J. Eur. Math. Soc. 24, 303–339 (2022) DOI 10.4171/JEMS/1101

Agelos Georgakopoulos · John Haslegrave · Richard Montgomery · Bhargav Narayanan

Spanning surfaces in 3-graphs

Received January 15, 2019; revised March 24, 2021

Abstract. We prove a topological extension of Dirac's theorem suggested by Gowers in 2005: for any connected, closed surface $\mathcal{S}$, we show that any two-dimensional simplicial complex on n vertices in which each pair of vertices belongs to at least $\frac{n}{3} + o(n)$ facets contains a homeomorph of $\mathcal{S}$ spanning all the vertices. This result is asymptotically sharp, and implies in particular that any 3-uniform hypergraph on n vertices with minimum codegree exceeding $\frac{n}{3} + o(n)$ contains a spanning triangulation of the sphere.

Keywords. Extremal simplicial topology, spanning structures in hypergraphs, Dirac's theorem, triangulated surfaces

1. Introduction

In this paper, we extend the classical graph-theoretic result of Dirac [3] on spanning cycles to the setting of simplicial 2-complexes, or equivalently, the setting of 3-uniform hypergraphs (or 3-*graphs* for short). Dirac's theorem determines the best-possible minimum degree condition which guarantees that an n-vertex graph contains a *Hamiltonian cycle*, i.e., a cycle spanning the entire vertex set of the graph. A natural generalisation is to treat a "spanning cycle in a 3-graph" as a triangulation of the 2-sphere spanning the vertex set, and here we determine asymptotically the best-possible minimum codegree condition which guarantees the existence of such an object in an n-vertex 3-graph.

Dirac's theorem is of central importance in graph theory, and a number of different extensions to 3-graphs have previously been shown; see [14] for a broad overview. All these results – see [1, 9, 21, 22], for example – treat a "spanning cycle in a 3-graph" as

Agelos Georgakopoulos: Mathematics Institute, University of Warwick, CV4 7AL, United Kingdom; a.georgakopoulos@warwick.ac.uk

John Haslegrave: Mathematics Institute, University of Warwick, CV4 7AL, United Kingdom; j.haslegrave@cantab.net

Richard Montgomery: School of Mathematics, University of Birmingham, B15 2TT, United Kingdom; r.h.montgomery@bham.ac.uk

Bhargav Narayanan: Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA; narayanan@math.rutgers.edu

Mathematics Subject Classification (2020): Primary 05E45; Secondary 05C65, 05C35

a rigid pattern of interlocking edges with respect to some cyclic ordering of the underlying vertex set, an inherently one-dimensional notion. Here, we shall take a topological point of view, and consider a two-dimensional extension to 3-graphs.

To motivate our point of view, already implicit in the work of Brown, Erdős and Sós [24], we start by observing that a Hamiltonian cycle in a graph G is a set of edges of G such that the simplicial complex induced by these edges is homeomorphic to $\mathbb{S}^1$, the one-dimensional sphere, where, additionally, the 0-skeleton of this complex is the entire vertex set of G. By analogy, we define *a copy of the sphere* in a 3-graph $\mathcal{H}$ to be a set of edges of $\mathcal{H}$ such that the simplicial complex induced by these edges is homeomorphic to $\mathbb{S}^2$, the two-dimensional sphere, and we say that a copy of the sphere in a 3-graph $\mathcal{H}$ is *spanning* or *Hamiltonian* if the 0-skeleton of the associated simplicial complex is the entire vertex set of $\mathcal{H}$. The following natural question, in the spirit of Dirac's theorem, was raised by Gowers [6].

Problem 1.1. *What degree conditions guarantee the existence of a spanning copy of the sphere in a* 3*-graph?*

Equivalently, Problem 1.1 asks for degree conditions that guarantee the existence of a homeomorphic copy of $\mathbb{S}^2$ containing all the vertices in a simplicial 2-complex. We will, however, adhere to the language of hypergraphs in what follows, as this is the language in which most related results have been formulated.

With regards to the origins of Problem 1.1, the question would appear to have first been raised by Gowers at a seminar given by Kühn [12] on one of the more traditional notions of a "cycle in a 3-graph" mentioned earlier. However, being a rather natural question, it has also been asked independently by other people subsequently (Conlon [2], for example), so it would perhaps be fair to say that the problem should more or less be considered folklore at this juncture. While our discussions with experts in the area suggest that Problem 1.1 has attracted a reasonable amount of attention, with the exception of an unpublished partial result of Conlon, Ellis and Keevash [2], almost nothing seems to be known about it.

Before we turn to answering the above question, let us place the problem in a more general context. While the circle $\mathbb{S}^1$ is, up to homeomorphism, the unique connected, closed 1-manifold, this is no longer the case in two dimensions. Hence, one can ask a more general question by replacing the sphere with an arbitrary connected, closed 2-manifold (or *surface* for short). The following more general question fits into the "higher-dimensional combinatorics" programme of Linial [16–18], and was also suggested by Gowers [7].

Problem 1.2. *Given a surface* $\mathcal{S}$*, what degree conditions guarantee the existence of a spanning copy of* $\mathcal{S}$ *in a* 3*-graph?*

Of course, by the classification theorem (see [23, 26], for example), every surface is homeomorphic to either the sphere $\mathbb{S}^2$, a connected sum of finitely many tori, or a connected sum of finitely many real projective planes. Also, to be clear, a copy of a surface $\mathcal{S}$ in a 3-graph $\mathcal{H}$ is, as before, a set of edges of $\mathcal{H}$ such that the simplicial complex induced by these edges is homeomorphic to $\mathcal{S}$, and we say that a copy of $\mathcal{S}$ in a 3-graph $\mathcal{H}$ is spanning if the 0-skeleton of the associated simplicial complex is the entire vertex set of $\mathcal{H}$.

Finally, instead of asking for a spanning copy of a specific surface, one might settle for a spanning copy of any surface whatsoever. In other words, we could ask the following weaker question.

Problem 1.3. *What degree conditions guarantee the existence of a spanning copy of some surface in a 3-graph?*

In this paper, we shall prove a result that gives an asymptotically sharp solution to Problems 1.1, 1.2 and 1.3. Recall that the codegree of a pair of vertices in a 3-graph $\mathcal{H}$ is the number of edges of $\mathcal{H}$ containing the vertex pair. Writing $\delta_2(\mathcal{H})$ for the minimum codegree of a 3-graph $\mathcal{H}$, our main result says the following.

Theorem 1.4. *For every surface $\mathcal{S}$ and every $\mu > 0$, the following holds for all sufficiently large $n \in \mathbb{N}$: any 3-graph $\mathcal{H}$ on n vertices with $\delta_2(\mathcal{H}) \geq \frac{n}{3} + \mu n$ contains a spanning copy of $\mathcal{S}$. Moreover, for each $n \in \mathbb{N}$, there exists a 3-graph $\mathcal{H}$ on n vertices with $\delta_2(\mathcal{H}) = \lfloor \frac{n}{3} \rfloor - 1$ such that there are at most $2\lceil \frac{n}{3} \rceil$ vertices in the 0-skeleton of a copy of any surface in $\mathcal{H}$.*

The second half of Theorem 1.4 follows from the simple construction shown in Figure 1. It will be helpful to have some notation to discuss this construction: we define a relation on the edge set of a 3-graph $\mathcal{H}$ by saying that two edges of $\mathcal{H}$ *touch* if they intersect in two vertices, and we call an equivalence class of edges in the transitive closure of this relation a *tight component* of $\mathcal{H}$. Observe that the set of edges constituting a copy of a surface in a 3-graph $\mathcal{H}$ must belong to a single tight component of $\mathcal{H}$ since all the surfaces under consideration are without boundary. Now, given $n \in \mathbb{N}$, let X, Y and Z be three disjoint sets of vertices, with sizes as equal as possible, such that $|X| + |Y| + |Z| = n$ and consider, as in Figure 1, the 3-graph $\mathcal{H}$ on the vertex set $X \cup Y \cup Z$ whose edge set consists of all triples either intersecting X in two vertices and Y in one, intersecting Y in two vertices and Z in one, or intersecting Z in two vertices and X in one. It is easy to see that $\delta_2(\mathcal{H}) = \lfloor \frac{n}{3} \rfloor - 1$. Furthermore, it is clear that the edge set of $\mathcal{H}$ consists of three tight components, with each tight component spanning two of the three vertex classes X, Y and Z. Thus, there are at most $2\lceil \frac{n}{3} \rceil$ vertices in the 0-skeleton of a copy of any surface in $\mathcal{H}$.

It is left then to prove the first half of Theorem 1.4. One main obstacle to doing this, in contrast to Dirac's theorem, is that, potentially, many of the edges inherently cannot contribute to a spanning copy of any surface. The construction above shows that for each

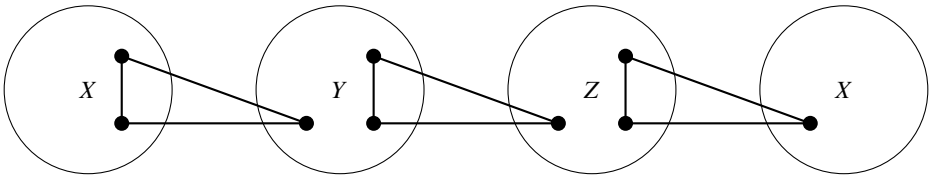


Fig. 1. An extremal construction.

$n \in \mathbb{N}$, there exists a 3-graph $\mathcal{H}$ with $\delta_2(\mathcal{H}) = \lfloor \frac{n}{3} \rfloor - 1$ which does not contain a spanning tight component. As we shall see (in Proposition 3.1), a 3-graph $\mathcal{H}$ on n vertices whose minimum codegree exceeds $\frac{n}{3}$ contains a spanning tight component. However, it may additionally have another non-spanning tight component whose edges are of no use whatsoever when trying to build a spanning copy of any surface. This already presents a challenge, but the situation is in fact more intricate: indeed, as we shall see (in Conjecture 5.3 and the discussion preceding it), even assuming that $\mathcal{H}$ has a unique spanning tight component does not make the problem at hand significantly easier.

It is perhaps worth mentioning that the construction depicted in Figure 1 is not uniquely extremal, and indeed, there exist other non-isomorphic families of constructions that demonstrate that Theorem 1.4 is tight. Given a surface $\mathcal{S}$ with Euler characteristic $\chi \in \mathbb{Z}$, and any sufficiently large natural number $n \in \mathbb{N}$, we construct a 3-graph on n vertices with no spanning copy of $\mathcal{S}$ as follows. Let X and Y be two disjoint sets of vertices, with $|X| + |Y| = n$ and $|X|$ the least integer exceeding $\frac{1}{3}(2n - 2\chi)$, and let $\mathcal{H}$ be the 3-graph on $X \cup Y$ whose edge set consists of all triples meeting X in an odd number of vertices. It is clear that

$$\delta_2(\mathcal{H}) = |Y| - 1 = \left\lceil \frac{n + 2\chi}{3} \right\rceil - 2,$$

and it is easily verified that this construction is not isomorphic to the one discussed earlier. Now, suppose for a contradiction that there is a spanning copy of $\mathcal{S}$ in $\mathcal{H}$. This copy of $\mathcal{S}$ cannot contain any triple contained entirely in X since these edges all lie in a single tight component that spans X but not Y. Next, view each edge of $\mathcal{H}$ in this copy of $\mathcal{S}$ as a facet of a triangulation of $\mathcal{S}$ and count, for each facet, the number of vertices of X on its boundary. Each facet contains exactly one vertex in X, so this quantity is equal to the number of facets, which by Euler's formula, is $2n - 2\chi$. On the other hand, each vertex in X is on the boundary of at least three facets since $\mathcal{S}$ has no boundary, so this quantity is at least $3|X|$. We conclude that $2n - 2\chi \geq 3|X|$, which is a contradiction.

A few comments about results in the vicinity of Theorem 1.4 are also in order. Questions in the spirit of Problem 1.1 have previously been asked for graphs: Kühn, Osthus and Taraz [15] proved a Dirac-type theorem for finding spanning planar triangulations in graphs; applying Theorem 1.4 to the 3-graph of all the triangles in a graph recovers their result (a sharper version of which is however obtained in [13]). In the probabilistic setting, it is natural to ask when a spanning copy of the sphere is likely to appear in the binomial random 3-graph: Luria and Tessler [19] recently established a sharp threshold result for this problem. Finally, it has also been brought to our attention that Conlon, Ellis and Keevash [2] earlier proved (in unpublished work, as previously mentioned) a weaker statement in the direction of Theorem 1.4, showing that any 3-graph $\mathcal{H}$ on n vertices with $\delta_2(\mathcal{H}) \geq \frac{2n}{3} + o(n)$ contains a spanning copy of the sphere.

Organisation of the paper. We establish some notation and collect together the various results required for the proof of our main result in Section 2. We give a short sketch of the proof of Theorem 1.4 highlighting the main obstacles and how we circumvent them in Section 3; the proof proper follows in Section 4. Finally, we conclude by discussing some open problems and directions for further research in Section 5.

2. Preliminaries

For $n \in \mathbb{N}$, let $[n] = \{1, 2, \dots, n\}$. For a set X and $r \in \mathbb{N}$, we write $X^{(r)}$ for the family of r-element subsets of X. In this language, an r-graph G is a pair (V, E) of finite sets with $E \subset V^{(r)}$. Here, we shall only be concerned with 2-graphs and 3-graphs; as usual, we refer to 2-graphs as graphs.

Let $G = (V, E)$ be a graph. For a vertex $x \in V(G)$, its *neighbourhood* $N_G(x)$ is the set of vertices adjacent to x, and its *degree* $\mathrm{d}_G(x)$ is the size of $N_G(x)$. The *minimum degree* $\delta(G)$ of G is

$$\delta(G) = \min\{\mathrm{d}_G(x) : x \in V(G)\}.$$

For any two sets of vertices $X, Y \subset V(G)$, we write $E_G(X, Y)$ for the set of edges with one endpoint in X and one endpoint in Y.

Let $\mathcal{H} = (V, E)$ be a 3-graph. For a pair of distinct vertices $x, y \in V(\mathcal{H})$, we define their *neighbourhood* $N_{\mathcal{H}}(x, y)$ to be the set of vertices $z \in V(\mathcal{H})$ such that $xyz \in E(\mathcal{H})$, and we define the *codegree* $\mathrm{d}_{\mathcal{H}}(x, y)$ of x and y to be size of $N_{\mathcal{H}}(x, y)$. The *minimum codegree* $\delta_2(\mathcal{H})$ of $\mathcal{H}$ is

$$\delta_2(\mathcal{H}) = \min\{\mathrm{d}_{\mathcal{H}}(x, y) : x, y \in V(\mathcal{H}) \text{ and } x \neq y\}.$$

For any vertex $v \in V(\mathcal{H})$, we define the link graph $L_{\mathcal{H}}(v)$ of v to be the graph on $V(\mathcal{H}) \setminus \{v\}$ where two vertices x and y are joined with an edge if $v \in N_{\mathcal{H}}(x, y)$. Finally, as mentioned earlier, we define a relation on the edge set of a 3-graph $\mathcal{H}$ by saying that two edges of $\mathcal{H}$ *touch* if they intersect in two vertices, and we call an equivalence class of edges in the transitive closure of this relation a *tight component* of $\mathcal{H}$.

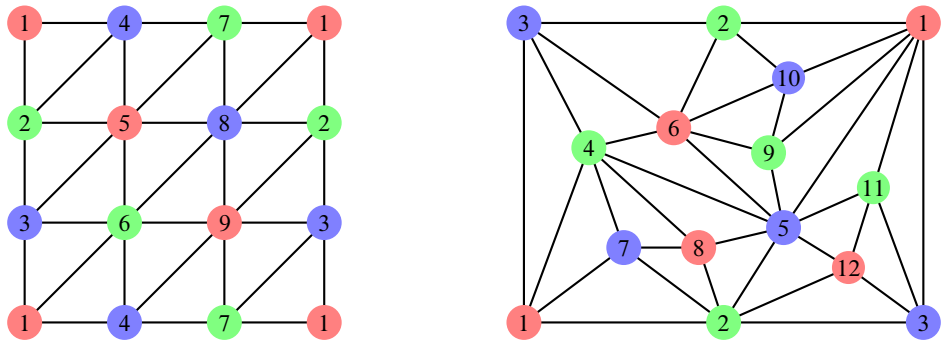
We will use the following classical result of Erdős [4] generalising an old result of Kövari, Sós and Turán [11]. Recall that an r-graph $\mathcal{G}$ is *degenerate* if there exists a colouring of $V(\mathcal{G})$ with r colours such that each edge of $\mathcal{G}$ meets each of the r colour classes.

Theorem 2.1. *For each degenerate r-graph $\mathcal{G}$, there exists $c > 0$ such that every r-graph $\mathcal{H}$ on n vertices with $|E(\mathcal{H})| \geq n^{r-c}$ contains a copy of $\mathcal{G}$ as a subgraph.*

In our proofs, we shall require the conclusion of Theorem 2.1 for some specific degenerate graphs and 3-graphs that we now define. For $k \in \mathbb{N}$, we denote a cycle of length k by C_k; recall that if k is even, then the k-cycle C_k is degenerate. Next, we define two degenerate 3-graphs as in Figure 2, namely a 3-graph $\mathcal{T}_9$ on the vertex set [9] and a 3-graph $\mathcal{P}_{12}$ on the vertex set [12]. It is clear that the simplicial complex induced by $\mathcal{T}_9$ is homeomorphic to the two-dimensional torus. It may also be verified that the simplicial complex induced by $\mathcal{P}_{12}$ is homeomorphic to the real projective plane; indeed, it is not hard to see that $\mathcal{P}_{12}$ is obtained from a simple modification of the standard 6-point triangulation of the real projective plane.

