3 \in E(\mathcal{H})$ with $x_i \in X_i$ for $i \in [3]$.

Theorem 6.1. *Let $n \gg \ell \gg 1$. Let $\mathcal{H}$ be a 3-uniform hypergraph on n vertices which contains no odd pseudocycles of length at most ℓ, and which maximises the number of edges under these conditions. Then there is a partition $\{A, B\}$ of the vertices of $\mathcal{H}$ into non-empty sets such that all AAB triples are edges of $\mathcal{H}$ (and there are no AAA and ABB triples).*

By iterating the above result, we prove Theorem 2.9, restated here, which gives an upper bound on the number of edges in a hypergraph with no short odd pseudocycles.

Theorem 2.9. *There exists $L > 0$ such that the following holds. If $\mathcal{H}$ is a 3-uniform hypergraph on n vertices which does not contain a pseudocycle of length ℓ for any $\ell \leq L$ with $3 \nmid \ell$, then $e(\mathcal{H}) \leq f(n) + O(1)$.*

Recall that Theorem 2.9 is tight, up to the additive $O(1)$ error term, as evidenced by $\mathcal{H}(x_1, \ldots, x_k)$ for a suitable choice of x_i's.

We next show how Theorem 2.9 implies our main result, Theorem 1.3, restated here.

Theorem 1.3. *Let ℓ be sufficiently large with $\ell \equiv 1$ or $2 \pmod 3$. Then $\pi(\mathcal{C}_\ell^3) = 2\sqrt{3} - 3$.*

Recall that the t-*blow-up* of an r-uniform hypergraph $\mathcal{H}$, denoted $\mathcal{H}[t]$, is the hypergraph with vertex set $V(\mathcal{H}) \times [t]$ and edges all r-sets $\{(x_1, i_1), \ldots, (x_r, i_r)\}$ such that $\{x_1, \ldots, x_r\} \in E(\mathcal{H})$. For a family $\mathcal{F}$ of hypergraphs, we denote by $\mathcal{F}[t]$ the family of t-blow-ups of members of $\mathcal{F}$. Recall that Theorem 2.1 (whose proof can be found in [19]) asserts that taking the t-blow-up of a hypergraph does not change its Turán density. The following generalisation for finite families of hypergraphs can be proved similarly.

Theorem 6.2 ([19], Theorem 2.2)**.** *Let s and t be integers, and let $\mathcal{F}$ be a family of r-graphs with $|\mathcal{F}| \leq s$. Then $\pi(\mathcal{F}[t]) = \pi(\mathcal{F})$.*

To prove Theorem 1.3, we will note that an odd cycle $C_m^{(3)}$ is contained in an m-*blow-up* of any odd pseudocycle of length at most $m/2$, and apply the last theorem.

Proof of Theorem 1.3 using Theorem 2.9. Let m be an integer with $m \geq 2L$ and $3 \nmid m$, where L is the constant from Theorem 2.9. Recall that $f(n) = (2\sqrt{3} - 3 + o(1))\binom{n}{3}$. Let $\varepsilon > 0$ and let $\mathcal{H}$ be an n-vertex 3-uniform hypergraph with $e(\mathcal{H}) \geq (2\sqrt{3} - 3 + \varepsilon)\binom{n}{3}$ and n sufficiently large. We claim that $\mathcal{H}$ contains a copy of $C_m^{(3)}$.

Theorem 2.9 and Theorem 6.2 imply that $\mathcal{H}$ contains $F[m]$ for some ℓ-pseudocycle F with $\ell \leq L$ and $3 \nmid \ell$. It suffices to show that F contains an m-pseudocycle, because then $C_m^{(3)}$ will be contained in $F[m]$. To see this, let $v_1 \dots v_\ell$ be an ordering of $V(F)$ such that $v_i v_{i+1} v_{i+2} \in E(F)$, with the indices taken modulo ℓ.

In case $m \equiv \ell \,(\mathrm{mod}\ 3)$, consider the sequence

$$(v_1 v_2 v_3)^{\frac{m-\ell}{3}} v_1 v_2 \dots v_\ell,$$

where $(v_1 v_2 v_3)^x$ stands for x repetitions of the sequence $v_1 v_2 v_3$. This is a sequence of order m certifying that F contains an m-pseudocycle.

Otherwise, if $m \equiv 2\ell \,(\mathrm{mod}\ 3)$, the same is certified for instance by the sequence

$$(v_1 v_2 v_3)^{\frac{m-2\ell}{3}} (v_1 v_2 \dots v_\ell)^2. \qquad \square$$

All that remains now is to prove Theorem 6.1. We will state and prove some preliminary results in the following subsection, and then prove the theorem in Section 6.2.

6.1 Preparation

The *diameter* of a hypergraph $\mathcal{H}$ is the minimum ℓ such that the following holds: for every $x, y, z, w \in V(\mathcal{H})$ (where x, y are distinct and z, w are distinct) whenever there is a pseudopath from xy to zw, there is such a pseudopath of order at most ℓ.

We have already shown that n-vertex hypergraphs with no odd pseudocycles have at most $f(n)$ edges. To prove the same for pseudocycles of bounded length, we will pass to a subhypergraph with bounded diameter, which is the purpose of the following two propositions.

Proposition 6.3. *Let $\mathcal{H}$ be a 3-uniform hypergraph of diameter $\ell \geq 4$. If $\mathcal{H}$ has an odd pseudocycle, then it has an odd pseudocycle of length at most 4ℓ.*

Proof. Let C be the shortest odd pseudocycle in $\mathcal{H}$. Assuming that its length is at least $3\ell + 4$, we may index it by $xyv_1 \dots v_k abu_1 \dots u_t$ with $t \geq 2\ell$, $k \geq \ell$. Note that the length of C is $k + t + 4 \not\equiv 0 \,(\mathrm{mod}\ 3)$.

Since $\mathcal{H}$ contains a pseudopath from xy to ab, it also contains such a pseudopath $P = xyw_1 \dots w_r ab$ with $r \leq \ell - 4$. The pseudocycle $xyw_1 \dots w_r abu_1 \dots u_t$ is shorter than C, so it must not be odd, that is, $r + t + 4 \equiv 0 \,(\mathrm{mod}\ 3)$.

