

# RESOLUTION OF THE KOHAYAKAWA–KREUTER CONJECTURE

MICHA CHRISTOPH, ANDERS MARTINSSON, RAPHAEL STEINER, AND YUVAL WIGDERSON

ABSTRACT. A graph  $G$  is said to be *Ramsey* for a tuple of graphs  $(H_1, \dots, H_r)$  if every  $r$ -coloring of the edges of  $G$  contains a monochromatic copy of  $H_i$  in color  $i$ , for some  $i$ . A fundamental question at the intersection of Ramsey theory and the theory of random graphs is to determine the threshold at which the binomial random graph  $G_{n,p}$  becomes a.s. Ramsey for a fixed tuple  $(H_1, \dots, H_r)$ , and a famous conjecture of Kohayakawa and Kreuter predicts this threshold. Earlier work of Mousset–Nenadov–Samotij, Bowtell–Hancock–Hyde, and Kuperwasser–Samotij–Wigderson has reduced this probabilistic problem to a *deterministic* graph decomposition conjecture. In this paper, we resolve this deterministic problem, thus proving the Kohayakawa–Kreuter conjecture. Along the way, we prove a number of novel graph decomposition results which may be of independent interest.

MSC classifications: 05C70, 05D10 (primary), 05C80 (secondary).

## 1. INTRODUCTION

The topic of this paper is that of graph decompositions, which we view from two different perspectives. From a graph-theoretic perspective, a graph decomposition result is a statement of the form “any graph  $G$  satisfying certain properties can be edge-partitioned into subgraphs with a certain structure”. A famous result of this type is Nash-Williams’ theorem [27], which gives a necessary and sufficient sparsity condition for when  $G$  can be decomposed into the edge-union of  $k$  forests.

Additionally, we also study graph decompositions from a Ramsey-theoretic perspective, in which a typical statement is of the form “if a graph  $G$  is sufficiently complex, then it cannot be edge-partitioned into subgraphs avoiding a certain structure”. The foundational result of this type is Ramsey’s theorem [29], which can be phrased as saying that any sufficiently large complete graph cannot be edge-partitioned into two subgraphs of bounded clique number. Thus, in some sense, the graph-theoretic perspective on decomposition asks when a decomposition into certain structures is *possible*, and the Ramsey-theoretic perspective asks when certain structures are *unavoidable* in *every* decomposition.

The main results of this paper interact with both perspectives. On the one hand, we resolve a conjecture of Kohayakawa and Kreuter [16] on Ramsey properties of random graphs. On the other hand, the proof uses the first perspective. For example, one of our main theorems, Theorem 1.7 below, states that if a graph satisfies a certain sparsity condition (closely related to the one appearing in Nash-Williams’ theorem), then it can be edge-partitioned into a forest and a subgraph satisfying a different, stronger sparsity condition. This statement and its proof are closely related to many known graph decomposition theorems, such as the main results of [4, 8, 11, 15, 23, 24, 27]. As mentioned above, the reason we care about the precise statement comes from the Ramsey-theoretic study of random graphs, a topic we now introduce.

**1.1. Ramsey properties of random graphs.** Given an  $r$ -tuple  $(H_1, \dots, H_r)$  of graphs, a graph  $G$  is said to be *Ramsey* for  $(H_1, \dots, H_r)$  if every  $r$ -coloring of the edges of  $G$  contains a monochromatic copy of  $H_i$  in color  $i$ , for some  $1 \leq i \leq r$ . Equivalently, in the language above,  $G$  is Ramsey for  $(H_1, \dots, H_r)$  if there is no decomposition of  $G$  into the edge-union of  $r$  graphs,

---

*Date:* August 19, 2024.

Research of M.C. and R.S. funded by SNSF Ambizione grant No. 216071.

Research of Y.W. supported by Dr. Max Rössler, the Walter Haefner Foundation, and the ETH Zürich Foundation.

the  $i$ th of which is  $H_i$ -free<sup>1</sup>. In general, the fundamental question of graph Ramsey theory is to understand which graphs  $G$  are Ramsey for a given  $r$ -tuple; for more information, see e.g. the survey [2]. An important special case of this question, which came to prominence in the late 1980s thanks to pioneering work of Frankl–Rödl [6] and Łuczak–Ruciński–Voigt [21], is the question of when a *random* graph is Ramsey for a given  $r$ -tuple with high probability. More precisely, if  $G_{n,p}$  denotes the binomial random graph with edge density  $p$ , then the question is for which values of  $p = p(n)$  one has that  $G_{n,p}$  is Ramsey for  $(H_1, \dots, H_r)$  a.a.s.<sup>2</sup> In the symmetric case  $H_1 = \dots = H_r$ , this question was completely resolved in seminal work of Rödl and Ruciński [30, 31, 32]. For a graph  $J$ , let us denote by  $v(J), e(J)$  its number of vertices and edges, respectively, and let us define the *maximal 2-density* of a graph  $H$  with<sup>3</sup>  $e(H) \geq 2$  to be

$$m_2(H) := \max \left\{ \frac{e(J) - 1}{v(J) - 2} : J \subseteq H, v(J) \geq 3 \right\}.$$

With this notation, the random Ramsey theorem of Rödl and Ruciński [30, 32] is as follows.

**Theorem 1.1** (Rödl–Ruciński). *Let  $H$  be a graph which is not a forest and let  $r \geq 2$  be an integer. There exist constants  $C > c > 0$  such that*

$$\lim_{n \rightarrow \infty} \Pr(G_{n,p} \text{ is Ramsey for } (\underbrace{H, \dots, H}_{r \text{ times}})) = \begin{cases} 1 & \text{if } p \geq Cn^{-1/m_2(H)}, \\ 0 & \text{if } p \leq cn^{-1/m_2(H)}. \end{cases}$$

Any theorem of this form really consists of two statements, which are usually called the 1- and 0-statements; the latter says that the asymptotic probability of an event is 0 in a certain regime, and the former says that the asymptotic probability is 1 in a different regime.

There is a simple heuristic explanation for why the threshold for the Ramsey property is controlled by the quantity  $m_2(H)$ . To explain it, let us suppose for simplicity that  $H$  is *strictly 2-balanced*, meaning that  $m_2(J) < m_2(H)$  for any proper subgraph  $J \subsetneq H$ . Then one can easily verify that at the regime  $p \asymp n^{-1/m_2(H)}$ , an average edge of  $G_{n,p}$  lies in a constant number of copies of  $H$ . Thus, if  $p \leq cn^{-1/m_2(H)}$  where  $c \ll 1$ , then a typical edge lies in no copy of  $H$ , and one expects the copies of  $H$  in  $G_{n,p}$  to be “well spread-out”. Thus, it is reasonable to expect that one can  $r$ -color the edges without creating monochromatic copies of  $H$ . On the other hand, if  $p \geq Cn^{-1/m_2(H)}$  where  $C \gg 1$ , then a typical edge lies in a large (constant) number of copies of  $H$ . In this regime, we expect a lot of interaction between different  $H$ -copies, and it should be difficult to avoid creating monochromatic copies.

The heuristic above suggests that the “reason” why  $G_{n,p}$  should be Ramsey for  $(H, \dots, H)$  is because of a “global” interaction between the copies of  $H$ . However, there is also a potential “local” reason: for any fixed graph  $G$  that is Ramsey for  $(H, \dots, H)$ , if  $G$  is a subgraph of  $G_{n,p}$ , then certainly  $G_{n,p}$  is Ramsey for  $(H, \dots, H)$ . Therefore, in order to prove the 0-statement in Theorem 1.1, one necessarily has to prove that when  $p \leq cn^{-1/m_2(H)}$ , then  $G_{n,p}$  a.a.s. does not contain any fixed  $G$  which is Ramsey for  $(H, \dots, H)$ . It is well-known (e.g. [14, Theorem 3.4]) that the threshold for appearance of  $G$  in  $G_{n,p}$  is determined by the *maximal density* of  $G$ , defined as

$$m(G) := \max \left\{ \frac{e(J)}{v(J)} : J \subseteq G, v(J) \geq 1 \right\}.$$

Thus, a necessary condition for the 0-statement in Theorem 1.1 is the following lemma, which Rödl and Ruciński [30] termed the *deterministic lemma*.

**Lemma 1.2** (Rödl–Ruciński). *Let  $H$  be a graph which is not a forest, and let  $r \geq 2$  be an integer. If  $G$  is Ramsey for the  $r$ -tuple  $(H, \dots, H)$ , then  $m(G) > m_2(H)$ .*

<sup>1</sup>We say that a graph is  $H$ -free if it does not contain a copy of  $H$  as a (not necessarily induced) subgraph.

<sup>2</sup>We say that an event happens *asymptotically almost surely* (a.a.s.) if its probability tends to 1 as  $n \rightarrow \infty$ , where the implicit parameter  $n$  will always be clear from context.

<sup>3</sup>By convention, one also defines  $m_2(H) = 0$  if  $e(H) = 0$  and  $m_2(H) = \frac{1}{2}$  if  $e(H) = 1$ .

Equivalently, the deterministic lemma can be phrased as a decomposition result: if  $m(G) \leq m_2(H)$ , then  $G$  can be edge-partitioned into  $r$  graphs, each of which is  $H$ -free. As discussed above, the deterministic lemma is certainly a necessary condition for the 0-statement in Theorem 1.1 to hold, but in fact, Rödl and Ruciński [30] proved a so-called *probabilistic lemma*, which states that this simple necessary condition is also sufficient. Informally, the probabilistic lemma guarantees that the “local” and “global” reasons discussed above are the only two reasons why  $G_{n,p}$  is Ramsey for  $(H, \dots, H)$ , and the deterministic lemma rules out the “local” reason.

Theorem 1.1 provides a very satisfactory answer to the question “when is  $G_{n,p}$  Ramsey for  $(H_1, \dots, H_r)$ ?” in the case that  $H_1 = \dots = H_r$ , but says nothing about the general case. However, nearly thirty years ago, Kohayakawa and Kreuter [16] formulated a conjecture for the threshold for an arbitrary  $r$ -tuple of graphs. Given two graphs  $H_1, H_2$  with  $m_2(H_1) \geq m_2(H_2)$ , they defined the *mixed 2-density* to be

$$m_2(H_1, H_2) := \max \left\{ \frac{e(J)}{v(J) - 2 + 1/m_2(H_2)} : J \subseteq H_1, v(J) \geq 2 \right\}.$$

It is well-known and easy to verify (see e.g. [19, Lemma 3.4] or [1, Proposition 3.1]) that  $m_2(H_1) \geq m_2(H_1, H_2) \geq m_2(H_2)$ , and that both inequalities are strict if one is.

**Conjecture 1.3** (Kohayakawa–Kreuter). *Let  $H_1, \dots, H_r$  be graphs, and suppose that  $m_2(H_1) \geq \dots \geq m_2(H_r)$  and  $m_2(H_2) > 1$ . There exist constants  $C > c > 0$  such that*

$$\lim_{n \rightarrow \infty} \Pr(G_{n,p} \text{ is Ramsey for } (H_1, \dots, H_r)) = \begin{cases} 1 & \text{if } p \geq Cn^{-1/m_2(H_1, H_2)}, \\ 0 & \text{if } p \leq cn^{-1/m_2(H_1, H_2)}. \end{cases}$$

We remark that the assumption  $m_2(H_2) > 1$  is equivalent to saying that  $H_2$  is not a forest; this condition was added by Kohayakawa, Schacht, and Spöhel [17] to rule out sporadic counterexamples (in analogy to the statement of Theorem 1.1). Just as in the case of Theorem 1.1, there is a simple heuristic explanation for why the function  $m_2(H_1, H_2)$  controls the threshold for the asymmetric Ramsey property of  $G_{n,p}$ . Roughly speaking, one can verify that at the density  $p \asymp n^{-1/m_2(H_1, H_2)}$ , the number of edges appearing in the union of the  $H_1$ -copies is of the same order as the number of copies of  $H_2$ . As the edges not participating in copies of  $H_1$  are irrelevant (they can be colored in color 1 with no adverse consequences), we find that an average “relevant” edge intuitively lies in a constant number of copies of  $H_2$ , and the heuristic above suggests that this is the threshold at which it becomes hard to avoid monochromatic copies of  $H_2$ . We remark that this heuristic, as well as the statement of Conjecture 1.3, does not involve  $H_3, \dots, H_r$  at all; thus, the intuition is that being Ramsey for a tuple is essentially as hard as being Ramsey for the two densest graphs in the tuple.

Conjecture 1.3 has received a great deal of attention over the past three decades [1, 9, 12, 16, 17, 18, 19, 22]. For many years, most papers on the topic aimed to prove the Kohayakawa–Kreuter conjecture for certain special families of  $H_1, \dots, H_r$ ; for example, Kohayakawa and Kreuter [16] proved it when every  $H_i$  is a cycle, Marciniszyn, Skokan, Spöhel, and Steger [22] proved it when every  $H_i$  is a clique, Liebenau, Mattos, Mendonça, and Skokan [20] proved it when  $H_1$  is a clique and  $H_2$  is a cycle, Hyde [13] proved it for most pairs of regular graphs, and Kuperwasser and Samotij [18] proved it when  $m_2(H_1) = m_2(H_2)$ . More recent works have proved results in greater generality. Notably, Mousset, Nenadov, and Samotij [25] established the 1-statement of Conjecture 1.3 for all  $(H_1, \dots, H_r)$ . Subsequently, Bowtell–Hancock–Hyde [1] and Kuperwasser–Samotij–Wigderson [19] independently proved a generalization of the Rödl–Ruciński probabilistic lemma in the asymmetric case.

**Theorem 1.4** (Bowtell–Hancock–Hyde, Kuperwasser–Samotij–Wigderson). *Suppose that for every pair of graphs  $(H_1, H_2)$  with  $m_2(H_1) > m_2(H_2) > 1$ , the following holds: if  $G$  is Ramsey for  $(H_1, H_2)$ , then  $m(G) > m_2(H_1, H_2)$ . Then Conjecture 1.3 is true<sup>4</sup>.*

<sup>4</sup>Strictly speaking, the result of [1, 19] only implies that the 0-statement of Conjecture 1.3 is true, but the 1-statement is already known to be true unconditionally, thanks to the result of Mousset–Nenadov–Samotij [25] mentioned above.