We will also use Szemerédi's regularity lemma [25]; to state the lemma, we need some more notation. Given a graph G, and two disjoint nonempty sets of vertices $X, Y \subset V(G)$, we define the density of the pair (X, Y) by

$$d_G(X, Y) = \frac{|E_G(X, Y)|}{|X||Y|},$$



$\mathcal{T}_9$ $\mathcal{P}_{12}$

Fig. 2. The edge sets of $\mathcal{T}_9$ and $\mathcal{P}_{12}$ consist of all the triangles in the respective figures.

and additionally, for $\varepsilon > 0$, we say that the pair (X, Y) is *ε-regular* if we have

$$|d_G(A, B) - d_G(X, Y)| \leq \varepsilon$$

for all $A \subset X$ and $B \subset Y$ with $|A| \geq \varepsilon|X|$ and $|B| \geq \varepsilon|B|$. A partition $V_0 \cup V_1 \cup \cdots \cup V_k$ of the vertex set of a graph G is said to be an *ε-regular partition* if

(1) $|V_0| \leq \varepsilon|V(G)|$,

(2) $|V_1| = |V_2| = \cdots = |V_k|$, and

(3) all but at most εk^2 pairs (V_i, V_j) with $1 \leq i < j \leq k$ are ε-regular.

In this language, the regularity lemma may be phrased as follows.

Theorem 2.2. *For every $\varepsilon > 0$ and each $t \in \mathbb{N}$, there exists an integer T such that every graph G on at least T vertices admits an ε-regular partition $V_0 \cup V_1 \cup \cdots \cup V_k$ of its vertex set with $t \leq k \leq T$.* ∎

It will also be convenient to have a few consequences of the regularity lemma. The following proposition follows from the fact that the degrees of vertices in regular pairs are typically well-behaved.

Proposition 2.3. *For every $\alpha > 0$, there exists $\beta > 0$ such that the following holds for all $n \in \mathbb{N}$. In each bipartite graph G between vertex classes X and Y with $\alpha n \leq |X|, |Y| \leq n$ and $|E(G)| \geq \alpha n^2$, there exists a subset $U \subset X$ with $|U| \geq \beta n$ such that for each $x \in U$, there exists a subset $U_x \subset U$ of size at least $\frac{3}{4}|U|$ with $|N_G(x) \cap N_G(y)| \geq \beta n$ for each $y \in U_x$.* ∎

We shall additionally use the following proposition establishing "supersaturation" for 4-cycles in dense graphs, a special case of the results of Erdős and Simonovits [5].

Proposition 2.4. *For every $\alpha > 0$, there exists $\beta > 0$ such that every n-vertex graph G with $|E(G)| \geq \alpha n^2$ contains at least βn^4 copies of the 4-cycle C_4.* ∎

We shall use the multiplicative Chernoff bound; see [8], for instance.

Proposition 2.5. *Let X be a binomial random variable with mean μ. Then for any fixed $\delta > 0$, we have*

$$\mathbb{P}(X > (1+\delta)\mu) < \left(\frac{\mathrm{e}^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$$

and

$$\mathbb{P}(X < (1-\delta)\mu) < \left(\frac{\mathrm{e}^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}.$$ ∎

For completeness, we also recall the classification theorem of surfaces; see [23, 26], for example.

Theorem 2.6. *Every surface is homeomorphic to either the sphere $\mathbb{S}^2$, a connected sum of finitely many tori, or a connected sum of finitely many real projective planes.* ∎

Finally, we need some notation for dealing with hierarchies of constants in our proofs. We shall say that a statement holds "for $\delta \ll \varepsilon$" if for any fixed $\varepsilon \in (0,1]$, there exists $\delta_\varepsilon \in (0,1]$ such that the statement in question holds for all $\delta \in (0, \delta_\varepsilon]$. Hierarchies with more constants are defined analogously and are to be read from the right to the left.

To avoid clutter, we shall frequently drop the subscript specifying the graph or 3-graph in the notation above when the graph or 3-graph in question is clear, abbreviating, for example, $E_G(X,Y)$ by $E(X,Y)$ or $L_{\mathcal{H}}(v)$ by $L(v)$. Additionally, we systematically omit floors and ceilings whenever they are not crucial.

3. Overview of our strategy

Let us briefly discuss our approach to establishing Theorem 1.4. To begin with, we discuss finding spanning copies of the sphere since this captures most of the difficulties involved. We shall then say a few words about how we find spanning copies of a general surface.

It is not hard to see that a 3-graph $\mathcal{H}$ on n vertices with $\delta_2(\mathcal{H}) \gtrsim \frac{3n}{4}$ contains a spanning copy of the sphere. This follows from Dirac's theorem; indeed, fix a pair of vertices of $\mathcal{H}$, say x and y, and consider a graph G on $V(\mathcal{H}) \setminus \{x,y\}$, where two vertices u and v are joined if $x, y \in N_{\mathcal{H}}(u,v)$, i.e., if $uvx, uvy \in E(\mathcal{H})$. It is easy to verify that $\delta(G) \gtrsim \frac{n}{2}$, so it follows from Dirac's theorem that G contains a Hamiltonian cycle. Of course, a Hamiltonian cycle in G translates back to a "spanning double pyramid" in $\mathcal{H}$ (see Figure 3), and this is of course a spanning copy of the sphere in $\mathcal{H}$.

While the above argument is not particularly efficient, it does contain the following useful idea: it is easy to find reasonably large spheres in any dense 3-graph. Indeed, if we repeat the argument above in a dense 3-graph $\mathcal{H}$, then the auxiliary graph $G_{x,y}$ that we construct will be dense for most pairs of vertices $x, y \in V(\mathcal{H})$; consequently, $G_{x,y}$ will typically contain long cycles, and these translate back into large double pyramids in $\mathcal{H}$. This idea may be used to show that a 3-graph $\mathcal{H}$ on n vertices with $\delta_2(\mathcal{H}) \gtrsim \frac{2n}{3}$ contains a spanning copy of the sphere: in such a graph $\mathcal{H}$, every edge is in a tetrahedron, and we may use this to build "absorbing structures" in the spirit of Rödl, Ruciński and

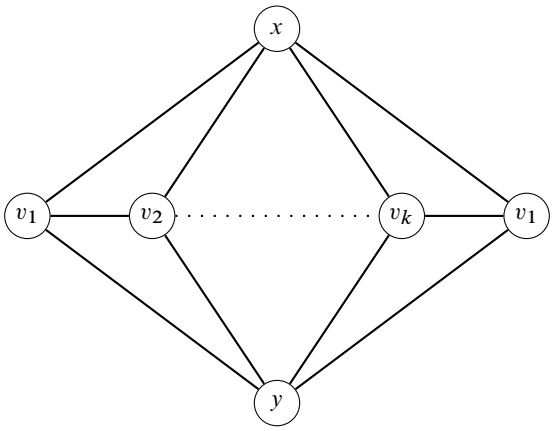


Fig. 3. A double pyramid with apexes x and y.

Szemerédi [21] that may be used to combine a small number of large spheres that almost span the vertex set into a single sphere spanning the vertex set.

Approaches based on the ideas discussed above however reach a natural barrier at the threshold of $\frac{n}{2}$. Indeed, above this threshold, all the edges of a 3-graph under consideration necessarily belong to a single tight component, but this is no longer true below this threshold, as shown by the following proposition.

Proposition 3.1. *If $\mathcal{H}$ is a 3-graph on n vertices with $\delta_2(\mathcal{H}) > \frac{1}{2}(n-3)$, then all the edges of $\mathcal{H}$ belong to a single tight component. Moreover, for each $n \in \mathbb{N}$, there exists a 3-graph $\mathcal{H}$ on n vertices with $\delta_2(\mathcal{H}) = \lfloor \frac{1}{2}(n-3) \rfloor$ whose edge set decomposes into two tight components.*

Proof. First, let $\mathcal{H}$ be a 3-graph on $[n]$ with $\delta_2(\mathcal{H}) > \frac{1}{2}(n-3)$. Let us induce an edge-colouring of the complete graph on $[n]$ by setting the colour of an edge e of the complete graph to be the tight component corresponding to all the edges of $\mathcal{H}$ containing e; that this colouring is well-defined follows immediately from the definition of a tight component. Now, observe that no two edges of the complete graph incident to the same vertex can be coloured differently; indeed, if xy and xz constitute such a pair of edges for some $x, y, z \in [n]$, then it is clear that $z \notin N_{\mathcal{H}}(x, y)$, $y \notin N_{\mathcal{H}}(x, z)$ and $N_{\mathcal{H}}(x, y) \cap N_{\mathcal{H}}(x, z) = \varnothing$, which leads to an easy contradiction since $\delta_2(\mathcal{H}) > \frac{1}{2}(n-3)$. It follows instantly that this induced edge-colouring of the complete graph uses only one colour since the complete graph is connected (and every edge is coloured). Therefore, all the edges of $\mathcal{H}$ belong to a single tight component, proving the first part of the claim.

Next, given $n \in \mathbb{N}$, let X and Y be two disjoint sets of $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$ vertices respectively, and let $\mathcal{H}$ be the 3-graph on the vertex set $X \cup Y$ whose edge set consists of all the triples meeting Y in an odd number of vertices. It is easy to see that $\delta_2(\mathcal{H}) = \lfloor \frac{1}{2}(n-3) \rfloor$, and it is not hard to verify that the edge set of $\mathcal{H}$ decomposes into two tight components, one consisting of all the triples meeting Y in three vertices and the other consisting of

all the triples meeting Y in one vertex. This construction proves the second part of the claim. ∎

To get down to the threshold of $\frac{n}{3}$, we need to work somewhat harder. As remarked earlier, it is clear that only those edges from a spanning tight component of a 3-graph $\mathcal{H}$ are of any use in constructing a spanning copy of the sphere in $\mathcal{H}$. An argument similar to the one used to prove Proposition 3.1 shows that a 3-graph $\mathcal{H}$ on n vertices whose minimum codegree exceeds $\frac{n}{3}$ contains a spanning tight component. However, to prove Theorem 1.4, it will be necessary to say something more about the structure of, and the interaction between, the tight components of a 3-graph $\mathcal{H}$ on n vertices with $\delta_2(\mathcal{H}) \gtrsim \frac{n}{3}$.

First, we shall demonstrate that at least one of the spanning tight components of such a 3-graph has reasonably good "connectibility properties"; this will help us in forming connected sums of smaller surfaces that we will build over the course of our proof. Next, we shall show that it is possible to find a small number of spheres that almost span the vertex set, while crucially ensuring that these spheres meet sufficiently many edges with the good connectibility properties. Finally, we shall use the fact that the 3-graphs $\mathcal{T}_9$ and $\mathcal{P}_{12}$, which respectively represent the torus and the real projective plane, are degenerate to either add handles or crosscaps as needed.

In order to implement the above ideas, we shall rely on a number of techniques from extremal and probabilistic combinatorics, including those of absorption, regularity and supersaturation.

4. Proof of the main result

We shall divide the proof of Theorem 1.4 into stages, each addressing different aspects of the strategy outlined in the previous section.

Let us point out two notational conventions that we adopt. Given a subgraph $\mathcal{G}$ of a 3-graph $\mathcal{H}$, the subset of the vertex set $V(\mathcal{H})$ spanned by $\mathcal{G}$ is denoted by $V(\mathcal{G})$; in the case where $\mathcal{G}$ consists of a single edge $e \in E(\mathcal{H})$, we abuse notation slightly and write $V(e)$ for $V(\mathcal{G})$. To avoid clutter, we will also talk about "spheres in a 3-graph" when, strictly speaking, we mean "copies of the sphere $\mathbb{S}^2$ in a 3-graph", and similarly for the torus and the real projective plane.

We shall make use of the following simple averaging lemma at several points.

Lemma 4.1. *For any $r \in \mathbb{N}$ and $\gamma \in (0,1)$, the following holds for all sufficiently large $m \in \mathbb{N}$. For any finite set $\mathcal{X}$, and any system $X_1, \dots, X_m$ of subsets of $\mathcal{X}$ satisfying*

$$\sum_{i=1}^{m} |X_i| \geq \gamma m |\mathcal{X}|,$$

we may find a set $K \subset [m]$ with $|K| = r$ such that

$$\left| \bigcap_{i \in K} X_i \right| \geq \frac{\gamma^r |\mathcal{X}|}{2}.$$

Proof. Choose a random subset $K \subset [m]$ of size r. For each $x \in \mathcal{X}$, x belongs to $\bigcap_{i\in K} X_i$ with probability $\binom{f(x)}{r}/\binom{m}{r}$, where $f(x)$ is the number of indices $i \in [m]$ such that $x \in X_i$. By Jensen's inequality, we have

$$\mathbb{E}\left(\left|\bigcap_{i\in K} X_i\right|\right) \geq |\mathcal{X}|\binom{\gamma m}{r}\binom{m}{r}^{-1}$$

provided that $\gamma m > r$. It is now straightforward to prove that the claim holds for all sufficiently large $m \in \mathbb{N}$. ∎

4.1. Double pyramids in dense 3-graphs

Recall that a 3-graph is said to be *degenerate* or 3-*partite* if its vertex set may be partitioned into three classes in such a way that each edge meets all three classes. The following lemmas make precise our earlier observation that it is easy to find large spheres in any dense 3-graph.

Lemma 4.2. *Let $\frac{1}{n} \ll \delta, \frac{1}{k} \ll 1$, let $\mathcal{H}$ be a 3-graph on n vertices, and let $T \subset V(\mathcal{H})$ meet at least δn^3 edges of $\mathcal{H}$. Then $\mathcal{H}$ contains a double pyramid on $2k+2$ vertices with apexes in T.*

Proof. For each vertex $v \in T$, recall that $L(v)$ denotes the graph on $V(\mathcal{H}) \setminus \{v\}$ with $xy \in E(L(v))$ whenever $vxy \in E(\mathcal{H})$. Applying Lemma 4.1 to these graphs – with $r = 2$ and $\mathcal{X} = V(\mathcal{H})^2$ – we can find a pair of vertices $v, w \in T$ such that there are at least $\frac{1}{2}\delta^2 n^2$ edges common to $L(v)$ and $L(w)$. By Theorem 2.1 – applied with $r = 2$ – we may find a copy of the cycle C_{2k} among these common edges. Such a cycle yields a double pyramid in $\mathcal{H}$ with apexes v and w and $2k+2$ vertices in total. ∎

Lemma 4.3. *Let $\frac{1}{n} \ll \delta, \varepsilon, \eta \ll 1$, and let $\mathcal{H}$ be a 3-partite 3-graph with vertex classes A, B and T with $|A| = |B| = n$ and $|T| \geq \varepsilon n$ such that for each $A' \subset A$ and $B' \subset B$ with $|A'|, |B'| \geq \varepsilon n$, there are at least $\delta|A'||B'||T|$ edges in $\mathcal{H}[T \cup A' \cup B']$. Then $\mathcal{H}$ contains at most ηn disjoint spheres that each have 2 vertices in T and together cover at least $(1-\varepsilon)n$ vertices from each of A and B.*

Proof. Fix an integer k such that $\frac{1}{n} \ll \frac{1}{k} \ll \delta, \varepsilon, \eta$. Iteratively, remove spheres S from $\mathcal{H}$ with $|V(S) \cap A| = |V(S) \cap B| \geq k$ and $|V(S) \cap T| = 2$, until it is not possible to find another sphere with these properties. This removes at most ηn disjoint spheres. Suppose that we are left with $A' \subset A$, $B' \subset B$ and $T' \subset T$.

We will show that $|A'| = |B'| < \varepsilon n$, completing the proof. Suppose for the sake of contradiction that $|A'| = |B'| \geq \varepsilon n$. Note that we have

$$|T'| \geq |T| - \frac{2n}{k} \geq \left(1 - \frac{\delta}{2}\right)|T|,$$

so there are at least $\frac{1}{2}\delta|A'||B'||T|$ edges in $\mathcal{H}[T' \cup A' \cup B']$. Applying Lemma 4.2 to $\mathcal{H}[T' \cup A' \cup B']$, we may find a double pyramid on $2k+2$ vertices with apexes in T'. Since $\mathcal{H}$ is 3-partite, the remaining $2k$ vertices are evenly divided between A' and B', yielding a contradiction. ∎

4.2. Colouring and connecting edges

Next, we detail how we will connect spheres together into a larger sphere. In this section, we work towards showing that the edges in our 3-graph can be (mostly) coloured red and green so that any disjoint edges of the same colour can be connected together into a sphere (and, moreover, in many different ways). Thus, two disjoint spheres which share some colour among their edges can be connected together into a larger sphere.

More precisely, given an integer $k \in \mathbb{N}$ and a 3-graph $\mathcal{H}$ on n vertices, we say two edges $e, f \in E(\mathcal{H})$ are (α, k)-*connectible* if, for some l with $1 \leq l \leq k$, there are at least αn^l sets $A \subset V(\mathcal{H})$ with $|A| = l$ for which there is a sphere in $\mathcal{H}$ containing the edges e and f with vertex set $A \cup V(e) \cup V(f)$. This will be the most convenient definition to use in proving connectibility properties, however, we shall then only use the following simple consequence of (α, k)-connectibility.

