Fractional clique decompositions of dense graphs

Richard Montgomery[*]

Abstract

For each $r \geq 4$, we show that any graph G with minimum degree at least $(1 - 1/(100r))|G|$ has a fractional K_r-decomposition. This improves the best previous bounds on the minimum degree required to guarantee a fractional K_r-decomposition given by Dukes (for small r) and Barber, Kühn, Lo, Montgomery and Osthus (for large r), giving the first bound that is tight up to the constant multiple of r (seen, for example, by considering Turán graphs).

In combination with work by Glock, Kühn, Lo, Montgomery and Osthus, this shows that, for any graph F with chromatic number $\chi(F) \geq 4$, and any $\varepsilon > 0$, any sufficiently large graph G with minimum degree at least $(1 - 1/(100\chi(F)) + \varepsilon)|G|$ has, subject to some further simple necessary divisibility conditions, an (exact) F-decomposition.

1 Introduction

Given a graph F, a graph G has an F-decomposition if we can cover the edges of G exactly with edge-disjoint copies of F. The study of such decompositions dates back at least as far as 1847, when Kirkman [14] showed that any complete graph K_n on n vertices has a triangle (K_3-)decomposition if and only if $n \equiv 1$ or $3 \mod 6$. As the edges in a triangle decomposition of K_n are partitioned into triangles, we must have $3|e(K_n)$. Furthermore, the $n-1$ neighbours of each vertex v are divided into pairs forming triangles with v, so we must have $2|n-1$. In combination, this gives rise to the necessary and sufficient condition on n for the existence of a triangle decomposition of K_n.

It was not until the 1970's that Kirkman's theorem was generalised by Wilson [16, 17, 18, 19] to consider F-decompositions of cliques for more general graphs F. If a graph G has an F-decomposition, then it follows immediately from the partition of $E(G)$ into copies of F that $e(F)$ divides $e(G)$. Furthermore, by considering the copies of F used in the decomposition around each vertex, we can see that the highest common factor of the degrees of the vertices in F must divide the degree of each vertex in G. If G satisfies both of these necessary conditions, then we say G is F-divisible. For each graph F, Wilson [16, 17, 18, 19] showed that any sufficiently large F-divisible complete graph has an F-decomposition. In 2014, the long-conjectured generalisation of Wilson's theorem to hypergraph cliques F was proved in a breakthrough result of Keevash [13]. Very recently, an independent, combinatorial, proof of this generalisation was given by Glock, Kühn, Lo and Osthus [10], who then built on their methods to give a full generalisation of Wilson's theorem to arbitrary hypergraphs F [9].

In general, we do not expect a simple characterisation of the F-divisible graphs which have an F-decomposition. For example, Dor and Rasi [3] have shown that determining whether a graph has an F-decomposition is NP-complete if F has a connected component with at least 3 edges. It is therefore natural to ask instead if all sufficiently dense F-divisible graphs have an F-decomposition. Much of the work in this area has been motivated by the following beautiful conjecture of Nash-Williams on triangle decompositions.

Conjecture 1.1 (Nash-Williams [15])**.** *There exists $N \in \mathbb{N}$ such that, if G is a K_3-divisible graph on at least N vertices with $\delta(G) \geq 3|G|/4$, then G has a K_3-decomposition.*

The first progress towards Conjecture 1.1 was given by Gustavsson [11], who showed, for each graph F, that there is some constant $\varepsilon(F) > 0$ such that every sufficiently large F-divisible graph G satisfying $\delta(G) \geq$

---

[*]School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK. r.h.montgomery@bham.ac.uk

$(1-\varepsilon(F))|G|$ has an F-decomposition. The bound on $\varepsilon(F)$ claimed by Gustavsson left much room for improvement, showing only that $\varepsilon(F) \leq 10^{-37}|F|^{-94}$. In recent years, a rich seam of progress was initiated by Barber, Kühn, Lo and Osthus [2] by relating F-decompositions in dense graphs to *fractional F-decompositions*.

Given a graph F, we say that a graph G has a *fractional F-decomposition* if we may (non-negatively) weight the copies of F in G so that each edge in $E(G)$ is in copies of F with total weight 1. That is, letting $\mathcal{F}(G)$ be the set of all copies of F in G, there is some function $\omega : \mathcal{F}(G) \to [0,1]$ so that, for each $e \in E(G)$, we have

$$\sum_{F' \in \mathcal{F}(G):e \in E(F')} \omega(F') = 1.$$

In 2014, Barber, Kühn, Lo and Osthus [2] introduced an innovative iterative absorption method capable of turning an *approximate* F-decomposition in an F-divisible graph G with high minimum degree into an F-decomposition, where an approximate F-decomposition is a disjoint set of copies of F in G covering most of the edges of G. Haxell and Rödl [12] had demonstrated that a large graph with a fractional F-decomposition must have an approximate F-decomposition. This allowed Barber, Kühn, Lo and Osthus [2] to give much improved bounds on $\varepsilon(F)$ by using bounds on the minimum degree required to guarantee an appropriate fractional F-decomposition.

The methods in [2] were subsequently developed and extended by Glock, Kühn, Lo, Osthus and the current author [8]. In order to describe this accumulated progress, for each graph F and integer n, let $\delta_F^*(n)$ be the least $\delta > 0$ such that any graph G with n vertices and $\delta(G) > \delta n$ has a fractional F-decomposition. For each graph F, let $\delta_F^* = \limsup_{n \to \infty} \delta_F^*(n)$. The main results in [8] imply the following.

Theorem 1.2 (Glock, Kühn, Lo, Montgomery, Osthus [8]). *Let F be a graph, let $\varepsilon > 0$, and let $\chi = \chi(F)$. Any sufficiently large F-divisible graph G with $\delta(G) \geq (\max\{\delta_{K_\chi}^*, 1 - 1/(\chi + 1)\} + \varepsilon)|G|$ has an F-decomposition.*

Aside from its own intrinsic interest, determining the value of $\delta_{K_r}^*$ for each r therefore has a significant part to play in the study of decompositions of dense graphs. In particular, to prove Conjecture 1.1 asymptotically it is sufficient to show that $\delta_{K_3}^* \leq 3/4$ (as already followed from the results in [2]).

In the particular case of triangles, increasingly good bounds on $\delta_{K_3}^*$ were given by Yuster [20], Dukes [5, 6] and Garaschuk [7], before Dross [4] gave an elegantly efficient proof that $\delta_{K_3}^* \leq 0.9$. For each $r \geq 4$, Yuster [20] showed that $\delta_{K_r}^* \leq 1 - 1/(9r^{10})$ and gave a construction showing that $\delta_{K_r}^* \geq (1 - 1/(r + 1))n$. Dukes [5, 6] used tools from linear algebra to show that $\delta_{K_r}^* \leq 1 - 2/(9r^2(r-1)^2)$, before Barber, Kühn, Lo, Osthus and the current author [1] were able to generalise and extend Dross's methods for fractional triangle decompositions to show that $\delta_{K_r}^* \leq 1 - 1/(10^4 r^{3/2})$.

In this paper, we show that, for each $r \geq 4$, $\delta_{K_r}^* \leq 1 - 1/(100r)$. This improves the known upper bound for $\delta_{K_r}^*$ for each $r \geq 4$, and confirms, up to the constant 100, the correct dependence of $\delta_{K_r}^*$ on r.

Theorem 1.3. *Let $r \geq 4$. If a graph G has minimum degree at least $(1 - 1/(100r))|G|$, then G has a fractional K_r-decomposition.*

In combination with Theorem 1.2, this implies the following.

Corollary 1.4. *Let F be a graph with $\chi(F) \geq 4$ and let $\varepsilon > 0$. Any sufficiently large F-divisible graph G with $\delta(G) \geq (1 - 1/(100\chi(F)) + \varepsilon)|G|$ has an F-decomposition.*

Consideration of Turán graphs with $\chi(F) - 1$ classes (for example) confirms Corollary 1.4 is tight up to the constant that appears before $\chi(F)$. We have not sought further small improvements in the corresponding constant in Theorem 1.3, where they would complicate the proof for little gain. Substantial new ideas appear needed to approach the conjectured value of $\delta_{K_r}^*$.

At the very highest level, our methods to prove Theorem 1.3 take the same form as those used in [1]. We find an initial weighting of some cliques in G which is *close* to a fractional K_r-decomposition, before making local adjustments to correct this to a fractional K_r-decomposition. Within this framework, however, our methods are entirely different. We introduce a simple and efficient way to make the adjustments to the initial weighting, and use a much improved initial weighting of the cliques in G (in comparison to the simple, (essentially) uniform, initial weighting used in [1]).

2

In particular, we find the following simple observation useful: A graph G has a fractional K_r-decomposition if we can randomly pick an r-clique in G (with some well-chosen probability distribution) so that the probability an edge e is in this r-clique is constant across all the edges $e \in E(G)$. This observation, seen by using a weighting on the r-cliques proportional to the probability distribution, is trivial. However, the perspective it brings allows us to use a random process to select a random r-clique of G, before translating this to an initial weighting which would be more difficult to conceive or describe directly.