Now consider the pseudocycle $C_1 = v_1 \dots v_k \tilde{P}$. Recall that $\tilde{P}$ is a $(2r + 6)$-vertex pseudopath from ab to xy (see (1)), so C_1 is indeed a pseudocycle. The length of C_1 is $k+2r+6 \equiv k-r \equiv k+t+4 \not\equiv 0 \,(\mathrm{mod}\ 3)$. Noting that $k + 2r + 6 \leq k + 2\ell - 2 \leq k + t$, this contradicts the minimality of C. $\square$

Proposition 6.4. *Let $1/\ell \ll \varepsilon \ll 1$, and let $\mathcal{H}$ be an n-vertex hypergraph. Then there is a subgraph $\mathcal{H}' \subseteq \mathcal{H}$ with $e(\mathcal{H}') \geq e(\mathcal{H}) - \varepsilon n^3$ whose diameter is at most ℓ.*

Proof. First we form a subgraph $\mathcal{H}' \subseteq \mathcal{H}$ in which each vertex pair has codegree either 0 or at least εn, as follows. If there are vertices u, v whose codegree in the *current* hypergraph is smaller than εn, delete all edges containing uv. Repeat this step until each pair has codegree degree either 0 or at least εn. Denote the resulting hypergraph by $\mathcal{H}'$. Observe that the number of deleted edges is at most $\varepsilon n \cdot \binom{n}{2}$ since the edges containing each pair were removed at most once. Hence $e(\mathcal{H}') \geq e(\mathcal{H}) - \varepsilon n^3$.

Given ordered pairs uv and $u'v'$ which are connected by a pseudopath in $\mathcal{H}'$, let $P = uvx_0x_1 \dots x_t u'v'$ be a shortest such pseudopath. For each i, let B_i be the set of ordered pairs ab such that $x_i x_{i+1} ab$ is a tight path in $\mathcal{H}'$. We claim that the sets B_{10i} are mutually disjoint for $0 \leq i < \frac{t}{10}$. Suppose not, and take $ab \in B_{10i} \cap B_{10j}$ for some $0 \leq i < j < \frac{t}{10}$. Then $x_i x_{i+1} ab x_{j+1} a x_j x_{j+1}$ is a pseudopath with only five vertices between x_i and x_j, which can be used to form a shorter pseudopath than P connecting uv and $u'v'$, contradiction. Now since $|B_i| \geq \varepsilon^2 n^2/2$ for every i (using the fact that the codegree of each pair in $\mathcal{H}'$ is either 0 or at least εn), we have

$$\left\lfloor \frac{t}{10} \right\rfloor \cdot \frac{\varepsilon^2 n^2}{2} \leq n^2,$$

so $t \leq \frac{20}{\varepsilon^2}$. Hence the diameter of $\mathcal{H}'$ is at most $\ell := \frac{20}{\varepsilon^2} + 4$, as required. □

As alluded to in Section 2, we can already prove Corollary 2.8, restated here, which is a weakening of Theorem 2.9, with only an asymptotic upper bound, which depends on ℓ, on the number of edges.

Corollary 2.8. *Let $1/n \ll 1/\ell \ll \epsilon \ll 1$, and let $\mathcal{H}$ be an n-vertex hypergraph with no odd pseudocycles of length at most ℓ. Then $e(\mathcal{H}) \leq f(n) + \epsilon n^3$.*

This bound will be used in the proof of Proposition 6.6. Note that the analogous bound on the extremal number of proper odd tight cycles follows from Theorem 6.2.

Proof of Corollary 2.8. Assume the opposite, that $e(\mathcal{H}) \geq f(n) + \varepsilon n^3$. Applying Proposition 6.4 with the parameters $\ell/4$ and $\varepsilon/2$, we obtain a hypergraph $\mathcal{H}' \subseteq \mathcal{H}$ with at least $f(n) + \varepsilon n^3/2$ edges whose diameter is at most $\ell/4$. $\mathcal{H}'$ contains no odd pseudocycles of length at most ℓ, so by Proposition 6.3, it contains no odd pseudocycles. Hence we may apply Theorem 2.4 to obtain a good colouring of $\partial\mathcal{H}'$ with $e(\mathcal{H}') > f(n)$ cherries, contradicting Theorem 2.5. □

The following proposition gives a near-optimal lower bound on the vertex degrees in a largest hypergraph on n vertices with no short odd pseudocycles.

Proposition 6.5. *Let $1/n \ll 1/\ell \ll \varepsilon \ll 1$, and let $\mathcal{H}$ be an n-vertex hypergraph with no odd pseudocycles of length at most ℓ, which maximises the number of edges under these conditions. Then $d(u) \geq (3\alpha - \varepsilon)n^2$ for every vertex u.*

Proof. Given vertices u and v in $\mathcal{H}$, consider the hypergraph $\mathcal{H}_{uv}$ obtained from $\mathcal{H}$ by removing all edges containing v and then adding the edge $e-u+v$, for each edge e that contains u but not v. Observe that $\mathcal{H}$ has no odd pseudocycles of length at most ℓ; indeed, if there were such a cycle then we could replace each instance of v by u to obtain an odd pseudocycle of the same length in $\mathcal{H}$ (whereby it is important that $\mathcal{H}_{uv}$ has no edges containing both u and v), a contradiction. Since $e(\mathcal{H}_{uv}) \geq e(\mathcal{H}) - d(v) + d(u) - n$ and by maximality of $\mathcal{H}$, we have $d(v) \geq d(u) - n$. Since u and v were arbitrary, this implies that the maximum and minimum degrees of $\mathcal{H}$ differ by at most n. In particular, using $e(\mathcal{H}) \geq f(n) = \alpha n^3 + o(n^3)$, which follows from the maximality of $\mathcal{H}$ and Proposition 4.1,

$$\delta(\mathcal{H}) \geq \frac{3e(\mathcal{H})}{n} - n \geq \frac{3f(n)}{n} - n \geq (3\alpha - \varepsilon)n^2. \qquad \square$$

Next, we prove a stability version of the previous proposition.