Lemma 4.4. *If two edges $e, f \in E(\mathcal{H})$ are (α, k)-connectible, then for some l with $1 \leq l \leq k$ there are at least $\frac{\alpha n}{l}$ pairwise disjoint sets $A \subset V(\mathcal{H})$ with $|A| = l$ for which there is a sphere with vertex set $A \cup V(e) \cup V(f)$ in $\mathcal{H}$ containing the edges e and f.*

Proof. Let l with $1 \leq l \leq k$ be a number witnessing the fact that e, f are (α, k)-connectible. Choose a maximal collection $A_1, A_2, \ldots, A_r$ of disjoint sets with the required properties and note that $U = \bigcup_{i=1}^{r} A_i$ comprises rl vertices. By the maximality of our choice, U intersects every set of vertices of size l which can be used to create a sphere including e and f. There are at least αn^l such sets, but at most rln^{l-1} sets of size l meet U, so we must have $rl \geq \alpha n$, as required. ∎

Next, we demonstrate that almost all touching pairs of edges in a dense 3-graph are easily connectible.

Lemma 4.5. *Let $\frac{1}{n} \ll \delta \ll \varepsilon < 1$, and let $\mathcal{H}$ be a 3-graph on n vertices. Then all but at most εn^4 touching pairs of edges of $\mathcal{H}$ are contained in at least δn^2 spheres with six vertices in $\mathcal{H}$.*

Proof. We need to show that if F is any set of εn^4 touching edge-pairs of $\mathcal{H}$, then at least one of these pairs is contained in at least δn^2 spheres with six vertices. To this end, fix a set F of εn^4 touching edge-pairs of $\mathcal{H}$. By averaging, there are two vertices $x, y \in V(\mathcal{H})$ such that $e \triangle f = \{x, y\}$ for at least εn^2 edge-pairs $\{e, f\} \in F$, where $\triangle$ denotes the symmetric difference.

Let $G = G_{x,y}$ be the graph on the vertex set $V(\mathcal{H})$ where we join two vertices u and v if $uvx, uvy \in E(\mathcal{H})$. As

$$|E(G)| \geq \varepsilon n^2,$$

we know from Proposition 2.4 that G contains δn^4 copies of the 4-cycle C_4. Consequently, there is some $uv \in E(G)$ such that there are at least δn^2 pairs $w, z \in V(\mathcal{H})$ such that the set $\{u, v, w, z\}$ induces a copy of C_4 in G containing the edge uv. For any such a copy of C_4, note that $\{x, y, u, v, w, z\}$ is the vertex set of a sphere in $\mathcal{H}$ containing the touching edge-pair $\{uvx, uvy\}$. Thus, we have shown that F contains a pair $\{uvx, uvy\}$ that is contained in at least δn^2 spheres in $\mathcal{H}$ with six vertices, proving the claim. ∎

Our next (rather coarse) lemma develops the connectibility properties of dense 3-graphs further.

Lemma 4.6. *Let $\frac{1}{n} \ll \alpha \ll \beta \ll \varepsilon \ll 1$, and let $\mathcal{H}$ be a 3-graph on n vertices with $\delta_2(\mathcal{H}) \geq \frac{n}{3}$. Then all but at most εn^3 edges of $\mathcal{H}$ can be coloured so that*

(1) *there are at least $\frac{1}{2}\beta n^3$ edges of each colour, and*

(2) *any two disjoint edges of the same colour are $(\alpha, 33)$-connectible.*

It is worth remarking that here and in what follows, when we speak of a pair edges of the same colour being connectible, we mean this in $\mathcal{H}$ (and not necessarily in their own colour class) unless we explicitly say otherwise.

Proof of Lemma 4.6. Take $\delta, \eta > 0$ so that $\beta \ll \delta \ll \eta \ll \varepsilon$ and call a pair of touching edges *good* if it can be completed into a sphere using two more vertices in at least δn^2 different ways and *bad* otherwise. Note that by Lemma 4.5, there are at most ηn^4 bad pairs in $\mathcal{H}$.

We iteratively colour the edges of $\mathcal{H}$ as follows. Suppose that we have a set $F \subset E(\mathcal{H})$ of at least εn^3 uncoloured edges. Write $\mathbf{E}$ and $\mathbf{F}$ for the sets of ordered 3-tuples (x, y, z) such that $xyz \in E(\mathcal{H})$ and $xyz \in F$, respectively. Given a 3-tuple $\mathbf{f} = (v_1, v_2, v_3) \in \mathbf{F}$, we define

$$A(\mathbf{f}) = \{(x, y, z) : v_2v_3x, v_3xy, xyz \in E(\mathcal{H})\},$$

noting that

$$|A(\mathbf{f})| \geq \left(\frac{n}{3} - 1\right)\left(\frac{n}{3} - 2\right)\left(\frac{n}{3} - 3\right) \geq \frac{n^3}{54}.$$

Clearly, there are at least $\frac{\varepsilon n^6}{9}$ ordered pairs $(\mathbf{f}, \mathbf{e})$ of 3-tuples with $\mathbf{f} \in \mathbf{F}$ and $\mathbf{e} \in A(\mathbf{f})$ since every edge of F corresponds to 6 different 3-tuples in $\mathbf{F}$.

For $\mathbf{f} \in \mathbf{F}$, let $B(\mathbf{f}) \subset A(\mathbf{f})$ be the set of 3-tuples $(x, y, z) \in A(\mathbf{f})$ such that the pairs $\{v_1v_2v_3, v_2v_3x\}$, $\{v_2v_3x, v_3xy\}$, and $\{v_3xy, xyz\}$ of touching edges are all good. Note that every bad pair $\{w'x'y', x'y'z'\}$ in $\mathcal{H}$ arises as one of the three pairs corresponding to (v_1, v_2, v_3, x, y, z) as above for at most $12n^2$ choices of this 6-tuple. Therefore, there are at most $12\eta n^6$ ordered pairs $(\mathbf{f}, \mathbf{e})$ with $\mathbf{f} \in \mathbf{F}$ and $\mathbf{e} \in A(\mathbf{f}) \setminus B(\mathbf{f})$.

Now, if $|B(\mathbf{f})| < \eta n^3$ for some $\mathbf{f} \in \mathbf{F}$, then $|A(\mathbf{f}) \setminus B(\mathbf{f})| \geq (\frac{1}{54} - \eta)n^3$, so our previous observation implies that there are at most $12\eta n^3/(\frac{1}{54} - \eta) \leq 3\varepsilon n^3$ different 3-tuples $\mathbf{f} \in \mathbf{F}$ for which $|B(\mathbf{f})| < \eta n^3$. Hence, there are at least $6\varepsilon n^3 - 3\varepsilon n^3 \geq 3\varepsilon n^3$ different 3-tuples $\mathbf{f} \in \mathbf{F}$ for which $|B(\mathbf{f})| \geq \eta n^3$, and correspondingly, there is a set $F' \subset F$ of at least $\frac{\varepsilon n^3}{2}$ edges for which at least one of the six 3-tuples corresponding to that edge has this property.

Next, we create an auxiliary bipartite graph G between F' and $E(\mathcal{H})$ with an edge between $f \in F'$ and $e \in E(\mathcal{H})$ if some pair $\{\mathbf{f}, \mathbf{e}\}$ of 3-tuples corresponding to these edges of $\mathcal{H}$ satisfies $\mathbf{e} \in B(\mathbf{f})$; note that G has at least $\frac{1}{12}\varepsilon\eta n^6$ edges. We appeal to Proposition 2.3 and find a subset $F'' \subset F'$ of size at least βn^3, such that for any two edges $f_1, f_2 \in F''$ there are at least $\frac{1}{2}\beta n^3$ edges $f_3 \in F''$ with the property that the pairs f_1, f_3 and f_2, f_3 each have at least βn^3 common neighbours in G.

Thus, for disjoint edges $f_1, f_2 \in F''$, there are at least $\frac{1}{4}\beta n^3$ edges $f_3 \in F''$ disjoint from f_1 and f_2 so that the pairs f_1, f_3 and f_2, f_3 each have at least βn^3 com-

mon neighbours in G. For each such f_3, note that there are at least $\frac{1}{4}(\beta n^3)^2$ choices of $e_1, e_2 \in E(\mathcal{H})$ which are disjoint from each other and from f_1, f_2 and f_3 and such that $f_1e_1, e_1f_3, f_3e_2, e_2f_2 \in E(G)$. For any such pair e_1 and e_2, we proceed as follows. There are at least $(\delta n^2)^3 - O(n^5) \geq \frac{1}{2}\delta^3 n^6$ ways to find a sphere S_1 joining f_1 and e_1 using six extra vertices and avoiding the vertices in f_3, e_2 and f_2; to see this, note that f_1 and e_1 are joined in G, which means that we may walk from f_1 to e_1 along three good touching pairs of edges in $\mathcal{H}$, so gluing three spheres, one sitting each of these good touching pairs, furnishes us with a candidate for S_1, and ignoring candidates that meet f_3, e_2 and f_2 gives us the required bound. A similar argument for each of the edges $e_1f_3, f_3e_2, e_2f_2 \in E(G)$ in turn shows that we can find spheres S_2, S_3 and S_4 containing the pairs e_1f_3, f_3e_2 and e_2f_2 in turn, each with six vertices in addition to those in the edge pair, and so that all the additional vertices are distinct and not in any of the edges f_1, f_2, f_3, e_1 and e_2. Furthermore, when choosing each sphere we have at least $\frac{1}{2}\delta^3 n^6$ choices. Thus, we can choose such spheres in at least $\frac{1}{16}\delta^{12}n^{24} - O(n^{23}) \geq \frac{1}{32}\delta^{12}n^{24}$ different ways so that gluing S_1, S_2, S_3 and S_4 together (i.e., taking their symmetric difference) results in a sphere containing f_1 and f_2 and 33 extra vertices.

There were at least $\frac{1}{4}\beta n^3$ choices of f_3 and $\frac{1}{4}(\beta n^3)^2$ choices of e_1 and e_2, so thus in total we can find at least $\frac{1}{512}\delta^{12}\beta^3 n^{33} \geq \alpha n^{33}$ spheres containing f_1, f_2 and 33 extra vertices. Thus, f_1 and f_2 are $(\alpha, 33)$-connectible. As f_1 and f_2 were arbitrary disjoint edges in F'', we can colour all the edges in F'' using a new colour.

It is clear that this procedure colours the edges of $\mathcal{H}$ as required, thereby proving the lemma. ∎

The next observation refines the previous lemma, and is the starting point of our characterisation of the connectibility properties of the tight components of a sufficiently dense 3-graph.

Lemma 4.7. *Let $\frac{1}{n} \ll \alpha, \frac{1}{k} \ll \varepsilon \ll 1$, and $\mathcal{H}$ be a 3-graph on n vertices with $\delta_2(\mathcal{H}) \geq \frac{n}{3}$. Then all but at most εn^3 edges of $\mathcal{H}$ can be coloured so that any two disjoint edges of the same colour are (α, k)-connectible and there are at most εn^4 pairs of touching edges with different colours.*

Proof. Take $\delta, \eta > 0$ so that $\alpha, \frac{1}{k} \ll \delta \ll \eta \ll \varepsilon$ and set $k_0 = 33$. By Lemma 4.5, there are at most $\frac{1}{2}\eta^2\varepsilon n^4$ touching edge-pairs in $\mathcal{H}$ which are not contained in at least δn^2 spheres with size 6 in $\mathcal{H}$. By Lemma 4.6, we can colour the edges of $\mathcal{H}$ so that at most εn^3 edges are uncoloured, there are at least ηn^3 edges of each colour, and any two disjoint edges of the same colour are (δ, k_0)-connectible. Let l be the number of colours used by this colouring C_0, and note that

$$l \leq \binom{n}{3}(\eta n^3)^{-1} < \frac{1}{\eta}.$$

We iteratively construct a sequence of colourings $(C_i)_{1 \leq i \leq m}$ as follows. For each $i \geq 0$, if there are at least $\eta^2\varepsilon n^4$ touching pairs of edges each using two fixed colours c_1, c_2 of C_i, then we define the colouring C_{i+1} by replacing c_1, c_2 with a common new colour (if there is more than one choice for such a pair of colours, we pick any acceptable

pair). Of course, these recolourings may be performed at most $l-1$ times, and we stop after $m < l$ steps when there are at most

$$\binom{l}{2}\eta^2\varepsilon n^4 < \varepsilon n^4$$

non-monochromatic touching edge-pairs in the colouring C_m.

For $0 \le i \le m$, we prove by induction that any two disjoint edges of the same colour in C_i are $(\delta^{3^i}, 3^i k_0)$-connectible in $\mathcal{H}$. This is certainly true for $i = 0$, so suppose that we have established this fact for i, and that $e, f \in E(\mathcal{H})$ have the same colour in C_{i+1} but not in C_i, where they are coloured, say, red and blue respectively. Then there are at least $\eta^2\varepsilon n^4$ touching red-blue edge pairs in C_i, and at least $\frac{1}{2}\eta^2\varepsilon n^4$ of these pairs are contained in at least δn^2 spheres with size 6 in $\mathcal{H}$. At least half of the latter pairs (e', f') satisfy the additional requirement that $V(e) \cup V(f)$ and $V(e') \cup V(f')$ are disjoint, since there are only $O(n^3)$ touching-edge pairs which fail to satisfy it. By the induction hypothesis and the pigeonhole principle, we observe that for some $k_1, k_2 \le 3^{i-1}k_0$, there are at least

$$\frac{\eta^2\varepsilon n^4}{4(3^{i-1}k_0)^2} > \delta n^4$$

of the above red-blue pairs of touching edges (e', f') such that there are

(1) at least $\delta^{3^{i-1}} n^{k_1}$ ways to complete e and e' into a sphere S_1 using k_1 extra vertices,

(2) at least $\delta^{3^{i-1}} n^{k_2}$ ways to complete f' and f into a sphere S_2 using k_2 extra vertices, and

(3) at least δn^2 ways to complete e' and f' into a sphere S_3 using two extra vertices.

Thus, there are at least $\delta^{2(1+3^{i-1})} n^{k_1+k_2+6} < 2\delta^{3^i} n^{k_1+k_2+6}$ ways to choose a combination of these objects, namely a pair (e', f') and three spheres S_1, S_2 and S_3 with the above properties. Note that for every such choice, the symmetric difference of the edge sets of S_1, S_2 and S_3 forms a sphere containing e and f and $k_1 + k_2 + 6 \le 2 \cdot 3^{i-1}k_0 + 6 < 3^i k_0$ additional vertices, provided the sets of extra vertices used in 1, 2 and 3 above are disjoint from each other and from $V(e) \cup V(e') \cup V(f') \cup V(f)$. As the number of ways to choose any triple of vertex sets, let alone spheres, with k_1, k_2 and 2 elements respectively, such that two of the sets intersect or one intersects a given 10-element set, is $O(n^{k_1+k_2+1})$, we deduce that there are at least $2\delta^{3^i} n^{k_1+k_2+6} - O(n^{k_1+k_2+5})$ spheres in $\mathcal{H}$ including e, f and $k_1 + k_2 + 6$ extra vertices, so it follows that e and f are $(\delta^{3^i}, 3^i k_0)$-connectible in $\mathcal{H}$.

It is now clear that our final colouring C_m has the required properties, completing the proof. ∎

4.3. Colour interaction

We say that a 3-graph $\mathcal{H}$ on n vertices is (ε, μ)-*coloured* if $\delta_2(\mathcal{H}) \ge (\frac{1}{3} + \mu)n$ and all but εn^3 of its edges are coloured red or green so that

(1) there are at most εn^4 pairs of touching red and green edges, and

(2) at least $\frac{\mu n}{4}$ vertices are contained in fewer than εn^2 red edges.

The following lemma turns the (ε, μ)-coloured property into a property of the link graphs $L(v)$ we will use in Section 4.6 to construct almost spanning spheres. For an edge-coloured 3-graph $\mathcal{H}$ as above, and $v \in V(\mathcal{H})$, we define the *green link graph* $\mathrm{GL}(v)$ to be the graph on $V(\mathcal{H}) \setminus \{v\}$ where two vertices x and y are joined with an edge if vxy is a *green edge* of $\mathcal{H}$.

Lemma 4.8. *Let $\frac{1}{n} \ll \varepsilon \ll \mu \ll 1$, and let $\mathcal{H}$ be an n-vertex (ε, μ)-coloured 3-graph. Then, for all but at most $\mu^4 n$ vertices $v \in V(\mathcal{H})$, the green link graph $\mathrm{GL}(v)$ has at least $(\frac{1}{3} + \frac{\mu}{2})n$ vertices with degree at least $(\frac{1}{3} + \frac{\mu}{2})n$.*

Proof. We construct an auxiliary red-green coloured graph G on the vertex set $V(\mathcal{H})$ as follows. For each pair $x, y \in V(\mathcal{H})$ such that there are at least $(\frac{1}{3} + \frac{3\mu}{4})n$ edges of one colour in $\mathcal{H}$ containing x and y, add the edge xy to G, giving it that colour, where if this pair satisfies this condition for both colours, we colour xy green. It suffices to show that all but at most $\mu^4 n$ vertices are in at least $(\frac{1}{3} + \frac{\mu}{2})n$ green edges of G, since if xy is a green edge of G then there are at least $(\frac{1}{3} + \frac{3\mu}{4})$ green edges of the form xyz, and it follows that y has degree at least $(\frac{1}{3} + \frac{3\mu}{4})$ in $\mathrm{GL}(x)$.