The rest of this paper is organised as follows. After describing our notation, we give a more detailed sketch of our methods in Section 2. In Section 3, we find an initial weighting of the cliques in our graph, before, in Section 4, correcting this to give a fractional K_r-decomposition. Aside from the elementary, but slightly involved, calculation required to prove a key lemma (Lemma 2.1), which we defer to Section 6, Sections 3 and 4 give the simplest exposition of our methods for finding a fractional K_r-decomposition using a minimum degree bound with, up to a constant, the correct dependence on r (proving Lemma 4.3). In Section 5, we then make some improvements to a central lemma in Section 3 (Lemma 3.6) to reduce the bounds used by our methods, and hence prove Theorem 1.3. In Section 6, we give the calculation required to prove Lemma 2.1.

1.1 Notation

A graph G has vertex set $V(G)$, edge set $E(G)$, and minimum degree $\delta(G)$. For each $r \geq 2$, we denote by $\mathcal{K}_r(G)$ the set of copies of K_r, the clique with r vertices, in G. Where the graph used is clear from context we will use $\mathcal{K}_r = \mathcal{K}_r(G)$. For a set $E \subseteq V(G)^{(2)}$, we denote by $G + E$ and $G - E$ the graphs with vertex set $V(G)$ and edge sets $E(G) \cup E$ and $E(G) \setminus E$ respectively. In particular, we often use this when E is a matching, i.e. a set of independent edges. For each $e \in V(G)^{(2)}$ we let $G - e = G - \{e\}$ and $G + e = G + \{e\}$.

For a graph G, we denote by $\bar{G}$ the *complement* of G – the graph with vertex set $V(G)$ and edge set $V(G)^{(2)} \setminus E(G)$. For each $v \in V(G)$, $N(v)$ is the set of neighbours of v, and $N^c(v) = V(G) \setminus N(v)$, the set of non-neighbours of v. When we have a weighting w_K, $K \in \mathcal{K}_r(G)$, of the r-cliques in a graph G, we say for each edge $e \in E(G)$ that the *weight over e* is $\sum_{K \in \mathcal{K}_r(G):e \in E(K)} w_K$. Given any event A, we let

$$\mathbf{1}_A = \begin{cases} 1 & \text{if } A \text{ occurs,} \\ 0 & \text{otherwise.} \end{cases}$$

2 Proof Sketch

In order to aid our sketch, let us recap very briefly the methods used by Barber, Kühn, Lo, Osthus and the current author in [1], where methods originated by Dross [4] were extended and generalised. In order to find a fractional K_r-decomposition of a graph G with high minimum degree in [1], an initial (essentially uniform) weighting was given to the subgraphs of G isomorphic to K_r, before a series of small local changes to this weighting were made (using structures called 'gadgets') to correct this weighting to a fractional K_r-decomposition. In contrast, here we make our initial weighting using a random process which is capable of getting far closer to a fractional decomposition than a simple uniform weighting of cliques (see Section 2.1). We use this, in fact, to get close to a fractional K_{2r+2}-decomposition of the graph G. We then make our corrections 'within the $(2r + 2)$-cliques' to convert this to a fractional K_r-decomposition of G (see Section 2.2).

For each part of the proof, it is crucial that we can show that any complete graph with at least $2r + 2$ vertices has a fractional K_r-decomposition, even if we remove an arbitrary matching. Such a clique with removed edges is sufficiently symmetric that we may give a fractional K_r-decomposition directly, but as the calculation is slightly involved we defer it to Section 6. We state the required result here, however, for reference.

Lemma 2.1. *Let $r \geq 3$ and $k \geq 2r + 2$. If $M \subseteq E(K_k)$ is a matching, then $K_k - M$ has a fractional K_r-decomposition.*

We will give our first initial weighting in Section 3, before returning to it in Section 5 to make improvements. We do this in order to give, in Sections 3 and 4, as clean as possible an exposition of our methods to demonstrate there is some $\varepsilon > 0$ such that, for all $r \geq 3$, any graph G with minimum degree $(1 - \varepsilon/r)|G|$ has a fractional K_r-decomposition (see Lemma 4.3).

In order to find our initial weighting we gain from (variations of) the following simple observation. If we can pick a random r-clique H from a graph G so that each edge in G is equally likely to appear in $E(H)$, then G has a fractional K_r-decomposition. Indeed, the weights $w_K = e(G) \cdot \mathbb{P}(H = K)/\binom{r}{2}$, $K \in \mathcal{K}_r$, will form a fractional K_r-decomposition of G.

We will use a random process to weight subgraphs $H \subseteq G$ with $\delta(H) \geq |H| - 6$ so that each edge is given total weight close to 1 (that is, for each edge the sum of the weight of the subgraphs containing that edge is close to 1). If each such subgraph H has at least $32r + 62$ vertices, then we can find a fractional K_{2r+2}-decomposition of H by partitioning the non-edges of H into 5 matchings (using a little extra structure in our particular graphs H) and repeatedly applying Lemma 2.1. We then improve our methods in Section 5, using more structure that naturally arises in the subgraphs H to reduce their size to only $18r + 18$ vertices. The method for initially weighting these subgraphs H remains the same and we will now give a representative sketch of this process, picking a random subgraph with $50r$ vertices.

2.1 Our initial weighting

Let G be a graph with $50rm$ vertices and $\delta(G) \geq (1 - 1/(100r))|G| \geq |G| - m/2$. We will sketch how to pick a random subgraph H of G with $50r$ vertices so that $\delta(H) \geq |H| - 6$ and every edge is approximately equally likely to appear in H (c.f. Lemma 3.2). We do so by describing a random process which chooses a subgraph H with $|H| = 50r$ and $\delta(H) \geq |H| - 3$ in which every vertex is equally likely to appear in H, before describing how to alter this process so that each edge is approximately equally likely to appear in H (at the expense of weakening the degree condition to $\delta(H) \geq |H| - 6$).

We choose the vertices $\{a_1, \ldots, a_{50r}\}$ of H randomly in $50r$ stages. At stage i, we choose a subset A_i of m vertices to *consider*, from vertices in G that have not yet been considered, and then pick a vertex $a_i \in A_i$ uniformly at random. By including in A_i all the non-neighbours of a_{i-1} that have not yet been considered (of which there are at most $m/2$), we ensure that every non-neighbour of a_{i-1} is considered at stage i or earlier. Thus, each vertex a_i has no non-neighbours in $V(H)$ except for itself and possibly a_{i-1} or a_{i+1}, and therefore the final graph H satisfies $\delta(H) \geq |H| - 3$. Precisely, we carry out the following process for each i, $1 \leq i \leq 50r$.

- Pick $A_i \subseteq V(G) \setminus (\bigcup_{j<i} A_j)$ uniformly at random subject to $|A_i| = m$ and $N^c(a_{i-1}) \setminus (\bigcup_{j<i} A_j) \subseteq A_i$ (possible as $|N^c(a_{i-1})| \leq m/2$).

- Pick $a_i \in A_i$ uniformly at random.

Each vertex is considered exactly once in this process, and added to $V(H)$ with probability $1/m$. If the two vertices in an edge e are considered in different stages, then the probability that e appears in H is $1/m^2$. If an edge appears within some set A_i, then it cannot appear in H. Some edges in G may be much more likely than others to appear within some set A_i, and thus be less likely to appear in H. The resulting variation in how likely each edge is to appear in H can (for some graphs G) be too much to later correct using our methods, and therefore we need alter this random process.

Consider then taking twice as many vertices in A_i (that is, $2m$ vertices) at each stage (for only $25r$ stages in total), and picking *two* vertices a_i and b_i from A_i to add to $V(H)$. By including any unconsidered non-neighbours of a_{i-1} or b_{i-1} in A_i, we can ensure H is still almost complete, in fact satisfying $\delta(H) \geq |H| - 6$. An edge that appears in some set A_i may now possibly appear in H, an improvement on the first process. However, this probability is too small (close to $1/2m^2$) compared to edges whose vertices are considered at different stages (still $1/m^2$).

Fortunately, an edge is only significantly more likely than others to appear in A_i if it appears in $(N^c(a_{i-1}) \cup N^c(b_{i-1})) \setminus (\bigcup_{j<i} A_j) \subseteq A_i$ (as the other vertices in A_i are chosen randomly from the remaining unconsidered vertices). As this set is a subset of A_i with size at most $m = |A_i|/2$, we will be able to choose a_i and b_i from A_i so that any edges in $(N^c(a_{i-1}) \cup N^c(b_{i-1})) \setminus (\bigcup_{j<i} A_j)$ appear in H with probability $1/m^2$, but so that each vertex in A_i still appears in $V(H)$ with probability $1/m$ (otherwise we will alter the probability an edge appears in H if its vertices are considered at different stages). Thus, we will be able to choose our subgraph H so that each edge is in H with roughly the same probability.

The precise process we use is given in the proof of Lemma 3.2, and is depicted in Figure 1.