Theorem 1.4 is a generalization of the Rödl–Ruciński probabilistic lemma discussed above. Indeed, the condition in Theorem 1.4 is clearly necessary for the 0-statement of Conjecture 1.3 to hold, as if there were some  $G$  with  $m(G) \leq m_2(H_1, H_2)$  which is Ramsey for  $(H_1, H_2)$ , then that  $G$  appears with positive probability in  $G_{n, cn^{-1/m_2(H_1, H_2)}}$  for any constant  $c > 0$ , and hence the 0-statement of Conjecture 1.3 would be false. Theorem 1.4 then says that this necessary condition is also sufficient. Thus, the validity of Conjecture 1.3 is reduced to a deterministic graph decomposition question. In both [1, 19], this deterministic condition was verified for most pairs  $(H_1, H_2)$ , but its verification for all pairs remained open.

**1.2. Our results.** As discussed above, the earlier works [1, 19, 25] reduced Conjecture 1.3 to verifying a certain deterministic condition. Our main result confirms that this condition always holds, thus completing the proof of Conjecture 1.3.

**Theorem 1.5.** *Let  $H_1, H_2$  be graphs with  $m_2(H_1) > m_2(H_2) > 1$ . If a graph  $G$  is Ramsey for  $(H_1, H_2)$ , then  $m(G) > m_2(H_1, H_2)$ .*

Equivalently, in the graph decomposition language, Theorem 1.5 states that if  $m(G) \leq m_2(H_1, H_2)$ , then  $G$  can be edge-partitioned into an  $H_1$ -free graph and an  $H_2$ -free graph.

In fact, we prove two more general graph decomposition theorems, Theorems 1.6 and 1.7 below, which we expect to be of independent interest. It is not hard to see (and we show this in the next section) that these two results, plus simple well-known arguments, suffice to prove Theorem 1.5.

Recall that a *pseudoforest* is a graph in which every connected component contains at most one cycle. If  $F$  is a subgraph of  $G$ , we denote by  $G - F$  the subgraph of  $G$  comprising all edges not in  $F$ .

**Theorem 1.6.** *Every graph  $G$  contains a pseudoforest  $F \subseteq G$  such that  $m_2(G - F) \leq m(G)$ .*

Equivalently, the conclusion of Theorem 1.6 can be phrased as saying that  $G$  can be edge-partitioned into a pseudoforest and a subgraph of bounded maximal 2-density. We remark that many results along the lines of Theorem 1.6 are known in the literature, starting with Hakimi's theorem [11] from 1965, which states that  $G$  can be decomposed into  $k$  pseudoforests if and only if  $m(G) \leq k$ . More recently, several papers (e.g. [4, 8, 24]) have proved strengthenings of Hakimi's theorem under the assumption that  $m(G) \leq m$ , where  $m$  is not necessarily an integer. The novelty in Theorem 1.6 (which does not follow from any of the results mentioned above) is that the condition we guarantee about  $G - F$  is that it is sparse with respect to  $m_2$ , which is a *different* density measure from the one we started with. We remark that Theorem 1.6 is a slight weakening of a conjecture of Kuperwasser, Samotij, and Wigderson [19, Conjecture 1.5]; see Section 7 for details.

On its own, Theorem 1.6 already suffices to prove Conjecture 1.3 in almost all cases, namely for all tuples  $(H_1, \dots, H_r)$  where  $H_2$  contains a strictly 2-balanced subgraph  $H'_2$  with  $m_2(H_2) = m_2(H'_2)$  such that  $H'_2$  is not a cycle; this includes almost all cases that were known before. However, Theorem 1.6 cannot be used to resolve the remaining cases, so to prove Theorem 1.5 we need another decomposition result. To state it, let us define the *maximal  $\frac{4}{3}$ -density* of  $G$  to be

$$m_{\frac{4}{3}}(G) := \max \left\{ \frac{e(J)}{v(J) - \frac{4}{3}} : J \subseteq G, v(J) \geq 2 \right\}.$$

**Theorem 1.7.** *Let  $m > \frac{3}{2}$  be a real number. Every graph  $G$  with  $m(G) \leq m$  contains a forest  $F \subseteq G$  such that  $m_{\frac{4}{3}}(G - F) < m$ .*

We remark that the choice of  $\frac{4}{3}$  is somewhat arbitrary, and our techniques can prove similar results where one replaces  $\frac{4}{3}$  by another real number. However, working with the maximal  $\frac{4}{3}$ -density is useful for combining Theorems 1.6 and 1.7 and proving Theorem 1.5.

The rest of this paper is organized as follows. In the next section, we show how to deduce Theorem 1.5 from Theorems 1.6 and 1.7, and sketch the strategy we employ in the proofs

of both Theorems 1.6 and 1.7. We prove various general lemmas common to both proofs in Section 3, then prove Theorem 1.6 in Section 4. The proof of Theorem 1.7 is more involved, and is split into Section 5, where we state and prove a key technical lemma, and Section 6, where we complete the proof. We end in Section 7 with some concluding remarks and open problems.

## 2. PROOF OVERVIEW

**2.1. Proof of Theorem 1.5 assuming Theorems 1.6 and 1.7.** In addition to the new ingredients, Theorems 1.6 and 1.7, we will need three additional lemmas, all of which are well-known in the literature. The first is Nash-Williams' theorem [27], which was mentioned above. For its statement, let us define the *maximal 1-density* (also known as the *fractional arboricity*) of a graph  $G$  to be

$$m_1(G) := \max \left\{ \frac{e(J)}{v(J) - 1} : J \subseteq G, v(J) \geq 2 \right\}.$$

**Lemma 2.1** (Nash-Williams). *Let  $k \geq 1$  be an integer. A graph  $G$  can be edge-partitioned into  $k$  forests if and only if  $m_1(G) \leq k$ .*

The second lemma we need is a simple inequality relating  $m(G)$  and  $m_1(G)$ ; its short proof can be found, for example, in [19, Lemma 5.4].

**Lemma 2.2.** *For any graph  $G$ , we have  $m_1(G) \leq m(G) + \frac{1}{2}$ .*

Finally, we quote a simple result proved independently in [19, Proposition 5.2(a)] and [1, Lemma 1.9(ii)], which establishes Theorem 1.5 whenever  $H_2$  is not bipartite.

**Lemma 2.3.** *Let  $H_1, H_2$  be graphs with  $m_2(H_1) > m_2(H_2) > 1$ . If  $H_2$  is strictly 2-balanced and not bipartite, then any graph  $G$  which is Ramsey for  $(H_1, H_2)$  satisfies  $m(G) > m_2(H_1, H_2)$ .*

With these preliminaries, we are ready to prove Theorem 1.5.

*Proof of Theorem 1.5.* Fix a pair of graphs  $H_1, H_2$  with  $m_2(H_1) > m_2(H_2) > 1$ . Recall that this implies that  $m_2(H_1) > m_2(H_1, H_2) > m_2(H_2)$ . We may assume without loss of generality that  $H_2$  is strictly 2-balanced; indeed, if it is not, we pass to a strictly 2-balanced subgraph  $H'_2 \subseteq H_2$  with  $m_2(H'_2) = m_2(H_2)$ , and observe that any  $H'_2$ -free graph is necessarily  $H_2$ -free. Thus, we henceforth assume that  $H_2$  is strictly 2-balanced. This in particular implies that  $H_2$  is connected and that the minimum degree of  $H_2$  is at least 2. (In fact, any strictly 2-balanced graph is 2-connected as shown in [28, Lemma 3.3].)

Let us fix a graph  $G$  with  $m(G) \leq m_2(H_1, H_2)$ . Our goal is to edge-partition  $G$  into an  $H_1$ -free graph and an  $H_2$ -free graph. We are done by Lemma 2.3 if  $H_2$  is not bipartite, so let us assume that  $H_2$  is bipartite.

If  $m(G) \leq \frac{3}{2}$ , then  $m_1(G) \leq 2$  by Lemma 2.2, and therefore Lemma 2.1 implies that  $G$  can be edge-partitioned into two forests. As  $m_2(H_1) > m_2(H_2) > 1$ , neither  $H_1$  nor  $H_2$  is a forest, and thus this is an edge-partition of  $G$  into an  $H_1$ -free graph and an  $H_2$ -free graph.

Now suppose that  $H_2$  contains at least two cycles. As  $H_2$  is connected, this implies that  $H_2$  is not contained in any pseudoforest. We now apply Theorem 1.6 to find a pseudoforest  $F \subseteq G$  with  $m_2(G - F) \leq m(G) \leq m_2(H_1, H_2) < m_2(H_1)$ . By the above,  $F$  is  $H_2$ -free. Moreover, as  $m_2(G - F) < m_2(H_1)$ , we see that  $G - F$  is  $H_1$ -free. This is our desired decomposition.

Finally, we may assume that  $m(G) > \frac{3}{2}$ , and that  $H_2$  is connected, bipartite, and contains only one cycle. As the minimum degree of  $H_2$  is at least 2, we find that  $H_2 = C_\ell$  is a cycle, for some even  $\ell \geq 4$ . As  $m_2(C_\ell) = (\ell - 1)/(\ell - 2)$ , we conclude that  $m_2(H_2) \leq m_2(C_4) = \frac{3}{2}$ .

Therefore,

$$\begin{aligned}
m_2(H_1, H_2) &= \max_{J \subseteq H_1} \frac{e(J)}{v(J) - 2 + 1/m_2(H_2)} \\
&\leq \max_{J \subseteq H_1} \frac{e(J)}{v(J) - 2 + 1/(\frac{3}{2})} \\
&= \max_{J \subseteq H_1} \frac{e(J)}{v(J) - \frac{4}{3}} = m_{\frac{4}{3}}(H_1).
\end{aligned}$$

We now apply Theorem 1.7 with  $m = m(G) > \frac{3}{2}$ . We find a forest  $F \subseteq G$  with  $m_{\frac{4}{3}}(G - F) < m$ . As  $H_2$  is not a forest, we see that  $F$  is  $H_2$ -free, whereas the above implies that  $m_{\frac{4}{3}}(G - F) < m \leq m_2(H_1, H_2) \leq m_{\frac{4}{3}}(H_1)$ , and thus  $G - F$  is  $H_1$ -free. This is the desired decomposition, and the proof is complete.  $\square$

**2.2. Proof strategy and sketch.** We now sketch the strategy we employ to prove Theorems 1.6 and 1.7. The same general strategy is used for both; we begin with Theorem 1.6, which is somewhat simpler both conceptually and technically.

So let us fix some real number  $m > 0$  and some graph  $G$  with  $m(G) \leq m$ . Our goal is to define a pseudoforest  $F \subseteq G$ , in such a way that we can control  $m_2(G - F)$ . Our main tool to do this is the concept of *allocations*, which are a fractional version of an edge-orientation of  $G$ .

**Definition 2.4.** Let  $G$  be a graph and  $m \geq 0$  a real number. An  $m$ -allocation of  $G$  is a mapping  $\theta : V(G)^2 \rightarrow [0, 1]$  such that the following hold.

- (1) If  $u, v \in V(G)$  and  $uv \notin E(G)$ , then  $\theta(u, v) = \theta(v, u) = 0$ .
- (2)  $\theta(u, v) + \theta(v, u) = 1$  for every edge  $uv \in E(G)$ .
- (3) For every vertex  $u \in V(G)$ , we have  $\sum_{v \in V(G)} \theta(u, v) \leq m$ .

Note that if  $\theta$  is integer-valued, it can be viewed as an edge-orientation of  $G$  (namely we orient an edge  $uv$  according to whether  $\theta(u, v)$  is 0 or 1). If  $\theta$  is integer-valued, then condition (3) above corresponds to saying that every vertex has outdegree at most  $m$ . In general, allocations can be thought of as fractional orientations, with an upper bound on the “fractional outdegree” of every vertex. The reason we care about allocations is that the existence of an  $m$ -allocation of  $G$  is equivalent to the statement  $m(G) \leq m$ , as stated in the next theorem.

**Theorem 2.5.** Let  $G$  be a graph and  $m > 0$  a real number. Then  $m(G) \leq m$  if and only if  $G$  admits an  $m$ -allocation.

Results of this type go back at least to work of Hakimi [11] and of Frank and Gyárfás [5], who proved Theorem 2.5 under the extra assumption that  $m$  is an integer (in which case they can get a true orientation, rather than an allocation). Their result actually implies Theorem 2.5 by passing to an appropriate auxiliary multigraph; one can also deduce Theorem 2.5 from the results of [10]. We present an alternative proof, which is closely modeled on the proof of [26, Lemma 2], in Section 3.1. For the moment we continue with the proof overview.

As we are given a graph  $G$  with  $m(G) \leq m$ , we apply Theorem 2.5 to find an  $m$ -allocation  $\theta : V(G)^2 \rightarrow [0, 1]$ . We say that an edge  $e = uv \in E(G)$  is *integral* if  $\theta(u, v), \theta(v, u) \in \{0, 1\}$ . Otherwise, if  $\theta(u, v), \theta(v, u) \in (0, 1)$ , then we say that the edge  $uv$  is *fractional*. We denote by  $E_{\text{int}}(G, \theta)$  and  $E_{\text{frac}}(G, \theta)$  the set of integer and fractional edges of  $G$ , respectively. Moreover, we can naturally associate with  $\theta$  a digraph  $D_\theta = (V(G), A_\theta)$ , where  $A_\theta := \{(u, v) : \theta(u, v) = 1\}$ . Clearly,  $D_\theta$  is an orientation of the integral edges in  $G$ . Similarly, we denote by  $G_\theta^{\text{frac}}$  the subgraph of  $G$  consisting of the fractional edges  $E_{\text{frac}}(G, \theta)$ .

There are a number of elementary operations one can perform on an  $m$ -allocation to get another  $m$ -allocation. By doing such operations, we may assume that  $\theta$  satisfies certain additional properties; for example, we can ensure that  $G_\theta^{\text{frac}}$  is a forest (see Lemma 3.4).

Having chosen  $\theta$  so that it satisfies these extra properties, we are ready to define the pseudoforest  $F$  that we need to output in order to prove Theorem 1.6. To do so, for every vertex

$u \in V(G)$  that is the source of at least one arc in  $D_\theta$ , we pick one such arc  $(u, v)$  arbitrarily, and define  $\vec{F}$  to consist of all chosen arcs. Thus,  $\vec{F}$  is a digraph with maximum outdegree 1, so its underlying graph  $F \subseteq G$  is a pseudoforest.