Proposition 6.6. *Let $1/n \ll 1/\ell \ll \varepsilon_1 \ll \varepsilon_2 \ll 1$, and let $\mathcal{H}$ be an n-vertex 3-uniform hypergraph with no odd pseudocycles of length at most ℓ. If $e(\mathcal{H}) \geq f(n) - \varepsilon_1 n^3$ then $d(u) \leq (3\alpha + \varepsilon_2)n^2$ for every vertex u.*

Proof. Let $\mu = \sqrt{\varepsilon_1} \leq \varepsilon_2/10$. Let X be the set of vertices x with $d(x) \leq 3(\alpha+\mu)n^2$. Then $e(\mathcal{H}) \geq (n - |X|)(\alpha+\mu)n^2$. By Corollary 2.8 (and the properties of $f(n)$) we also have $e(\mathcal{H}) \leq (\alpha + \varepsilon_1)n^3$. Putting the two inequalities together, we get

$$\begin{aligned} &(\alpha+\varepsilon_1)n^3 \geq (n - |X|)(\alpha+\mu)n^2 \\ \implies \quad &|X| \geq \frac{(\alpha+\mu)n - (\alpha+\varepsilon_1)n}{\alpha+\mu} = \frac{\mu - \varepsilon_1}{\alpha+\mu} \cdot n \geq \mu n. \end{aligned}$$

Let u be a vertex of maximum degree in $\mathcal{H}$, and let X' be a subset of X of size $t := \mu n$. We may assume $u \notin X'$ because otherwise $d_{\mathcal{H}}(u) \leq (3\alpha + 3\mu)n^2 \leq (3\alpha + \varepsilon_2)n^2$, as required. Now consider the hypergraph $\mathcal{H}_1$ formed in two steps as follows. First, define $\mathcal{H}_0 = \mathcal{H} \setminus X'$; then $e(\mathcal{H}_0) \geq e(\mathcal{H}) - t \cdot 3(\alpha+\mu)n^2$ and $d_{\mathcal{H}_0}(u) \geq d_{\mathcal{H}}(u) - tn$. Second, let $\mathcal{H}_1$ be the hypergraph obtained by adding $|X'|$ copies of u to $\mathcal{H}_0$. Then

$$\begin{aligned} e(\mathcal{H}_1) &\geq e(\mathcal{H}_0) + t \cdot d_{\mathcal{H}_0}(u) \\ &\geq e(\mathcal{H}) - t \cdot 3(\alpha+\mu)n^2 + t \cdot (d_{\mathcal{H}}(u) - tn) \\ &\geq f(n) - \varepsilon_1 n^3 + t \cdot (d_{\mathcal{H}}(u) - tn - 3(\alpha+\mu)n^2) \\ &= f(n) - \varepsilon_1 n^3 + \mu n \cdot (d_{\mathcal{H}}(u) - (3\alpha + 4\mu)n^2). \end{aligned}$$

Notice that $\mathcal{H}_1$ has no odd pseudocycles of length at most ℓ. Thus, by Corollary 2.8, we have $e(\mathcal{H}_1) \leq f(n) + \varepsilon_1 n^3$. Hence, using $\mu = \sqrt{\varepsilon_1} \leq \varepsilon_2/10$,

$$d_{\mathcal{H}}(u) \leq (3\alpha + 4\mu)n^2 + (2\varepsilon_1/\mu)n^2 \leq (3\alpha + \varepsilon_2)n^2,$$

as required. $\square$

6.2 The structure of odd-pseudocycle-free graphs

We now prove the main result in the section, Theorem 6.1. The starting point of the proof uses the relation between hypergraphs with no odd pseudocycles and good colourings of K_n, as well as the stability result about cherries from the previous section, to conclude the following: there is a coloured graph G with a nice structure such that almost all cherries in G are triples in $\mathcal{H}$ and vice versa. This readily implies the existence of a partition $\{A, B\}$ of the vertices such that $|A| \approx \beta n$ and for almost every vertex u in $\mathcal{H}$ the following holds: almost all vertices in A are joined to almost all $A \times B$ pairs, and almost all vertices in B are joined to almost all $A^{(2)}$ pairs. The main difficulty of the proof lies in showing that there is such a partition for which *every* vertex in A is joined to almost all pairs in $A \times B$, and similarly for vertices in B. This is achieved in Claim 6.7 and the main idea is to compare several graphs obtained by modifying the triples containing a given vertex. Given a partition as above, to conclude the proof, we argue (using the fact that $\mathcal{H}$ has no short odd pseudocycles) that the number of AAB "non-edges" exceeds the number of AAA and ABB edges, unless all of these numbers are 0. The maximality of $\mathcal{H}$ implies that all these numbers are indeed 0, meaning that $\mathcal{H}$ has all AAB edges and no AAA, ABB edges.

Proof of Theorem 6.1. Let $\varepsilon_7 = 0.1$ and let $\varepsilon_1, \ldots, \varepsilon_6$, and ℓ satisfy

$$0 < 1/\ell \ll \varepsilon_1 \ll \ldots \ll \varepsilon_7.$$

Let $\mathcal{H}'$ be a subgraph of $\mathcal{H}$ on the same vertex set with at least $e(\mathcal{H}) - \varepsilon_1 n^3$ edges, that has diameter at most $\ell/4$; such $\mathcal{H}'$ exists by Proposition 6.4. By Proposition 6.3, $\mathcal{H}'$ has no odd pseudocycles, so by Theorem 2.4, there is a good colouring of $\partial\mathcal{H}'$.

Extending the good colouring of $\partial\mathcal{H}'$ arbitrarily to also cover vertex pairs which are not in the shadow, we obtain a coloured graph (recall that this is a complete graph whose edges are either blue or oriented and red) G' on vertex set $V := V(\mathcal{H})$, such that every edge in $\mathcal{H}'$ is a cherry in G'. By maximality of $\mathcal{H}$, we have $c(G') \geq e(\mathcal{H}') \geq e(\mathcal{H}) - \varepsilon_1 n^3 \geq f(n) - \varepsilon_1 n^3$.