2.2 Correcting the weighting

Suppose we have a graph G along with a weighting w_K, $K \in \mathcal{K}_{2r+2}$, of the $(2r+2)$-cliques in G so that each edge in G has total weight at least 1 and at most $1 + 1/(2r)$. (Such will be the result of our initial weighting.) We will convert this into a fractional decomposition of G into subgraphs with $2r + 2$ vertices which are complete except for some independent non-edges – say the set of such subgraphs in G is $\mathcal{M}$. As each of the subgraphs in $\mathcal{M}$ has a fractional K_r-decomposition by Lemma 2.1, we can combine fractional K_r-decompositions of the subgraphs in $\mathcal{M}$, and the fractional decomposition of G into subgraphs in $\mathcal{M}$, to get a fractional K_r-decomposition of G.

Suppose then an edge $e \in E(G)$ is given weight $1 + \lambda$ by the weighting of $(2r + 2)$-cliques, with $0 \leq \lambda \leq 1/2r$. Consider what would happen if, for each clique $K \in \mathcal{K}_{2r+2}$ containing e, we replaced the weight w_K on K by $w_K/(1 + \lambda)$ and added weight $(1 - 1/(1 + \lambda))w_K$ to $K - e$. We get a weighting of the graphs in $\mathcal{M}$ such that the weight on e is 1, and the weight on the other edges is unchanged, while only weight $(1 - 1/(1 + \lambda))w_K \leq w_K/(2r+1)$ has been moved from any clique K containing e.

For each $K \in \mathcal{K}_{2r+2}$, we could similarly adjust the weight on any independent set of edges M in $E(K)$ simultaneously by moving weight from K to $K - M$. Furthermore, by moving weight from K to $K - E$ for different subsets $E \subseteq M$, we can make different corrections to the weight on different edges in M. As each such clique $K \in \mathcal{K}_{2r+2}$ can be covered by $2r + 1$ sets of independent edges, we will see that we can make controlled adjustments of up to $w_K/(2r+1)$ to the weight on the edges in K without decreasing the weight on K to become negative. Making appropriate such adjustments for each clique in $K \in \mathcal{K}_{2r+2}$ will allow us to decrease the weight on each edge until it is exactly 1 by moving weight off the cliques $K \in \mathcal{K}_{2r+2}$ and onto other subgraphs in $\mathcal{M}$. This is carried out in Section 4, to prove Lemma 4.2, which in combination with the initial weighting lemma (Lemma 3.6, and its improvement Lemma 5.4) proves Theorem 1.3.

3 Choosing an almost-complete random subgraph

In this section we will show that, given some minimum degree condition in a graph G, we can find a weighting w_K, $K \in \mathcal{K}_r$, so that, for each $e \in E(G)$, $1 - 1/r \leq \sum_{K \in \mathcal{K}_r : e \in E(K)} w_K \leq 1$. As sketched in Section 2, we will use a random process to pick a subgraph H in G with $\delta(H) \geq |H| - 6$ so that every edge in G is roughly equally likely to appear within H (proving Lemma 3.2). Each subgraph H will be large enough that we can show H has a fractional K_r-decomposition, and we can combine this with the weighting from the probability distribution on such graphs H to get the required approximate fractional K_r-decomposition of G.

In fact, the random graph H will have the stronger property that it contains a spanning subgraph H' which is isomorphic to a graph of the following form.

Definition 3.1. *For each integer $r \geq 1$, let M_r be the graph with vertex set $\{a_1, b_1, \ldots, a_r, b_r\}$ which has every edge present between sets $\{a_i, b_i\}$ and $\{a_j, b_j\}$ if $i \notin \{j - 1, j, j + 1\}$, and no other edges.*

Note that, for each $r \geq 3$, $\delta(M_r) = 2r - 6$. Using Lemma 2.1 repeatedly, we can show that any graph H with $32r + 62$ vertices which contains a copy of M_{16r+31} has a fractional K_r-decomposition (see Lemma 3.6). Later, in Section 5, we will carry out some more work to allow us to use graphs H with only $18r + 18$ vertices (see Lemma 5.4).

3.1 Our initial probability distribution

We are now ready to give our random process. Note that the following lemma will be applied eventually with r replaced by a larger function of r. We will also use a divisibility condition on the number of vertices in G in this lemma; this, as we shall see, we can assume by duplicating vertices in our initial graph.

Lemma 3.2. *Let $r \geq 3$ and let G be a graph with $n = 2rm$ vertices and $\delta(G) \geq n - m/2$. Let $\mathcal{M}$ be the set of induced subgraphs of G with $2r$ vertices which contain a copy of M_r. Then, with an appropriate probability distribution, we may randomly select a graph $M \in \mathcal{M}$ so that, for each $e \in E(G)$,*

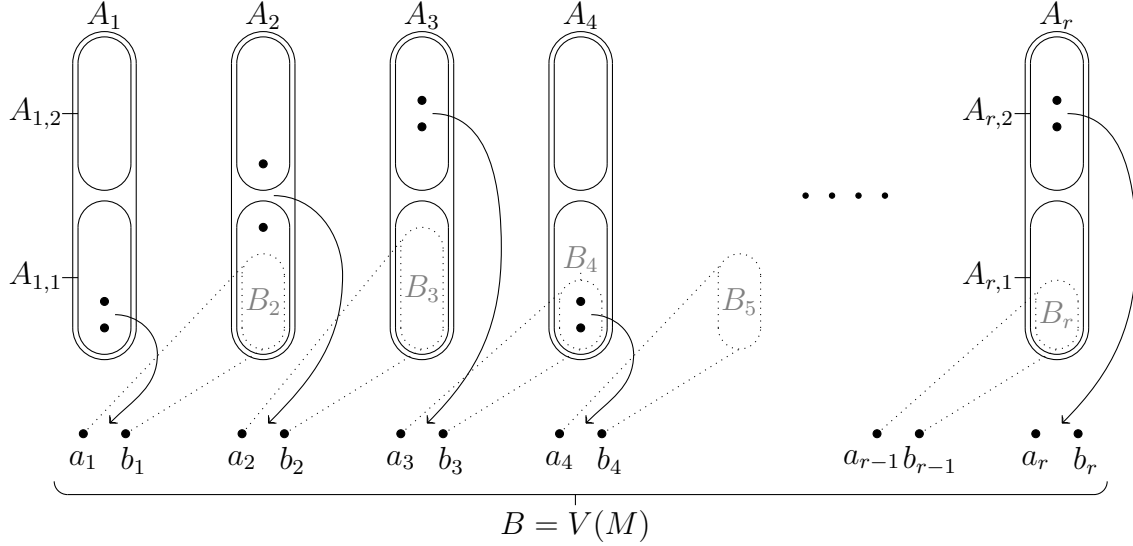$$1 - \frac{4}{r} \leq m^2 \cdot \mathbb{P}(e \in E(M)) \leq 1. \tag{3.1}$$

Figure 1: The random process used to select the set $B = \{a_1, b_1, \ldots, a_r, b_r\}$ which forms the vertex set of the random subgraph M. At each stage i, $1 \le i \le r$, a set of vertices $A_i = A_{i,1} \cup A_{i,2}$ is considered, from which we select two vertices a_i and b_i. For $i \ge 2$, $A_{i,1}$ is chosen to contain any non-neighbours of a_{i-1} and b_{i-1} which have not yet been considered (the vertices in the set B_i).

Proof. Let $B = \{a_1, b_1, \ldots, a_r, b_r\}$ be a random subset of $V(G)$ picked according to the following method, which is depicted in Figure 1: Let $B_1 = \emptyset$ and carry out the following steps for $1 \le i \le r$.

S1 If $i > 1$, let $B_i = (N^c(a_{i-1}) \cup N^c(b_{i-1})) \setminus (\bigcup_{j<i} A_j)$, noting that $|B_i| \le m$.

S2 Let $A_{i,1}$ be the set B_i combined with an $(m - |B_i|)$-sized subset of $V(G) \setminus ((\bigcup_{j<i} A_j) \cup B_i)$ selected independently and uniformly at random.

S3 Let $A_{i,2}$ be an m-sized subset of $V(G) \setminus ((\bigcup_{j<i} A_j) \cup A_{i,1})$ selected independently and uniformly at random, and let $A_i = A_{i,1} \cup A_{i,2}$.

S4 Let $\{a_i, b_i\}$ be a pair of distinct vertices from A_i chosen independently at random so that

$$\mathbb{P}(\{a_i, b_i\} = \{a, b\}) = \begin{cases} 1/m^2 & \text{if } \{a,b\} \subseteq A_{i,1} \text{ or } \{a,b\} \subseteq A_{i,2} \\ 1/m^3 & \text{otherwise.} \end{cases} \tag{3.2}$$

Let $M = G[B]$. We will show that M is a random subgraph in $\mathcal{M}$ which satisfies (3.1) for each $e \in E(G)$.

In the process above, we found a random partition $(A_i)_{i=1}^r$, and selected 2 vertices from each set A_i to form B, so that $|B| = 2r$. Note further that, for each $i < r$, $N^c(a_i) \cup N^c(b_i) \subseteq B_{i+1} \cup (\bigcup_{j \le i} A_j)$. Therefore, for each $i < r$ and $j > i + 1$, $a_j, b_j \in A_j$ are neighbours of a_i and b_i. Thus, $M \in \mathcal{M}$.