Having defined  $F$ , it remains to prove the main claim in Theorem 1.6, namely that  $m_2(G - F) \leq m$ . Equivalently, we wish to prove that for every  $U \subseteq V(G)$  with  $|U| \geq 3$ , we have  $e_{G-F}(U) - 1 \leq m(|U| - 2)$ , where we use the notation  $e_{G-F}(U) := e((G - F)[U])$ . So we assume for contradiction that this is not the case, and proceed to analyze the structure of a violating set  $U$ . The proof is somewhat involved and we defer the details to Section 4, but the basic idea is to examine how  $U$  interacts with the digraph  $D_\theta$ , and to use both the definition of  $F$  and the extra properties we imposed on the  $m$ -allocation  $\theta$ .

The proof of Theorem 1.7 follows a similar strategy, but is more involved for a few different reasons. The first reason is that in Theorem 1.7, we wish to output a *forest*  $F \subseteq G$ . So we cannot simply pick an out-arc in  $D_\theta$  for every vertex of  $G$ , as this creates cycles. However, by examining the structure of the digraph  $D_\theta$  (specifically, by partitioning it into strongly connected components), we can pick an out-arc for *almost* every vertex, and control the set of vertices with no out-arcs, while ensuring that the subgraph we so define is a forest. Having done this, we again assume for contradiction that  $m_{\frac{4}{3}}(G - F) \geq m$ , and work with a violating set  $U \subseteq V(G)$ , namely a set  $U$  with  $|U| \geq 2$  satisfying  $e_{G-F}(U) \geq m(|U| - \frac{4}{3})$ . The analysis of  $U$  in this proof is substantially more complicated, and we need to split into a number of different cases, depending on the intersection pattern of  $U$  with the strongly connected components of  $D_\theta$ . Most of these cases can be dealt with via rather simple (but tedious) arguments; this is done in Section 5, where we state and prove a technical lemma dealing with most of these cases. For the remaining cases, some more care needs to be taken in the definition of  $\theta$  and the choice of  $F$ ; by imposing some further conditions on  $\theta$  and  $F$ , we are able to dispense with the remaining cases in Section 6.

### 3. SETUP AND GENERAL LEMMAS

In this section, we state and prove various general lemmas that we will use in the proofs of both Theorems 1.6 and 1.7.

**3.1. Proof of Theorem 2.5.** We begin with the proof of Theorem 2.5 (which we deferred in Section 2), following the proof approach of [26, Lemma 2]. As discussed above, there are a number of other ways of proving this result, including a reduction to [5, Theorem 1] or an application of Hall's matching theorem.

*Proof of Theorem 2.5.* Associated to the graph  $G$  and the parameter  $m$ , we define a flow network  $F$  as follows.  $F$  has a single source vertex  $s$  and a single sink vertex  $t$ , as well as two other sets of vertices  $V, E$ . The vertices in  $V$  are identified with  $V(G)$ , and the vertices in  $E$  are identified with  $E(G)$ . There are directed edges  $s \rightarrow V \rightarrow E \rightarrow t$ , defined as follows. For every  $v \in V$ , we place a directed edge of capacity  $m$  from  $s$  to  $v$ . For every  $v \in V$  and  $e \in E$  such that  $v$  is incident to  $e$  in  $G$ , we place an edge of infinite capacity from  $v$  to  $e$ . Finally, for every  $e \in E$ , we place a directed edge of capacity 1 from  $e$  to  $t$ .

The max-flow min-cut theorem implies that the maximum  $s-t$  flow in  $F$  equals the minimum weight of an  $s-t$  cut in  $F$ . As all edges between  $V$  and  $E$  have infinite capacity, they do not participate in any minimum cut. Hence any minimum cut must consist of some edges between  $s$  and  $V$ , and some edges between  $E$  and  $t$ . Fix such a minimum cut, and let  $A \subseteq V, B \subseteq E$  be the set of vertices in  $V, E$ , respectively, incident to a cut edge. The weight of this cut is  $m|A| + |B|$ . For this to be a true  $s-t$  cut in  $F$ , we must have no directed edges from  $V \setminus A$  to  $E \setminus B$ ; by the definition of the edges  $V \rightarrow E$ , this implies that  $E \setminus B$  consists only of edges with both endpoints in  $A$ . Therefore,  $|E \setminus B| \leq e(G[A])$ , and thus  $|B| \geq e(G) - e(G[A])$ . Therefore the weight of the cut is at least  $m|A| + e(G) - e(G[A])$ . Conversely, we see that for any  $A \subseteq V$ , there is a cut of weight exactly  $m|A| + e(G) - e(G[A])$ , obtained by defining  $E \setminus B$  to be the set of edges entirely contained in  $A$ .

Suppose first that  $m(G) \leq m$ . In particular, we find that  $\frac{e(G[A])}{|A|} \leq m$ , implying that the weight of the minimum cut is at least  $e(G)$ . The cut separating  $t$  from the other vertices has weight  $e(G)$ , as there are  $e(G)$  edges of capacity 1 incident to  $t$ , and is thus a minimum cut. Therefore, there exists an  $s-t$  flow in  $F$  of weight exactly  $e(G)$ . In such a flow every vertex in  $E$  must receive weight exactly 1. This means that for any  $e = uv \in E(G)$ , we have a total flow of 1 into the node  $e$ , coming from its neighbours  $u, v \in V$ . Defining  $\theta(u, v)$  to be the flow  $u$  sends to  $e$ , and similarly  $\theta(v, u)$  the flow  $v$  sends to  $e$ , we find that  $\theta(u, v) + \theta(v, u) = 1$ . Moreover, the fact that every  $u \in V$  is incident to an edge of capacity  $m$  from  $s$  shows that this defines a valid  $m$ -allocation.

Conversely, suppose that there is an  $m$ -allocation of  $G$ . As in the previous paragraph, this defines an  $s-t$  flow of weight  $e(G)$  in  $F$ . Thus, every cut must have weight at least  $e(G)$ . Therefore, for any  $A \subseteq V$ , we have that  $m|A| + e(G) - e(G[A]) \geq e(G)$ , implying that  $m(G) \leq m$ .  $\square$

**3.2. Strong components in digraphs.** Let us fix the following terminology. For a directed graph  $D$ , we denote by  $A(D)$  the set of arcs (directed edges) in  $D$ . Given a directed graph  $D$ , a *strong component* of  $D$  is an inclusion-wise maximal subset  $X \subseteq V$  of vertices such that the induced subdigraph  $D[X]$  is strongly connected. It is well-known that the strong components of a digraph  $D$  partition its vertex-set. We call a strong component  $X$  of  $D$  a *terminal component* if there are no arcs in  $D$  that start in a vertex in  $X$  and end in a vertex outside of  $X$ . We will need the following statements about how changing a digraph affects the number and structure of its terminal components.

**Lemma 3.1.**

- (a) Let  $D_1$  and  $D_2$  be digraphs on the same vertex-set such that  $A(D_1) \subseteq A(D_2)$ . Then the number of terminal components of  $D_2$  is at most as large as the number of terminal components of  $D_1$ . Furthermore, if there exist distinct terminal components  $U, W$  of  $D_1$  and a directed path in  $D_2$  that starts in  $U$  and ends in  $W$ , then the number of terminal components of  $D_2$  is strictly smaller than the number of terminal components of  $D_1$ .
- (b) Let  $X$  be a terminal component of a digraph  $D$  and let  $(u, v) \in A(D[X])$ . Then the collection of terminal components of  $D - (u, v)$  is obtained from the collection of terminal components of  $D$  by replacing  $X$  with a subset of it that contains  $u$ .
- (c) Let  $D$  be a digraph and let  $u$  be a vertex of  $D$  that does not belong to any terminal component. If there exists a directed path in  $D$  that starts in  $u$ , ends in a terminal component of  $D$  and does not use the arc  $(u, v)$ , then the digraphs  $D - (u, v)$  and  $D$  have the same terminal components.
- (d) Let  $D$  be a digraph and let  $u$  be a vertex that does not belong to any terminal component. Let  $D + (u, v)$  be the digraph obtained by adding a new arc  $(u, v)$  to  $D$ . Then  $D$  and  $D + (u, v)$  have the same terminal components.
- (e) Let  $D$  be a digraph and  $\{u\}$  a singleton terminal component of  $D$ , and let  $D + (u, v)$  be obtained from  $D$  by adding a new arc  $(u, v)$  to  $D$ . Then one of the following holds.
  - The number of terminal components of  $D + (u, v)$  is smaller than that of  $D$ .
  - The collection of terminal components of  $D + (u, v)$  is obtained from the collection of terminal components of  $D$  by replacing  $\{u\}$  with a terminal component containing  $\{u, v\}$ .

*Proof.*

- (a) For every strong component  $X$  of  $D_1$ , we have that  $D_1[X]$  and thus  $D_2[X]$  is strongly connected, and hence  $X$  is fully contained in one of the strong components of  $D_2$ . This directly implies that every strong component of  $D_2$  is a disjoint union of some strong components of  $D_1$ . Now let  $Y$  be any terminal component of  $D_2$ . Let  $X_1, \dots, X_k$  be the strong components of  $D_1$  such that  $Y = \bigcup_{i=1}^k X_i$ . Consider the digraph  $D_1[Y]$ , whose strong components are  $X_1, \dots, X_k$ . It is well-known (and easy to see) that every digraph contains at least one terminal component. Hence, there exists  $i \in \{1, \dots, k\}$  such that



there are no edges in  $D_1[Y]$  that leave  $X_i$ . But since  $Y$  is a terminal component of  $D_2 \supseteq D_1$ , this means that there are also no edges in the whole of  $D_1$  that leave  $X_i$ , and hence  $X_i$  is a terminal component of  $D_1$  that is contained in  $Y$ .

We have shown that every terminal component of  $D_2$  contains a terminal component of  $D_1$ , and hence the number of terminal components of  $D_2$  is at most the number of terminal components of  $D_1$ .

Let  $P$  be a directed path in  $D_2$  that starts in  $U$  and ends in  $W$ . We claim that either  $U$  and  $W$  are contained in the same terminal component of  $D_2$ , or at least one of  $U, W$  is not contained in any terminal component of  $D_2$ . In both cases, we can immediately conclude that there must be strictly more terminal components in  $D_1$  than in  $D_2$ , as desired.

Suppose that contrary to the above claim, there exist two distinct terminal components  $Y_1, Y_2$  of  $D_2$  such that  $U \subseteq Y_1$  and  $W \subseteq Y_2$ . Since  $P$  starts in  $U$  and ends in  $W$ , this implies that there exists an arc in  $D_2$  leaving  $Y_1$ , a contradiction to it being a terminal component. This concludes the proof.

- (b) Let  $Y$  be a terminal component of  $D - (u, v)$ . If  $u \notin Y$ , then  $Y$  is a terminal component of  $D$  that is disjoint from  $X$ . So suppose  $u \in Y$ . Since every other vertex in  $Y$  is reachable via a directed path from  $u$  in  $D - (u, v)$ , and since no arc in  $D - (u, v)$  leaves  $Y$  and no arc in  $D$  leaves  $X$ , it follows that  $Y$  is equal to the set of vertices reachable in  $D - (u, v)$  via a directed path starting in  $u$ , and that  $Y \subseteq X$ . In fact, it can be easily checked that the latter set of vertices always induces a strongly connected subdigraph of  $D - (u, v)$  (since every vertex in  $X$  can reach  $u$  via a directed path in  $D - (u, v)$ ). Hence, we have shown that the set of vertices reachable from  $u$  in  $D - (u, v)$  forms the unique terminal component of  $D - (u, v)$  contained in  $X$ .

In the other direction, note that any terminal component of  $D$  distinct from  $X$  is also disjoint from  $X$  and hence remains a terminal component also in  $D - (u, v)$ . This shows that the terminal components of  $D$  and  $D - (u, v)$  are the same, apart from the replacement of  $X$  by a subset containing  $u$ .

- (c) Let  $X$  be a terminal component of  $D - (u, v)$ . Unless  $u \in X$  and  $v \notin X$ , also in  $D$  there are no arcs leaving  $X$  and thus  $X$  forms a terminal component also of  $D$ . Now suppose  $u \in X$  and  $v \notin X$ . Then  $(u, v)$  is the only arc in  $D$  leaving  $X$ . Let  $P$  be a directed path in  $D - (u, v)$  starting in  $u$  and ending in a terminal component  $Y$  of  $D$ . Then, since  $P$  cannot leave  $X$ , we have  $V(P) \subseteq X$  and thus  $X \cap Y \neq \emptyset$ . Since  $D[X]$  is strongly connected and  $Y$  is a strong component of  $D$ , by definition, we have  $X \subseteq Y$ . In particular, we have  $u \in Y$ , which contradicts our assumption that  $u$  is not part of any terminal component of  $D$ . This contradiction shows that our assumption above was wrong, and hence  $X$  is indeed a terminal component also of  $D$ .

In the other direction, suppose  $X$  is some terminal component of  $D$ . We assumed that  $u$  is not contained in any terminal component of  $D$ , hence  $u \notin X$ . But then necessarily  $(u, v) \notin A(D[X])$ , and thus  $X$  is also a terminal component of  $D - (u, v)$ .

- (d) Let  $X$  be a terminal component of  $D$ . Then by assumption  $u \notin X$ , and hence also in  $D + (u, v)$  no arc leaves  $X$ . Hence,  $X$  is also a terminal component of  $D + (u, v)$ . In the other direction, suppose that  $X$  is a terminal component of  $D + (u, v)$ . If  $(u, v) \notin A(D[X])$ , then clearly  $X$  is also a terminal component of  $D$ . So now suppose that  $(u, v) \in A(D[X])$ . We can now apply item (b) of this lemma to find that there exists a terminal component of  $D + (u, v) - (u, v) = D$  contained in  $X$  that contains  $u$ . This is a contradiction, since we assumed that  $u$  is not part of any terminal component of  $D$ . Hence our assumption was false, we indeed have  $(u, v) \notin A(D[X])$ , and hence  $X$  is also a terminal component of  $D$ .

- (e) Suppose that the number of terminal components of  $D + (u, v)$  is not smaller than that of  $D$ . Together with the fact that every strongly connected (terminal) component of  $D + (u, v)$ , which does not contain  $u$ , is also a strongly connected (terminal) component of  $D$ , we get that the collection of terminal components of  $D + (u, v)$  is obtained from

the collection of terminal components of  $D$  by replacing  $\{u\}$  with some  $X \supseteq \{u\}$ . By definition of a terminal component, we find that  $X$  contains  $v$  as well.  $\square$

**3.3. Properties of and operations on allocations.** A natural way of changing a given  $m$ -allocation  $\theta$  of a graph to another is to “shift” weights along a cycle. As this operation will be used in several places of our proof, we isolate the elementary facts about it in the following statement.