If xy is not an edge of G, then the edges of $\mathcal{H}$ containing x and y either include at least $\frac{\mu n}{8}$ uncoloured edges or include at least $\frac{\mu n}{8}$ edges of each colour. In the latter case, there are at least $\frac{1}{64}\mu^2 n^2$ pairs of differently-coloured edges in $\mathcal{H}$ which touch along xy; since there are at most εn^4 pairs of differently-coloured touching edges in $\mathcal{H}$, no more than $64\varepsilon\frac{n^2}{\mu^2}$ non-edges xy can occur for this reason. Similarly, since there are at most εn^3 uncoloured edges in $\mathcal{H}$, at most $24\varepsilon\frac{n^2}{\mu}$ non-edges xy can occur in G for the former reason. Hence there are at most

$$88\varepsilon\frac{n^2}{\mu^2} < \mu^6 n^2$$

non-edges in G.

Since $\mathcal{H}$ is (ε, μ)-coloured, there are at least $\frac{\mu n}{4}$ vertices in fewer than εn^2 red edges of $\mathcal{H}$; let A be a set of $\frac{\mu n}{4}$ such vertices. Since each red edge of G extends to more than $\frac{n}{3}$ red edges of $\mathcal{H}$, each vertex in A is in fewer than $6\varepsilon n$ red edges of G, so there are fewer than $\frac{3}{2}\varepsilon\mu n^2 < \mu^6 n^2$ red edges of G meeting A.

Let B be the set of vertices with fewer than $\frac{\mu n}{8}$ green edges in G to A. Since for each such vertex there are at least $\frac{\mu n}{8}$ non-edges or red edges with the other end in A, and there are fewer than $2\mu^6 n^2$ pairs which are non-edges or red edges meeting A, each of which is counted at most twice, we deduce that $|B| < 32\mu^5 n$. Let B' be the set of vertices with degree less than $(1 - \frac{\mu}{8})n$ in G. Since G has fewer than $\mu^6 n^2$ non-edges, we have $|B'| < 8\mu^5 n$.

Now, fix a vertex $v \in V(\mathcal{H}) \setminus (B \cup B')$, and suppose that v is in fewer than $(\frac{1}{3} + \frac{\mu}{2})n$ green edges in G. Since $v \notin B$, there are at least $\frac{\mu n}{8}$ choices of $u \in A$ for which uv is a green edge of G; by definition, for each choice of u there are at least $(\frac{1}{3} + \frac{3\mu}{4})n$ choices of w for which uvw is a green edge of $\mathcal{H}$. Since v is in fewer than $(\frac{1}{3} + \frac{\mu}{2})n$ green edges in G, and $v \notin B'$, it must be the case that vw is a red edge of G for at least $\frac{\mu n}{8}$ of these choices of w, so there are at least $\frac{1}{64}\mu^2 n^2$ pairs (e, f) where e is a red edge of G, f is a green edge of $\mathcal{H}$ and $v \in e \subset f$.

However, there are at most $3\varepsilon n^3 < \mu^7 n^3$ pairs (e, f) for which e is a red edge of G, f is a green edge of $\mathcal{H}$ and $e \subset f$, since each such pair corresponds to at least $\frac{n}{3}$ differently-coloured touching edge pairs in $\mathcal{H}$, of which there are at most εn^4. Each such pair corresponds to at most two choices of v, so there are at most $128\mu^5 n$ vertices in $V(\mathcal{H}) \setminus (B \cup B')$ which are in fewer than $(\frac{1}{3} + \frac{\mu}{2})n$ green edges of G; it follows that, in total, at most $168\mu^5 n < \mu^4 n$ vertices are in fewer than $(\frac{1}{3} + \frac{\mu}{2})n$ green edges of G, as required. ∎

Crucial for our purposes is the fact that any sufficiently dense 3-graph admits an (ε, μ)-colouring with strong connectibility properties within each colour class.

Lemma 4.9. *Let $\frac{1}{n} \ll \alpha, \frac{1}{k} \ll \varepsilon \ll \mu$, and let $\mathcal{H}$ be a 3-graph on n vertices with $\delta_2(\mathcal{H}) \geq (\frac{1}{3} + \mu)n$. Then it is possible to (ε, μ)-colour $\mathcal{H}$ so that any monochromatic pair of disjoint edges of $\mathcal{H}$ is (α, k)-connectible.*

Proof. Choose $\eta > 0$ so that $\alpha, \frac{1}{k} \ll \eta \ll \varepsilon$. By Lemma 4.7, all but at most ηn^3 edges of $\mathcal{H}$ can be coloured so that any two disjoint edges of the same colour are (α, k)-connectible and there are at most ηn^4 pairs of touching edges with different colours. It only remains to show then that some two colours account for all but $(\varepsilon - \eta)n^3$ of the coloured edges, and that at least $\frac{\mu n}{4}$ vertices appear in fewer than εn^2 edges of one of these colours.

To prove this, we shall first show that for all but a few pairs of vertices, there are a large number of edges of a single colour containing that pair which also meet a large number of edges of the same colour along each of the other pairs they contain. In order to consider pairs of vertices in this way, we will find it convenient to define an auxiliary coloured graph G. We then show that there is only room in $\mathcal{H}$ for such large monochromatic structures in at most two colours; this is similar to proving that a 3-graph with minimum codegree exceeding $\frac{n}{3}$ can have at most two tight components. Finally, we shall use a double counting argument on the interactions between colours to show that there is a similar obstacle to nearly all vertices meeting many edges of both colours.

We start by defining an auxiliary graph G on $V(\mathcal{H})$ where we join a pair $x, y \in V(\mathcal{H})$ of vertices if there is a unique colour c so that there are at least $(\frac{1}{3} + \mu - \varepsilon)n$ edges with colour c in $\mathcal{H}$ containing x and y; moreover, we assign the colour c to such an edge. If xy is not an edge of G, then it follows that at least

$$(\delta_2(\mathcal{H}) - \varepsilon n)\varepsilon n > \frac{\varepsilon n^2}{3}$$

pairs of differently-coloured touching edges of $\mathcal{H}$ have intersection $\{x, y\}$. Thus, G is "nearly complete", missing at most $\frac{3\eta n^2}{\varepsilon}$ edges.

Claim 4.10. *At most $\varepsilon^2 n^2$ pairs of vertices are in fewer than $(\frac{1}{3} + \mu - \varepsilon)n$ monochromatic triangles of G.*

Proof. Since G has at most $\frac{3\eta n^2}{\varepsilon} < \frac{1}{6}\varepsilon^3 n^2$ non-edges, fewer than $\frac{1}{6}\varepsilon^3 n^3$ edges of $\mathcal{H}$ contain a non-edge of G. For every coloured edge of $\mathcal{H}$ which contains an edge of G coloured differently, there are at least $(\frac{1}{3} + \mu - \varepsilon)n$ pairs of differently-coloured touching edges of $\mathcal{H}$, so we conclude that at most $\frac{1}{6}\varepsilon^3 n^3$ such edges exist. Therefore, all but at

most $\frac{1}{3}\varepsilon^3 n^3$ edges of $\mathcal{H}$ correspond to monochromatic triangles in G. Consequently, at most $\varepsilon^2 n^2$ pairs of vertices lie in more than εn edges of $\mathcal{H}$ which do not correspond to a monochromatic triangle, so all other pairs of vertices lie in at least $(\frac{1}{3} + \mu - \varepsilon)n$ monochromatic triangles in G, proving our claim. ∎

Now, let us remove all the edges of G which are in fewer than $(\frac{1}{3} + \mu - \varepsilon)n$ monochromatic triangles of G, resulting in a graph missing at most

$$\frac{3\eta n^2}{\varepsilon} + \varepsilon^2 n^2 < 2\varepsilon^2 n^2$$

edges, and then remove, iteratively, at most $\frac{1}{2}\varepsilon n$ vertices with minimum degree from this graph to obtain a graph G' of minimum degree at least $(1 - 4\varepsilon)n$. In G', every edge is in at least $(\frac{1}{3} + \frac{\mu}{2})n$ monochromatic triangles, since any edge of G' was in at least $(\frac{1}{3} + \mu - \varepsilon)n$ monochromatic triangles of G, and in passing to G', edges have been removed from at most $8\varepsilon n$ such triangles and vertices have been removed from at most $\frac{1}{2}\varepsilon n$ of these triangles. Calling a triangle *rainbow* if its edges have three different colours, we then have the following claim.

Claim 4.11. *There are at most $\varepsilon^3 n^3$ rainbow triangles in G'.*

Proof. If xyz is a rainbow triangle in G', say with xy red, yz green and xz blue, then write W_r for the set of vertices $w \in V(\mathcal{H})$ such that wxy is a red edge of $\mathcal{H}$, and define W_g and W_b analogously. We clearly have

$$|W_r|, |W_g|, |W_b| \geq \left(\frac{1}{3} + \mu - \varepsilon\right)n,$$

so we can find μn vertices in more than one set, and hence μn pairs of differently-coloured touching edges. Each such pair arises from at most two rainbow triangles, so it follows that there are at most $\frac{2\eta n^3}{\mu} < \varepsilon^3 n^3$ rainbow triangles in G'. ∎

Next, suppose there are at least three colours in G', say red, green and blue. Each colour meets at least $(\frac{1}{3} + \frac{\mu}{2})n$ vertices by the definition of G', and at most εn vertices are in $\varepsilon^2 n^2$ rainbow triangles by Claim 4.11, so we may pick a vertex v incident to edges of at least two colours, red and green say, which is in fewer than $\varepsilon^2 n^2$ rainbow triangles. Let V_r and V_g be the red and green neighbourhoods of v in G'. Since $|V_r \cup V_g| \geq (\frac{2}{3} + \mu)n$, at least $\frac{3}{2}\mu n > 2\varepsilon n$ vertices in $V_r \cup V_g$ meet blue edges, and at least one of these vertices is in fewer than εn rainbow triangles containing v. Pick such a vertex w, and assume without loss of generality that $w \in V_r$. Note that, as $vw \in E(G')$ is red, and thus in at least $(\frac{1}{3} + \frac{\mu}{2})n$ red triangles in G', w has at least $(\frac{1}{3} + \frac{\mu}{2})n$ red neighbours in V_r. Furthermore, as w meets a blue edge, it must be in at least $(\frac{1}{3} + \frac{\mu}{2})n$ blue edges, at most εn of which are with a vertex in V_g (for each such edge gives rise to a rainbow triangle containing vw). In total, w is in at least $(\frac{2}{3} + \mu - \varepsilon)n$ blue or green edges without a vertex in V_g, which contradicts $|V_g| \geq (\frac{1}{3} + \frac{\mu}{2})n$.

Therefore, G' only has two colours, say, red and green. Write B_r and B_g for the set of vertices spanned by the red and green edges respectively, and suppose without loss of generality that $|B_r| \leq |B_g|$.

Claim 4.12. *We have*

$$|B_r| \leq \left(1 - \frac{\mu}{4} - 6\varepsilon\right)n,$$

whence $|V(G') \setminus B_r| \geq (\frac{\mu}{4} + 3\varepsilon)n$.

Proof. Suppose, for the sake of a contradiction, that

$$|B_g| \geq |B_r| \geq \left(1 - \frac{\mu}{4} - 6\varepsilon\right)n.$$

At least $(1 - \frac{\mu}{2} - 12\varepsilon)n$ vertices are in both B_r and B_g, and each such vertex is incident to at least $(\frac{1}{3} + \frac{\mu}{2})n$ edges of each colour, and to at least $(1 - 4\varepsilon)n$ edges in total by the definition of G'. Consequently, for each such vertex v, there are at least

$$\left(\frac{1}{3} + \frac{\mu}{2}\right)\left(\frac{2}{3} - \frac{\mu}{2} - 4\varepsilon\right)n^2 \geq \frac{n^2}{5}$$

pairs $x, y \in V(G')$ with vx red and vy green. Each such pair produces a triple which is not a monochromatic triangle in G', and no such triple is counted more than twice, so there are at least $\frac{n^3}{11}$, say, distinct triples of this kind. However, G' has at least $(1 - \frac{\varepsilon}{2})(1 - 4\varepsilon)\frac{n^2}{2}$ edges and each edge of G' is in at least $(\frac{1}{3} + \frac{\mu}{2})n$ monochromatic triangles, so the number of triples inducing monochromatic triangles is at least

$$\left(1 - \frac{\varepsilon}{2}\right)(1 - 4\varepsilon)\left(\frac{1}{3} + \frac{\mu}{2}\right)\frac{n^3}{3} > \frac{n^3}{11}.$$

We therefore find at least

$$\frac{n^3}{11} + \frac{n^3}{11} > \binom{n}{3}$$

distinct triples on $V(G')$, which is a contradiction. ∎

We now obtain the colouring we seek as follows: we un-colour all the edges of $\mathcal{H}$ not coloured red or green, as well as all the edges of $\mathcal{H}$ containing either a non-edge of G' or a vertex not in $V(G')$. Let us count the number of edges of $\mathcal{H}$ that are no longer coloured after we do this. At most ηn^3 edges of $\mathcal{H}$ were not coloured in the original colouring, at most $2\varepsilon^2 n^3$ edges of $\mathcal{H}$ contain a non-edge of G', and at most $\frac{1}{2}\varepsilon n^3$ edges of $\mathcal{H}$ contain a vertex not in $V(G')$. Finally, any edge of $\mathcal{H}$ containing an edge of G' not coloured either red or green contributes at least $\frac{n}{3}$ pairs of differently-coloured touching edges, so there are at most $3\eta n^3$ such edges. Therefore, the number of edges of $\mathcal{H}$ not coloured in our final colouring is at most

$$\eta n^3 + 2\varepsilon^2 n^3 + \frac{\varepsilon n^3}{2} + 3\eta n^3 < \varepsilon n^3.$$

Finally, notice that any red edge in our final colouring of $\mathcal{H}$ incident to a vertex in $V(G') \setminus B_r$ is in at least $\frac{n}{3}$ pairs of differently-coloured touching edges, so the number of such edges is at most $3\eta n^3 < \varepsilon^2 n^3$. It follows from Claim 4.12 that at least $\frac{1}{4}\mu n$ vertices in $V(G') \setminus B_r$ meet fewer than εn^2 red edges, as required. All the other properties that we need are inherited from the original colouring of $\mathcal{H}$, proving the statement. ∎

4.4. Green-tinged absorbers

We now build absorbing structures in an (ε, μ)-coloured 3-graph. These structures will be crucial in transforming an "almost spanning" copy of a surface into a spanning one.

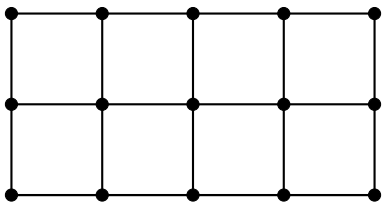
Fix an (ε, μ)-coloured 3-graph $\mathcal{H}$. We say a subgraph of $\mathcal{H}$ is *green-tinged* if it has at least two green edges. Given a subset U of the vertices of $\mathcal{H}$, a *green-tinged absorber for* U is a sphere S in $\mathcal{H}[V(\mathcal{H}) \setminus U]$ which contains two green edges e and f so that for every $U' \subset U$, there is a sphere in $\mathcal{H}$ on the vertex set $V(S) \cup U'$ containing both e and f.

Lemma 4.13. *Let $\frac{1}{n} \ll \varepsilon \ll \mu \ll 1$, $\frac{1}{n} \ll \eta$, and suppose that $\mathcal{H}$ is an (ε, μ)-coloured 3-graph with n vertices. Then for any $R \subset V(\mathcal{H})$ with $|R| \leq \frac{1}{72}\mu n$, there is a collection of $l \leq \eta n$ vertex-disjoint spheres in $\mathcal{H}[V(\mathcal{H}) \setminus R]$, spanning at most $8|R|$ vertices in total, which are respectively green-tinged absorbers for some pairwise disjoint sets $R_1, R_2, \ldots, R_l \subset R$ with $R_1 \cup R_2 \cup \cdots \cup R_l = R$.*

Proof. Fix an integer $k \geq \frac{\mu}{36\eta}$. We shall use a greedy procedure to iteratively find "large" absorbers for subsets of R of size k until fewer than $\frac{1}{2}\eta n$ vertices of R remain, and then find "small" absorbers for the remaining vertices one by one. We shall also ensure that the absorber that we construct for a subset $X \subset R$ at any particular iteration of our procedure uses at most $8|X|$ vertices. Observe that our choice of k ensures that at most $\frac{1}{2}\eta n$ large absorbers and $\frac{1}{2}\eta n$ small absorbers are built.

Let us now describe a step of our iterative procedure. Write A for the union of the vertices in the vertex-disjoint absorbers selected so far, and R' for the remaining vertices in R for which we have yet to built an absorber; in particular, we initially have $A = \varnothing$ and $R' = R$. Note that we shall always ensure that $|A \cup R| < 9|R| \leq \frac{\mu n}{8}$.

We first describe how we proceed when $|R'| \geq \frac{1}{2}\eta n$. Let $B \subset V(\mathcal{H})$ be the set of vertices contained in at most εn^2 red edges, so that $|B| \geq \frac{1}{4}\mu n$, and choose a set $B' \subset B \setminus (A \cup R)$ with $|B'| = \frac{\mu n}{8}$. For each $v \in R'$, let L_v be the subgraph of the link-graph $L(v)$ induced by the set $V(\mathcal{H}) \setminus (A \cup R)$. Let $L'_v \subset L_v$ be the subgraph of L_v consisting of those edges of L_v which are additionally in at least $\frac{n}{3}$ green edges of $\mathcal{H}$. We shall show that L'_v contains many edges incident with B'.



k

Fig. 4. The double ladder DL_k.