Therefore, it is only left to show that (3.1) holds for each $e \in E(G)$. We will show this by conditioning on how the vertices of e appear in the random partition.

Fix an edge $e = uv \in E(G)$, and note that if u appears in A_i, but v has not appeared in A_j, $j \le i$, then the probability that $e \in E(M)$ is $1/m^2$. Indeed, in this case u and v were selected to be in B independently at random, with (by symmetry) probability $1/m$, so that

$$\mathbb{P}\big(uv \in E(M) | \exists i \text{ s.t. } u \in A_i \text{ and } v \notin \bigcup_{j \le i} A_j\big) = 1/m^2. \tag{3.3}$$

This also holds with u and v switched, so we only need consider the case when u and v first appear in the same set A_i. In this case, the edge uv appears in M with the distribution in (3.2), so that

$$\mathbb{P}(uv \in E(M)|\exists i \text{ and } j \text{ s.t. } \{u, v\} \subseteq A_{i,j}) = 1/m^2, \tag{3.4}$$

and

$$\mathbb{P}(uv \in E(M)|\exists i \text{ s.t. } u \in A_{i,1} \text{ and } v \in A_{i,2}) = 1/m^3. \tag{3.5}$$

By (3.3), (3.4) and (3.5), and the symmetric cases with u switched with v, we therefore have

$$\mathbb{P}(e \in E(M)|\nexists i \text{ s.t. } V(e) \cap A_{i,1}, V(e) \cap A_{i,2} \neq \emptyset) = 1/m^2$$

and

$$\mathbb{P}(e \in E(M)|\exists i \text{ s.t. } V(e) \cap A_{i,1}, V(e) \cap A_{i,2} \neq \emptyset) = 1/m^3.$$

Thus,

$$\mathbb{P}(e \in E(M)) = \mathbb{P}(\nexists i \text{ s.t. } V(e) \cap A_{i,1}, V(e) \cap A_{i,2} \neq \emptyset) \cdot \frac{1}{m^2} + \mathbb{P}(\exists i \text{ s.t. } V(e) \cap A_{i,1}, V(e) \cap A_{i,2} \neq \emptyset) \cdot \frac{1}{m^3}$$

$$= \frac{1}{m^2} - \left(\frac{1}{m^2} - \frac{1}{m^3}\right) \cdot \mathbb{P}(\exists i \text{ s.t. } V(e) \cap A_{i,1}, V(e) \cap A_{i,2} \neq \emptyset),$$

so that

$$1 \geq m^2 \cdot \mathbb{P}(e \in E(M)) \geq 1 - \mathbb{P}(\exists i \text{ s.t. } V(e) \cap A_{i,1}, V(e) \cap A_{i,2} \neq \emptyset). \tag{3.6}$$

We will show the following claim.

Claim 3.3. *For each $i \in [r]$, $\mathbb{P}(v \in A_{i,2}|u \in A_{i,1}) \leq 2/r$.*

Claim 3.3 implies that (3.1) holds for the edge uv, as required. Indeed, we have

$$\mathbb{P}(\exists i \text{ s.t. } v \in A_{i,2} \text{ and } u \in A_{i,1}) = \sum_{i=1}^{r} \mathbb{P}(v \in A_{i,2}|u \in A_{i,1}) \cdot \mathbb{P}(u \in A_{i,1})$$

$$\leq (2/r) \cdot \sum_{i=1}^{r} \mathbb{P}(u \in A_{i,1}) \leq 2/r.$$

In combination with the symmetric case with u switched with v, and (3.6), this implies (3.1). Therefore, it is left only to prove Claim 3.3.

Proof of Claim 3.3. Firstly, if $i \leq 3r/4$, we have

$$\mathbb{P}(v \in A_{i,2}|u \in A_{i,1}) \leq \mathbb{P}\left(v \in A_{i,2}|u \in A_{i,1} \wedge v \notin \left(\bigcup_{j<i} A_j\right) \cup A_{i,1}\right) = \frac{m}{2(r-i)m+m} \leq \frac{2}{r}.$$

We therefore can assume that $i > 3r/4$. It is plausible that $\mathbb{P}(v \in A_{i,2}|u \in A_{i,1}) \leq 2/r$, as, if $u \in A_{i,1}$, then in the above process it seems far more likely that v appears in $\bigcup_{j<i}(A_j \setminus B_j)$ than in $A_{i,2}$, as the former set is at least $(i-1)$ times as large as the latter. Conditioning on $u \in A_{i,1}$, however, might reduce the probability v appears in the former set, so we will prove carefully that $\mathbb{P}(v \in A_{i,2}|u \in A_{i,1}) \leq 2/r$.

We need to consider the random process up to and including the choice of $A_{i,2}$ (when we will certainly know whether $v \in A_{i,2}$ or not). In particular, we consider the space of all the possible choices that can be made during this part of the process. That is, all possible choices of

$$X := (A_{1,1}, A_{1,2}, a_1, b_1, \ldots, A_{i-1,1}, A_{i-1,2}, a_{i-1}, b_{i-1}, A_{i,1}, A_{i,2}).$$

Note that

$$Y = (\bar{A}_{1,1}, \bar{A}_{1,2}, \bar{a}_1, \bar{b}_1, \ldots, \bar{A}_{i-1,1}, \bar{A}_{i-1,2}, \bar{a}_{i-1}, \bar{b}_{i-1}, \bar{A}_{i,1}, \bar{A}_{i,2}) \tag{3.7}$$

is a possible value for X if

- $\bar{A}_{1,1}, \bar{A}_{1,2}, \ldots, \bar{A}_{i,1}, \bar{A}_{i,2}$ are disjoint sets with size m in $V(G)$,

- for each $j \in [i-1]$, $\bar{a}_j, \bar{b}_j \in \bar{A}_{j,1} \cup \bar{A}_{j,2}$, and

- for each $j \in [i-1]$, $N^c(\bar{a}_j) \cup N^c(\bar{b}_j) \subseteq (\bigcup_{j' < j}(\bar{A}_{j',1} \cup \bar{A}_{j',2})) \cup \bar{A}_{j,1}$.

To calculate $\mathbb{P}(v \in A_{i,2} | u \in A_{i,1})$ we need to consider all possible such Y for which, in addition, $u \in \bar{A}_{i,1}$ – let $\mathcal{Z}$ be the set of such sequences. Furthermore, let $\mathcal{Z}'$ be the set of such sequences in $\mathcal{Z}$ for which $v \in \bar{A}_{i,2}$. We therefore have

$$\mathbb{P}(v \in A_{i,2} | u \in A_{i,1}) = \frac{\mathbb{P}(X \in \mathcal{Z}')}{\mathbb{P}(X \in \mathcal{Z})}.$$

Calculating this fraction is complicated by the fact that $\mathbb{P}(X = Y)$ is likely to differ across $Y \in \mathcal{Z}$. Thus, let us partition $\mathcal{Z}$ into classes $\mathcal{Z}_1, \ldots, \mathcal{Z}_\ell$, for the smallest possible ℓ, so that there are $p_1, \ldots p_k > 0$ such that, for each $k \in [\ell]$ and $Y \in \mathcal{Z}_k$, $\mathbb{P}(X = Y) = p_k$. For each $k \in [\ell]$, let $\mathcal{Z}'_k = \mathcal{Z}_k \cap \mathcal{Z}'$. We will show the following claim.

Claim 3.4. *For each $k \in [\ell]$, $|\mathcal{Z}'_k| \leq 2|\mathcal{Z}_k|/r$.*

If Claim 3.4 holds, then

$$\mathbb{P}(v \in A_{i,2} | u \in A_{i,1}) = \frac{\mathbb{P}(X \in \mathcal{Z}')}{\mathbb{P}(X \in \mathcal{Z})} = \frac{\sum_{k \in [\ell]} p_k |\mathcal{Z}'_k|}{\sum_{k \in [\ell]} p_k |\mathcal{Z}_k|} \leq \max_{k \in [\ell]} \left\{ \frac{|\mathcal{Z}'_k|}{|\mathcal{Z}_k|} \right\} \leq \frac{2}{r},$$

and hence Claim 3.3 holds. It is sufficient then to prove Claim 3.4.

Proof of Claim 3.4. Fix $k \in [\ell]$. Create an auxilliary bipartite graph H with vertex classes $\mathcal{Z}_k \setminus \mathcal{Z}'_k$ and $\mathcal{Z}'_k$, where YY' is an edge for $Y \in \mathcal{Z}_k \setminus \mathcal{Z}'_k$ and $Y' \in \mathcal{Z}'_k$ exactly when Y can be transformed into Y' by switching v with some other vertex.