**Observation 3.2.** *Let  $G$  be a graph, let  $m > 0$  be a real number, and let  $\theta : V(G)^2 \rightarrow \mathbb{R}_+$  be an  $m$ -allocation on  $G$ . Let  $C$  be a cycle in  $G$ , and let  $u_0, u_1, u_2, \dots, u_k = u_0$  for some  $k \geq 3$  be the cyclic sequence of vertices along  $C$ . Let*

$$\varepsilon := \min\{\theta(u_{i-1}, u_i) : 1 \leq i \leq k\}.$$

*Suppose that  $\varepsilon > 0$ , and define a new mapping  $\theta' : V(G)^2 \rightarrow [0, 1]$  by setting*

$$\theta'(x, y) := \begin{cases} \theta(x, y) - \varepsilon, & \text{if } \exists i \in \{1, \dots, k\} \text{ such that } (x, y) = (u_{i-1}, u_i), \\ \theta(x, y) + \varepsilon, & \text{if } \exists i \in \{1, \dots, k\} \text{ such that } (x, y) = (u_i, u_{i-1}), \\ \theta(x, y), & \text{otherwise.} \end{cases}$$

*for every  $x, y \in V(G)$ . Then  $\theta'$  is still an  $m$ -allocation of  $G$ . Furthermore,  $D_{\theta'}$  is obtained from  $D_\theta$  by removing the arcs*

$$\{(u_{i-1}, u_i) : 1 \leq i \leq k, \theta(u_{i-1}, u_i) = 1\}$$

*and adding the arcs*

$$\{(u_i, u_{i-1}) : 1 \leq i \leq k, \theta(u_{i-1}, u_i) = \varepsilon\}.$$

Given a graph  $G$ , a cycle  $C$  equipped with a cyclic orientation and an  $m$ -allocation  $\theta$  on  $G$ , the above forms a well-defined way to generate a new  $m$ -allocation  $\theta'$  on  $G$ , which we from now on will refer to as the  $m$ -allocation obtained from  $\theta$  by shifting along  $C$ .

In our proofs, we will work with special allocations which are chosen to have certain extra properties. The following definition captures the way we choose these special allocations

**Definition 3.3.** Given a real number  $m$  and a graph  $G$  for which there exists an  $m$ -allocation, we say that an  $m$ -allocation is *optimal* if it is chosen such that:

- (1) The number of terminal components of  $D_\theta$  is minimum among all possible choices of  $m$ -allocations  $\theta$ , and
- (2) subject to point (1), the number of fractional edges is minimized (equivalently, the number of integral edges is maximized), and
- (3) subject to points (1) and (2), the total number of vertices contained in terminal components is maximized, and
- (4) subject to points (1), (2) and (3), the number of arcs in  $D_\theta$  that end in a vertex forming a singleton terminal component of  $D_\theta$  is maximized.

The following lemmas establish some properties that follow from this choice of  $\theta$ . Recall that  $G_\theta^{\text{frac}}$  is the spanning subgraph of  $G$  comprising all the fractional edges.

**Lemma 3.4.** *If  $\theta$  is an optimal  $m$ -allocation, then  $G_\theta^{\text{frac}}$  is a forest.*

*Proof.* Suppose for a contradiction that there exists a cycle  $C$  in  $G_\theta^{\text{frac}}$ , and let  $u_0, u_1, \dots, u_k = u_0$  be its cyclic sequence of vertices. Let  $\theta'$  be the  $m$ -allocation obtained from  $\theta$  by shifting along  $C$ . By Observation 3.2,  $\theta'$  is still an  $m$ -allocation on  $G$ . Let  $i_0 \in \{1, \dots, k\}$  be an index that minimizes  $\theta(u_{i_0-1}, u_{i_0})$ . By Observation 3.2 and since all edges along  $C$  are fractional in  $\theta$ , we have  $A(D_{\theta'}) \supseteq A(D_\theta) \cup \{(u_{i_0}, u_{i_0-1})\}$ . This implies by Lemma 3.1(a) that the number of terminal components in  $D_{\theta'}$  is at most as large as the number of terminal components in  $D_\theta$ . Together with the fact that  $|E_{\text{frac}}(G, \theta')| < |E_{\text{frac}}(G, \theta)|$  (since  $A(D_{\theta'}) \supsetneq A(D_\theta)$ ) this contradicts the optimality of  $\theta$ . Hence our assumption was wrong, and the graph  $G_\theta^{\text{frac}}$  is indeed a forest.  $\square$

**Lemma 3.5.** *Let  $\theta$  be an optimal  $m$ -allocation, let  $v, w$  be vertices of a terminal component  $C$  of  $D_\theta$  such that  $(v, w) \in A(D_\theta)$ , and let  $P = (w = u_0, \dots, u_\ell = v)$  be a path in  $G_\theta^{\text{frac}}$  such that all the inner vertices are in some terminal component of  $D_\theta$  different from  $C$ . Then there exists  $0 \leq i \leq \ell - 1$  such that  $u_i$  and  $u_{i+1}$  lie in the same terminal component of  $D_\theta$ .*

*In particular, if  $U$  and  $W$  are distinct terminal components of  $D_\theta$ , then for every  $u \in U$  the set  $N_G(u) \cap W$  is independent in  $G$ .*

*Proof.* Suppose for contradiction that for all  $0 \leq i \leq \ell - 1$ , it holds that  $u_i$  and  $u_{i+1}$  are in different terminal components of  $D_\theta$ . Let  $\theta'$  be the  $m$ -allocation on  $G$  that is obtained from  $\theta$  by shifting along the cycle obtained by adding  $(v, w)$  to  $P$ . By Observation 3.2 and since by assumption  $\theta(v, w) = 1$ ,  $A(D_{\theta'})$  is obtained from  $A(D_\theta)$  by removing the arc  $(v, w)$  and adding at least one of the arcs  $(u_i, u_{i-1})$  for some  $1 \leq i \leq \ell$ . Let  $1 \leq j \leq \ell$  be such that  $(u_j, u_{j-1})$  is added. Let  $D' = D_\theta - (v, w)$ . By Lemma 3.1(b), it follows that the terminal components of  $D'$  are the same as the terminal components of  $D_\theta$ , except that  $C$  is replaced by some  $X \subseteq C$  with  $v \in X$ . Observe that there exists a path  $Q$  from  $w$  to  $v$  in  $D'$  since  $D_\theta[C]$  is strongly connected and  $(v, w)$  is not part of any path from  $w$  to  $v$ .

Recall that by assumption, if  $2 \leq j \leq \ell - 1$  then  $u_j$  and  $u_{j-1}$  are in different terminal components of  $D_\theta$ , and these terminal components are distinct from  $C$ . Hence they also lie in distinct terminal components of  $D'$ . If  $j = \ell$ , then  $u_j = u_\ell = v$  is contained in the terminal component  $X \subseteq C$  of  $D'$ , and  $u_{j-1} = u_{\ell-1}$  is in a terminal component of  $D_\theta$  distinct from  $C$ . Hence,  $u_j, u_{j-1}$  are also in different terminal components in this case. Finally, if  $j = 1$ , then adding the arc  $(u_1, u_0 = w)$  introduces a path (namely  $(u_1, w) + Q$ ) from  $u_1$  to  $v$  in  $D'$ , and  $u_1$  and  $v$  are also in different terminal components of  $D'$ . In each case, we conclude by Lemma 3.1(a) that  $D' + (u_j, u_{j-1})$  and thus also  $D_{\theta'} \supseteq D' + (u_j, u_{j-1})$  has strictly fewer terminal components than  $D_\theta$ , contradicting the optimality of  $\theta$ .

For the second statement, note first that all edges between  $u$  and  $W$  must be fractional, as both  $U$  and  $W$  are terminal components of  $D_\theta$ . Therefore,  $N_G(u) \cap W$  cannot contain any fractional edges, since this would produce a triangle in  $G_\theta^{\text{frac}}$ , contradicting Lemma 3.4. On the other hand, if  $N_G(u) \cap W$  contains an integral edge  $(v, w) \in A(D_\theta)$ , we obtain a contradiction to the first statement of this lemma by setting  $P = (w, u, v)$ .  $\square$

**3.4. Spines.** In order to define the pseudoforest and forest guaranteed in Theorems 1.6 and 1.7, respectively, we use the notion of a *spine* of a digraph, which we now define.

**Definition 3.6.** Let  $D$  be a digraph. A spanning subdigraph  $H \subseteq D$  is called a *spine* of  $D$  if the following hold.

- Every vertex in a non-terminal component of  $D$  has out-degree exactly 1 in  $H$ .
- In every terminal component of  $D$ , all vertices have out-degree 1 in  $H$ , apart from at most one vertex of out-degree 0.
- The underlying graph of  $H$  is a forest.

**Lemma 3.7.** *Let  $D$  be a digraph and  $R \subseteq V(D)$  a set of vertices containing exactly one vertex from each terminal component of  $D$ . Then  $D$  has a spine  $H$  such that every vertex in  $V(D) \setminus R$  has out-degree exactly one in  $H$ .*

*Proof.* Let  $X$  be any strong component of  $D$ . Then  $D[X]$  is strongly connected, and hence for every vertex  $x \in X$  it contains an in-tree with root vertex  $x$  that contains all the vertices in  $X$ .

To build  $H$ , do the following. For every non-terminal component  $X$ , pick an arc  $(x, y) \in A(D)$  such that  $x \in X, y \notin X$  and a spanning in-tree  $T_X$  of  $D[X]$  rooted at  $x$ . Then add all the arcs of  $T_X$  as well as the arc  $(x, y)$  to the arc-set of  $H$ . Next, for every terminal component  $X$ , pick the unique vertex  $r \in R \cap X$  and an in-tree rooted at  $r$  that spans  $D[X]$  and add all the arcs of this tree to  $H$ .

At the end of this process,  $H$  defines a spanning subdigraph of  $D$  in which every vertex has exactly one out-arc, except for one vertex in each terminal component (namely, the vertices in  $R$ ). Finally, suppose towards a contradiction that the underlying graph  $F$  of  $H$  contains a

cycle. Then since  $H$  has maximum out-degree 1, this cycle would have to be directed in  $H$ , and thus it would be fully contained in a strong component of  $D$ . However, it follows directly by definition of  $H$  that  $F$  restricted to any strong component of  $D_\theta$  is a tree, yielding the desired contradiction. Thus,  $F$  is indeed a forest, as desired.  $\square$

**Lemma 3.8.** *Let  $\theta$  be an optimal  $m$ -allocation, and let  $H$  be a digraph containing a spine of  $D_\theta$ . Let further  $u, v, w$  be vertices such that  $\{u\}$  is a singleton terminal component of  $D_\theta$ ,  $uv, uw \in E(G_\theta^{\text{frac}})$ , and  $(v, w) \in A(D_\theta) \setminus A(H)$ . Then  $\{w\}$  is also a singleton terminal component of  $D_\theta$ .*

*Proof.* Towards a contradiction, suppose that  $\{w\}$  is not a singleton terminal component of  $D_\theta$ .

Let  $\theta'$  be the  $m$ -allocation on  $G$  that is obtained from  $\theta$  by shifting along the triangle  $u, v, w$  (in this cyclic order). By Observation 3.2,  $A(D_{\theta'})$  is obtained from  $A(D_\theta)$  by removing the arc  $(v, w)$  and adding one (or both) of the arcs  $(v, u), (u, w)$ .

If  $v$  is contained in any terminal component of  $D_\theta$ , then the same terminal component of  $D_\theta$  also must contain  $w$ , and hence we are in the setting of Lemma 3.5, which yields a contradiction to this case (note that since  $u$  is contained in a singleton terminal component, it is necessarily contained in a distinct terminal component from  $v, w$ ).

So suppose in the following that  $v$  is not contained in any terminal component of  $D_\theta$ . Let  $P$  be the directed path in the spine of  $D_\theta$  contained in  $H$  that one obtains by starting at  $v$  and repeatedly following the unique out-arc in the spine until one reaches a terminal component of  $D_\theta$ . Since we assumed  $(v, w) \notin A(H)$ , the path  $P$  does not use  $(v, w)$ , and so we can apply Lemma 3.1(c) to find that the digraphs  $D_\theta - (v, w)$  and  $D_\theta$  have the same terminal components. Since  $D_{\theta'} \supseteq D_\theta - (v, w)$ , by Lemma 3.1(a), we find that  $D_{\theta'}$  has at most as many terminal components as  $D_\theta$ . Also, since it is obtained from  $D_\theta - (v, w)$  by adding at least one of  $(v, u), (u, w)$ , it has at least as many arcs as  $D_\theta$ . If both of the arcs are added, then  $\theta'$  induces strictly more integral edges than  $\theta$ , which contradicts the optimality of  $\theta$ . Thus, in the following suppose that exactly one of the arcs  $(u, w), (v, u)$  is added.

For a first case, suppose that  $(u, w) \in A(D_{\theta'})$ . It then follows from Lemma 3.1(e) that one of the following holds.

- $D_{\theta'}$  has strictly fewer terminal components than  $D_\theta - (v, w)$  (and thus than  $D_\theta$ ).
- In  $D_{\theta'}$  we have exactly as many terminal components as in  $D_\theta - (v, w)$ , but strictly more vertices that are contained in terminal components than in  $D_\theta - (v, w)$  (and thus than in  $D_\theta$ ).

In both cases, this yields a contradiction to the optimality of  $\theta$ .