Thus, by Theorem 5.2, there is a graph G satisfying (a)–(c) in Theorem 5.2 on vertex set V. That is, G has at least as many cherries as G', all but at most $\varepsilon_2 n^3$ cherries in G are cherries in G', and V can be partitioned into sets $X_1, \ldots, X_k$ such that: $G[X_i]$ is blue for $i \in [k]$; $|X_i| = (\beta \pm \varepsilon_2)n \cdot (|X_i| + \ldots + |X_k|)$ for $i \in [k]$; and all $X_i \times X_j$ pairs in G are red and oriented towards X_i, for $1 \leq i < j \leq k$. Recall that $\beta = \frac{3-\sqrt{3}}{2}$ was defined in (5).

Define $X_{>i} := X_{i+1} \cup \ldots \cup X_k$, and define $X_{\geq i}$ analogously. Let H be the subgraph of G whose edges are either pairs in $X_i \times X_i$ that are in at least $(|X_{i+1}| + \ldots + |X_k|) - \varepsilon_3 n$ triples in $(X_i \times X_i \times X_{>i}) \cap E(\mathcal{H})$, or pairs in $X_i \times X_j$, where $i < j$, that are in at least $|X_i| - \varepsilon_3 n$ triples in $(X_i \times X_i \times X_j) \cap E(\mathcal{H})$.

Denoting the number of non-edges in H by $\bar{e}(H)$, we have that the number of cherries in G that are not edges in $\mathcal{H}$ is at least $\bar{e}(H) \cdot \varepsilon_3 n/3$. Recall that $e(\mathcal{H}') \geq e(\mathcal{H}) - \varepsilon_1 n^3 \geq f(n) - \varepsilon_1 n^3$ and that all edges in $\mathcal{H}'$ are cherries in G'. But $c(G') \le f(n)$ (by Theorem 2.5), so all but $\varepsilon_1 n^3$ cherries in G' are edges in $\mathcal{H}'$ and thus in $\mathcal{H}$. Since there are at most $\varepsilon_2 n^3$ cherries in G that are not cherries in G', it follows that all but at most $(\varepsilon_1 + \varepsilon_2)n^3 \le 2\varepsilon_2 n^3$ cherries in G are edges in $\mathcal{H}$. Hence $\bar{e}(H) \cdot \varepsilon_3 n/3 \le 2\varepsilon_2 n^3$, showing $\bar{e}(H) \le (6\varepsilon_2/\varepsilon_3)n^3 \le \varepsilon_3 n^2$.

Let k_0 be the maximum i such that $|X_i| \ge \varepsilon_4 n$. Define subsets $X_i' \subseteq X_i$ as follows: if $i < k_0$ let X_i' be the set of vertices in X_i that have degree at least $|X_i| - \varepsilon_4 n$ in $H[X_i]$ and degree at least $|X_{>i}| - \varepsilon_4 n$ in $H[X_i, X_{>i}]$; if $i \ge k_0$, define $X_i' := \emptyset$. Since $(\varepsilon_4 n/2)\sum_{i<k_0} |X_i \setminus X_i'| \le \bar{e}(H) \le \varepsilon_3 n^2$ and $|X_{\ge k_0}| \le 10\varepsilon_4 n$ (using c(iv)), we have

$$\sum_{i\in[k]} |X_i \setminus X_i'| \le 10\varepsilon_4 n + (2\varepsilon_3/\varepsilon_4)n \le 20\varepsilon_4 n.$$

Let $X := X_1' \cup \ldots \cup X_k'$ and $Y := V \setminus X$. We have seen that $|Y| \le 20\varepsilon_4 n \le \varepsilon_5 n$.

For $v \in V$, let $N(v)$ be the *link* of v, namely the graph spanned by pairs uw such that $uvw \in E(\mathcal{H})$. Write $A := X_1'$ and $B := X \setminus X_1'$.

Claim 6.7. *One of the graphs $N(u)[A]$ and $N(u)[A, B]$ has at most $\varepsilon_6 n^2$ non-edges, for every $u \in V$.*

Proof. Let $\varepsilon_5 \ll \mu \ll \varepsilon_6$. Note that the claim holds for all $u \in X$, so it suffices to prove it for $u \in Y$. Fix such u.

Let $\mathcal{F}$ be the hypergraph on vertex set X whose edges are all $X_i'X_i'X_j'$ triples with $1 \le i < j \le k_0$. We will construct two hypergraphs $\mathcal{F}_i^+$ (for $i \in \{1, 2\}$), that consist of $\mathcal{F}$ with one additional vertex u_i, which is a suitable modification of u, and that have no odd pseudocycles of length at most $\ell/10$. We will argue that if both $N(u)[A]$ and $N(u)[A, B]$ have at least $\varepsilon_6 n^2$ non-edges then $d_{\mathcal{F}_i^+}(u_i) > (3\alpha + \mu)n^2$ for some $i \in [2]$, contradicting Proposition 6.6.

Let F_0 be the graph on vertex set X with edges $E(H) \cap E(N(u))$. Recall that vertices in X_i' have at most $2\varepsilon_4 n$ non-neighbours in $H[X_{>i}']$. Thus, using Proposition 6.5 for a lower bound on $d_{\mathcal{H}}(u)$, we have $e(F_0) \ge d_{\mathcal{H}}(u) - |Y| \cdot n - |X| \cdot 2\varepsilon_4 n \ge (3\alpha - 10\varepsilon_5)n^2$. We modify F_0 as follows, while possible: remove each edge xy satisfying: $x, y \in A$ and x has degree 1 in A; or $x \in A$, $y \in B$, and x has degree 1 into B or y has degree 1 into A. Call the resulting graph F and notice that $|E(F_0) \setminus E(F)| \le 2n$, implying that

$$e(F) \ge (3\alpha - 20\varepsilon_5)n^2. \tag{12}$$

Recall that $\mathcal{F}$ is the hypergraph on vertex set X whose edges are all $X_i'X_i'X_j'$ triples with $1 \le i < j \le k_0$, and let $\mathcal{F}^+$ be the hypergraph obtained by adding the vertex u to $\mathcal{F}$ along with all edges uvw such that $vw \in E(F)$. We argue that $\mathcal{F}^+$ has no odd pseudocycles of length at most $\ell/10$. To do so, we prove the following.