Note that, for each $Y \in \mathcal{Z}_k \setminus \mathcal{Z}'_k$ labelled as in (3.7), v has to be switched with some vertex in $\bar{A}_{i,2}$ in order to get a sequence in $\mathcal{Z}'_k$. Thus, for each $Y \in \mathcal{Z}_k \setminus \mathcal{Z}'_k$, $d_H(Y) \leq m$. We will show, for each $Y' \in \mathcal{Z}'_k$, $d_H(Y') \geq mr/2$, whence

$$(mr/2) \cdot |\mathcal{Z}'_k| \leq \sum_{Y' \in \mathcal{Z}'_k} d_H(Y') = e(H) = \sum_{Y \in \mathcal{Z}_k \setminus \mathcal{Z}'_k} d_H(Y) \leq m|\mathcal{Z}_k \setminus \mathcal{Z}'_k| \leq m|\mathcal{Z}_k|,$$

and thus Claim 3.4 follows.

Fix then $Y' = (\bar{A}'_{1,1}, \bar{A}'_{1,2}, \bar{a}'_1, \bar{b}'_1, \ldots, \bar{A}'_{i-1,1}, \bar{A}'_{i-1,2}, \bar{a}'_{i-1}, \bar{b}'_{i-1}, \bar{A}'_{i,1}, \bar{A}'_{i,2}) \in \mathcal{Z}'_k$. Recall the steps **S1**–**S4** in the random process. Note that **S1** is deterministic and, for each $j \leq i$, the probability of each possible choice at **S3** is always the same. For each $j \leq i$, the probability of each possible choice at **S2** depends exactly on the size of $(N^c(a_j) \cup N^c(b_j)) \setminus (\bigcup_{j' < j} A_{j'})$. For each $j < i$, the probability of each possible choice at **S4** depends on whether each of a_j and b_j appears in $A_{j,1}$ or $A_{j,2}$.

Thus, noting that, for each $j < i$, $\{\bar{a}'_j, \bar{b}'_j\} \subseteq N^c(\bar{a}'_j) \cup N^c(\bar{b}'_j)$, we can switch any vertices not in $\{u\} \cup (\bigcup_{j < i}(N^c(\bar{a}'_j) \cup N^c(\bar{b}'_j)))$ in Y' and get a sequence with the same probability of occurence as Y. Therefore, if we take Y' and switch v with any vertex in $(\bigcup_{j < i}(\bar{A}'_{j,1} \cup \bar{A}'_{j,2})) \setminus (\bigcup_{j < i}(N^c(\bar{a}'_j) \cup N^c(\bar{b}'_j)))$ then we get a sequence in $\mathcal{Z}_k \setminus \mathcal{Z}'_k$. Therefore,

$$d_H(Y') \geq \left| (\bigcup_{j < i}(\bar{A}'_{j,1} \cup \bar{A}'_{j,2})) \setminus (\bigcup_{j < i}(N^c(\bar{a}'_j) \cup N^c(\bar{b}'_j))) \right| \geq 2(i-1)m - 2(i-1)m/2 = (i-1)m \geq mr/2,$$

as required. This completes the proof of Claim 3.4, and hence the lemma. □ □

3.2 Turning to our initial weighting

We will now show our (appropriately-sized) random subgraph H has a fractional K_r-decomposition using the following lemma, before later improving our methods in Section 5.

Lemma 3.5. *Let $r \geq 3$ and $\ell \geq 1$. Let G be a graph on at least $2^\ell r + 2^{\ell+1} - 2$ vertices so that $E(\bar{G})$ can be split into ℓ matchings. Then, G has a fractional K_r-decomposition.*

Proof. We will prove the lemma for all $r \geq 3$ using induction on ℓ. When $\ell = 1$, the lemma follows directly from Lemma 2.1. Let $\ell > 1$ and assume then that the lemma holds for $\ell - 1$, and let G be a graph on at least $2^\ell r + 2^{\ell+1} - 2 = 2^{\ell-1}(2r+2) + 2^\ell - 2$ vertices so that $E(\bar{G})$ can be split into ℓ matchings.

Let $E(\bar{G}) = E_1 \cup E_2$, where E_1 is a matching and E_2 can be split into $\ell - 1$ matchings. Let $\mathcal{M}$ be the set of induced subgraphs of G with $2r + 2$ vertices and no two vertices in the same edge in E_2. By the inductive hypothesis, $G + E_1$ has a fractional K_{2r+2}-decomposition. Each copy of K_{2r+2} in $G + E_1$ has the same vertex set as a graph in $\mathcal{M}$, and therefore G has a fractional decomposition into graphs in $\mathcal{M}$. By Lemma 2.1, each subgraph in $\mathcal{M}$ has a fractional K_r-decomposition, and thus G has as well. □

We can now combine Lemmas 3.2 and 3.5 to give our initial weighting.

Lemma 3.6. *Let $r \geq 3$ and let G be a graph with $n = (32r + 62)m$ vertices and $\delta(G) \geq n - m/2$. Then, there is a set of weights w_K, $K \in \mathcal{K}_r$, such that, for each edge $e \in E(G)$,*

$$1 - \frac{1}{r} \leq \sum_{K \in \mathcal{K}_r : e \in E(K)} w_K \leq 1. \tag{3.8}$$

Proof. Let $\mathcal{M}$ be the set of induced subgraphs of G with $32r + 62$ vertices which contain a copy of M_{16r+31}. By Lemma 3.2, we may find non-negative weights p_M, $M \in \mathcal{M}$, so that, for each $e \in E(G)$,

$$1 - \frac{1}{r} \leq 1 - \frac{4}{16r + 31} \leq \sum_{M \in \mathcal{M} : e \in E(M)} p_M \leq 1. \tag{3.9}$$

The set $E(\bar{M}_{16r+31})$ can be covered by 5 matchings, and hence, for each $M \in \mathcal{M}$, $E(\bar{M})$ can as well. Therefore, by Lemma 3.5, each $M \in \mathcal{M}$ has a fractional K_r-decomposition, so there exist non-negative weights $w_{M,K}$, $K \in \mathcal{K}_r$, such that, for each $e \in E(G)$,

$$\sum_{K \in \mathcal{K}_r : e \in E(K)} w_{M,K} = \mathbf{1}_{\{e \in E(M)\}}.$$

For each $K \in \mathcal{K}_r$, let $w_K = \sum_{M \in \mathcal{M}} p_M \cdot w_{M,K} \geq 0$. Then, for each $e \in E(G)$,

$$\sum_{K \in \mathcal{K}_r : e \in E(K)} w_K = \sum_{M \in \mathcal{M}} p_M \cdot \sum_{K \in \mathcal{K}_r : e \in E(K)} w_{M,K} = \sum_{M \in \mathcal{M}} p_M \cdot \mathbf{1}_{\{e \in E(M)\}}. \tag{3.10}$$

Combining (3.9) and (3.10) shows that the weights w_K, $K \in \mathcal{K}_r$, satisfy (3.8). □

4 Tidying up the approximate fractional decomposition

The aim of this section is to prove the following lemma, which shows that a graph with an approximate fractional K_{2r+2}-decomposition has a fractional K_r-decomposition.

Lemma 4.1. *Let $r \geq 3$ and let G be a graph for which there is a set of non-negative weights w_K, $K \in \mathcal{K}_{2r+2}$, such that, for each $e \in E(G)$,*

$$1 - \frac{1}{2r + 1} \leq \sum_{K \in \mathcal{K}_{2r+2} : e \in E(K)} w_K \leq 1. \tag{4.1}$$

Then, G has a fractional K_r-decomposition.

We then use this lemma to deduce from Lemma 3.6 that any graph G with minimum degree at least $(1 - 1/(128r + 252))|G|$ has a fractional K_r-decomposition (see Lemma 4.3). In order to prove Lemma 4.1, we first show that we can weight the r-cliques of K_{2r+2} to achieve any particular weight over each edge, as long as these weights lie in $[1 - 1/(2r + 1), 1]$. The proof of the following lemma was sketched in Section 2.2.