For the second case, suppose that  $(v, u) \in A(D_{\theta'})$  and  $(u, w) \notin A(D_{\theta'})$ . By what was said above and Lemma 3.1(a), we have that  $D_{\theta'}$  has at most as many terminal components as  $D_\theta$ , and at least as many integral edges as  $D_\theta$ . Furthermore, since we assumed that  $v$  is not part of any terminal component of  $D_\theta$ , it is also not part of any terminal component of  $D_\theta - (v, w)$ , since we found above that these two digraphs have the same terminal components. Hence, Lemma 3.1(d) implies that also the digraph  $D_{\theta'} = (D_\theta - (v, w)) + (v, u)$  has the same terminal components as  $D_\theta$ , and hence the number of vertices that are in terminal components is the same in  $D_\theta$  and  $D_{\theta'}$ . However, we claim that the number of edges of  $D_{\theta'}$  ending in singleton terminal components is strictly larger than in  $D_\theta$ , which will yield the desired contradiction to the optimality of  $\theta$ . To see this, note that by what was said above the set of singleton terminal components of  $D_\theta$  and  $D'_\theta$  is identical. However, adding the edge  $(v, u)$  increases the in-degree of the singleton terminal component  $u$ , and no in-degree of any other singleton terminal component is reduced, since we assumed that  $w$  is not a singleton terminal component. This contradiction concludes the proof.  $\square$

#### 4. PROOF OF THEOREM 1.6

Given these preliminaries, we are now ready to prove Theorem 1.6.

*Proof of Theorem 1.6.* Let  $m = m(G)$  and, by Theorem 2.5, let  $\theta$  be an optimal  $m$ -allocation of  $G$ . Let  $H$  be a spine of  $D_\theta$ , which exists by Lemma 3.7. Let  $H'$  be a digraph obtained by adding an arbitrary out-arc to every vertex of  $H$  which does not yet have an out-arc in  $H$  but does have at least one out-arc in  $D_\theta$ . Finally, let  $F$  denote the underlying graph of  $H'$ , and note that  $F$  is a pseudoforest (since it admits an orientation of maximum out-degree 1). Note that the only vertices of  $H'$  which do not have an out-arc are the vertices that form singleton terminal components of  $D_\theta$ .

We now claim that  $m_2(G - F) \leq m$ . Towards a contradiction, suppose that  $m_2(G - F) > m$ . Then there exists a set  $U \subseteq V(G)$  such that  $|U| \geq 3$  and

$$(1) \quad m(|U| - 2) + 1 < e_{G-F}(U).$$

We make the following observation.

**Claim 4.1.**  $|U| > 2m - 1$ .

*Proof.* Suppose  $|U| \leq 2m - 1$ . Together with (1), we get

$$\binom{|U|}{2} \geq e_{G-F}(U) > m(|U| - 2) + 1 \geq \frac{|U| + 1}{2}(|U| - 2) + 1.$$

But  $|U|(|U| - 1) = (|U| + 1)(|U| - 2) + 2$ , a contradiction.  $\square$

Let  $d$  denote the number of terminal components of  $D_\theta$  that intersect  $U$  in exactly one vertex.

**Claim 4.2.**  $d \leq 1$ .

*Proof.* As  $G_\theta^{\text{frac}}$  forms a forest by Lemma 3.4, there are at most  $|U| - 1$  fractional edges in  $G[U]$ . For every vertex  $u \in U$  such that there exists a terminal component  $W$  of  $D_\theta$  satisfying  $W \cap U = \{u\}$ , there are no arcs in  $D_\theta$  from  $u$  to  $U$ . Note that there are exactly  $d$  such vertices in  $U$ , by definition of  $d$ . Furthermore, every other vertex in  $U$  (for which such a terminal component does not exist) has out-degree at most  $m - 1$  in  $D_\theta - H'$ , since each such vertex has at most  $\lfloor m \rfloor \leq m$  out-arcs in  $D_\theta$  by definition of  $\theta$ , and exactly one out-arc is included in  $H'$ . Bounding the number of integral edges in  $(G - F)[U]$  by the sum over the out-degrees of vertices in  $(D_\theta - H')[U]$ , we get

$$e_{G-F}(U) \leq (|U| - 1) + (m - 1)(|U| - d).$$

Comparing this with (1), we obtain

$$-2m + 2 < -(m - 1)d,$$

implying that  $d < 2$ .  $\square$

Since for every edge  $uv \in E(G)$  it holds that  $\theta(u, v) + \theta(v, u) = 1$ , we find that

$$(2) \quad e_{G-F}(U) = \sum_{uv \in E((G-F)[U])} (\theta(u, v) + \theta(v, u)) = \sum_{u \in U} \sum_{v \in N_{G-F}(u) \cap U} \theta(u, v).$$

Let  $X$  be the set of vertices that form singleton terminal components of  $D_\theta$ . Then, by definition of an  $m$ -allocation and since every vertex outside of  $X$  has out-degree 1 in  $H'$ , we find that for every  $u \in V(G) \setminus X$ , we have  $\sum_{v \in N_{G-F}(u) \cap U} \theta(u, v) \leq m - 1$ , whereas for every  $u \in X$ , we have  $\sum_{v \in N_{G-F}(u) \cap U} \theta(u, v) \leq m$ .

For now, suppose that for all  $u \in U$ , we have  $\sum_{v \in N_{G-F}(u) \cap U} \theta(u, v) \leq m - 1$ . Then (2) together with (1) gives

$$m(|U| - 2) + 1 < (m - 1)|U|.$$

Simplifying, we have  $|U| < 2m - 1$ , contradicting Claim 4.1. Therefore, there exists at least one vertex  $u \in U$  with  $\sum_{v \in N_{G-F}(u) \cap U} \theta(u, v) > m - 1$ . Fix such a vertex and denote it by  $u^*$ . Observe that  $u^* \in U \cap X$ , which implies by Claim 4.2 that  $d = 1$ . In particular, we find that  $u^*$  is the unique vertex in  $U$  satisfying  $\sum_{v \in N_{G-F}(u) \cap U} \theta(u, v) > m - 1$ . Note that  $u^*$  does not have any out-arcs in  $D_\theta$ . We conclude that  $u^*$  is incident to at least  $\lfloor m \rfloor$  edges in  $G_\theta^{\text{frac}}[U]$ . Picking

$k \in \mathbb{N}$  and  $0 \leq \varepsilon < 1$  such that  $m = k + \varepsilon$ , we obtain that  $u^*$  is incident to at least  $k$  edges in  $G_\theta^{\text{frac}}[U]$ .

**Claim 4.3.**  $2m - 1 < |U| < 2m$ .

*Proof.* The lower bound follows from Claim 4.1. Again combining (1) and (2), we get

$$m(|U| - 2) + 1 < e_{G-F}(U) \leq (m - 1)|U| + 1,$$

where we use the fact that  $u^*$  is the only vertex in  $U$  satisfying  $\sum_{v \in N_{G-F}(u) \cap U} \theta(u, v) > m - 1$ . The upper bound follows directly.  $\square$

**Claim 4.4.**  $\varepsilon < 1/2$ .

*Proof.* Suppose towards a contradiction that  $\varepsilon \geq 1/2$ . Since by Claim 4.3,  $2m - 1 < |U| < 2m$ , we get that  $\varepsilon > 1/2$  and  $|U| = 2k + 1$ . It follows that

$$\left(k + \frac{1}{2}\right)(2k - 1) + 1 < (k + \varepsilon)(2k - 1) + 1 = m(|U| - 2) + 1 < e_{G-F}(U),$$

where we make use of (1) to get the last inequality. On the other hand, every edge in  $(G - F)[U]$  corresponds to either an edge of  $G_\theta^{\text{frac}}$  or an arc of  $D_\theta - H'$ . By Lemma 3.4,  $G_\theta^{\text{frac}}$  is a forest and thus, contributes at most  $|U| - 1$  edges to  $(G - F)[U]$ . Since every vertex in  $U \setminus \{u^*\}$  has at most  $k - 1$  integral out-arcs in  $D_\theta - H'$ , and  $u^*$  has no out-arcs at all in  $D_\theta - H'$ , we get that at most  $(|U| - 1) \cdot (k - 1)$  edges of  $(G - F)[U]$  correspond to an edge of  $D_\theta - H'$ . Using these two observations, we get

$$e_{G-F}(U) \leq (|U| - 1) + (|U| - 1) \cdot (k - 1) = 2k + 2k \cdot (k - 1) = 2k^2.$$

Combining the above lower and bound bound on  $e_{G-F}(U)$ , we get

$$\left(k + \frac{1}{2}\right)(2k - 1) + 1 < e_{G-F}(U) \leq 2k^2,$$

which is easily seen to be a contradiction by subtracting  $2k^2$  from both sides.  $\square$

So suppose  $\varepsilon < 1/2$ . By Claim 4.3, we get that  $|U| = 2k$ . Recall that  $u^*$  is incident to at least  $k$  fractional edges in  $G_\theta^{\text{frac}}[U]$ . By Lemma 3.8 and the fact that  $d \leq 1$ , there do not exist  $v, w \in U$  such that both  $u^*v$  and  $u^*w$  are fractional edges, and such that  $(v, w) \in A(D_\theta) \setminus A(H')$ . Similarly, by Lemma 3.4, there are no fractional edges between fractional neighbours of  $u^*$ . Thus, the fractional neighbourhood of  $u^*$  in  $U$  induces an independent set in  $G - F$ , and therefore

$$e_{G-F}(U) \leq \binom{2k}{2} - \binom{k}{2}.$$

Together with (1), we have

$$k(2k - 2) + 1 \leq m(2k - 2) + 1 < k(2k - 1) - \frac{k(k - 1)}{2}.$$

Simplifying, we get

$$2 < k(3 - k).$$

This is a contradiction for every  $k \in \mathbb{N}$ , concluding the proof.  $\square$

## 5. MAIN TECHNICAL LEMMA FOR THEOREM 1.7

We now begin the proof of Theorem 1.7, which is the content of this section and Section 6. In this section, we prove the following technical lemma, which allows us to argue that a minimal violating set  $U$  must have certain structural properties with respect to the terminal components of  $D_\theta$ , and moreover allows us to dispense of most values of  $m$ .

**Lemma 5.1.** *Let  $G$  be a graph with an optimal  $m$ -allocation  $\theta$  for some  $m > 3/2$ . Let  $H$  be any spine of  $D_\theta$  and  $F$  the underlying undirected graph of  $H$ . Let  $U \subseteq V(G)$  be an inclusion-wise minimal set of vertices such that  $|U| \geq 2$  and  $e_{G-F}(U) \geq m(|U| - 4/3)$  (assuming such a set exists). Then the following hold.*

- Either  $m \leq \frac{9}{5}$  or  $m = \frac{9}{4}$ .
- $U$  intersects exactly one terminal component of  $D_\theta$ , and it intersects this component in at least two vertices.
- $|U| \leq \frac{4}{3}m + 1$ .

*Proof.* By assumption, we have that  $|U| \geq 2$  and

$$(3) \quad e_{G-F}(U) \geq m \left( |U| - \frac{4}{3} \right).$$

Note that the above inequality together with  $m > 3/2$  implies that  $|U| > 2$ . Let us consider the collection  $\mathcal{C}$  consisting of all strong components of  $D_\theta$  that intersect  $U$  in at least one vertex. Let us further split  $\mathcal{C}$  into three parts, namely  $\mathcal{C}_1$ , consisting of all non-terminal components in  $\mathcal{C}$ ,  $\mathcal{C}_2$ , consisting of all terminal components in  $\mathcal{C}$  that contain at least two vertices of  $U$ , and  $\mathcal{C}_3$ , consisting of all terminal components in  $\mathcal{C}_3$  that intersect  $U$  in precisely one vertex. Clearly, the intersections  $(X \cap U)_{X \in \mathcal{C}}$  form a partition of  $U$ . In the following, let us denote  $U' := \bigcup_{X \in \mathcal{C}_1} (X \cap U)$  and let  $c := |\mathcal{C}_2|$ ,  $d := |\mathcal{C}_3|$ .

**5.1. Upper bounds on  $e_{G-F}(U)$ .** An essential strategy towards proving Lemma 5.1 is to compare various upper bounds on  $e_{G-F}(U)$  to find a contradiction with (3). In this subsection we give several upper bounds on  $e_{G-F}(U)$  which will be useful later.

Recall that the edges of  $G$  are partitioned into  $E_{\text{frac}}(G, \theta)$  and  $E_{\text{int}}(G, \theta)$ . Therefore,

$$(4) \quad e_{G-F}(U) = |E(G_\theta^{\text{frac}}[U])| + |A((D_\theta - H)[U])|.$$

Since  $G_\theta^{\text{frac}}$  is a forest by Lemma 3.4, there are at most  $|U| - 1$  fractional edges. Additionally, we know that every vertex has at most  $\lfloor m \rfloor$  out-arcs in  $D_\theta$  by definition of  $\theta$ , at most  $\lfloor m \rfloor - 1$  out-arcs in  $(D_\theta - H)[U]$  if it is not the root of a terminal component in  $\mathcal{C}_2$  by Definition 3.6, and no out-arcs in  $D_\theta[U]$  if it is a vertex in  $\mathcal{C}_3$ . Thus, we get

$$(5) \quad e_{G-F}(U) \leq |U| - 1 + (\lfloor m \rfloor - 1)(|U| - d) + c.$$

Yet another way of counting  $e_{G-F}(U)$ , similarly as in (2) in the proof of Theorem 1.6, is by summing up  $\theta(u, v) + \theta(v, u) = 1$  for each  $uv \in E((G - F)[U])$ . That is,

$$(6) \quad e_{G-F}(U) = \sum_{(u,v) \in U^2, uv \in E(G-F)} \theta(u, v) = \sum_{u \in U} \sum_{v \in N_{G-F}(u) \cap U} \theta(u, v).$$

**5.2. Upper and lower bounds on  $|U|$ .** Comparing the inequalities of the previous subsection, we can obtain bounds on the size of  $U$ . Our next few claims provide such bounds, in terms of  $c, d, m$ , and the size of  $U'$ .

**Claim 5.2.**  $|U| > 2m - 1$ .

*Proof.* Suppose  $|U| \leq 2m - 1$ . Together with (3), we get

$$\binom{|U|}{2} \geq e_{G-F}(U) \geq m \left( |U| - \frac{4}{3} \right) \geq \frac{|U| + 1}{2} \left( |U| - \frac{4}{3} \right),$$

which simplifies to  $|U| \leq 2$ , a contradiction.  $\square$

The following claim uses the above inequality to get an upper bound on the size of  $U$ .