Let $xy, vw \in E(H)$, and let P be a pseudopath in $\mathcal{F}$ from xy to vw on t vertices. Then there is a pseudopath P' in $\mathcal{H}$ from xy to vw of order t (if $t \in \{2, 3\}$) or $t + 3$ (otherwise). (13)

We prove (13) by induction on t. If $t = 2$ we can take $P' = P$. Suppose that $t = 3$, so $P = xyw$. Let i_1, i_2, i_3 be such that $x \in X'_{i_1}$, $y \in X'_{i_2}$ and $w \in X'_{i_3}$. Since $xy \in E(H)$, we know that for almost every $a \in X'_{i_3}$ the following holds: $xya \in E(\mathcal{H})$ and $ya \in E(H)$; pick such an a with $a \neq w$. Similarly, yab and ywb are edges in $\mathcal{H}$ for almost every $b \in X'_{i_1}$; pick such b. The path $xyabyw$ satisfies the requirements.

Next, suppose that $t = 4$, so $P = xyvw$. Let i_1, i_2, i_3, i_4 be such that $x \in X'_{i_1}$, $y \in X'_{i_2}$, $v \in X'_{i_3}$ and $w \in X'_{i_4}$. As $xy \in E(H)$, almost all $a \in X'_{i_3}$ satisfy $xya \in E(\mathcal{H})$ and $ya \in E(H)$; fix such a. Similarly, almost all $c \in X'_{i_2}$ satisfy $cvw \in E(\mathcal{H})$, $cv \in E(H)$ and $ac \in E(H)$; fix such c. Finally, almost every $b \in X'_{i_3}$ satisfies $yab, abc, bcv \in E(\mathcal{H})$; fix such b. Then $P' = xyabcvw$ satisfies the requirements.

Finally, suppose that $t \geq 5$, and write $P = v_1 \dots v_t$, so $x = v_1$, $y = v_2$, $v = v_{t-1}$ and $w = v_t$. Let i_j be such that $v_j \in X'_{i_j}$ for $j \in [t]$. As usual, since $xy = v_1v_2 \in E(H)$, almost all $a \in X'_{i_3}$ satisfy: $v_2a \in E(H)$ and $v_1v_2a \in E(\mathcal{H})$. Let $Q = v_2av_4 \dots v_t$. Then Q is a pseudopath in $\mathcal{F}$ of order $t-1$ that starts and ends with edges in H. By induction, there is a pseudopath Q' in $\mathcal{H}$ from v_2a to $v_{t-1}v_t$ of order $t+2$. Then we can take $P' = v_1Q'$, completing the proof of (13).

Now suppose that $C = v_1 \dots v_t$ is a pseudocycle in $\mathcal{F}^+$, where $t \leq \ell/10$. We need to show that t is divisible by 3. If C does not go through u, then C is in $\mathcal{F}$, implying that t is indeed divisible by 3. So we may assume that C goes through u at least once. This shows that C can be written as $uP_1u \dots uP_k$, where P_i is a pseudopath in $\mathcal{F}$ whose first two vertices and last two vertices form edges in F. It follows from (13) that for each $i \in [k]$ there is a pseudopath P'_i in $\mathcal{H}$ whose first two vertices and last two vertices match those of P_i and whose order satisfies $|P'_i| - |P_i| \in \{0, 3\}$. Then $C' := uP'_1u \dots uP'_k$ is a cycle in $\mathcal{H}$ with $|C'| \leq |C| + 3k \leq 4|C| \leq \ell$ and $|C'| \equiv |C| \pmod 3$. By the properties of $\mathcal{H}$, we have that $|C'|$ is divisible by 3, implying that $|C|$ is divisible by 3, as required.

Let A_0 and A_1 be the sets of vertices in A incident with AA and AB edges in F, respectively (that is, $a_0 \in A_0$ if F contains an edge a_0x with $x \in A$). To show that A_0 and A_1 are disjoint, assume that $a_1 \in A_0 \cap A_1$, so that there is a path $a_0a_1b_0$ in F with $a_0, a_1 \in A$ and $b_0 \in B$. By construction of F, b_0 has an F-neighbour $a_2 \in A - A_1$, so $a_0a_1b_0a_2$ is a path in F. Let a_3 and b_1 be arbitrary vertices in A and B, respectively (distinct from previously chosen vertices). Then $a_0a_1ub_0a_2a_3b_1$ is cycle of length 7 in $\mathcal{F}^+$, a contradiction.

Let B_1 be the set of vertices in B incident with AB edges in F. We claim that B_1 is independent in F. Indeed, otherwise there is a path $a_1b_1b_2a_2$ in F, using a similar argument to the above paragraph. Now, choosing $a_3, a_4 \in A$ and $b_3 \in B$ to be arbitrary unused vertices, we obtain a cycle $a_1b_1ub_2a_2a_3b_3a_4$ of length 8 in $\mathcal{F}^+$ and reach a contradiction.

Let F_1 and F_2 be graphs on vertex set X, defined as follows: $E(F_1) = A \times B$ and $E(F_2) = A^{(2)} \cup E(F[B])$. Now define $\mathcal{F}_i^+$ to be the graph obtained from $\mathcal{F}$ by adding a new vertex u_i and edges u_ivw such that $vw \in E(F_i)$, for $i \in [2]$. Thus $\mathcal{F}_i^+$ and $\mathcal{F}^+$ differ only on edges touching u_i or u. We claim that $\mathcal{F}_i^+$ has no odd pseudocycles of length at most $\ell/10$. Indeed, this is easy to see for $i = 1$, because we can think of $\mathcal{F}_1^+$ as obtained by extending X'_1 by one vertex. To see that this also holds for $i = 2$, notice that in $\mathcal{F}_2^+$, the AAB and BBB triples are in different strong components, so any pseudocycle C in $\mathcal{F}_2^+$ is either a pseudocycle in $\mathcal{F}^+$ or consists only of edges containing exactly two vertices from A.

Notice that $e(\mathcal{F}_i^+) \geq c(G) - |Y|n^2 \geq f(n) - (\varepsilon_1 + \varepsilon_5)n^3 \geq f(n) - 2\varepsilon_5 n^3$, because all cherries in G that do not touch Y are edges in $\mathcal{F}$ and $c(G) \geq c(G') \geq f(n) - \varepsilon_1 n^3$. Using this lower bound and the fact that $\mathcal{F}_i^+$ has no odd pseudocycles of length at most $\ell/10$, Proposition 6.6 implies that $d_{\mathcal{F}_i^+}(u_i) \leq (3\alpha + \mu)n^2$. Since $d_{\mathcal{F}^+}(u) = e(F) \geq (3\alpha - 20\varepsilon_5)n^2$ (see (12)), we have $e(F_i) - e(F) = d_{\mathcal{F}_i^+}(u_i) - d_{\mathcal{F}^+}(u) \leq (\mu + 20\varepsilon_5)n^2 \leq 2\mu n^2$ for $i \in [2]$.