Since $|A \cup R| \leq \frac{\mu n}{8}$, it follows that L_v has minimum degree $(\frac{1}{3} + 7\frac{\mu}{8})n$. Every vertex $x \in B'$ has at least $(\frac{1}{3} + \frac{3\mu}{4})n$ edges in L_v to $V(\mathcal{H}) \setminus (A \cup R \cup B')$, and by the definition of B', at most $\frac{4\varepsilon n}{\mu}$ of these edges in L_v are contained in more than $\frac{1}{4}\mu n$ red edges of $\mathcal{H}$, for otherwise we would get

$$\left(\frac{4\varepsilon n}{\mu}\right)\left(\frac{1}{4}\mu n\right) = \varepsilon n^2$$

red edges incident with $x \in B'$. A similar calculation shows that at most $12\varepsilon\frac{n^2}{\mu}$ edges of L_v are in more than $\frac{1}{4}\mu n$ uncoloured edges of $\mathcal{H}$. Easily, if e is an edge of L_v contained in at most $\frac{1}{2}\mu n$ edges of $\mathcal{H}$ that are either red or uncoloured, then e lies in L'_v. From these facts, we conclude that L'_v contains more than one third of all edges between B' and $V(\mathcal{H}) \setminus (A \cup R \cup B')$ for each $v \in R'$.

For a set $U \subset R'$ of size k, write $L_U = \cap_{v\in U} L'_v$. Now, by Lemma 4.1 – applied with $\mathcal{X}$ being the set $B' \times (V(\mathcal{H}) \setminus (A \cup R \cup B'))$ of possible edges of L_U, $r = k$ and $\gamma = \frac{1}{3}$ – we may find a set $U \subset R'$ of size k such that L_U contains more than a $\frac{3^{-k}}{2}$ proportion of the pairs in $B' \times (V(\mathcal{H}) \setminus (A \cup R \cup B'))$, so

$$|E(L_U)| \geq \frac{3^{-k}\mu n^2}{17}.$$

Each edge in L_U is contained in at least $\frac{n}{3}$ green edges of $\mathcal{H}$ by definition, so averaging using Lemma 4.1 again – now with $\mathcal{X} = E(L_U)$, $m = |V(\mathcal{H}) \setminus (A \cup R)|$, each set X_i being the subset of $E(L_U)$ which form green edges with a given vertex in $V(\mathcal{H})\setminus(A\cup R)$, $r = k + 1$ and $\gamma = \frac{1}{4}$ – we can find a set $T = \{t_1, t_2, \ldots, t_{k+1}\}$ of $k + 1$ vertices disjoint from U, and a subgraph L'_U of L_U with at least $\frac{1}{136}12^{-k}\mu n^2$ edges with the property that each edge in L'_U and each vertex in T together form a green edge of $\mathcal{H}$.

We can now find a copy of the double ladder DL_k with k spaces, depicted in Figure 4, in L'_U; indeed, since DL_k is bipartite, Theorem 2.1 guarantees the existence of such a copy. The green edges consisting of vertices of T and edges of the copy of DL_k contain an absorbing sphere S_U for U, shown in Figure 5. Since all the edges of S_U containing t_{k+1} are green, and are present no matter which vertices are absorbed, S_U is a green-tinged absorber; it has $4(k + 1) \leq 8k$ vertices, as required.

We now turn to dealing with how to proceed when $|R'| < \frac{1}{2}\eta n$. In this case, we find absorbing spheres for the vertices in R' one by one as follows. We pick $v \in R'$, construct the graph L'_v as above, and use Lemma 4.1 with $\mathcal{X} = E(L'_v)$, $r = 2$ and $\gamma = \frac{1}{4}$, to find a pair of vertices $T = \{t_1, t_2\}$ disjoint from $A \cup R$ and a graph $L''_v \subset L'_v$ with at least

$$\frac{|E(L'_v)|}{32} > \frac{n^2}{800}$$

edges, each of which forms a green edge of $\mathcal{H}$ with both t_1 and t_2. We may find a copy of DL_1 in L''_v using Theorem 2.1, and the corresponding green edges of $\mathcal{H}$ contain a green-tinged absorber for $\{v\}$ of order 8, obtained by repeating the construction of Figure 5 with $k = 1$.

The iterative procedure described above clearly produces the absorbers we require, proving the lemma. ∎

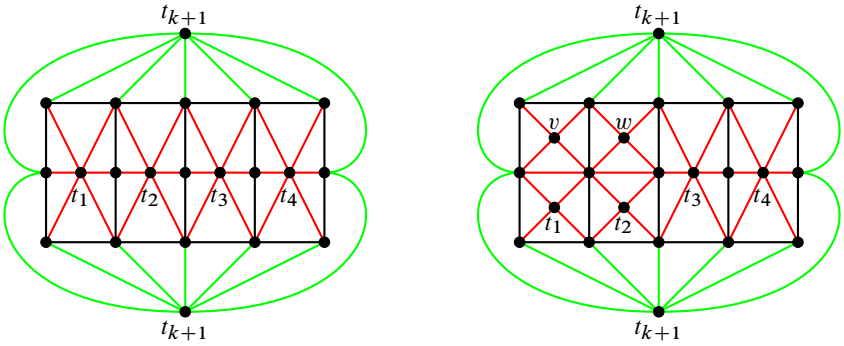


Fig. 5. The absorbing sphere S_U for U is shown on the left; each triangle in the figure corresponds to an edge of $\mathcal{H}$, which exists because the underlying double ladder lies in L'_U. On the right, we show how S_U can absorb two arbitrary vertices $v, w \in U$.

By letting the spheres we construct absorb all of R in Lemma 4.13, we immediately obtain the following corollary.

Corollary 4.14. *Let $\frac{1}{n} \ll \varepsilon \ll \mu \ll 1$, $\frac{1}{n} \ll \eta$ and suppose that $\mathcal{H}$ is a (ε, μ)-coloured 3-graph on n vertices. Then for any $R \subset V(\mathcal{H})$ with $|R| \leq \frac{\mu n}{72}$, there is a collection of at most ηn vertex-disjoint green-tinged spheres in $\mathcal{H}$ which cover R.* ∎

4.5. Spheres covering twice as many bad vertices as good vertices

The following lemma will allow us to create spheres that use twice as many vertices from some specific set of "bad" vertices as from a specific set of "good" vertices. This will prove useful in Section 4.6 since the connectibility properties of our 3-graphs will impose constraints of precisely this nature.

Lemma 4.15. *Let $\frac{1}{N} \ll \beta, \gamma, \frac{1}{k}$ and suppose that $\mathcal{H}$ is a 3-graph on vertex classes A and B, where A has at least N vertices and every edge has one vertex in B and two vertices in A. Suppose that for at least βN vertices $v \in B$, the link-graph $L(v)$ has at least $(\frac{1}{2}+\gamma)|A|$ vertices with degree at least $(\frac{1}{2}+\gamma)|A|$. Then there is a sphere S in $\mathcal{H}$ which has $4k$ vertices in B and $2k+2$ vertices in A.*

Proof. Write B' for the set of vertices $v \in B$ for which $L(v)$ has at least $(\frac{1}{2}+\gamma)|A|$ vertices with degree at least $(\frac{1}{2}+\gamma)|A|$. For each $v \in B'$, let H_v be the 3-graph of edges which are triangles in $L(v)$. Choose a vertex v_1 with degree at least $(\frac{1}{2}+\gamma)|A|$ in $L(v)$, and note that at least $2\gamma|A|$ of its neighbours must have degree at least $(\frac{1}{2}+\gamma)|A|$ in $L(v)$ also. Thus, choosing v_2 to be such a neighbour, there are at least $2\gamma|A|$ triangles containing v_1 and v_2. As there were at least $2\gamma(\frac{1}{2}+\gamma)|A|^2$ choices for the pair (v_1, v_2), $L(v)$ contains at least $\frac{1}{3}\gamma^2|A|^3$ triangles. Hence, H_v has at least $\frac{1}{3}\gamma^2|A|^3$ edges.

By Lemma 4.1 we can now find a set $K \subset B'$ of size $4k$ for which the 3-graph $H_K = \bigcap_{v\in K} H_v$ has at least $\frac{1}{12}(2\gamma^2)^{4k}|A|^3$ edges. By Lemma 4.2, applied with $T = A$

and $\delta = \frac{1}{12}(2\gamma^2)^{4k}2$, we can find a double pyramid on exactly $2k+2$ vertices in H_K. This double pyramid has $4k$ faces, and by putting a vertex from K into each face of this double pyramid, we get a sphere in $\mathcal{H}$ with $4k$ vertices from B and $2k+2$ vertices from A. ∎

4.6. Almost spanning green spheres

We now show that we can find a reasonably small set of disjoint spheres which span all but a small fraction of the vertices of an (ε, μ)-coloured 3-graph $\mathcal{H}$. The proof uses Szemerédi's regularity lemma, and the following lemma will be used to partition the reduced graph.

Lemma 4.16. *Let $\frac{1}{n} \ll \varepsilon$, and suppose that G is an n-vertex graph. Then we can find a partition of $V(G)$ into sets Z, B, C and D such that*

(1) *there is a perfect matching in Z,*

(2) *there is a perfect matching between C and D, and*

(3) $|E(B \cup D, Z \cup B \cup D)| \leq \varepsilon n^2$.

Proof. Let M be a matching of maximum size in G, and for a set of vertices X, write $N_M(X)$ for the neighbourhood of X in M, i.e., for the set of vertices joined to X by edges of M. Let B be the set of vertices which are not in M. We will find sequences of disjoint sets $C_1, \dots, C_r$ and $D_1, \dots, D_r$ as follows: iteratively, for each $i \geq 1$, if there are at least $\frac{1}{2}\varepsilon n$ vertices in

$$V(M) \setminus \bigcup_{j<i}(C_i \cup D_i)$$

each with at least $\frac{1}{2}\varepsilon n$ neighbours in

$$B \cup \left(\bigcup_{j<i} D_j\right),$$

then let C_i be the set of those vertices and set $D_i = N_M(C_i)$.

Note that the sets $C_1, C_2, \dots, C_r$ as constructed above are disjoint, and each such set has size at least $\frac{1}{2}\varepsilon n2$; this ensures that $r \leq \frac{2}{\varepsilon}$. We may also assume that $r \geq 1$, since if not, we are done by setting $Z = V(M)$, $B = V(G) \setminus Z$ and $C = D = \varnothing$. Indeed, then $E(B \cup D, Z \cup B \cup D) = E(B, Z)$ by the maximality of M. Since no iterations of our procedure were possible, there are at most $\frac{1}{2}\varepsilon n$ vertices in Z with more than $\frac{1}{2}\varepsilon n$ neighbours in B, contributing at most $\frac{1}{2}\varepsilon n^2$ edges in total, and the remaining vertices contribute at most $\frac{1}{2}\varepsilon n^2$ edges, giving the required result.

Now, let $C = \bigcup_{i=1}^{r} C_i$ and let $D = \bigcup_{i=1}^{r} D_i = N_M(C)$, and note that $1 \leq r \leq \frac{2}{\varepsilon}$. We shall now deduce some properties of the sets of vertices we have constructed thus far.

Claim 4.17. *For any vertex $v \in B \cup D$ and any set $U \subset B \cup C \cup D$ with size at most $\frac{1}{8}\varepsilon n$ such that $v \notin U \cup N_M(U)$, there is, for some $0 \leq s \leq \frac{1}{\varepsilon}$, a path of length $2s$ from v into B in $G[(B \cup C \cup D) \setminus U]$ which contains s edges from M.*

Proof. Starting with $x_0 = v$, find the longest sequence of distinct vertices

$$x_0, y_0, x_1, y_1, \ldots, y_{s-1}, x_s$$

in $(B \cup C \cup D) \setminus U$ such that

(i) for each $0 \leq i < s$, $x_i y_i \in M$,

(ii) for each $0 \leq i \leq s$, $x_i \in (B \cup (\bigcup_{j \leq r-i} D_j)) \setminus (U \cup N_M(U))$, and

(iii) for each $0 \leq i < s$, $y_i x_{i+1} \in E(G)$.

Note that such a sequence exists as the one-term sequence x_0 satisfies these conditions. Note also that if $x_i \in B$ for some i, then $i = s$ and the sequence stops at that point since there are no subsequent choices for y_i. In particular, since x_r, if it exists, must be in B, we have $s \leq r \leq \frac{1}{\varepsilon}$. We claim that $x_s \in B$, and that this sequence consequently gives us the desired path.

Suppose for the sake of a contradiction that $x_s \notin B$. Then

$$x_s \in \left(\bigcup_{j \leq r-s} D_j \right) \setminus (U \cup N_M(U)),$$

so choosing y_s such that $x_s y_s \in M$, we have

$$y_s \in \left(\bigcup_{j \leq r-s} C_j \right) \setminus U.$$

Note that since $x_i y_i \in M$ for each $i < s$, we have that y_s is distinct from $y_1, \ldots, y_{s-1}$. We know that $y_s \in \bigcup_{j \leq r-s} C_j$, so, by definition, y_s has at least $\frac{1}{2}\varepsilon n$ neighbours in $B \cup (\bigcup_{j \leq r-(s+1)} D_j)$. Therefore, since

$$|U \cup N_M(U) \cup \{x_0, y_0, x_1, y_1, \ldots, y_{s-1}, x_s\}| \leq \frac{\varepsilon n}{4} + \frac{2}{\varepsilon} < \frac{1}{2}\varepsilon n,$$

we can find a vertex x_{s+1} which is a neighbour of y_s satisfying

$$x_{s+1} \in B \cup \left(\bigcup_{j \leq r-(s+1)} D_j \right) \setminus (U \cup N_M(U) \cup \{x_0, y_0, x_1, y_1, \ldots, y_{s-1}, x_s\}),$$

contradicting maximality of the sequence, and proving the claim. ∎

Claim 4.18. *At most one vertex in each edge of M has at least three neighbours in $B \cup D$.*

Proof. Suppose to the contrary that for some edge $u_0 v_0 \in M$, each of u_0 and v_0 has at least three neighbours in $B \cup D$. Then there exist distinct vertices u_1 and v_1 in the set $(B \cup D) \setminus \{u_0, v_0\}$ with $u_0 u_1, v_0 v_1 \in E(G)$. Note that $B \cup D$ contains no edges of M by definition, so we necessarily have $u_1 \neq N_M(v_1)$.

Appealing to Claim 4.17, we now find an integer s_1 with $0 \leq s_1 \leq r$ and a path P_1 in $G[(B \cup C \cup D) \setminus \{u_0, v_0, v_1, N_M(v_1)\}]$ with length $2s_1$ from u_1 into B which contains s_1 edges in M. Note that $N_M(V(P_1)) \subset V(P_1)$, since this path contains edges of M covering all but the last vertex of the path, which is in B and hence not in $V(M)$.

Using Claim 4.17 again, we find another integer s_2 with $0 \leq s_2 \leq r$ and a path P_2 in $G[(B \cup C \cup D) \setminus (\{u_0, v_0\} \cup V(P_1))]$ with length $2s_2$ from v_1 into B which contains s_2 edges in M.

Switching edges into the matching M along P_1 and P_2 creates a matching in G with the same number of edges as M which does not contain either u_1 or v_1. Removing u_0v_0 from M and adding u_0u_1 and v_0v_1 thus creates a matching larger than M in G, which is a contradiction. ∎

By construction, every vertex in C has at least $\frac{1}{2}\varepsilon n$ neighbours in the set $B \cup D$, so Claim 4.18 implies that C contains no edges of M. Since $C \subset V(M)$, and $D = N_M(C)$, we have that $C \cap D = \varnothing$ and that M contains a perfect matching between C and D. Now, let $Z = V(M) \setminus (C \cup D)$ and note that M also contains a perfect matching on Z, and that Z, B, C and D partition $V(G)$. It therefore suffices to show that

$$|E(B \cup D, Z \cup B \cup D)| \leq \varepsilon n^2.$$

To do so, we first make the following simple observation.

Claim 4.19. *There are no edges of G within the set $B \cup D$.*

Proof. Suppose not, so that there is an edge $uv \in E(G)$ with $u, v \in B \cup D$. Note that $uv \notin M$ since every edge of M meets $Z \cup C$.

Appealing to Claim 4.17, we first find an integer s_1 with $0 \leq s_1 \leq r$ and a path P_1 in $G[B \cup C \cup D \setminus \{v, N_M(v)\}]$ with length $2s_1$ from u into B which contains s_1 edges in M. As in the proof of Claim 4.18, we have $N_M(V(P_1)) \subset V(P_1)$, and consequently $v, N_M(v) \notin N_M(P_1)$

Using Claim 4.17 again, we find another integer s_2 with $0 \leq s_2 \leq r$ and a path P_2 in $G[B \cup C \cup D \setminus V(P_1)]$ with length $2s_2$ from v into B which contains s_2 edges in M.

Switching edges into the matching M along P_1 and P_2 creates a matching in G with the same number of edges as M which does not contain u or v. Adding uv creates a matching larger than M in G, which is a contradiction. ∎

We know that there are at most $\frac{1}{2}\varepsilon n$ vertices in Z with at least $\frac{1}{2}\varepsilon n$ neighbours in G in $B \cup D$ (for otherwise, the iterative process we used above would have continued). We know by Claim 4.19 that the set $B \cup D$ induces no edges, so we deduce that

$$|E(B \cup D, Z \cup B \cup D)| \leq \left(\frac{1}{2}\varepsilon n\right)|Z| + \left(\frac{1}{2}\varepsilon n\right)|B \cup D| \leq \varepsilon n^2,$$

as required. ∎

We are now in a position to prove the "almost covering" lemma that we require.