**Claim 5.3.**  $|U| \leq \frac{4}{3}m + c + d$ .

*Proof.* Consider a vertex  $u \in U$ . Since  $\theta$  is an  $m$ -allocation, we have  $\sum_{v \in N_G(u)} \theta(u, v) \leq m$ . Therefore, if  $u$  has out-degree one in  $H$ , then

$$\sum_{v \in N_{G-F}(u)} \theta(u, v) \leq m - 1.$$

By Definition 3.6, every vertex in  $U'$  has out-degree 1 in  $H$ , and at most one vertex in each of the sets  $(U \cap X)_{X \in \mathcal{C}_2 \cup \mathcal{C}_3}$  can have out-degree 0 in  $H$ , so in total there can be at most  $c + d$  such vertices in  $U$ . Together with (6), this implies

$$e_{G-F}(U) \leq (m-1)(|U| - c - d) + m(c + d) = (m-1)|U| + c + d.$$

Combining with (3), we find

$$m \left( |U| - \frac{4}{3} \right) \leq (m-1)|U| + c + d,$$

and rearranging yields  $|U| \leq \frac{4}{3}m + c + d$ , as claimed.  $\square$

Observe that the previous claim implies that  $|U| \leq \frac{4}{3}m + 1$  if  $(c, d) = (1, 0)$ , which was the third claimed result in the statement of Lemma 5.1. Therefore, to complete the proof, it remains to prove that  $(c, d) = (1, 0)$  and either  $m < 9/5$  or  $m = 9/4$ .

**Claim 5.4.** *If  $(c, d) \neq (1, 0)$ , then  $|U| \geq \frac{4}{3}mc + md + |U'| - \frac{4}{3}m + 1$ . If additionally  $c > 0$ , then  $|U| > \frac{4}{3}mc + md + |U'| - \frac{4}{3}m + 1$ .*

*Proof.* Suppose that  $(c, d) \neq (1, 0)$ , and let us prove the stated inequality. By the minimality of  $U$ , we have that  $e_{G-F}(U) \geq m(|U| - \frac{4}{3})$  but  $e_{G-F}(Y) < m(|Y| - \frac{4}{3})$  for every  $Y \subsetneq U$  with  $|Y| \geq 2$ . Thus, we have  $e_{G-F}(X \cap U) < m(|X \cap U| - \frac{4}{3})$  for every  $X \in \mathcal{C}_2$  since no  $X \in \mathcal{C}_2$  fully contains  $U$  as otherwise  $(c, d) = (1, 0)$ . We furthermore trivially have  $e_{G-F}(X \cap U) = 0$  for every  $X \in \mathcal{C}_3$ .

Our goal in the following will be to give an upper bound on the number of all edges in  $e_{G-F}(U)$  based on (4). As discussed earlier,  $G_\theta^{\text{frac}}[U]$  is a forest by Lemma 3.4 and hence contains at most  $|U| - 1$  edges.

To bound the number of arcs in  $(D_\theta - H)[U]$ , we split the arc-set of this digraph into those arcs that start in the set  $U'$  and those arcs that start in  $U \setminus U'$ . Since  $H$  is a spine, every vertex in a non-terminal component has out-degree 1 in  $H$  and out-degree at most  $m$  in  $D_\theta$  by definition of  $\theta$ . Thus, we obtain that the number of arcs in  $(D_\theta - H)[U]$  starting in  $U'$  is upper-bounded by  $(m-1)|U'|$ . Secondly, since by definition all components in  $\mathcal{C}_2$  and  $\mathcal{C}_3$  are terminal and thus have no arcs in  $D_\theta$  leaving them, the number of arcs in  $(D_\theta - H)[U]$  starting in  $U \setminus U'$  is at most  $\sum_{X \in \mathcal{C}_2 \cup \mathcal{C}_3} e_{G-F}(X \cap U)$ . Altogether, we obtain the following upper bound on the number of edges in  $(G - F)[U]$ :

$$e_{G-F}(U) \leq (|U| - 1) + (m-1)|U'| + \sum_{X \in \mathcal{C}_2 \cup \mathcal{C}_3} e_{G-F}(X \cap U).$$

Plugging in the upper bounds on  $e_{G-F}(X \cap U)$  for  $X \in \mathcal{C}_2 \cup \mathcal{C}_3$  mentioned above and comparing with (3), it follows that

$$\begin{aligned} m \left( |U| - \frac{4}{3} \right) &\leq (|U| - 1) + (m-1)|U'| + \sum_{X \in \mathcal{C}_2} m \left( |X \cap U| - \frac{4}{3} \right) \\ &= (|U| - 1) + (m-1)|U'| + m(|U| - |U'| - d) - \frac{4}{3}mc \\ &= (m+1)|U| - |U'| - \frac{4}{3}mc - md - 1. \end{aligned}$$

Note that the inequality above is strict if  $c > 0$ . Rearranging now yields that  $|U| \geq \frac{4}{3}mc + md + |U'| - \frac{4}{3}m + 1$  with strict inequality if  $c > 0$ , as desired.  $\square$

**5.3. Consequences.** Combining Claims 5.3 and 5.4 now directly yields the following.

**Claim 5.5.** *If  $(c, d) \neq (1, 0)$  then  $(\frac{4}{3}m - 1)c + (m-1)d + |U'| \leq \frac{8}{3}m - 1$ , with a strict inequality if  $c > 0$ .*



*Proof.* Combining the lower bound on  $|U|$  from Claim 5.4 and the upper bound on  $|U|$  from Claim 5.3 yields that

$$\frac{4}{3}mc + md + |U'| - \frac{4}{3}m + 1 \leq |U| \leq \frac{4}{3}m + c + d.$$

Rearranging now implies

$$\left(\frac{4}{3}m - 1\right)c + (m - 1)d + |U'| \leq \frac{8}{3}m - 1.$$

Furthermore, if  $c > 0$  then the inequality in Claim 5.4 is strict, yielding a strict inequality.  $\square$

Using the previous arguments, we are able to obtain a case distinction into a finite number of options for the pair  $(c, d)$ . The rest of the proof will involve dealing with each case in turn.

**Claim 5.6.** *We have that  $c \leq 2$ . Furthermore, the following hold:*

- If  $c = 0$ , then  $d \leq 1$ .
- If  $c = 1$ , then  $d \leq 3$  and  $d \leq 2$  provided  $m \geq 9/5$ .
- If  $c = 2$ , then  $d \leq 1$  and  $d = 0$  provided  $m \geq 2$ .

*Proof.* The claim trivially holds if  $(c, d) = (1, 0)$ , so assume in the following that  $(c, d) \neq (1, 0)$ .

Suppose first towards a contradiction that  $c = 0$  and  $d \geq 2$ . By Claim 5.4 we have  $|U| \geq md + |U'| - \frac{4}{3}m + 1$ . Since we also have  $|U| = |U'| + d$ , it follows that  $md + |U'| - \frac{4}{3}m + 1 \leq |U'| + d$ , and thus  $(m - 1)d \leq \frac{4}{3}m - 1$ . This is a contradiction, since  $(m - 1)d \geq 2(m - 1) > \frac{4}{3}m - 1$  for  $m > \frac{3}{2}$ .

Next suppose that  $c > 0$ . This implies, by Claim 5.5 that

$$\left(\frac{4}{3}m - 1\right)c + (m - 1)d + |U'| < \frac{8}{3}m - 1,$$

and thus

$$(7) \quad d < \frac{\frac{4}{3}m(2 - c) + c - 1}{m - 1}.$$

Suppose first that  $c = 1$ . Then (7) implies  $d < \frac{\frac{4}{3}m}{m - 1} < \frac{2}{\frac{3}{2} - 1} = 4$ , so  $d \leq 3$ , as claimed. If  $m > \frac{9}{5}$ , we obtain the stronger bound  $d < \frac{12/5}{4/5} = 3$ , so  $d \leq 2$ .

Similarly, if  $c = 2$ , then  $d < \frac{1}{m - 1} < \frac{1}{\frac{3}{2} - 1} = 2$ , and so  $d \leq 1$ . If  $m \geq 2$ , we obtain the stronger bound  $d < 1$  and thus  $d = 0$ .

If  $c \geq 3$ , we obtain  $d < \frac{2 - \frac{4}{3}m}{m - 1} < 0$ , a contradiction, and thus we must have  $c \leq 2$ .  $\square$

Let us now assume towards a contradiction that one of  $(c, d) \neq (1, 0)$ ,  $9/5 < m < 9/4$  or  $m > 9/4$  holds. Based on this and the previous claim, the proof now splits into the following cases, which we list in the order that we resolve them.

- (a)  $(c, d) = (0, 0)$ ,
- (b)  $(c, d) = (1, 3)$  and  $m < 9/5$ ,
- (c)  $(c, d) = (1, 0)$  and  $m > 9/4$ ,
- (d)  $(c, d) = (1, 0)$  and  $9/5 < m < 9/4$ ,
- (e)  $(c, d) = (2, 1)$  and  $m < 2$ ,
- (f)  $(c, d) = (2, 0)$ ,
- (g)  $(c, d) = (1, 2)$ ,
- (h)  $(c, d) = (0, 1)$ ,
- (i)  $(c, d) = (1, 1)$ .

**5.4. Easy arguments based on set sizes.** We now begin the case distinction to handle all the cases delineated in the above list. Many of the cases can be dealt with very directly by simply comparing the upper and lower bounds we have on the sizes of various sets. As all these arguments are simple and similar, we collect them all in this subsection.

**Case (a).** If  $(c, d) = (0, 0)$ , then from Claims 5.2 and 5.3 we obtain that

$$2m - 1 < |U| \leq \frac{4}{3}m + 0 + 0,$$

which simplifies to  $m < \frac{3}{2}$ , a contradiction.

**Case (b).** Suppose  $(c, d) = (1, 3)$  and  $3/2 < m < 9/5$ . Then  $|U| \geq 2c + d = 5$ . Since there are no arcs in  $D_\theta$  between different terminal components and  $m < 2$ , the only vertex in  $U$  with an out-arc in  $D_\theta - H$  is the root of the component in  $\mathcal{C}_2$  if it is in  $U$ . Otherwise there are no integral edges at all. Using that  $G_\theta^{\text{frac}}$  is a forest by Lemma 3.4, it follows from (3) and (4) that

$$\frac{3|U|}{2} - 2 < m \left( |U| - \frac{4}{3} \right) \leq e_{G-F}(U) \leq (|U| - 1) + 1 = |U|,$$

where the first inequality uses that  $m > 3/2$ . Rearranging yields that  $|U| < 4$ , a contradiction.

**Case (c).** Suppose  $(c, d) = (1, 0)$  and  $m > \frac{9}{4}$ . By Claim 5.3 we then have  $|U| \leq \frac{4}{3}m + 1$ . Using (3), it follows that

$$m \left( |U| - \frac{4}{3} \right) \leq e_{G-F}(U) \leq \binom{|U|}{2} = \frac{|U| - 1}{2} |U| \leq \frac{2}{3}m|U|.$$

Rearranging now yields that  $|U| \leq 4$ . If  $|U| = 4$ , then from the above we obtain  $m(4 - \frac{4}{3}) \leq 6$  and thus  $m \leq \frac{9}{4}$ , a contradiction. Similarly, if  $|U| = 3$ , then  $m(3 - \frac{4}{3}) \leq 3$ , contradicting  $m > \frac{9}{4}$ .

**Case (d).** Suppose that  $(c, d) = (1, 0)$  and  $9/5 < m < 9/4$ . By Claim 5.3, we have  $|U| \leq \frac{4}{3}m + 1 < 4$ , so  $|U| \leq 3$  and thus  $|U| = 3$ . Using (3), we find

$$e_{G-F}(U) \geq m \left( |U| - \frac{4}{3} \right) = \frac{5m}{3} > 3,$$

which gives a contradiction since  $e_{G-F}(U)$  can contain at most 3 edges.

**Case (e).** Suppose  $(c, d) = (2, 1)$  and  $3/2 < m < 2$ . Note that  $|U| \geq 2c + d = 5$ . By Claim 5.3, we get

$$|U| \leq \frac{4}{3}m + 3 < 6.$$

It follows that  $|U| = 5$ . Note that this implies that  $U$  shares exactly two vertices with two terminal components of  $D_\theta$  and exactly one vertex with a third terminal component of  $D_\theta$ . Observe that there can be at most two arcs in  $D_\theta[U]$  since there are no arcs between different terminal components. By (4) together with the fact that  $G_\theta^{\text{frac}}$  is a forest by Lemma 3.4, we get

$$e_{G-F}(U) \leq 4 + 2 = 6.$$

By (3), we get

$$\frac{11}{2} < m \left( 5 - \frac{4}{3} \right) \leq e_{G-F}(U).$$

It follows that  $(G - F)[U]$  contains exactly two integral edges and four fractional edges, that is, the fractional edges are a tree. Let  $(u, v) \in A(D_\theta[U])$  be an arc corresponding to one of those two integral edges. Since the fractional edges form a tree on  $U$ , there exists a path  $P$  in  $G_\theta^{\text{frac}}[U]$  from  $u$  to  $v$ . Note that each of these fractional edges connects two different terminal components of  $D_\theta$ , since all the edges in  $G[U]$  within a terminal component are contained in  $D_\theta$ . The existence of such a path contradicts the conclusion of Lemma 3.5.

**5.5. Trickier arguments involving more careful counting.** In this section we resolve the remaining cases, which turn out to be more involved. Let us start with the following claim.

**Claim 5.7.** *If  $(c, d) = (1, 2)$  or  $(c, d) = (2, 0)$ , then  $U' = \emptyset$ .*

*Proof.* By Claim 5.5 we have  $|U'| < \frac{8}{3}m - 1 - (\frac{4}{3}m - 1)c - (m - 1)d$ . For  $(c, d) = (1, 2)$  and  $(c, d) = (2, 0)$  we have that  $|U'| < 2 - \frac{2}{3}m < 2 - \frac{2}{3} \cdot \frac{3}{2} = 1$  and  $|U'| < 1$  respectively, so in either case we may conclude  $U' = \emptyset$ .  $\square$

**Case (f).** Suppose that  $(c, d) = (2, 0)$ . By Claim 5.7 we then have  $U' = \emptyset$ . Let  $C_1, C_2$  denote the two terminal components that intersect  $U$ , and let  $U_1 = U \cap C_1$  and  $U_2 = U \cap C_2$ . Note that, by definition of  $c$ , we have  $|U_1|, |U_2| \geq 2$  and since  $d = 0$  and  $U' = \emptyset$  we have  $U_1 \cup U_2 = U$ . If we denote by  $e_{G-F}(U_1, U_2)$  the number of edges with one endpoint in each  $U_i$ , we have that

$$e_{G-F}(U) = e_{G-F}(U_1) + e_{G-F}(U_2) + e_{G-F}(U_1, U_2) \leq \binom{|U_1|}{2} + \binom{|U_2|}{2} + (|U| - 1),$$

where we used that all the edges in  $G - F$  going between  $U_1$  and  $U_2$  connect the two terminal components  $C_1$  and  $C_2$  and are thus fractional, and that by Lemma 3.4 the fractional edges form a forest.