To finish, suppose first that $|A_0| \geq |B_1|$. Recalling that F and F_1 coincide on B, and that F has no edges in $(A_1 \cup B_1) \times A_0$ or $A_1^{(2)}$, we have

$$\begin{aligned} 2\mu n^2 \geq e(F_2) - e(F) &\geq -|A_1||B_1| + |A_0||A_1| + \binom{|A_1|}{2} + \bar{e}(F[A_0]) \\ &\geq \frac{|A_1|^2}{2} + \bar{e}(F[A_0]) + O(n). \end{aligned}$$

It follows that $|A_1| \leq 5\mu^{1/2}n$ and $\bar{e}(F[A_0]) \leq 5\mu n^2$. Altogether $\bar{e}(F[A]) \leq |A_1|\,n + \bar{e}(F[A_0]) \leq 10\mu^{1/2}n^2 \leq \varepsilon_6 n^2$. Since $F[A] \subseteq N(u)[A]$, Claim 6.7 is proved in this case.

Now we consider the remaining case, namely that $|A_0| \leq |B_1|$. Let $B_0 = B \setminus B_1$, and recall that F has no edges in $B_1^{(2)}$ or in $A_0 \times B_1$. Using $|A| \geq |B| = |B_0| + |B_1|$,

$$\begin{aligned} 2\mu n^2 &\geq e(\mathcal{F}_1^+) - e(\mathcal{F}^+) \\ &\geq -\binom{|A_0|}{2} - \binom{|B_0|}{2} - |B_0||B_1| + |A||B_0| + |A_0||B_1| + \bar{e}(F[A_1, B_1]) \\ &\geq |A_0|(|B_1| - |A_0|) + |B_0|(|A| - |B_0| - |B_1|) + \frac{|A_0|^2}{2} + \frac{|B_0|^2}{2} + \bar{e}(F[A_1, B_1]) + O(n) \\ &\geq \frac{|A_0|^2}{2} + \frac{|B_0|^2}{2} + \bar{e}(F[A_1, B_1]) + O(n). \end{aligned}$$

Thus, we have $|A_0|, |B_0| \leq 5\mu^{1/2}n$ and $\bar{e}(F[A_1, B_1]) \leq 5\mu n^2$. This implies that $\bar{e}(F[A, B]) \leq |A_0|\,n + |B_0|\,n + \bar{e}(F[A_1, B_1]) \leq \varepsilon_6 n^2$, proving Claim 6.7. □

Let A^* be the set of vertices u such that $N(u)[A, B]$ has at most $\varepsilon_6 n^2$ non-edges, and let $B^* := V \setminus A^*$. Note that $A \subseteq A^*$, and by Claim 6.7, for every $u \in B^*$ the graph $N(u)[A]$ has at most $\varepsilon_6 n^2$ non-edges. Let t_1 be the number of $A^*A^*A^*$ triples in $\mathcal{H}$, let t_2 be the number of $A^*B^*B^*$ triples in $\mathcal{H}$, and let s be the number of $A^*A^*B^*$ triples that are not edges in $\mathcal{H}$. Let $\mathcal{H}^*$ be the hypergraph obtained from $\mathcal{H}$ by removing all $A^*A^*A^*$ and $A^*B^*B^*$ triples and adding all missing $A^*A^*B^*$ triples. Then $\mathcal{H}^*$ has no odd pseudocycle of length at most ℓ; this follows from observing that every pseudocycle in $\mathcal{H}^*$ is either a pseudocycle in $\mathcal{H}$ or each of its edges has exactly two vertices in A^*. Moreover, $e(\mathcal{H}^*) - e(\mathcal{H}) = s - (t_1 + t_2)$. By maximality of $\mathcal{H}$ we have $s \leq t_1 + t_2$.

Claim 6.8. *$t_1 \leq \varepsilon_7 s$.*

Proof. Let $\varepsilon_6 \ll \mu \ll \varepsilon_7$.

We first show that for every distinct $u, v \in A^*$, there are at most μn vertices $w \in A^*$ such that $uvw \in E(\mathcal{H})$.

Suppose there exist $u, v \in A^*$ violating this. Let W be the set of vertices $w \in A^*$ such that $uvw \in E(\mathcal{H})$, so $|W| \geq \mu n$. Consider the graph $(N(u) \cap N(v))[W, B]$; its edges are pairs wb such that $w \in W$, $b \in B$, and $uwb, vwb \in E(\mathcal{H})$. This graph has at most $2\varepsilon_6 n^2$ non-edges, by Claim 6.7. Thus there exists $b \in B$ with at least $\frac{1}{2}\mu n$ neighbours in the aforementioned graph; denote its set of neighbours by W'. Now, by Claim 6.7, b is adjacent in $\mathcal{H}$ to all but at most $\varepsilon_6 n^2$ pairs in W', so there exists a triple $w_1 w_2 b \in E(\mathcal{H})$ with $w_1, w_2 \in W'$. Thus uvw_1bw_2 is a pseudocycle of length 5, contradiction.

To finish the argument, we count the four-tuples

$$Q := \{\{u, v, w, z\} : u, v, w \in A^*, z \in B^*, uvw \in E(\mathcal{H}), uvz \notin E(\mathcal{H})\}$$

in two different ways. For each vertex $b \in B^*$ and $A^*A^*A^*$ triple $uvw \in E(\mathcal{H})$, at least one of the triples uvb, uwb, vwb is not in $E(\mathcal{H})$ (since otherwise $\mathcal{H}$ has a 4-cycle), so $|Q| \geq t_1|B^*|$. On the other hand, it follows from the above paragraph that any $A^*A^*B^*$ triple $uvz \notin E(\mathcal{H})$ extends to at most μn elements of Q, so $|Q| \leq s\mu n$. Hence

$$t_1 \leq \frac{|Q|}{|B^*|} \leq \frac{s\mu n}{|B^*|} \leq \varepsilon_7 s,$$

as claimed. □

Claim 6.9. $t_2 \leq 2s/3$.