Lemma 4.2. *Let $r \geq 3$ and $f : E(K_{2r+2}) \to [1 - 1/(2r + 1), 1]$. Let $\mathcal{K}_r = \mathcal{K}_r(K_{2r+2})$. Then, there is a set of non-negative weights w_K, $K \in \mathcal{K}_r$, such that, for each $e \in E(K_{2r+2})$,*

$$\sum_{K \in \mathcal{K}_r : e \in E(K)} w_K = f(e).$$

Proof. As is well-known, we can find disjoint matchings $M_1, \ldots, M_{2r+1}$ such that $\bigcup_i M_i = E(K_{2r+2})$. Let $\mathcal{M}$ be the set of matchings $M \subseteq E(K_{2r+2})$. By Lemma 2.1, for each $M \in \mathcal{M}$, we can choose weights $w_{M,K} \geq 0$, $K \in \mathcal{K}_r$, so that, for each $e \in E(K_{2r+2})$,

$$\sum_{K \in \mathcal{K}_r : e \in E(K)} w_{M,K} = \mathbf{1}_{\{e \notin M\}}. \tag{4.2}$$

Now, for each $i \in [2r + 1]$, label the edges in M_i as $e_{i,1}, \ldots, e_{i,r+1}$, so that $f(e_{i,1}) \leq f(e_{i,2}) \leq \ldots \leq f(e_{i,r+1})$. For each $M \in \mathcal{M}$, let

$$w_M = \begin{cases} f(e_{i,j+1}) - f(e_{i,j}) & \text{if } M = \{e_{i,1}, \ldots, e_{i,j}\} \text{ for some } i \in [2r + 1], j \in [r] \\ 1 - f(e_{i,r+1}) & \text{if } M = \{e_{i,1}, \ldots, e_{i,r+1}\} \text{ for some } i \in [2r + 1] \\ 1 - \sum_{i \in [2r+1]} (1 - f(e_{i,1})) & \text{if } M = \emptyset \\ 0 & \text{otherwise.} \end{cases} \tag{4.3}$$

As $f(e_{i,1}) \geq 1 - 1/(2r + 1)$ for each $i \in [2r + 1]$, we have $w_\emptyset \geq 0$. As $f(e_{i,r+1}) \leq 1$, and due to the ordering of the edges in each matching M_i, the weights w_M, $M \in \mathcal{M}$, give a non-negative weighting of the matchings in $\mathcal{M}$. For each $K \in \mathcal{K}_r$, let

$$w_K = \sum_{M \in \mathcal{M}} w_M \cdot w_{M,K}, \tag{4.4}$$

so that w_K, $K \in \mathcal{K}_r$, is a non-negative weighting of $\mathcal{K}_r$.

Note that, for each $i \in [2r + 1]$ and $j \in [r + 1]$, we have

$$\sum_{j' < j} w_{\{e_{i,1}, \ldots, e_{i,j'}\}} = f(e_{i,j}) - f(e_{i,1}), \tag{4.5}$$

and, for each $i \in [2r + 1]$, we have

$$\sum_{j \in [r+1]} w_{\{e_{i,1}, \ldots, e_{i,j}\}} = 1 - f(e_{i,1}). \tag{4.6}$$

Therefore, for each $e \in E(G)$, letting i and j be such that $e = e_{i,j}$, we have

$$\sum_{K \in \mathcal{K}_r : e \in E(K)} w_K \overset{(4.4)}{=} \sum_{M \in \mathcal{M}} w_M \cdot \sum_{K \in \mathcal{K}_r : e \in E(K)} w_{M,K} \overset{(4.2)}{=} \sum_{M \in \mathcal{M}} w_M \cdot \mathbf{1}_{\{e \notin M\}}$$

$$= w_\emptyset + \sum_{i' \neq i} \sum_{j' \in [r+1]} w_{\{e_{i',1}, \ldots, e_{i',j'}\}} + \sum_{j' < j} w_{\{e_{i,1}, \ldots, e_{i,j'}\}}$$

$$\overset{(4.6),(4.5)}{=} w_\emptyset + \sum_{i' \neq i} (1 - f(e_{i',1})) + f(e_{i,j}) - f(e_{i,1})$$

$$\overset{(4.3)}{=} 1 - (1 - f(e_{i,1})) + f(e_{i,j}) - f(e_{i,1}) = f(e_{i,j}),$$

as required. $\qquad \square$

We can now prove Lemma 4.1 by using Lemma 4.2 on each $(2r + 2)$-clique of G simultaneously.

Proof of Lemma 4.1. For each $e \in E(G)$, let $z_e = (1 + 1/(2r)) \cdot \sum_{K \in \mathcal{K}_{2r+2} : e \in E(K)} w_K$, so that, using (4.1), we have $1 - 1/(2r + 1) \leq 1/z_e \leq 1$. For each $K \in \mathcal{K}_{2r+2}$, using Lemma 4.2, take non-negative weights $w_{K,K'}$, $K' \in \mathcal{K}_r = \mathcal{K}_r(G)$, so that, for each $e \in E(G)$,

$$\sum_{K' \in \mathcal{K}_r : e \in E(K')} w_{K,K'} = (1/z_e) \cdot \mathbf{1}_{\{e \in E(K)\}}. \tag{4.7}$$

For each $K' \in \mathcal{K}_r$, let $w_{K'} = \sum_{K \in \mathcal{K}_{2r+2}} (1 + 1/(2r)) w_K \cdot w_{K,K'}$, so that we have a non-negative weighting of $\mathcal{K}_r$. Then, for each $e \in E(G)$, we have

$$\sum_{K' \in \mathcal{K}_r : e \in E(K')} w_{K'} = \sum_{K \in \mathcal{K}_{2r+2}} (1 + 1/(2r)) w_K \cdot \sum_{K' \in \mathcal{K}_r : e \in E(K')} w_{K,K'}$$

$$\overset{(4.7)}{=} \sum_{K \in \mathcal{K}_{2r+2}} (1 + 1/(2r)) w_K \cdot (1/z_e) \cdot \mathbf{1}_{\{e \in E(K)\}}$$

$$= \Big((1 + 1/(2r)) \sum_{K \in \mathcal{K}_{2r+2} : e \in E(K)} w_K \Big) / z_e = 1.$$

Thus, the weights $w_{K'}$, $K' \in \mathcal{K}_r$, form a fractional K_r-decomposition of G. $\qquad\square$

We can now conclude from Lemma 3.6 and Lemma 4.1 the following weakened form of Theorem 1.3.

Lemma 4.3. *Let $r \geq 3$. If a graph G has minimum degree at least $(1 - 1/(128r + 252))|G|$, then G has a fractional K_r-decomposition.*

Proof. Let $m = |G|$ and form a graph G' by copying each vertex of G $64r + 126$ times, where two vertices in G' have an edge between them if and only if the two original vertices did in G. Then, we have $\delta(G') \geq |G'| - (64r + 126)|G|/(128r + 252) = |G'| - m/2$.

By Lemma 3.6, there is a set of weights w_K, $K \in \mathcal{K}_{2r+2} := \mathcal{K}_{2r+2}(G')$, such that, for each $e \in E(G')$,

$$1 - \frac{1}{2r+2} \leq \sum_{K \in \mathcal{K}_{2r+2} : e \in E(K)} w_K \leq 1.$$

Thus, by Lemma 4.1, G' has a fractional K_r-decomposition.

Each r-clique in G' has vertices which are copied from some shared r-clique in G. Let $\tau : \mathcal{K}_r(G') \to \mathcal{K}_r(G)$ be the natural projection so that, for each $K \in \mathcal{K}_r(G')$, $V(K)$ is a set of copies of vertices in $V(\tau(K))$. For each $K \in \mathcal{K}_r(G)$, letting $w_K = (1/k^2) \cdot \sum_{K' \in \mathcal{K}_r(G') : \tau(K') = K} w_{K'}$, with $k = 64r + 126$, then forms a fractional K_r-decomposition of G. $\qquad\square$

5   Improving on Lemma 3.6

To improve on Lemma 3.6, we essentially consider the following question: For how small a value of ℓ can we show that any graph G with 2ℓ vertices containing a copy of M_ℓ has a K_r-decomposition (see Definition 3.1)? For Lemma 3.6, where we used $\ell \geq 16r + 31$, we simply used Lemma 2.1 applied several times (via Lemma 3.5). In this section, we will improve this to use $\ell \geq 9r + 8$ in Lemma 5.3 (taking also that ℓ is even to simplify the proof). We will first define a new graph W_k and show that, for sufficiently large k, any graph with a spanning copy of W_k has a fractional K_r-decomposition.

Definition 5.1. *Let W_k be a complete k-partite graph with classes of size 4 on the vertex set $[4k]$.*

Lemma 5.2. *Let $r \geq 3$, $k \geq (3r + 2)/2$, and let G be a graph with $4k$ vertices which contains a copy of W_k. Then, G has a fractional K_r-decomposition.*

Proof. Using that G contains a copy of W_k, divide the set $V(G)$ into $A_1, \ldots, A_k$ so that each set A_i has size 4 and G contains all the edges between different sets A_i. Let $E \subseteq E(G)$ be the set of edges of G which lie within some set A_i. Let $\mathcal{M}$ be the set of all subgraphs $M \subseteq G$ with $|M| = 2r + 2$ vertices and $\delta(M) \geq |M| - 2$. By picking $M \in \mathcal{M}$ randomly, we will weight the subgraphs in $\mathcal{M}$ so that each edge in G is given total weight 1. As, by Lemma 2.1, each subgraph $M \in \mathcal{M}$ has a fractional K_r-decomposition, this implies that G has a fractional K_r-decomposition.

Pick $I \subseteq [k]$ randomly so that $|I| = r + 1$, and, starting with $V_0 = \emptyset$, for each $i \in I$, pick uniformly and independently a random subset of A_i with size 2 and add it to V_0. Let $M = G[V_0]$. Note that $M \in \mathcal{M}$.