Furthermore, we claim that the three estimates  $e_{G-F}(U_1) \leq \binom{|U_1|}{2}$ ,  $e_{G-F}(U_2) \leq \binom{|U_2|}{2}$  and  $e_{G-F}(U_1, U_2) \leq |U| - 1$  cannot all be equalities simultaneously. Indeed, if  $e_{G-F}(U_1, U_2) = |U| - 1$  then the edges going across  $U_1$  and  $U_2$  form a tree on  $|U| \geq 4$  vertices, meaning there exists a vertex in either  $U_1$  or  $U_2$  incident to at least two such edges. But if both  $U_1$  and  $U_2$  are cliques, this would imply that there exists a triangle in  $G - F$  formed by two fractional edges that cross between  $U_1$  and  $U_2$  and one integral edge inside  $U_1$  or  $U_2$ , a contradiction to Lemma 3.5. We therefore obtain

$$e_{G-F}(U) \leq \binom{|U_1|}{2} + \binom{|U_2|}{2} + (|U| - 2).$$

By convexity of  $\binom{x}{2}$  we now get

$$\begin{aligned} e_{G-F}(U) &\leq \binom{2}{2} + \binom{|U| - 2}{2} + |U| - 2 \\ &= \frac{(|U| - 2)(|U| - 3)}{2} + |U| - 1 \\ &= \frac{(|U| - 2)(|U| - 1)}{2} + 1. \end{aligned}$$

By (3), we have  $e_{G-F}(U) \geq m(|U| - \frac{4}{3}) = m(|U| - 2) + \frac{2}{3}m > m(|U| - 2) + 1$ . The above then implies that  $(|U| - 1)/2 > m$ , or, equivalently,  $|U| > 2m + 1$ . This contradicts Claim 5.3, which guarantees that  $|U| \leq \frac{4}{3}m + 2 = 2m + 1 + (1 - \frac{2}{3}m) < 2m + 1$ , using that  $m > \frac{3}{2}$ .

**Case (g).** Suppose  $(c, d) = (1, 2)$ . Let  $k \in \mathbb{N}$  and  $0 < \varepsilon \leq 1$  be such that  $m = k + \varepsilon$ . By Claim 5.4, we have that  $|U| > 2m + 1$ . But we also have  $|U| \leq \frac{4}{3}m + 3$  by Claim 5.3, which implies that  $|U| < 2m + 2$  since  $m > 3/2$ . We thus get that

$$(8) \quad 2m + 1 < |U| < 2m + 2.$$

This is impossible if  $m$  is an integer, so we may assume that  $0 < \varepsilon < 1$ , and thus that  $k = \lfloor m \rfloor$ . By Claim 5.7, we know that  $U'$  is empty. Let  $u_1, u_2$  be the two vertices from components in  $C_3$  and let  $X$  be the remaining vertices in  $U$ , which are vertices of the component in  $C_2$ . Note that  $u_1, u_2$  are incident to no arcs in  $D_\theta[U]$  since  $U'$  is empty and no arcs of  $D_\theta$  connect distinct terminal components.

Suppose first that  $1/2 < \varepsilon < 1$ . By (8), we have that  $|U| = 2k + 3$ . Using (3) and (5), we get

$$\left(k + \frac{1}{2}\right) \left(|U| - \frac{4}{3}\right) < m \left(|U| - \frac{4}{3}\right) \leq (|U| - 1) + (k - 1)(|U| - 2) + 1.$$

Simplifying, we arrive at

$$\frac{|U| - 4/3}{2} + \frac{2k}{3} < 2.$$

Using now that  $|U| = 2k + 3$ , we get  $5k/3 + 5/6 < 2$ , which is equivalent to  $k < 7/10$ . But  $k \geq 1$ , a contradiction.

Suppose then that  $0 < \varepsilon \leq 1/2$ . By (8), we get  $|U| = 2k + 2$ . If  $u_1$  is incident to at most  $k$  fractional edges in  $U$ , then  $U \setminus \{u_1\}$  contradicts the minimality of  $U$  with respect to  $e_{G-F} \geq m(|U| - 4/3)$  (and the same holds for  $u_2$ ). So we have that  $u_1, u_2$  are each incident to at least  $k + 1$  edges in  $G_\theta^{\text{frac}}[U]$ . Recall that by Lemma 3.4  $G_\theta^{\text{frac}}$  is a forest. First, suppose

that  $u_1 u_2 \notin E(G_\theta^{\text{frac}})$ . Thus, they both have at least  $k + 1$  edges to  $X$  in  $G_\theta^{\text{frac}}$ . Using that  $|U| - 2 = |X| = 2k$ , we conclude that  $|N_{G_\theta^{\text{frac}}}(u_1) \cap N_{G_\theta^{\text{frac}}}(u_2)| \geq 2$ , yielding a copy of  $C_4$  in  $G_\theta^{\text{frac}}$  and contradicting that  $G_\theta^{\text{frac}}$  is a forest. So we may assume that  $u_1 u_2$  is an edge of  $G_\theta^{\text{frac}}$ . Then  $|N_{G_\theta^{\text{frac}}}(u_1) \cap N_{G_\theta^{\text{frac}}}(u_2)| = 0$ , for otherwise  $G_\theta^{\text{frac}}$  would contain a triangle, again contradicting Lemma 3.4. Therefore, each of the  $2k$  vertices of  $X$  is a neighbour of either  $u_1$  or  $u_2$ . Since by Lemma 3.5 the sets  $N_G(u_1) \cap X, N_G(u_2) \cap X$  must be independent (and of size  $k$ ), we obtain that  $e_{G-F}(X) \leq k^2$ . It follows that  $e_{G-F}(U) \leq k^2 + 2k + 1$ . Comparing this to (3), we get

$$k \left( 2k + \frac{2}{3} \right) \leq k^2 + 2k + 1,$$

which is not true for any  $k \geq 2$ . But since  $k + \varepsilon = m > \frac{3}{2} \geq 1 + \varepsilon$ , we have that  $k \geq 2$ , a contradiction.

**Case (h).** Suppose  $(c, d) = (0, 1)$ . By Claims 5.2 and 5.3, we get

$$2m - 1 < |U| \leq \frac{4}{3}m + 1 < 2m,$$

where we use  $m > 3/2$ . Let us fix  $k \in \mathbb{N}$  and  $0 < \varepsilon \leq 1$  such that  $m = k + \varepsilon$ . Note that in fact, the bounds  $2m - 1 < |U| < 2m$  are impossible to achieve if  $m$  is an integer, and therefore, we may assume that  $\varepsilon < 1$  and thus that  $k = \lfloor m \rfloor$ . Let  $u^*$  be the vertex in the terminal component of  $D_\theta$  which is intersected by  $U$ .

Suppose first that  $1/2 \leq \varepsilon < 1$ . Since  $2m - 1 < |U| < 2m$ , we get that  $\varepsilon > 1/2$  and  $|U| = 2k + 1$ . Using (3) and (5), we get

$$\left( k + \frac{1}{2} \right) \left( 2k - \frac{1}{3} \right) \leq m \left( 2k - \frac{1}{3} \right) < 2k + (k - 1)2k = 2k^2,$$

which we simplify to

$$\frac{2k - 1/3}{2} \leq \frac{k}{3}.$$

But this is a contradiction, since  $(2k - 1/3)/2 = k - \frac{1}{6} \geq \frac{5}{6}k$ .

So we may assume that  $0 < \varepsilon < 1/2$ . Then  $k \geq 2$  and we find that  $|U| = 2k$ . Suppose that  $\sum_{v \in N_{G-F}(u^*) \cap U} \theta(u^*, v) \leq m - 1$ . Using (3) together with (6), we get

$$m \left( 2k - \frac{4}{3} \right) \leq (m - 1)2k.$$

Simplifying yields

$$2k \leq \frac{4}{3}m < \frac{4}{3}k + \frac{2}{3},$$

which does not hold for any positive integer  $k$ . It follows that  $\sum_{v \in N_{G-F}(u^*) \cap U} \theta(u^*, v) > m - 1 \geq k - 1$ . Note that  $u^*$  does not have an out-arc in  $D_\theta[U]$  since it is part of a terminal component of  $D_\theta$  intersecting  $U$  in a single vertex. By definition of  $\theta$ , it follows therefore that if  $u^*$  has an out-arc in  $D_\theta$  then  $u^*$  has at most  $m - 1$  outgoing weight to  $U$  in  $\theta$ , contradicting the computation above. Thus,  $u^*$  does not have any out-arcs in  $D_\theta$  and hence, forms a singleton terminal component. Thus, we obtain that  $u^*$  must be incident to at least  $k$  fractional edges in  $G_\theta^{\text{frac}}[U]$ . By Lemma 3.8, there cannot exist any arc of  $(D_\theta - H)[U]$  between two neighbours of  $u^*$  in  $G_\theta^{\text{frac}}$ , since this would imply that one of those two neighbours also forms a singleton terminal component contained in  $U$ , contradicting our assumption that  $d = 1$ . Also, no two neighbours of  $u^*$  in  $G_\theta^{\text{frac}}[U]$  can be joined by a fractional edge, since  $G_\theta^{\text{frac}}$  is a forest. Hence, the endpoints of those  $k$  fractional edges incident to  $u^*$  in  $G_\theta^{\text{frac}}[U]$  form an independent set in  $(G - F)[U]$ . We conclude that

$$e_{G-F}(U) \leq \binom{2k}{2} - \binom{k}{2}.$$

Together with (3), we have

$$k \left( 2k - \frac{4}{3} \right) < m \left( 2k - \frac{4}{3} \right) \leq k(2k - 1) - \binom{k}{2}.$$

Simplifying, we get

$$\binom{k}{2} < \frac{k}{3}.$$

This is a contradiction for every  $k \geq 2$ .

**Case (i).** Suppose  $(c, d) = (1, 1)$ . By Claims 5.2 and 5.3, we get

$$2m - 1 < |U| \leq \frac{4}{3}m + 2.$$

Using additionally that  $m > 3/2$ , we get that  $\frac{4}{3}m + 2 < 2m + 1$  and hence,

$$(9) \quad 2m - 1 < |U| < 2m + 1.$$

For each  $u \in U$ , let  $I(u)$  denote the out-degree of  $u$  in  $(D_\theta - H)[U]$ . Let  $X$  be the intersection of  $U$  with the component in  $\mathcal{C}_2$ . Let  $u^* \in U$  be the vertex in the component in  $\mathcal{C}_3$  and  $r \in X$  chosen such that  $I(r)$  is maximal. Note that  $X$  has no out-arcs in  $D_\theta[U]$ , as it is the intersection of  $U$  with a terminal component. Fix  $k \in \mathbb{N}$  and  $0 \leq \varepsilon < 1$  such that<sup>5</sup>  $m = k + \varepsilon$ . Then,  $I(u^*) = 0$  and  $I(r) \leq \lfloor m \rfloor = k$ . Furthermore, we have  $I(u) \leq k - 1$  for every  $u \in U \setminus \{u^*, r\}$ . Indeed, otherwise there would be at least two vertices in  $U \setminus \{u^*\}$  that have out-degree  $k = \lfloor m \rfloor$  in  $(D_\theta - H)[U]$ . However, this is impossible, since by Definition 3.6 all but at most one vertex of  $U \setminus \{u^*\}$  have out-degree 1 in  $H$ . Let

$$I := k - I(r) + \sum_{u \in U \setminus \{u^*, r\}} (k - 1 - I(u)).$$

Intuitively,  $I$  is the number of integral edges that are “missing”. Now, using that the number of integral edges in  $(G - F)[U]$  is equal to  $\sum_{u \in U} I(u)$ , we can rewrite (4) as

$$\begin{aligned} \mathbf{e}_{G-F}(U) &\leq |U| - 1 + \sum_{u \in U} I(u) \\ &\leq |U| - 1 + k + \left( \sum_{u \in U \setminus \{u^*, r\}} (k - 1) \right) - I \\ &= |U| - 1 + (k - 1)(|U| - 1) + 1 - I \\ &= k(|U| - 1) + 1 - I. \end{aligned}$$

Together with (3), we get

$$m \left( |U| - \frac{4}{3} \right) \leq k(|U| - 1) + 1 - I.$$

Plugging in  $m = k + \varepsilon$ , we simplify to

$$(10) \quad I \leq \frac{k}{3} - \varepsilon \left( |U| - \frac{4}{3} \right) + 1.$$

Let  $M$  denote the number of edges in the complement of  $(G - F)[U]$ , so that

$$\mathbf{e}_{G-F}(U) = \binom{|U|}{2} - M.$$

Combining this with (3), we get

$$(11) \quad M \leq \binom{|U|}{2} - m \left( |U| - \frac{4}{3} \right).$$

---

<sup>5</sup>Note that this is a different choice from the one made in cases (g) and (h), as we now allow  $\varepsilon = 0$  and earlier allowed  $\varepsilon = 1$ .

We now split into subcases depending on whether  $0 \leq \varepsilon \leq 1/2$  or  $1/2 < \varepsilon < 1$ . First, suppose that  $1/2 < \varepsilon < 1$ . By (9), we get that  $|U| \in \{2k+1, 2k+2\}$ . However, plugging  $\varepsilon > 1/2$  and  $|U| \in \{2k+1, 2k+2\}$  into (10) (and using the fact that  $I$  must be non-negative) implies that the only possibility is  $I = 0$ ,  $k = 1$  and  $|U| = 2k+1 = 3$ . This in turn implies that  $U \setminus \{u^*\} = X$ . By Lemma 3.5,  $(G-F)[U]$  cannot be a triangle, since this would imply that  $u^*$  has two adjacent neighbours in the terminal component containing  $X$ . Hence, we have  $m(3 - \frac{4}{3}) \leq e_{G-F}(U) \leq 2$ , a contradiction.