Lemma 4.20. *Let $\frac{1}{n} \ll \eta \ll \varepsilon \ll \mu \ll 1$, and suppose that $\mathcal{H}$ is an (ε, μ)-coloured 3-graph on n vertices. Then we can find a collection of at most ηn vertex-disjoint green spheres in $\mathcal{H}$ which cover at least $(1 - \mu^3)n$ vertices of $\mathcal{H}$.*

Proof. Our primary strategy is to set aside a small set T of vertices which will be used as apexes of double pyramids covering most of the vertices of $\mathcal{H}$; in fact, we shall find it con-

venient to divide T into three parts to be used at different stages of the proof. We also set aside a larger set A of vertices that lie in few red edges; these vertices will be used to create additional spheres that cover vertices that are difficult to cover using double pyramids. Having set aside these sets, we define an auxiliary graph on the remaining vertices whose edges are those pairs which form green edges of $\mathcal{H}$ with a reasonable number of vertices in T. We then use the regularity lemma to partition this graph. Regularity properties will then ensure that we may find double pyramids with apexes in T covering almost all of any given regular pair of reasonable density. Applying Lemma 4.16 to the reduced graph, we shall find disjoint regular pairs covering all but a few "bad" parts, and the additional conditions imposed by Lemma 4.16 on these bad parts will allow us to use Lemma 4.15 to find spheres covering most of the vertices from these bad parts with the help of additional "good" vertices, typically from A. We make this outline precise below.

Let k and l be integers such that $\frac{1}{n} \ll \frac{1}{k} \ll \eta$ and $\frac{1}{k} \ll \frac{1}{l} \ll \varepsilon$. Here l will be the number of parts of a regularity partition, and k will correspond to the size of spheres used in Lemma 4.15.

Pick random disjoint sets $T_1, T_2, T_3 \subset V(\mathcal{H})$ by including each vertex independently at random in T_i with probability $\frac{\mu^3}{10}$ for each $i \in [3]$, and in none of them with the remaining probability $1 - \frac{3}{10}\mu^3$. Observe that, by a simple application of Proposition 2.5, the properties **(A1)** and **(A2)** below hold with high probability for such random choices; thus, we can find sets $T_1, T_2, T_3 \subset V(\mathcal{H})$ satisfying the following conditions:

(A1) For each $i \in [3]$, $\frac{\mu^3 n}{20} \le |T_i| \le \frac{\mu^3 n}{8}$.

(A2) For each $i \in [3]$, every pair of vertices contained in at least $\varepsilon^2 n$ green edges of $\mathcal{H}$ is in at least $\varepsilon^3 n$ green edges with a third vertex in T_i.

Let A be the set of vertices in $V(\mathcal{H}) \setminus (T_1 \cup T_2 \cup T_3)$ that are in at most εn^2 red edges. Let G be the graph with vertex set $V(\mathcal{H}) \setminus (A \cup T_1 \cup T_2 \cup T_3)$ where for each pair of vertices $x, y \in V(G)$, we have $xy \in E(G)$ if and only if for each $i \in [3]$, there are at least $\varepsilon^3 n$ green edges in $\mathcal{H}$ of the form xyt with $t \in T_i$. In particular, by **(A2)**, we have $xy \in E(G)$ for every pair $x, y \in V(G)$ contained in at least $\varepsilon^2 n$ green edges of $\mathcal{H}$.

By Theorem 2.2, we can find a partition

$$V(G) = Y_0 \cup Y_1 \cup \cdots \cup Y_l$$

so that $|Y_0| \le \frac{\varepsilon^2 n}{100}$, $|Y_1| = |Y_2| = \cdots = |Y_l|$, and all but at most $\frac{1}{100}\varepsilon^2 l^2$ pairs of these sets are $(\frac{1}{100}\varepsilon^2)$-regular in G. Let R be the reduced graph on the vertex set $\{Y_1, Y_2, \ldots, Y_l\}$ with edges corresponding to regular pairs with density at least $\frac{1}{10}\varepsilon^2$ in G. Using Lemma 4.16, we obtain, for some m, a partition of $V(R)$ into sets $\mathcal{Z}, \mathcal{B}, \mathcal{C} = \{C_1, C_2, \ldots, C_m\}$ and $\mathcal{D} = \{D_1, D_2, \ldots, D_m\}$ so that

(1) $R[\mathcal{Z}]$ has a perfect matching,

(2) $C_i D_i \in E(R)$ for each $i \in [m]$, and

(3) $|E_R(\mathcal{B} \cup \mathcal{D}, \mathcal{Z} \cup \mathcal{B} \cup \mathcal{D})| \le \frac{1}{10}\varepsilon^2 l^2$.

Now, let Z, B, C and D be, respectively, the vertices of G contained in the classes in $\mathcal{Z}, \mathcal{B}, \mathcal{C}$ and $\mathcal{D}$. Note that the sets $T_1, T_2, T_3, A, B, C, D, Z$ and Y_0 partition $V(\mathcal{H})$. We begin with some simple properties of these sets.

Claim 4.21. *There are at most εn vertices $u \in B \cup D$ for which there are more than εn^2 green edges of $\mathcal{H}$ containing u and at least one other vertex in $Z \cup B \cup D$.*

Proof. Suppose to the contrary that there are more than εn such vertices. For each such vertex u, let X_u be the set of vertices $v \in Z \cup B \cup D$ for which at least $\frac{1}{2}\varepsilon n$ green edges of $\mathcal{H}$ contain u and v. We must have $|X_u| \geq \frac{1}{2}\varepsilon n$, since otherwise fewer than $\frac{1}{2}\varepsilon n^2$ green edges contain u and some vertex in X_u, and fewer than $\frac{1}{2}\varepsilon n^2$ contain u but no vertex in X_u. Every ordered pair of the form (u, v) with $v \in X_u$ corresponds to an edge of G, since there are more than $\varepsilon^2 n$ green edges containing $\{u, v\}$, and $u, v \in Z \cup B \cup D \subset V(G)$. There are at least $\sum_u |X_u| \geq \frac{1}{2}\varepsilon^2 n^2$ such ordered pairs, yielding

$$|E_G(B \cup D, Z \cup B \cup D)| \geq \frac{1}{4}\varepsilon^2 n^2.$$

Now, G has at most $\frac{n^2}{l^2}$ edges within Y_i for each i, at most $\frac{1}{100}\varepsilon^2 n^2$ edges between irregular pairs, and at most $\frac{1}{10}\varepsilon^2 n^2$ edges between pairs of density below $\frac{1}{10}\varepsilon^2$. Thus, all but at most $\frac{1}{8}\varepsilon^2 n^2$ edges of G lie between regular pairs corresponding to an edge in R, so we also have

$$\begin{aligned}|E_G(B \cup D, Z \cup B \cup D)| &< \left(\frac{n}{l}\right)^2 |E_R(\mathcal{B} \cup \mathcal{D}, \mathcal{Z} \cup \mathcal{B} \cup \mathcal{D})| + \frac{\varepsilon^2 n^2}{8} \\ &< \frac{\varepsilon^2 n^2}{4},\end{aligned}$$

contradicting the bound above and establishing the claim. ∎

Claim 4.22. *For all but at most $\frac{1}{64}\mu^3 n$ vertices $v \in B \cup D$, the green link graph $\mathrm{GL}(v)$ has at least $(\frac{1}{3} + \frac{\mu}{4})n$ vertices in $A \cup C$ with at least $(\frac{1}{3} + \frac{\mu}{4})n$ neighbours in $A \cup C$.*

Proof. By Claim 4.21 and Lemma 4.8, there is a set $U \subset B \cup D$ of size at least

$$|B \cup D| - \varepsilon n - \mu^4 n > |B \cup D| - \frac{\mu^3 n}{64}$$

such that for each $v \in U$, the green link graph $\mathrm{GL}(v)$ has the following properties:

(B1) The number of edges of $\mathrm{GL}(v)$ with a vertex in $Z \cup B \cup D$ is at most εn^2.

(B2) At least $(\frac{1}{3} + \frac{\mu}{2})n$ vertices of $\mathrm{GL}(v)$ have degree at least $(\frac{1}{3} + \frac{\mu}{2})n$.

Fixing $v \in U$, we shall show the property in the claim holds for this vertex v, thereby completing the proof of the claim.

Write W_v for the set of vertices with degree at least $(\frac{1}{3} + \mu 2)n$ in $\mathrm{GL}(v)$. Let

$$W_1 = W_v \cap (Z \cup B \cup D).$$

There are at least $|W_1|(\frac{1}{3} + \mu 2)\frac{n}{2}$ edges in $\mathrm{GL}(v)$ with a vertex in $Z \cup B \cup D$, so by **(B1)**, we have $|W_1| < 6\varepsilon n < \mu^2 n$.

Let $W_2 \subset A \cup C$ be the set of vertices which, in $\mathrm{GL}(v)$, have more than $\frac{1}{6}\mu n$ neighbours in $Z \cup B \cup D$. There are at least $|W_2|\frac{1}{6}\mu n$ edges in $\mathrm{GL}(v)$ with a vertex in the set $Z \cup B \cup D$, so by **(B1)**, we again have $|W_2| < \mu^2 n$.

Every vertex in $W_v \setminus (W_1 \cup W_2 \cup T_1 \cup T_2 \cup T_3)$ is in $A \cup C$ and has at least $(\frac{1}{3} + \frac{\mu}{3})n$ neighbours in $A \cup C \cup T_1 \cup T_2 \cup T_3$, and consequently has more than $(\frac{1}{3} + \frac{\mu}{4})n$ neighbours in $A \cup C$. Since

$$|W_v \setminus (W_1 \cup W_2 \cup T_1 \cup T_2 \cup T_3)| > |W_v| - 5\mu^2 n > \left(\frac{1}{3} + \frac{\mu}{4}\right)n,$$

the required property in the claim holds for v; this completes the proof. ∎

We are now in a position to start constructing the "almost spanning" collection of spheres that we require. Let $\mathfrak{S}_0$ be a collection of vertex-disjoint green spheres in the set $\mathcal{H}[A \cup B \cup C \cup D]$ so that $|V(\mathfrak{S}_0)|$ is maximal, subject to the following conditions:

(C1) Each sphere in $\mathfrak{S}_0$ has $4k$ vertices in $B \cup D$ and $2k + 2$ vertices in $A \cup C$.

(C2) $|D_i \cap V(\mathfrak{S}_0)| \leq |C_i \cap V(\mathfrak{S}_0)|$ for each $i \in [m]$.

First, note that $\mathfrak{S}_0$ consists of at most $\frac{n}{6k+2} \leq \eta\frac{n}{4}$ spheres, since each sphere in $\mathfrak{S}_0$ spans $6k + 2$ vertices and we chose $\frac{1}{k} \ll \eta$. Second, note that if $\mathfrak{S}_0 \neq \varnothing$, then, by **(C1)**, we have

$$\frac{|(A \cup C) \cap V(\mathfrak{S}_0)|}{|V(\mathfrak{S}_0)|} = \frac{(2k+2)|\mathfrak{S}_0|}{(6k+2)|\mathfrak{S}_0|} < \frac{1}{3} + \frac{1}{k} < \frac{1}{3} + \frac{\varepsilon}{3}; \tag{1}$$

therefore, we have the following:

$$3|(A \cup C) \cap V(\mathfrak{S}_0)| < |V(\mathfrak{S}_0)| + \varepsilon n. \tag{2}$$

Second, since $A \cup C$ and $B \cup D$ are disjoint sets in $V(\mathcal{H})$, we have

$$|(B \cup D) \setminus V(\mathfrak{S}_0)| \leq n - |V(\mathfrak{S}_0)| - |(A \cup C) \setminus V(\mathfrak{S}_0)|. \tag{3}$$

We now argue that (2) and (3), along with the maximality of $\mathfrak{S}_0$, imply the following.

Claim 4.23. *We have*

$$|(B \cup D) \setminus V(\mathfrak{S}_0)| \leq |(A \cup C) \setminus V(\mathfrak{S}_0)| + \frac{\mu^3 n}{16}.$$

Proof. Assume for the sake of a contradiction that

$$|(B \cup D) \setminus V(\mathfrak{S}_0)| > |(A \cup C) \setminus V(\mathfrak{S}_0)| + \frac{\mu^3 n}{16}. \tag{4}$$

We will distinguish between values of i for which **(C2)** holds with enough extra room that it cannot be the barrier to adding another sphere to $\mathfrak{S}_0$ and values of i for which **(C2)** is closer to equality. To this end, let $X \subset [m]$ be the set of values of i for which

$$|D_i \cap V(\mathfrak{S}_0)| > |C_i \cap V(\mathfrak{S}_0)| - 4k,$$

or equivalently for which

$$|D_i \setminus V(\mathfrak{S}_0)| < |C_i \setminus V(\mathfrak{S}_0)| + 4k,$$

and let $\tilde{D} = (\bigcup_{i \notin X} D_i) \setminus V(\mathfrak{S}_0)$. It follows from **(C2)** that

$$\begin{aligned} |\tilde{D}| &= |D \setminus V(\mathfrak{S}_0)| - \sum_{i \in X} |D_i \setminus V(\mathfrak{S}_0)| \\ &> |D \setminus V(\mathfrak{S}_0)| - \sum_{i \in X} (|C_i \setminus V(\mathfrak{S}_0)| + 4k) \\ &\geq |D \setminus V(\mathfrak{S}_0)| - |C \setminus V(\mathfrak{S}_0)| - 4mk \\ &> |D \setminus V(\mathfrak{S}_0)| - |C \setminus V(\mathfrak{S}_0)| - \varepsilon n. \end{aligned} \tag{5}$$

Finally, let $\tilde{B} = B \setminus V(\mathfrak{S}_0)$ and note that by (4) and (5), we have

$$|\tilde{B} \cup \tilde{D}| \geq |(B \cup D) \setminus V(\mathfrak{S}_0)| - |C \setminus V(\mathfrak{S}_0)| - \varepsilon n > \frac{\mu^3 n}{16} - \varepsilon n > \frac{\mu^3 n}{32}.$$

Thus, by Claim 4.22, there are at least $\frac{1}{64}\mu^3 n$ vertices $x \in \tilde{B} \cup \tilde{D}$ for which there are at least

$$\left(\frac{1}{3} + \frac{\mu}{4}\right)n - |(A \cup C) \cap V(\mathfrak{S}_0)|$$

vertices in $(A \cup C) \setminus V(\mathfrak{S}_0)$ with at least $(\frac{1}{3} + \frac{\mu}{4})n - |(A \cup C) \cap V(\mathfrak{S}_0)|$ neighbours in $(A \cup C) \setminus V(\mathfrak{S}_0)$ in the green link graph $\mathrm{GL}(x)$. Since $|(A \cup C) \cap V(\mathfrak{S}_0)| < (\frac{1}{3} + \frac{\varepsilon}{3})n$ by (1), we must have $|(A \cup C) \setminus V(\mathfrak{S}_0)| > \frac{\mu n}{8}$.

By the maximality of $\mathfrak{S}_0$, there are no green spheres with exactly $4k$ vertices in $\tilde{B} \cup \tilde{D}$ and $2k+2$ vertices in $(A \cup C) \setminus V(\mathfrak{S}_0)$, since such a sphere would be disjoint from all other spheres in $\mathfrak{S}_0$, and, by the choice of $\tilde{D}$, adding it to $\mathfrak{S}_0$ would not violate **(C2)**. By Lemma 4.15 applied to the green edges of $\mathcal{H}$ meeting two vertices of $(A \cup C) \setminus V(\mathfrak{S}_0)$ and one vertex of $\tilde{B} \cup \tilde{D}$, with $N = \frac{\mu n}{8}$, $\beta = \frac{\mu^2}{8}$ and $\gamma = \frac{\mu}{4}$, we have

$$\begin{aligned} \left(\frac{1}{3} + \frac{\mu}{4}\right)n - |(A \cup C) \cap V(\mathfrak{S}_0)| &\leq \left(\frac{1}{2} + \frac{\mu}{4}\right)|(A \cup C) \setminus V(\mathfrak{S}_0)| \\ &\leq \frac{1}{2}|(A \cup C) \setminus V(\mathfrak{S}_0)| + \frac{\mu n}{4}; \end{aligned}$$

in a little more detail, with these values of N, β and γ, there are at least βN vertices in $\tilde{B} \cup \tilde{D}$ whose link-graphs have at least $(\frac{1}{3} + \frac{\mu}{4})n - |(A \cup C) \cap V(\mathfrak{S}_0)|$ vertices with degrees at least $(\frac{1}{3} + \frac{\mu}{4})n - |(A \cup C) \cap V(\mathfrak{S}_0)|$, so for Lemma 4.15 to not apply, we need this common quantity to be at most $(\frac{1}{2} + \gamma)|(A \cup C) \setminus V(\mathfrak{S}_0)|$. We therefore have the following property:

$$n - 3|(A \cup C) \cap V(\mathfrak{S}_0)| \leq \frac{3}{2}|(A \cup C) \setminus V(\mathfrak{S}_0)|. \tag{6}$$

Adding (2), (3) and (6) together, and rearranging, we have

$$\begin{aligned} |(B \cup D) \setminus V(\mathfrak{S}_0)| &\leq \frac{1}{2}|(A \cup C) \setminus V(\mathfrak{S}_0)| + \varepsilon n \\ &\leq |(A \cup C) \setminus V(\mathfrak{S}_0)| + \frac{\mu^3 n}{16}, \end{aligned}$$

contradicting (4) and completing the proof of the claim. ∎

Now, for each $i \in [m]$, let $C_i' = C_i \setminus V(\mathfrak{S}_0)$ and choose a set $D_i' \subset D_i \setminus V(\mathfrak{S}_0)$ be a set of size $|C_i'|$, noting that such a choice is made possible by **(C2)**. Let $A' = A \setminus V(\mathfrak{S}_0)$ and

$$B' = (B \cup D) \setminus \left(\left(\bigcup_{i \in [m]} D_i' \right) \cup V(\mathfrak{S}_0) \right),$$

and note that by Claim 4.23, we have

$$\begin{aligned} |B'| &= |(B \cup D) \setminus V(\mathfrak{S}_0)| - \sum_{i \in [m]} |D_i'| \\ &\leq |(A \cup C) \setminus V(\mathfrak{S}_0)| - \sum_{i \in [m]} |C_i'| + \frac{\mu^3 n}{16} \\ &= |A'| + \frac{\mu^3 n}{16}. \end{aligned} \tag{7}$$

Furthermore, note that the sets

$$T_1, T_2, T_3, Y_0, Z, V(\mathfrak{S}_0), A', B', C_1', \ldots, C_l', D_1', \ldots, D_l' \tag{8}$$

constitute a partition of $V(\mathcal{H})$.