11

If $e \in E(G) \setminus E$ lies between A_i and A_j, then

$$\mathbb{P}(e \in E(M)) = \mathbb{P}(i, j \in I) \cdot 1/4 = \frac{\binom{r+1}{2}}{4\binom{k}{2}},$$

and, if $e \in E$ lies within A_i, then

$$\mathbb{P}(e \in E(M)) = \mathbb{P}(i \in I) \cdot 1/6 = (r+1)/(6k) \geq \frac{\binom{r+1}{2}}{4\binom{k}{2}},$$

as $k \geq (3r+2)/2$. Note that, if we delete any set of edges in E from M, then the resulting subgraph is still in $\mathcal{M}$. Therefore, we can alter our probability distribution to get a random subgraph $M' \in \mathcal{M}$ so that, for every $e \in E(G)$, we have $\mathbb{P}(e \in E(M')) = \binom{r+1}{2}/4\binom{k}{2}$. Thus, normalising appropriately, we can find weights w_M, $M \in \mathcal{M}$, so that, for each $e \in E(G)$, $\sum_{M \in \mathcal{M}: e \in E(M)} w_M = 1$. As every graph in $\mathcal{M}$ has a fractional K_r-decomposition, G itself then has a fractional K_r-decomposition. $\qquad \square$

We can now show that, if $\ell \geq (9r+8)/2$, then any graph containing a spanning copy of $M_{2\ell}$ has a fractional K_r-decomposition. We do this by fractionally decomposing such a graph into subgraphs with a spanning copy of W_k or $\bar{C}_{4k}$ (for some appropriate k), where C_{4k} is the cycle on $4k$ vertices. By using Lemma 5.2 and Lemma 3.5 respectively to find a fractional K_r-decomposition of these subgraphs, we will get a fractional K_r-decomposition of the original graph.

Lemma 5.3. *Let $r \geq 3$ and $\ell \geq (9r+8)/2$. If G is a graph with 4ℓ vertices which contains a copy of $M_{2\ell}$, then G has a fractional K_r-decomposition.*

Proof. Let $k = \lceil (3r+2)/2 \rceil$. Label $V(G)$ as $V = \{a_1, b_1, \ldots, a_{2\ell}, b_{2\ell}\}$ so that G has all edges between sets $\{a_i, b_i\}$ and $\{a_j, b_j\}$ if $i \notin \{j-1, j, j+1\}$, which is possible as G contains a copy of $M_{2\ell}$. Let $\mathcal{M}$ be the set of all subgraphs $H \subseteq G$ with $|H| = 4k$ which contain a copy of either $\bar{C}_{4k}$ or W_k.

Note that, for any graph H with $4k \geq 4r+6$ vertices which contains a copy of $\bar{C}_{4k}$, the set $E(\bar{H})$ can be covered by two matchings. Therefore, by Lemma 3.5, any such graph H has a fractional K_r-decomposition. In combination with Lemma 5.2, then, each graph in $\mathcal{M}$ has a fractional K_r-decomposition.

For each $j \in [\ell]$, let $A_j = \{a_{2j-1}, b_{2j-1}, a_{2j}, b_{2j}\}$. For each $i \in \{0, 1, 2\}$, pick a random induced subgraph C_i of G with $|C_i| = 4k$ according to the following procedure:

- Select a subset $I \subseteq [\ell]$ with $|I| = 2k$ uniformly and independently.

- For each $j \in I$,

  - if $i = 0$, let $V(C_i) \cap A_j = \{a_{2j-1}, b_{2j-1}\}$ or $\{a_{2j}, b_{2j}\}$ with probability $1/2$,
  - if $i = 1$, let $V(C_i) \cap A_j = \{a_{2j-1}, a_{2j}\}$ or $\{b_{2j-1}, b_{2j}\}$ with probability $1/2$, and
  - if $i = 2$, let $V(C_i) \cap A_j = \{a_{2j-1}, b_{2j}\}$ or $\{b_{2j-1}, a_{2j}\}$ with probability $1/2$.

- For each $j \notin I$, let $V(C_i) \cap A_j = \emptyset$.

Then, pick $i \in \{0, 1, 2\}$ uniformly and independently, and let $C = C_i$. Note that C_0 contains a copy of W_k and C_1 and C_2 both contain a copy of $\bar{C}_{4k}$. Therefore, $C \in \mathcal{M}$.

Let $E \subseteq E(G)$ be the set of edges of G which each lie within some set A_i. If $e \in E(G) \setminus E$ lies between A_i and A_j, then

$$\mathbb{P}(e \in E(C)) = \mathbb{P}(i, j \in I) \cdot 1/4 = \frac{\binom{2k}{2}}{4\binom{\ell}{2}},$$

and if $e \in E$ lies within A_i, then

$$\mathbb{P}(e \in E(C)) = \mathbb{P}(i \in I) \cdot 1/6 = k/(3\ell) \geq \frac{\binom{2k}{2}}{4\binom{\ell}{2}}.$$

12

Note that deleting any set of edges in E from C gives another subgraph in $\mathcal{M}$. Therefore, we can alter our probability distribution to give a random subgraph $C' \in \mathcal{M}$ so that $\mathbb{P}(e \in E(C')) = \binom{2k}{2}/4\binom{\ell}{2}$.

That is, normalising appropriately, we can find weights w_M, $M \in \mathcal{M}$, so that, for each $e \in E(G)$, we have $\sum_{M \in \mathcal{M}:e \in E(M)} w_M = 1$. Recalling that every graph in $\mathcal{M}$ has a fractional K_r-decomposition, G then has a fractional K_r-decomposition. $\qquad\square$

We can now use this improved fractional decomposition of graphs containing a spanning copy of $M_{2\ell}$ to improve on Lemma 3.6.

Lemma 5.4. *Let $r \geq 3$, $\ell = \lceil (9r+8)/2 \rceil$ and let G be a graph with $n = 4\ell m$ vertices and $\delta(G) \geq n - m/2$. Then, there is a set of weights w_K, $K \in \mathcal{K}_r$, such that, for each edge $e \in E(G)$,*

$$1 - \frac{1}{r} \leq \sum_{K \in \mathcal{K}_r:e \in E(K)} w_K \leq 1. \tag{5.1}$$

Proof. Let $\mathcal{M}$ be the set of induced subgraphs of G with 4ℓ vertices which contain a copy of $M_{2\ell}$. By Lemma 3.2, we can find non-negative weights p_M, $M \in \mathcal{M}$, so that, for each $e \in E(G)$,

$$1 - \frac{1}{r} \leq 1 - \frac{4}{2\ell} \leq \sum_{M \in \mathcal{M}:e \in E(M)} p_M \leq 1.$$

By Lemma 5.3, each graph in $\mathcal{M}$ has a fractional K_r-decomposition. Similarly to the proof of Lemma 3.6, we can combine these fractional K_r-decompositions with the weighting of the graphs in $\mathcal{M}$ to get a weighting w_K, $K \in \mathcal{K}_r$, which satisfies (5.1). $\qquad\square$

Subject only to the remaining proof of Lemma 2.1, we can now complete the proof of Theorem 1.3.

Proof of Theorem 1.3. Let $\ell = \lceil (18r+26)/2 \rceil$. Note that we can assume $n = 4\ell m \leq (36r+54)m \leq 50rm$ by copying each vertex 4ℓ times. Then, we have $\delta(G) \geq |G| - m/2$. By Lemma 5.4, there is a set of weights w_K, $K \in \mathcal{K}_{2r+2}$, such that, for each edge $e \in E(G)$,

$$1 - \frac{1}{2r+2} \leq \sum_{K \in \mathcal{K}_{2r+2}:e \in E(K)} w_K \leq 1.$$

Thus, by Lemma 4.1, G has a fractional K_r-decomposition. $\qquad\square$

6 Proof of Lemma 2.1

It remains only to show that any complete graph K_k on $k \geq 2r+2$ vertices with any set of independent edges M removed, $K_k - M$, has a fractional K_r-decomposition, thus proving Lemma 2.1. We will easily be able to show that we may assume that $k = 2r+2$ and $1 \leq |M| \leq r$. Then, up to symmetry, there are three types of edges in $K_{2r+2} - M$: edges with 0, 1, or 2 vertices in edges in M. We will consider three types of r-clique in $K_k - M$ defined by the number of vertices they contain in M. By adding weight uniformly to each clique of one type we can add weight to the edges of $K_k - M$ so that the same weight is added to edges of the same type. By choosing three different weights to add to the three different types of cliques, we can control the amount of weight added to each type of edge and gain a fractional K_r-decomposition of $K_k - M$.

Proof of Lemma 2.1. Note that we can assume that G has $2r+2$ vertices. Indeed, if G has more than $2r+2$ vertices, then by giving the induced $2r+2$ vertex subgraphs of G an appropriate uniform weight we can fractionally decompose G into graphs with $2r+2$ vertices each of which only lacks edges in some matching.

Let G then be a graph with vertex set $[2r+2]$ so that $M := E(\bar{G})$ is a matching. Let $k = |M|$, and let A be the set of vertices in some edge in M, so that $|A| = 2k$. If $k = r+1$ or 0, then adding weight 1 to every r-clique

13

in G weights the edges of G uniformly (due to the symmetry in G), and hence, by normalising these weights appropriately, we can find a fractional K_r-decomposition of G. Let us assume then that $1 \le k \le r$.