Proof. Let $\varepsilon_6 \ll \mu \ll \varepsilon_7$.

To begin with, we show that if uvw is an $A^*B^*B^*$ triple in $\mathcal{H}$ (with $u \in A^*$) then one of the pairs uv and vw is in at most μn triples of form $A^*A^*B^*$ in $\mathcal{H}$. Fix an $A^*A^*B^*$ triple $uvw \in E(\mathcal{H})$.

Let W' (resp. V') be the set of vertices $a \in A^*$ such that $uwa \in E(\mathcal{H})$ (resp. $uva \in E(\mathcal{H})$). Suppose that $|W'|, |V'| \geq \mu n$. Consider the graph $(N(v) \cap N(w))[W', V']$. By Claim 6.7, this graph contains an edge a_1a_2, i.e. we have a_1a_2w, $a_1a_2v \in E(\mathcal{H})$. By definition of W' and V', the triples uwa_1 and uva_2 are in $\mathcal{H}$. Hence uwa_1a_2v is a cycle of length 5, contradiction.

Let F be an auxiliary bipartite graph with parts A^* and B^* such that uv is an edge of F whenever (i) there is an $A^*B^*B^*$ triple in $\mathcal{H}$ containing uv, and (ii) the number of $A^*A^*B^*$ triples containing uv is at most μn. By the previous paragraph, each $A^*B^*B^*$ triple in $\mathcal{H}$ contains an edge of F, so

$$t_2 \leq |B^*| \cdot e(F) \leq 0.4n \cdot e(F).$$

Moreover, we claim that $d_F(v) \leq \mu n$ for every $v \in B^*$. Indeed, by (ii), the graph $N(v)[A^*]$ has at least $d_F(v)(|A^*| - \mu n)/2$ non-edges. If $d_F(v) > \mu n$ then this quantity is larger than $2\varepsilon_6 n^2$, contradicting Claim 6.7. Also using (ii), we conclude that

$$s \geq \sum_{v \in B^*} d_F(v) \cdot (|A^*| - d_F(v) - \mu n) \geq 0.6n \cdot e(F).$$

It follows that $t_2 \leq 2s/3$, as claimed. □

The last two claims, and the choice $\varepsilon_7 = 0.1$, say, show that $(t_1 + t_2) \leq 0.8s$. Since $s \leq t_1 + t_2$ this implies that $t_1 = t_2 = s = 0$. That is, all $A^*A^*B^*$ triples are edges in $\mathcal{H}$ (and there are no $A^*A^*A^*$ or $A^*B^*B^*$ edges). This proves Theorem 6.1. □

7 Open problems

There are several natural extensions of our result. Firstly, one could prove Conjecture 1.1, or perhaps determine the density of $C_\ell^{(3)}$ for smaller values of ℓ, say $\ell \leq 100$. Although we do not state our bound on ℓ explicitly, this would not be too cumbersome, since it is a polynomial in ε_7, and we set $\varepsilon_7 = 0.1$.

Of course our result should not extend to *all* values of $\ell \equiv 1$ or $2 \pmod 3$, since for $\ell = 4$, the tight cycle $C_4^{(3)}$ is the same as the tetrahedron $K_4^{(3)}$. Here the famous conjecture of Turán says that $\pi(K_4^{(3)}) = 5/9$, which is attained by a wide family of extremal constructions [4, 12, 21, 30]. Curiously, Fon-Der-Flaass showed that the conjectured extremal constructions $K_4^{(3)}$-free graphs can be constructed from oriented graphs in a manner reminiscent of Definition 2.3. Specifically Fon-Der-Flass [12] showed that if D is an oriented graph with no induced directed 4-cycles, then the 3-graph formed by induced copies of $\{ab, ac\}$ and $\{ab, bc, ca\}$ will be $K_4^{(3)}$-free.

A second interesting direction is determining the Turán density of r-uniform tight cycles for $r \geq 4$. For this, we do not even know of a conjectured optimal construction. Moreover, our characterisation of odd-pseudocycle-free hypergraphs (Theorem 2.4) does not have an obvious extension, as the straightforward extension of Definition 2.3 is too strong.

Recall that in Theorem 2.9 we prove an almost tight result (up to a constant additive error) for the Turán number of the family of pseudocycles of length ℓ, for all $\ell \leq L$ which are not divisible by 3, and large enough L. It is plausible that the same could be proved for $\mathcal{C}_\ell$ for large enough ℓ which is not divisible by 3. Namely, it is likely that $\mathrm{ex}(n, \mathcal{C}_\ell) \leq f(n) + O(1)$ for such ℓ? To tackle this, one is likely to require stability arguments, perhaps like those we used in Section 6. We remark that Liu, Mubayi, and Reiher [24] present a unified framework for tackling stability problems for a large class of hypergraph families. Unfortunately, this does not seem to be applicable in our case.

As mentioned in the introduction, there are many other specific 3-uniform hypergraphs for which determining the Turán density would be very interesting. Let us point out one conjecture which is perhaps less well known, and which can be found for instance in [26].

Conjecture 7.1. *Let $\mathcal{C}_5^-$ be the 3-uniform hypergraph obtained from the tight 5-cycle $\mathcal{C}_5^3$ by removing one edge. The Turán density of $\mathcal{C}_5^-$ is $\frac{1}{4}$.*

As in our case, one conjectured extremal hypergraph is an iterated construction; one may take a complete 3-partite 3-uniform hypergraph and then repeat the same construction recursively within each of the three parts. Similarly to our result, Balogh and Haoran [2] recently proved that the Turán density of $\mathcal{C}_\ell^-$, for sufficiently large ℓ which is not divisible by 3, is $\frac{1}{4}$.

Acknowledgements

We would like to thank Jozsef Balogh, Xizhi Liu, Dhruv Mubayi, Yuejian Peng, Oleg Pikhurko, and Alexander Sidorenko for bringing to our attention several important references. We would also like to thank the anonymous referee for their helpful comments.