Claim 4.24. *There is a set $\mathfrak{S}_1$ of at most $\frac{1}{4}\eta n$ vertex-disjoint green spheres in $\mathcal{H}$ with vertices in $T_1 \cup Z \cup Y_0$ which cover all but at most $\frac{1}{4}\mu^3 n$ vertices in $T_1 \cup Z \cup Y_0$.*

Proof. Set $n' = |Y_1| = \cdots = |Y_l|$ and note that we have

$$\left(1 - \frac{\varepsilon^2}{100}\right)\frac{n}{l} \leq n' \leq \frac{n}{l}.$$

Because $R[\mathcal{Z}]$ has a perfect matching, there exist distinct integers $a_1, a_2, \ldots, a_r$ and $b_1, b_2, \ldots, b_r$ such that

$$Z = \bigcup_{i=1}^{r} (Y_{a_i} \cup Y_{b_i}),$$

where each pair (Y_{a_i}, Y_{b_i}) is $(\frac{1}{100}\varepsilon^2)$-regular with density at least $\frac{1}{10}\varepsilon^2$. In particular, for any subsets $P \subset Y_{a_i}$ and $Q \subset Y_{b_i}$ with $|P|, |Q| \geq \frac{1}{100}\varepsilon^2 n'$, there are at least $\frac{9}{100}\varepsilon^2 |P||Q|$ pairs $p \in P, q \in Q$ for which there are at least $\varepsilon^3 n$ vertices $t \in T_1$ so that pqt is a green edge of $\mathcal{H}$. Consequently, if T_1' is any set obtained by removing at most $\varepsilon^4 n$ vertices from T_1, the induced 3-partite 3-graph of green edges of $\mathcal{H}$ between P, Q and T_1' contains at least $\varepsilon^5 |P||Q||T_1'|$ edges.

For each $i \in [r]$ in turn, we use Lemma 4.3 to find at most $\frac{1}{2}\eta n'$ green spheres in $Y_{a_i} \cup Y_{b_i} \cup T_1'$, each with two vertices in T_1', covering at least $(1 - \frac{1}{100}\varepsilon^2)n'$ vertices in each of Y_{a_i} and T_{b_i}, where $T_1' \subset T_1$ is the dynamically-updated set of vertices not yet used; at each stage, note that we have

$$|T_1 \setminus T_1'| \leq \frac{1}{2}\eta n \leq \varepsilon^4 n.$$

The iterative procedure above gives us a total of at most $\frac{1}{4}\eta|Z| \le \frac{1}{4}\eta n$ spheres covering all but at most $\frac{1}{100}\varepsilon^2|Z|$ vertices of Z. Since $|T_1| \le \frac{1}{8}\mu^3 n$ and $|Y_0| \le \frac{1}{100}\varepsilon^2 n$, all but at most

$$\frac{\mu^3 n}{8} + \frac{\varepsilon^2 n}{50} < \frac{\mu^3 n}{4}$$

vertices in $T_1 \cup Z \cup Y_0$ are covered, as required. ∎

Claim 4.25. *There is a set $\mathfrak{S}_2$ of at most $\frac{1}{4}\eta n$ vertex-disjoint green spheres with vertices in the set $T_2 \cup (\bigcup_{i\in[m]}(C_i' \cup D_i'))$ which cover all but at most $\frac{1}{4}\mu^3 n$ vertices in the set $T_2 \cup (\bigcup_{i\in[m]}(C_i' \cup D_i'))$.*

Proof. As we did in the proof of Claim 4.24, for each $i \in [m]$ in turn, we may find at most $\frac{1}{2}\eta n'$ green spheres covering all but at least $(1 - \frac{1}{100}\varepsilon^2)n'$ vertices in each of C_i' and D_i', using only vertices from those sets and vertices from T_2 which have not previously been used. This covers all but at most

$$\frac{\mu^3 n}{8} + \frac{\varepsilon^2 n}{100} < \frac{\mu^3 n}{4}$$

vertices in $T_2 \cup (\bigcup_{i\in[m]}(C_i' \cup D_i'))$, using at most $\frac{1}{4}\eta n$ vertex-disjoint green spheres, establishing the claim. ∎

Claim 4.26. *There is a set $\mathfrak{S}_3$ of at most $\frac{1}{4}\eta n$ vertex-disjoint green spheres with vertices in $T_3 \cup A' \cup B'$ which cover all but at most $\frac{1}{4}\mu^3 n$ vertices in $T_3 \cup A' \cup B'$.*

Proof. Choose partitions $A' = A_1 \cup A_2 \cup A_3$ and $B' = B_1 \cup B_2$ such that $|A_1| = |B_1|$, $|A_2| = |A_3|$ and $|B_2| \le \frac{1}{16}\mu^3 n$, noting that this is made possible by (7).

First we show that we may find at most $\frac{1}{8}\eta n$ vertex-disjoint green spheres in $T_3 \cup A_1 \cup B_1$ covering all but at most $\frac{1}{64}\mu^3 n$ vertices in each of A_1 and B_1, with each sphere using two vertices from T_3. We may of course assume that $|A_1| = |B_1| \ge \frac{1}{64}\mu^3 n$ as there is nothing to prove otherwise. Suppose $A_1' \subset A_1$ and $B_1' \subset B_1$ each have size at least $\mu^4 n$. By the definition of A, there are at most εn^3 red edges of $\mathcal{H}$ meeting A_1', and since $\mathcal{H}$ is (ε, μ)-coloured, there are at most εn^3 uncoloured edges in $\mathcal{H}$. Hence, at most $6\varepsilon n^2$ pairs (a, b) with $a \in A_1'$ and $b \in B_1'$ are in at least $\frac{n}{3}$ non-green edges of $\mathcal{H}$, so at least $|A_1'||B_1'| - 6\varepsilon n^2 > \frac{1}{2}|A_1'||B_1'|$ pairs are in at least $\varepsilon^3 n$ green edges with the third vertex in T_3. By Lemma 4.3, we can find the required spheres.

Similarly, since we have $|T_3 \setminus T_3'| \le \frac{1}{4}\eta n < \varepsilon^4 n$, where T_3' is the set of unused vertices from T_3 after almost covering $A_1 \cup B_1$ as above, we may find at most $\frac{1}{8}\eta n$ vertex-disjoint green spheres in $T_3' \cup A_2 \cup A_3$ covering all but at most $\frac{1}{64}\mu^3 n$ vertices in each of A_2 and A_3. In total, these spheres cover all but at most

$$\frac{\mu^3 n}{8} + \frac{\mu^3 n}{16} + \frac{4\mu^3 n}{64} = \frac{\mu^3 n}{4}$$

vertices in $T_3 \cup A' \cup B'$. ∎

Therefore, recalling the partition in (8), it is now clear that $\mathfrak{S}_0 \cup \mathfrak{S}_1 \cup \mathfrak{S}_2 \cup \mathfrak{S}_3$ is a set of at most ηn vertex-disjoint green spheres covering all but at most $\mu^3 n$ vertices in $\mathcal{H}$, as required. ∎

4.7. Proof of the main result

We have now gathered all the ingredients we require to complete the proof of our main result.

Proof of Theorem 1.4. By Theorem 2.6, if the surface $\mathcal{S}$ we wish to find a spanning copy of is not homoeomorphic to $\mathbb{S}^2$, then it is homeomorphic either to a connected sum of g tori or to a connected sum of g projective planes for some integer $g \geq 1$; if $\mathcal{S}$ is homeomorphic to $\mathbb{S}^2$, then we set $g = 0$.

Let $\frac{1}{n} \ll \eta \ll \alpha$, $\frac{1}{k} \ll \varepsilon \ll \mu \ll \frac{1}{g}$ (with the convention that if $g = 0$, we replace $\frac{1}{g}$ by 1), and suppose that $\mathcal{H}$ is a 3-graph on n vertices with $\delta_2(\mathcal{H}) \geq \frac{n}{3} + \mu n$. We construct a spanning copy of $\mathcal{S}$ in $\mathcal{H}$ as follows.

Colour the edges of $\mathcal{H}$. By Lemma 4.9, for $\frac{1}{n} \ll \alpha$, $\frac{1}{k} \ll \varepsilon \ll \mu$, we can colour some of the edges of $\mathcal{H}$ red and green to get an (ε, μ)-coloured 3-graph in which any monochromatic disjoint pair of edges is (α, k)-connectible.

Construct reservoirs. Next, we choose two disjoint subsets $R_1, R_2 \subset V(\mathcal{H})$ with

$$|R_1| = |R_2| = \frac{\mu n}{144}$$

and the following additional properties:

(D1) For any pair (e, f) of disjoint green edges of $\mathcal{H}$, and any subset $F \subset R_1$ with $|F| \leq 4\eta k n$, there exists a set $A \subset R_1 \setminus F$ with $|A| \leq k$ for which there is a sphere with vertex set $V(e) \cup V(f) \cup A$ containing e and f.

(D2) For each pair of vertices $x, y \in V(\mathcal{H})$, there are at least $(\frac{1}{3} + \mu 2)|R_2|$ vertices u in R_2 such that $uxy \in E(\mathcal{H})$.

(D3) There are at least $\frac{1}{8}\mu|R_2|$ vertices in R_2 contained in fewer than εn^2 red edges each.

To show that such a choice of R_1 and R_2 is possible, we employ the probabilistic method: we pick two disjoint subsets $R_1', R_2' \subset V(\mathcal{H})$ with $|R_1'| = |R_2'| = \frac{1}{144}\mu n$ uniformly at random, and show that they satisfy each of the above properties with probability at least $\frac{3}{4}$.

First, let us verify that R_1' satisfies **(D1)** with probability at least $\frac{3}{4}$. Let e and f be two arbitrary disjoint green edges. Since both these edges are green, e and f are (α, k)-connectible, so by Lemma 4.4, for some l with $1 \leq l \leq k$, there are $r \geq \frac{\alpha n}{l}$ pairwise disjoint sets $A_1, A_2, \ldots, A_r \subset V(\mathcal{H})$ for which there is a sphere with vertex set $V(e) \cup V(f) \cup A_i$ containing e and f. Each of these sets independently has probability $(\frac{\mu}{144})^l$ of being contained in R_1', so Proposition 2.5 implies that at least $\alpha(\frac{\mu}{144})^l \frac{n}{2l}$ of the A_i are contained in R_1' with probability $1 - o(n^{-6})$. In this case, for any $F \subset R_1'$ with $|F| \leq 4\eta k n$, since each vertex in F is in at most one of the disjoint sets A_i, at least

$$\alpha\left(\frac{\mu}{144}\right)^l \frac{n}{2l} - |F| > 0$$

of the A_i lie entirely within $R_1' \setminus F$. Since there are $O(n^6)$ possible pairs (e, f), it follows that **(D1)** holds with high enough probability.

Next, we prove that R_2' satisfies **(D2)** with probability at least $\frac{3}{4}$. For a given pair of vertices $x, y \in V(\mathcal{H})$, at least $(\frac{1}{3}+\mu)n$ vertices form an edge with x and y in $\mathcal{H}$. Each such vertex independently has probability $\frac{1}{n}|R_2'|$ of being in R_2'. Consequently, Proposition 2.5 implies that at least $(\frac{1}{3}+\mu 2)|R_2'|$ such vertices end up in R_2' with probability $1-o(n^{-2})$; it follows from the union bound that **(D2)** holds with high enough probability.

Finally, we prove that R_2' satisfies **(D3)** with probability at least $\frac{3}{4}$. By virtue of being (ε,μ)-coloured, there are at least $\frac{1}{4}\mu n$ vertices in fewer than εn^2 red edges in $\mathcal{H}$, and each such vertex ends up in R_2' with probability $\frac{1}{n}|R_2'|$, so Proposition 2.5 again implies that at least $\frac{1}{8}\mu|R_2'|$ such vertices are in R_2' with very high probability, proving that **(D3)** also holds with high enough probability.

Going forward, we fix these two disjoint sets R_1 and R_2 with the properties described above and set $R = R_1 \cup R_2$.

Build absorbers. Appealing to Lemma 4.13, we find a collection $\mathfrak{S}_1$ of at most ηn vertex-disjoint green-tinged spheres disjoint from R, which can absorb any subset of R and which have at most $\frac{1}{9}\mu n$ vertices in total; let U be the set of vertices belonging to spheres in $\mathfrak{S}_1$.

Find an almost-spanning collection of spheres. Let $\mathcal{G} = \mathcal{H}[V(\mathcal{H}) \setminus (R \cup U)]$ with colours inherited from $\mathcal{H}$. Note that because $|R \cup U| \le \frac{\mu n}{8}$, we have $|V(\mathcal{G})| \ge (1-\frac{\mu}{8})n$, so

$$\delta_2(\mathcal{G}) \ge \left(\frac{1}{3}+7\frac{\mu}{8}\right)n > \left(\frac{1}{3}+\mu 2\right)|V(\mathcal{G})|.$$

Additionally, the 3-graph $\mathcal{G}$ has at most $\varepsilon n^3 < 2\varepsilon|V(\mathcal{G})|^3$ uncoloured edges and at most $\varepsilon n^4 < 2\varepsilon|V(\mathcal{G})|^4$ pairs of touching red and green edges. Finally, $\mathcal{G}$ has at least $\frac{\mu n}{8}$ vertices in fewer than $\varepsilon n^2 < 2\varepsilon|V(\mathcal{G})|^2$ red edges. Consequently, $\mathcal{G}$ is $(2\varepsilon,\mu 2)$-coloured; in particular it contains at least $\frac{1}{100}\mu n^3$ green edges. If $\mathcal{S}$ is not homeomorphic to $\mathbb{S}^2$, then we employ Theorem 2.1 to obtain a collection $\mathfrak{G}$ of g green copies of $\mathcal{T}_9$ or $\mathcal{P}_{12}$ in $\mathcal{G}$ as appropriate, using at most $12g$ vertices; if $\mathcal{S}$ is homeomorphic to $\mathbb{S}^2$, then we set $\mathfrak{G} = \varnothing$.

Easily, the 3-graph $\mathcal{G}'$ obtained by removing vertices of surfaces in $\mathfrak{G}$ from $\mathcal{G}$ is $(3\varepsilon, \frac{\mu}{3})$-coloured, so by Lemma 4.20, we can find a collection $\mathfrak{S}_2$ of at most ηn vertex-disjoint green spheres in $\mathcal{G}'$ which cover all but at most $\mu^3 n$ of the vertices of $\mathcal{G}'$. Let B be the set of vertices in $V(\mathcal{H}) \setminus (R \cup U)$ which are not in any surface in $\mathfrak{G} \cup \mathfrak{S}_2$, and write W for $V(\mathfrak{G} \cup \mathfrak{S}_2)$.

To summarise, we have now decomposed the vertices of $\mathcal{H}$ into sets $R = R_1 \cup R_2$, $U = V(\mathfrak{S}_1)$, $W = V(\mathfrak{G} \cup \mathfrak{S}_2)$ and B, as shown in Figure 6.

Mop up uncovered vertices using the reservoir R_2. Next, we show that we can mop up the set B of uncovered vertices with green-tinged spheres using vertices in $B \cup R_2$ alone.

Claim 4.27. *$\mathcal{H}[B \cup R_2]$, with colours inherited from $\mathcal{H}$, is $(144^4\frac{\varepsilon}{\mu^4}, \frac{\mu}{3})$-coloured.*

Proof. Indeed, this 3-graph has m vertices where $m = |B \cup R_2| \ge |R_2|$, so

$$\frac{\mu n}{144} \le m < \frac{\mu n}{144} + \frac{\mu^2 n}{144^2}.$$

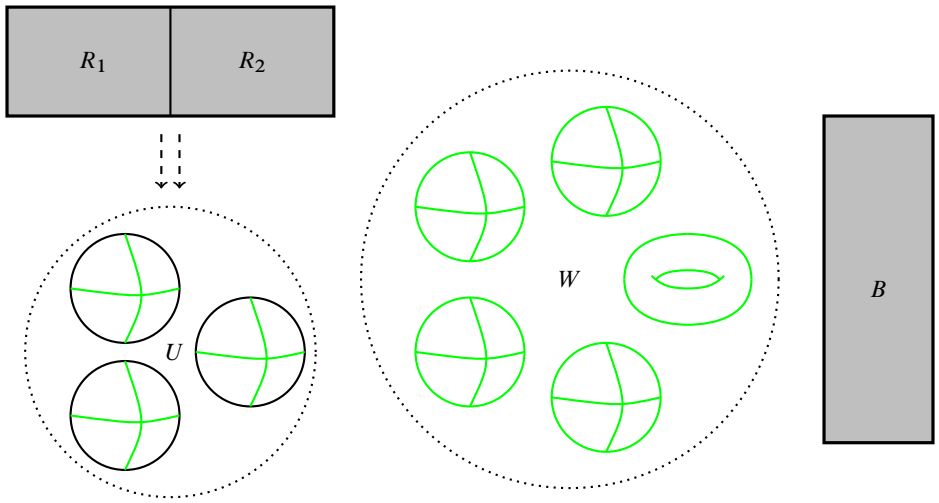


Fig. 6. The situation after finding an almost-spanning collection of spheres; the arrows signify the ability of the spheres in U to absorb any subset of R.