For each $i \in \{0, 1, 2\}$, let $E_i = \{e \in E(G) : |V(e) \cap A| = i\}$. If we can find a non-negative weighting w_K, $K \in \mathcal{K}_r$, so that, for each $i \in \{0, 1, 2\}$,

$$\sum_{K \in \mathcal{K}_r} |E(K) \cap E_i| \cdot w_K = |E_i|, \tag{6.1}$$

then, by the symmetry in G, we can easily convert this into a fractional K_r-decomposition of G. Indeed, given such a weighting w_K, $K \in \mathcal{K}_r$, for each $K \in \mathcal{K}_r$ let

$$w'_K = \frac{1}{|\{K' \in \mathcal{K}_r : |V(K') \cap A| = |V(K) \cap A|\}|} \sum_{K' \in \mathcal{K}_r : |V(K') \cap A| = |V(K) \cap A|} w_{K'}. \tag{6.2}$$

Note that, for each $\ell \in \{0, 1, \ldots, k\}$, w'_K is the same for each $K \in \mathcal{K}_r$ with $|V(K) \cap A| = \ell$. Due to the symmetry in G, then, for each $i \in \{0, 1, 2\}$, $z_e := \sum_{K \in \mathcal{K}_r : e \in E(K)} w'_K$ is the same, z_i say, for each $e \in E_i$. Furthermore, for each $\ell \in \{0, 1, \ldots, k\}$ and $i \in \{0, 1, 2\}$, $|E_i \cap E(K)|$ is the same for each $K \in \mathcal{K}_r$ with $|V(K) \cap A| = \ell$. Thus, for each $\ell \in \{0, 1, \ldots, k\}$ and $i \in \{0, 1, 2\}$, we have from (6.2) that

$$\sum_{K \in \mathcal{K}_r : |V(K) \cap A| = \ell} |E_i \cap E(K)| \cdot w'_K = \sum_{K \in \mathcal{K}_r : |V(K) \cap A| = \ell} |E_i \cap E(K)| \cdot w_K. \tag{6.3}$$

Therefore, for each $i \in \{0, 1, 2\}$,

$$|E_i| \cdot z_i = \sum_{e \in E_i} z_e = \sum_{e \in E_i} \sum_{K \in \mathcal{K}_r : e \in E(K)} w'_K = \sum_{K \in \mathcal{K}_r} |E(K) \cap E_i| \cdot w'_K \overset{(6.3)}{=} \sum_{K \in \mathcal{K}_r} |E(K) \cap E_i| \cdot w_K \overset{(6.1)}{=} |E_i|,$$

and hence $z_i = 1$. That is, for any $e \in E(G)$, $\sum_{K \in \mathcal{K}_r : e \in E(K)} w'_K = 1$. Thus, it is sufficient to find a non-negative weighting that satisfies (6.1).

For each k', $0 \le k' \le k$, if an r-clique has k' vertices in A, then it contains $k'(k'-1)/2$ edges in E_2, $k'(r - k')$ edges in E_1 and $(r - k')(r - k' - 1)/2$ edges in E_0. Furthermore, $|E_2| = 2k(2k - 2)/2$, $|E_1| = 2k(2r + 2 - 2k)$ and $|E_0| = (2r + 2 - 2k)(2r + 1 - 2k)/2$. Note that there exists an r-clique in G with k' vertices in A if $0 \le k' \le k$ and $r - k' \le 2r + 2 - |A| = 2r + 2 - 2k$, or, equivalently, $\max\{0, 2k - r - 2\} \le k' \le k$.

Thus, to prove the lemma it is sufficient to find $\max\{0, 2k - r - 2\} \le k_1, k_2, k_3 \le k$ and $x, y, z \ge 0$ for which we have

$$\begin{pmatrix} \frac{1}{2}k_1(k_1 - 1) & \frac{1}{2}k_2(k_2 - 1) & \frac{1}{2}k_3(k_3 - 1) \\ k_1(r - k_1) & k_2(r - k_2) & k_3(r - k_3) \\ \frac{1}{2}(r - k_1)(r - k_1 - 1) & \frac{1}{2}(r - k_2)(r - k_2 - 1) & \frac{1}{2}(r - k_3)(r - k_3 - 1) \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
$$= \begin{pmatrix} k(2k - 2) \\ 2k(2r + 2 - 2k) \\ (r + 1 - k)(2r + 1 - 2k) \end{pmatrix}. \tag{6.4}$$

Now, if $2k - r - 2 \ge 0$, then taking $k_1 = k$, $k_2 = k - 1$, $k_3 = 2k - r - 2$,

$$x = \frac{2(r + 1 - k)(2(r - k)^2 + kr + 5r - 4k + 2)}{r(r - 1)(r + 2 - k)},$$

$$y = \frac{2k(3r - 2k + 2)}{r(r - 1)} \text{ and } z = \frac{2k(k - 1)}{r(r - 1)(r + 2 - k)}$$

satisfies (6.4). As $k \le r$, we have $k - 1 > 2k - r - 2$, so that $2k - r - 2 \le k_1, k_2, k_3 \le r$ and, as $r \ge 3$ and $1 \le k \le r$, we have $x, y, z \ge 0$.

If $2k - r - 2 < 0$, then taking $k_1 = k$, $k_2 = k - 1$, $k_3 = 0$,

$$x = \frac{4(r + 1 - k)}{r - 1}, \quad y = \frac{4k}{r - 1}, \text{ and } z = \frac{2(r + 1 - k)}{r(r - 1)}$$

satisfies (6.4), where $x, y, z \ge 0$ as $k \le r$ and $r \ge 3$. Thus, for all possible values of k, G has a fractional K_r-decomposition. $\qquad\square$

References

[1] B. Barber, D. Kühn, A. Lo, R.H. Montgomery, and D. Osthus. Fractional clique decompositions of dense graphs and hypergraphs. *J. Combin. Theory Ser. B*, 127 (2017), 148–186.

[2] B. Barber, D. Kühn, A. Lo, and D. Osthus. Edge-decompositions of graphs with high minimum degree. *Adv. Math.*, 288 (2016), 337–385.

[3] D. Dor and M. Tarsi. Graph decomposition is NP-complete: a complete proof of Holyer's conjecture. *SIAM J. Comput.*, 26(4):1166–1187, 1997.

[4] F. Dross. Fractional triangle decompositions in graphs with large minimum degree. *SIAM Journal on Discrete Mathematics*, 30(1), 3642, 2016.

[5] P. Dukes. Rational decomposition of dense hypergraphs and some related eigenvalue estimates. *Linear Algebra Appl.*, 436(9):3736–3746, 2012.

[6] P. Dukes. Corrigendum to "Rational decomposition of dense hypergraphs and some related eigenvalue estimates" [Linear Algebra Appl. 436 (9) (2012) 3736–3746]. *Linear Algebra Appl.*, 467:267–269, 2015.

[7] K. Garaschuk. *Linear methods for rational triangle decompositions*. PhD thesis, University of Victoria, 2014.

[8] S. Glock, D. Kühn, A. Lo, R.H. Montgomery and D. Osthus. On the decomposition threshold of a given graph. *arXiv:1603.04724*, 2016.

[9] S. Glock, D. Kühn, A. Lo and D. Osthus. Hypergraph F-designs for arbitrary F. *arXiv:1706.01800*, 2017.

[10] S. Glock, D. Kühn, A. Lo and D. Osthus. The existence of designs via iterative absorption. *arXiv:1611.06827*, 2016.

[11] T. Gustavsson. *Decompositions of large graphs and digraphs with high minimum degree*. PhD thesis, Univ. of Stockholm, 1991.

[12] P.E. Haxell and V. Rödl. Integer and fractional packings in dense graphs. *Combinatorica*, 21(1):13–38, 2001.

[13] P. Keevash. The existence of designs. *arXiv:1401.3665*, 2014.

[14] T.P. Kirkman. On a problem in combinations. *Cambridge and Dublin Math. J.*, 2:191–204, 1847.

[15] C.St.J.A. Nash-Williams. An unsolved problem concerning decomposition of graphs into triangles. In *Combinatorial Theory and its Applications III*, 1179–183. North Holland, 1970.

[16] R.M. Wilson. An existence theory for pairwise balanced designs. I. Composition theorems and morphisms. *J. Combin. Theory Ser. A*, 13:220–245, 1972.

[17] R.M. Wilson. An existence theory for pairwise balanced designs. II. The structure of PBD-closed sets and the existence conjectures. *J. Combin. Theory Ser. A*, 13:246–273, 1972.

[18] R.M. Wilson. An existence theory for pairwise balanced designs. III. Proof of the existence conjectures. *J. Combin. Theory Ser. A*, 18:71–79, 1975.

[19] R.M. Wilson. Decompositions of complete graphs into subgraphs isomorphic to a given graph. In *Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975)*, 647–659. Congressus Numerantium, No. XV, Utilitas Math., Winnipeg, Man., 1976.

[20] R. Yuster. Asymptotically optimal K_k-packings of dense graphs via fractional K_k-decompositions. *J. Combin. Theory Ser. B*, 95(1):1–11, 2005.