References

[1] R. Baber and J. Talbot, *New Turán densities for 3-graphs*, Electron. J. Combin. **19** (2012), Paper P40. 2, 4

[2] J. Balogh and H. Luo, *Turán density of long tight cycle minus one hyperedge*, arXiv:2303.10530 (2023). 32

[3] B. Bollobás, I. Leader, and C. Malvenuto, *Daisies and other Turán problems*, Combin. Probab. Comput. **20** (2011), no. 5, 743–747. 2

[4] W. G. Brown, *On an open problem of Paul Turán concerning* 3*-graphs*, Studies in pure mathematics, Birkhäuser, Basel, 1983, pp. 91–93. 2, 31

[5] M. Bucić, J. Cooper, D. Král', S. Mohr, and D. Munhá Correia, *Uniform turán density of cycles*, Trans. Amer. Math. Soc. (2023). 4

[6] F. Chung and R. Graham, *Erdős on graphs*, A K Peters, Ltd., Wellesley, MA, 1998, His legacy of unsolved problems. 4

[7] D. De Caen and Z. Füredi, *The maximum size of 3-uniform hypergraphs not containing a Fano plane*, J. Combin. Theory Ser. B **78** (2000), no. 2, 274–276. 2

[8] P. Erdős, *On some extremal problems on r-graphs*, Discrete Math. **1** (1971), 1–6. 2

[9] P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), no. 12, 1087–1091. 2

[10] P. Erdős and V. T. Sós, *On Ramsey—Turán type theorems for hypergraphs*, Combinatorica **2** (1982), no. 3, 289–295. 4

[11] V. Falgas-Ravry and E. R. Vaughan, *Turán H-densities for 3-graphs*, Electron. J. Combin. (2012), P40. 2, 5, 6, 11

[12] D. G. Fon-Der-Flaass, *A method for constructing* $(3,4)$*-graphs*, Mat. Zametki **44** (1988), no. 4, 546–550, 559. 2, 31

[13] P. Frankl and Z. Füredi, *An exact result for 3-graphs*, Discrete Math. **50** (1984), no. 2-3, 323–328. 2

[14] Z. Füredi and M. Simonovits, *Triple systems not containing a Fano configuration*, Combin. Probab. Comput. **14** (2005), no. 4, 467–484. 2

[15] R. Glebov, D. Kráľ, and J. Volec, *A problem of Erdős and Sós on 3-graphs*, Israel J. Math. **211** (2016), no. 1, 349–366. 4

[16] H. Huang, *On the maximum induced density of directed stars and related problems*, SIAM J. Discrete Math. **28** (2014), no. 1, 92–98. 6, 11

[17] H. Huang and J. Ma, *On tight cycles in hypergraphs*, SIAM J. Discrete Math. **33** (2019), no. 1, 230–237. 4

[18] B. Janzer, *Large hypergraphs without tight cycles*, Comb. Theory **1** (2021), Paper No. 12, 4. 4

[19] P. Keevash, *Hypergraph Turán problems*, Surveys in combinatorics 2011, London Math. Soc. Lecture Note Ser., vol. 392, Cambridge Univ. Press, Cambridge, 2011, pp. 83–139. 2, 4, 22

[20] P. Keevash and B. Sudakov, *The Turán number of the Fano plane*, Combinatorica **25** (2005), 561–574. 2

[21] A. V. Kostochka, *Lower bound of the Hadwiger number of graphs by their average degree*, Combinatorica **4** (1984), no. 4, 307–316. 2, 31

[22] S. Letzter, *Hypergraphs with no tight cycles*, Proc. Amer. Math. Soc. **151** (2023), no. 02, 455–462. 4

[23] H. Liu, O. Pikhurko, M. Sharifzadeh, and K. Staden, *Stability from graph symmetrisation arguments with applications to inducibility*, J. London Math. Soc. **108** (2023), 1121–1162. 11

[24] X. Liu, D. Mubayi, and C. Reiher, *A unified approach to hypergraph stability*, J. Combin. Theory Ser. B **158** (2023), 36–62. 31

[25] W. Mantel, *Problem 28*, Wiskundige Opgaven **10** (1907), 60–61. 1

[26] D. Mubayi, O. Pikhurko, and B. Sudakov, *Hypergraph Turán problem: some open questions*, AIM workshop problem lists (2011), manuscript, homepages.warwick.ac.uk/~maskat/Papers/TuranQuestions.pdf. 2, 31

[27] D. Mubayi and V. Rödl, *On the Turán number of triple systems*, J. Combin. Theory Ser. A **100** (2002), no. 1, 136–152. 10

[28] O. Pikhurko, *On possible Turán densities*, Israel J. Math. **201** (2014), 415–454. 4

[29] A. A. Razborov, *On 3-hypergraphs with forbidden 4-vertex configurations*, SIAM J. Discrete Math. **24** (2010), no. 3, 946–963. 2

[30] A. A. Razborov, *On the Fon-der-Flaass interpretation of extremal examples for Turán's* $(3,4)$-*problem*, Tr. Mat. Inst. Steklova **274** (2011), no. Algoritmicheskie Voprosy Algebry i Logiki, 269–290. 2, 31

[31] A. A. Razborov, *Flag algebras: an interim report*, The Mathematics of Paul Erdős II, Springer, 2013, pp. 207–232. 2

[32] C. Reiher, V. Rödl, and M. Schacht, *On a Turán problem in weakly quasirandom 3-uniform hypergraphs*, J. Eur. Math. Soc. (JEMS) **20** (2018), no. 5, 1139–1159. 4

[33] B. Sudakov and I. Tomon, *The extremal number of tight cycles*, Int. Math. Res. Not. IMRN (2022), no. 13, 9663–9684. 4

[34] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok **48** (1941), 436–452. 1

[35] ______, *Research problem*, Közl MTA Mat. Kutató Int. **6** (1961), 417–423. 2

[36] B. Wu, *An Irrational Turán Density via Hypergraph Lagrangian Densities*, Electron. J. Combin. **29** (2022), no. 3, Paper No. 3.62. 3

[37] Z. Yan and Y. Peng, *An irrational Lagrangian density of a single hypergraph*, SIAM J. Discrete Math. **36** (2022), no. 1, 786–822. 3