By **(D2)**, each pair of vertices in $B \cup R_2$ has codegree at least $(\frac{1}{3} + \mu 2)\frac{\mu n}{144} > (\frac{1}{3} + \frac{\mu}{3})m$ in $\mathcal{H}[B \cup R_2]$. Also, $\mathcal{H}[B \cup R_2]$ certainly contains at most

$$\varepsilon n^3 \leq \frac{144^3 \varepsilon m^3}{\mu^3} < \frac{144^4 \varepsilon m^3}{\mu^4}$$

uncoloured edges, and at most

$$\varepsilon n^4 \leq \frac{144^4 \varepsilon m^4}{\mu^4}$$

pairs of differently-coloured touching edges. Finally, condition **(D3)** ensures that at least $\frac{1}{8}\mu|R_2| > \frac{\mu m}{12}$ vertices are in fewer than

$$\varepsilon n^2 \leq \frac{144^2 \varepsilon m^2}{\mu^2}$$

red edges. ∎

Consequently, it follows from Corollary 4.14 that there is a collection $\mathfrak{S}_3$ of at most ηn vertex-disjoint green-tinged spheres in $\mathcal{H}[B \cup R_2]$ which cover all the vertices in B.

Connect into a single spanning surface using the reservoir R_1. Order the surfaces in $\mathfrak{S}_1 \cup \mathfrak{S}_2 \cup \mathfrak{S}_3 \cup \mathfrak{G}$ as $S_1, S_2, \ldots, S_l$ where $\mathfrak{S}_1 = \{S_1, S_2 \ldots, S_q\}$; since $\mathfrak{S}_i$ contains at most ηn spheres for each $1 \leq i \leq 3$, and $|\mathfrak{G}| \leq g$, we have $l \leq 4\eta n$.

Since $V(\mathfrak{S}_1) = U$, $V(\mathfrak{G})$ and $V(\mathfrak{S}_2)$ are disjoint and $W = (V(\mathfrak{G}) \cup V(\mathfrak{S}_2)) \subset V(\mathcal{H}) \setminus (R \cup U)$, and $V(\mathfrak{S}_3) \subset B \cup R_2$, where $B = V(\mathcal{H}) \setminus (R \cup U \cup W)$, all the above surfaces are disjoint both from each other and from R_1. By the choice of $\mathfrak{S}_1$, for each

sphere S_i with $1 \le i \le q$, there are two green edges $e_i, f_i \in S_i$ and a set of vertices $Q_i \subset R$ such that for any set $Q' \subset Q_i$, there is a sphere on the vertex set $V(S_i) \cup Q'$ containing e_i and f_i, where the Q_i are pairwise disjoint and $\bigcup_{i=1}^{q} Q_i = R$. Every sphere in $\mathfrak{S}_2 \cup \mathfrak{S}_3$ is green-tinged, as is every surface in $\mathfrak{G}$, so for each i with $q < i \le l$, we again choose two different green edges $e_i, f_i \in S_i$.

For each $1 \le i \le l-1$ in turn, by **(D1)**, we find a sphere $\hat{S}_i$ containing f_i and e_{i+1} on the vertex set $V(f_i) \cup V(e_{i+1}) \cup A_i$, where $|A_i| \le k$ and $A_i \subset R_1 \setminus F_i$, where F_i is a nested sequence of sets given by $F_i = \bigcup_{j<i} A_j$; note that $|F_i| < ik < lk \le 4\eta nk$, so this is indeed possible. The role of these spheres $\hat{S}_i$ will be to connect the S_i together into a single surface. This works as follows: if we remove f_i and e_{i+1} from $\hat{S}_i$ as well as from S_i and S_{i+1} respectively, then the rest of $\hat{S}_i$ yields a cylindrical tube that helps us form the connected sum of S_i and S_{i+1}.

However, in addition to gluing the S_i together, we still need to take care of some uncovered vertices in R. Here, we rely on the properties of our absorbers, which allows us to incorporate them into the S_i without affecting our gluing plans. Indeed, writing $F = \bigcup_{j<l} A_j$, let $R_1' = R_1 \setminus F$ be the set of vertices in R_1 not used in any of the connecting spheres $\hat{S}_i$ for $1 \le i \le l-1$, and let R_2' be the set of vertices in R_2 not used in any sphere in $\mathfrak{S}_3$, i.e., $R_2' = B \cup R_2 \setminus V(\mathfrak{S}_3)$. For each $i \le q$, let $Q_i' = Q_i \cap (R_1' \cup R_2')$, and let S_i' be a sphere on vertex set $V(S_i) \cup Q_i'$ containing e_i and f_i. For each $i > q$, set $S_i' = S_i$.

To finish, we observe that the symmetric difference of the union of the edge-sets of all the S_i' and all the $\hat{S}_i$, or in other words, the edge set

$$\left(\bigcup_{j=1}^{l} S_j'\right) \cup \left(\bigcup_{j=1}^{l-1} \hat{S}_j\right) \setminus \{e_2, e_3, \dots, e_l, f_1, f_2, \dots, f_{l-1}\}$$

yields a surface homeomorphic to $\mathcal{S}$. The vertex set of this copy of $\mathcal{S}$ is

$$V' = V(\mathfrak{S}_1) \cup R' \cup V(\mathfrak{S}_2 \cup \mathfrak{G}) \cup V(\mathfrak{S}_3) \cup F.$$

Finally, since

$$\begin{aligned} V(\mathfrak{S}_1) \cup R' &= U \cup R_1' \cup R_2', \\ V(\mathfrak{S}_2 \cup \mathfrak{G}) &= V(\mathcal{H}) \setminus (R_1 \cup R_2 \cup U \cup B), \\ V(\mathfrak{S}_3) &= B \cup R_2 \setminus R_2', \\ F &= R_1 \setminus R_1', \end{aligned}$$

we see that $V' = V(\mathcal{H})$, so the copy of $\mathcal{S}$ that we have found is spanning, proving the result. ∎

5. Conclusion

In this paper, our main result, Theorem 1.4, asymptotically determines the minimum codegree guaranteeing the existence of a spanning copy of the sphere in a 3-graph. Several

natural questions remain; to begin with, it would be nice to eliminate the error term in our result, and in this direction, we conjecture the following.

Conjecture 5.1. *Every 3-graph $\mathcal{H}$ on $n > 3$ vertices with $\delta_2(\mathcal{H}) > \frac{n}{3}$ contains a spanning copy of the sphere $\mathbb{S}^2$.*

Just as natural is to ask what happens in "higher dimensions", or in other words, what the analogue of our main result for r-graphs ought to be for an arbitrary $r \in \mathbb{N}$. For any integer $r \geq 2$, the codegree of a set of $r-1$ vertices in an r-graph $\mathcal{H}$ is the number of edges of $\mathcal{H}$ containing the set in question, and writing $\delta_{r-1}(\mathcal{H})$ for the minimum codegree of an r-graph $\mathcal{H}$, we conjecture the following.

Conjecture 5.2. *For each $r \geq 2$, any r-graph $\mathcal{H}$ on $n > r$ vertices with $\delta_{r-1}(\mathcal{H}) > \frac{n}{r}$ contains a spanning copy of the $(r-1)$-dimensional sphere $\mathbb{S}^{r-1}$.*

That such a bound would be asymptotically best-possible is seen by the following construction generalising the one presented in Figure 1. Given a positive integer n divisible by r, let $X_1, X_2, \ldots, X_r$, be r disjoint sets of vertices of size $\frac{n}{r}$ each, and consider an r-graph $\mathcal{H}$ on the vertex set $V = X_1 \cup X_2 \cup \cdots \cup X_r$ constructed as follows. For a vertex $v \in V$, write $i(v)$ for the index such that $v \in X_{i(v)}$, and for a set Y of r vertices, let $i(Y) = \sum_{y \in Y} i(y) \pmod r$; we then take the edge set of $\mathcal{H}$ to consist precisely of those sets of Y of r vertices for which $i(Y) = 1$. Note that each "pattern" comprising r choices among the X_i (with repetitions allowed) gives rise to a distinct tight component in $\mathcal{H}$, none of which are spanning because $\sum_{1 \leq i \leq r} i \neq 1 \pmod r$; here, as before, we say that two edges in an r-graph touch if they meet in $r-1$ vertices, and a tight component is, once again, an equivalence class of the transitive closure of this relation. In particular, this r-graph $\mathcal{H}$ does not contain a spanning copy of any closed manifold.

In this paper, we addressed all two-dimensional surfaces simultaneously in Theorem 1.4. In the same spirit, one can also ask about other higher-dimensional manifolds than spheres, and in particular, whether there exist manifolds for which the codegree threshold differs qualitatively from that of the corresponding sphere; however, it is also perhaps worth remembering that unlike in the two-dimensional setting, not all higher-dimensional manifolds are triangulable; see [10, 20] for example.

Given that a minimum codegree of $\frac{n}{3}$ is the threshold at which an n-vertex 3-graph is guaranteed to both have a spanning tight component and a spanning copy of the sphere, it is natural to wonder to what extent the main obstacle to finding a spanning copy of the sphere is the existence of a spanning tight component. In particular, one could ask whether the global codegree condition in the statement of Theorem 1.4 of $\delta_2(\mathcal{H}) \geq (\frac{1}{3} + \mu)n$ for an n-vertex 3-graph $\mathcal{H}$ can be relaxed if we assume that $\mathcal{H}$ has a spanning tight component T and, say, that the codegrees of pairs inside T are large. Perhaps surprisingly, this is not the case! The following 3-graph $\mathcal{H}$ on n vertices has minimum codegree almost $\frac{n}{2}$ in its unique tight component, and no spanning copy of any surface. The vertex set of $\mathcal{H}$ consist of two designated vertices u and v and disjoint sets X and Y of size $\frac{1}{2}(n-2)$ each, and the edge set of $\mathcal{H}$ is obtained by starting with the complete 3-graph on its vertex set and subsequently removing every edge meeting both X and Y. Assuming $\mathcal{H}$ contains a spanning copy of any surface $\mathcal{S}$, remove any edges containing both u and v from $\mathcal{S}$

to split it into two tight components; this easily leads to a contradiction by elementary topological arguments.

The construction above, and its generalisation to r-graphs obtained by replacing the set $\{u, v\}$ by a set of $r-1$ vertices, leads to the following natural conjecture, made particularly attractive by the fact that it predicts a threshold at the codegree density of $\frac{1}{2}$, independent of the dimension.

Conjecture 5.3. *For each $r \geq 2$, if $\mathcal{H}$ is an r-graph on $n > r$ vertices comprising a single tight component such that, for every set Z of $r-1$ vertices of $\mathcal{H}$ that is contained in some edge of $\mathcal{H}$, there are at least $\frac{n}{2}$ edges of $\mathcal{H}$ containing Z, then $\mathcal{H}$ contains a spanning copy of the $(r-1)$-dimensional sphere $\mathbb{S}^{r-1}$.*

Let us close by remarking that many extremal results about cycles in graphs ought to have natural generalisations formulated in terms of triangulations of spheres in hypergraphs. We shall resist the temptation to list further open problems of this kind here, although we would certainly be very interested to see more results of this nature. Perhaps the main message of this paper is that it is possible to do "extremal simplicial topology" with a flavour similar to extremal graph theory, and that some of the major techniques in the latter field, like regularity and absorption, can be brought to bear on the former.

Acknowledgements. Some of the research in this paper was carried out while the third and fourth authors were visiting the University of Warwick, and was continued while the first and second authors were visiting the University of Cambridge; we are grateful for the hospitality of both universities. We would also like to thank Tim Gowers, Richard Mycroft and István Tomon for helpful discussions at various stages of this project.

Funding. The first and second authors were supported by the European Research Council grant 639046 under the European Union's Horizon 2020 research and innovation programme. The second author was also partially supported by UK Research and Innovation Future Leaders Fellowship MR/S016325/1. The third author was supported by the European Research Council grant 947978 under the European Union's Horizon 2020 research and innovation programme, and the Leverhulme Trust. The fourth author was partially supported by NSF grants CCF-1814409 and DMS-1800521.

References

[1] Bermond, J.-C., Germa, A., Heydemann, M.-C., Sotteau, D.: Hypergraphes hamiltoniens. In: Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS 260, CNRS, Paris, 39–43 (1978) Zbl 0413.05041 MR 539937

[2] Conlon, D.: Personal communication

[3] Dirac, G. A.: Some theorems on abstract graphs. Proc. Lond. Math. Soc. (3) **2**, 69–81 (1952) Zbl 0047.17001 MR 47308

[4] Erdős, P.: On extremal problems of graphs and generalized graphs. Israel J. Math. **2**, 183–190 (1964) Zbl 0129.39905 MR 183654

[5] Erdős, P., Simonovits, M.: Supersaturated graphs and hypergraphs. Combinatorica **3**, 181–192 (1983) Zbl 0529.05027 MR 726456

[6] Gowers, W. T.: Question posed at the Cambridge Combinatorics Seminar (2005)

[7] Gowers, W. T.: Personal communication

[8] Hagerup, T., Rüb, C.: A guided tour of Chernoff bounds. Inform. Process. Lett. **33**, 305–308 (1990) Zbl 0702.60021 MR 1045520

[9] Katona, G. Y., Kierstead, H. A.: Hamiltonian chains in hypergraphs. J. Graph Theory **30**, 205–212 (1999) Zbl 0924.05050 MR 1671170

[10] Kirby, R. C., Siebenmann, L. C.: On the triangulation of manifolds and the Hauptvermutung. Bull. Amer. Math. Soc. **75**, 742–749 (1969) Zbl 0189.54701 MR 242166

[11] Kövari, T., Sós, V. T., Turán, P.: On a problem of K. Zarankiewicz. Colloq. Math. **3**, 50–57 (1954) Zbl 0055.00704 MR 65617

[12] Kühn, D.: Personal communication

[13] Kühn, D., Osthus, D.: Spanning triangulations in graphs. J. Graph Theory **49**, 205–233 (2005) Zbl 1066.05057 MR 2145509

[14] Kühn, D., Osthus, D.: Hamilton cycles in graphs and hypergraphs: An extremal perspective. In: Proceedings of the International Congress of Mathematicians – Seoul 2014. Vol. IV, Kyung Moon Sa, Seoul, 381–406 (2014) Zbl 1373.05100 MR 3727617

[15] Kühn, D., Osthus, D., Taraz, A.: Large planar subgraphs in dense graphs. J. Combin. Theory Ser. B **95**, 263–282 (2005) Zbl 1075.05045 MR 2171366

[16] Linial, N., Luria, Z.: An upper bound on the number of high-dimensional permutations. Combinatorica **34**, 471–486 (2014) Zbl 1340.05007 MR 3259813

[17] Linial, N., Morgenstern, A.: On high-dimensional acyclic tournaments. Discrete Comput. Geom. **50**, 1085–1100 (2013) Zbl 1280.05052 MR 3138147

[18] Linial, N., Peled, Y.: On the phase transition in random simplicial complexes. Ann. of Math. (2) **184**, 745–773 (2016) Zbl 1348.05193 MR 3549622

[19] Luria, Z., Tessler, R. J.: A sharp threshold for spanning 2-spheres in random 2-complexes. Proc. Lond. Math. Soc. (3) **119**, 733–780 (2019) Zbl 1423.05094 MR 3960667

[20] Manolescu, C.: Pin(2)-equivariant Seiberg–Witten Floer homology and the triangulation conjecture. J. Amer. Math. Soc. **29**, 147–176 (2016) Zbl 1343.57015 MR 3402697

[21] Rödl, V., Ruciński, A., Szemerédi, E.: A Dirac-type theorem for 3-uniform hypergraphs. Combin. Probab. Comput. **15**, 229–251 (2006) Zbl 1082.05057 MR 2195584

[22] Rödl, V., Ruciński, A., Szemerédi, E.: Dirac-type conditions for Hamiltonian paths and cycles in 3-uniform hypergraphs. Adv. Math. **227**, 1225–1299 (2011) Zbl 1220.05083 MR 2799606

[23] Seifert, H., Threlfall, W.: Seifert and Threlfall: A textbook of topology. Pure Appl. Math. 89, Academic Press, New York (1980) Zbl 0469.55001 MR 575168

[24] Sós, V. T., Erdős, P., Brown, W. G.: On the existence of triangulated spheres in 3-graphs, and related problems. Period. Math. Hungar. **3**, 221–228 (1973) Zbl 0269.05111 MR 323647

[25] Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. Acta Arith. **27**, 199–245 (1975) Zbl 0303.10056 MR 369312

[26] Thomassen, C.: The Jordan–Schönflies theorem and the classification of surfaces. Amer. Math. Monthly **99**, 116–130 (1992) Zbl 0773.57001 MR 1144352