So we may now assume that  $0 \leq \varepsilon \leq 1/2$ . Since  $m > 3/2$ , we get  $k \geq 2$ . By (9), we get that  $|U| \in \{2k, 2k+1\}$ . Suppose first that  $|U| = 2k+1$ . Plugging this into (11), and using that  $k \leq m$ , we get

$$M \leq k(2k+1) - k \left(2k - \frac{1}{3}\right) = \frac{4}{3}k.$$

Next, let us look at (10). We get  $I \leq k/3 + 1 < k$ , where we used  $k \geq 2$ . Hence,  $I \leq k-1$ . It follows that

$$(12) \quad e_{G-F}(X) \geq \sum_{u \in X} I(u) \geq (k-1)|X| + 1 - I \geq (k-1)(|X| - 1) + 1.$$

It follows that  $|X| \geq 2k-1$ . Suppose now that  $\sum_{v \in N_{G-F}(u^*) \cap U} \theta(u^*, v) \leq m-1$ . Using (6), we get

$$e_{G-F}(U) \leq (m-1)|U| + 1.$$

Comparing this with (3), we get

$$m \left(|U| - \frac{4}{3}\right) \leq (m-1)|U| + 1.$$

Simplifying gives  $|U| \leq \frac{4}{3}m + 1 < 2m$ , contradicting  $|U| = 2k+1 \geq 2m$ . So we may assume that

$$(13) \quad \sum_{v \in N_{G-F}(u^*) \cap U} \theta(u^*, v) > m-1.$$

Recall that  $u^*$  does not have any out-arcs in  $D_\theta[U]$ . Additionally, the assumption (13) implies that  $u^*$  cannot have an out-arc in  $D_\theta$  to a vertex outside of  $U$ . It follows that  $u^*$  does not have any out-arcs in  $D_\theta$ . In particular,  $u^*$  is a singleton terminal component of  $D_\theta$ . Additionally,  $u^*$  is incident to at least  $k$  edges in  $G_\theta^{\text{frac}}[U]$ . Suppose that  $u^*$  is incident to at least two edges in the complement of  $(G-F)[U]$ . The fact that  $d = 1$  implies by Lemmas 3.4 and 3.8 that the fractional neighbourhood of  $u^*$  in  $U$  is an independent set, so we find that

$$M \geq \binom{k}{2} + 2.$$

It follows that

$$\binom{k}{2} + 2 \leq \frac{4}{3}k,$$

which holds for no  $k \in \mathbb{N}$ . So we may assume that  $u^*$  has at most one non-neighbour in  $(G-F)[U]$ . Then, using that there are no integral edges between  $u^*$  and  $X$ , we get that  $u^*$  is connected by at least  $|X| - 1$  fractional edges to  $X$ . Since by Lemma 3.5 the neighbours of  $u^*$  in  $X$  form an independent set in  $G$ , it follows that  $G[X]$  contains at most  $|X| - 1$  edges. On the other hand,  $(G-F)[X]$  must contain at least  $(k-1)(|X| - 1) + 1$  integral edges by (12). So we get

$$(k-1)(|X| - 1) + 1 \leq |X| - 1,$$

implying that  $k \leq 1$ , a contradiction.

We move on to the last case, that is  $|U| = 2k$ . We again plug this into (11) and (10) to get, using  $k \geq 2$ ,

$$M \leq k(2k-1) - k \left(2k - \frac{4}{3}\right) = \frac{k}{3} < k-1,$$

and

$$I \leq \frac{k}{3} + 1 < k.$$

As above, this implies that  $X$  has at least  $2k - 1$  vertices. Since there are at most  $k - 2$  missing edges, we get that  $u^*$  is adjacent to at least  $k + 1$  edges into  $|X|$ . Using Lemma 3.5, we get that there are at least  $\binom{k+1}{2}$  missing edges and thus,

$$\binom{k+1}{2} \leq M \leq k - 2,$$

a contradiction. This concludes the proof of the last remaining case and hence, of Lemma 5.1.  $\square$

## 6. PROOF OF THEOREM 1.7

Given Lemma 5.1, we now proceed with the proof of Theorem 1.7. We use Lemma 5.1 to deal with the majority of cases, whereas a more careful selection of the forest  $F$  is needed to resolve the final cases.

*Proof of Theorem 1.7.* We begin by picking an optimal  $m$ -allocation  $\theta$  of  $G$ , which we may do by Theorem 2.5. We now split the proof into three cases, depending on the value of  $m$ .

**Case 1:**  $9/5 < m < 9/4$  or  $m > 9/4$ . By Lemma 3.7,  $D_\theta$  contains a spine  $H$  with underlying undirected graph  $F$ , which is a forest. It then follows by Lemma 5.1 that for every  $U \subseteq V(G)$  with  $|U| \geq 2$ , it holds that  $e_{G-F}(U) < m(|U| - 4/3)$  and the statement follows.

**Case 2:**  $m = 9/4$ . Again, by Lemma 3.7, there exists a spine of  $D_\theta$ . Suppose that  $H$  is a spine of  $D_\theta$  minimizing the number of copies of  $K_4$  in  $G - F$ , where  $F$  is the underlying undirected graph of  $H$ . We are done if there is no  $U \subseteq V(G)$  with  $|U| \geq 2$  and  $e_{G-F}(U) \geq m(|U| - 4/3)$ , so we may assume that  $U$  is an inclusion-wise minimal such set. By Lemma 5.1, we get that  $U$  intersects exactly one terminal component  $X$  of  $D_\theta$  and it intersects  $X$  in at least two vertices as well as that  $|U| \leq \frac{4}{3}m + 1 = 4$ . We note that if  $|U| = 2$ , then our assumption reads  $e_{G-F}(U) \geq m(|U| - 4/3) = 2m/3 = 3/2$ , which is impossible since  $e_{G-F}(U) \leq 1$  if  $|U| = 2$ . Similarly, the case  $|U| = 3$  can be ruled out as our assumption reads  $e_{G-F}(U) \geq m(3 - 4/3) = 5m/3 = 15/4$ , which is impossible as  $15/4 > 3$ . So we may conclude that  $|U| = 4$ , and moreover we have that  $e_{G-F}(U) \geq m(4 - 4/3) = 8m/3 = 6$ , which implies that  $U$  induces a  $K_4$  in  $G - F$ .

Let  $W$  be any set of vertices of  $G$ . As in (2), we have that

$$(14) \quad e_{G-F}(W) = \sum_{(u,v) \in W^2, uv \in E(G-F)} \theta(u,v) = \sum_{u \in W} \sum_{v \in N_{G-F}(u) \cap W} \theta(u,v).$$

Let us apply the above equality to  $U$ . Let  $r$  denote the unique vertex in  $X$  that has out-degree 0 in  $H$  (i.e., the root of the in-tree that  $H$  induces on  $X$ ). Since  $H$  contains an out-arc for every vertex in  $U \setminus \{r\}$ , it follows that  $\sum_{v \in N_{G-F}(r) \cap U} \theta(r,v) \leq m$  and that  $\sum_{v \in N_{G-F}(u) \cap U} \theta(u,v) \leq m - 1$  for all  $u \in U \setminus \{r\}$ . This implies that  $r \in U$ , as otherwise,  $e_{G-F}(U) \leq (m - 1)|U| = 5$ . For the same reason,  $U$  must contain all out-neighbours of  $r$  in  $D_\theta$ . Let  $u \in U$  be an out-neighbour of  $r$  and observe that  $u \in X$  since  $X$  is a terminal component. Thus, by Definition 3.6, there is a directed path from  $u$  to  $r$  in  $H$ . Therefore,  $H + (r, u)$  contains a unique directed cycle, which necessarily includes the edge  $(r, u)$ . Let  $v$  denote the successor of  $u$  along this cycle. As  $U$  induces a  $K_4$  in  $G - F$ , and as  $uv \notin E(G - F)$ , we have that  $v \notin U$ . Moreover,  $H + (r, u) - (u, v)$  is a spine of  $D_\theta$  with root  $u$  in  $X$ . As  $H$  was chosen to minimize the number of copies of  $K_4$  in  $G - F$ , it must hold that there exists some vertex set  $\tilde{U}$  that induces a  $K_4$  in  $G - (F + ru - uv)$  but not in  $G - F$ . The same argument as above implies that  $u, v \in \tilde{U}$  and  $r \notin \tilde{U}$ . Moreover, we have  $e_{G-F}(\tilde{U}) \geq e_{G-(F+ru-uv)}(\tilde{U}) - 1 = 6 - 1 = 5$ . Observe that  $\tilde{U}$  is also a minimal set satisfying  $|\tilde{U}| \geq 2$  and  $e_{G-(F+ru-uv)}(\tilde{U}) \geq m(|\tilde{U}| - 4/3)$ . By Lemma 5.1, it follows that  $\tilde{U}$  only intersects one terminal component of  $D_\theta$ ; this terminal component is necessarily  $X$ , since  $u \in X \cap \tilde{U}$ .

On the one hand, applying (14) to  $W = U \cup \tilde{U}$ , we get  $e_{G-F}(U \cup \tilde{U}) \leq (m-1)|U \cup \tilde{U}| + 1$ , since every vertex in  $U \cup \tilde{U}$  distinct from  $r$  has out-degree 1 in  $H$ . On the other hand,  $e_{G-F}(U \cup \tilde{U}) \geq e_{G-F}(U) + e_{G-F}(\tilde{U}) - e_{G-F}(U \cap \tilde{U}) \geq 11 - \binom{|U \cap \tilde{U}|}{2}$ . The size of  $|U \cap \tilde{U}|$  is at least 1 as both sets contain  $u$ , and is at most 3 since  $r \in U \setminus \tilde{U}$ . Hence,  $m = 9/4$  must satisfy one of the inequalities

$$11 - \binom{1}{2} \leq 7(m-1) + 1,$$

$$11 - \binom{2}{2} \leq 6(m-1) + 1,$$

or

$$11 - \binom{3}{2} \leq 5(m-1) + 1.$$

But none of these inequalities hold, a contradiction which concludes the proof of this case.

**Case 3:**  $3/2 < m \leq 9/5$ . We begin by noting that since  $m < 2$ , every vertex of  $G$  has out-degree at most one in  $D_\theta$ . The resolution of this case now rests on the following crucial claim.

**Claim 6.1.** *Each non-singleton terminal component of  $D_\theta$  contains an arc  $(u, v)$  such that  $u$  and  $v$  have no common neighbours in  $G_\theta^{\text{frac}}$ .*

Before we prove the above claim, let us first show how to use it to complete the argument. Suppose then that every non-singleton terminal component contains an arc  $(u, v) \in A(D_\theta)$  such that  $u$  and  $v$  have no common neighbour in  $G_\theta^{\text{frac}}$ . Let  $R$  be a set of vertices containing exactly one such vertex  $u$  for each non-singleton terminal component, as well as all the vertices in singleton terminal components of  $D_\theta$ . By Lemma 3.7, there exists a spine  $H$  of  $D_\theta$  with underlying forest  $F$  such that every vertex in  $V(D) \setminus R$  has out-degree exactly one in  $H$  and therefore, out-degree zero in  $D_\theta - H$ . Suppose towards a contradiction that  $m \frac{4}{3}(G - F) \geq m$ . Then, there exists a minimal  $U \subseteq V(G)$  with  $|U| \geq 2$  and  $e_{G-F}(U) \geq m(|U| - 4/3)$ . By Lemma 5.1, it follows that  $U$  intersects a unique terminal component  $X$  of  $D_\theta$  and that  $U$  intersects  $X$  in at least two vertices. Additionally, we have that  $|U| \leq \frac{4}{3}m + 1 < 4$ . Observe that  $|U| > 2$  since  $m(2 - 4/3) > 1$ , so we conclude that  $|U| = 3$ . Additionally,  $e_{G-F}(U) \geq m(3 - 4/3) > 2$ , implying that  $(G - F)[U]$  is a triangle. By Lemma 3.4,  $G_\theta^{\text{frac}}$  is a forest and therefore  $(D_\theta - H)[U]$  contains at least one arc. But recall that the only vertices with out-arcs in  $(D_\theta - H)[U]$  are vertices of  $R$ , and that  $U$  intersects only one terminal component  $X$ . Let  $r$  be the root of  $X$ , so that  $U \cap R = \{r\}$ . Let  $(r, v)$  be the out-arc of  $r$  in  $(D_\theta - H)[U]$ , and note that by the argument above this is the only arc in  $(D_\theta - H)[U]$ . But by our choice of the roots,  $r$  and  $v$  do not share any neighbours in  $G_\theta^{\text{frac}}$  and hence  $(G - F)[U]$  can not be a triangle. This contradiction completes the proof, and all that remains is to prove Claim 6.1.

*Proof of Claim 6.1.* Let us start by proving the following lemma.

**Lemma 6.2.** *Let  $T$  be a forest,  $\ell \geq 2$  an integer, and let  $v_1, v_2, \dots, v_\ell$  be pairwise distinct vertices in  $T$  such that each of the vertex pairs  $v_1v_2, v_2v_3, \dots, v_\ell v_1$  have a common neighbour in  $T$ . Then there exists a vertex  $u$  that is adjacent to all vertices  $v_1, \dots, v_\ell$ .*

*Proof.* We prove the lemma by induction on  $\ell$ . It trivially holds if  $\ell = 2$ , so suppose  $\ell \geq 3$  and that we have proved the lemma with value  $\ell - 1$ . For each  $i = 1, \dots, \ell$ , let  $w_i$  be a common neighbour of  $v_i$  and  $v_{i+1}$  (index addition modulo  $\ell$ ). Let  $W := \{v_1, \dots, v_\ell, w_1, \dots, w_\ell\}$ . Since  $T[W]$  is a forest, it contains a vertex of degree at most 1, which must be one of  $v_i, i = 1, \dots, \ell$ . Without loss of generality say that  $v_\ell$  has degree 1 in  $T[W]$ . This implies that  $w_{\ell-1} = w_\ell$ , as both are neighbours of  $v_\ell$ . Hence,  $v_1, \dots, v_{\ell-1}$  also has the property that any pair of two cyclically consecutive vertices have a common neighbour in  $T$ . By the inductive hypothesis, there is a common neighbour  $w$  of  $v_1, \dots, v_{\ell-1}$ . Since both  $w$  and  $w_{\ell-1} = w_\ell$  are common neighbours of  $v_1$  and  $v_{\ell-1}$ , and since  $T$  contains no cycles, it follows that  $w = w_\ell = w_{\ell-1}$ , and thus  $w$  is a common neighbour of  $v_1, \dots, v_\ell$ .  $\square$



Note that, as the maximum out-degree of  $D_\theta$  is at most 1, each non-singleton terminal component of  $D_\theta$  is the vertex-set of a directed cycle. Suppose towards a contradiction that there exists a non-singleton terminal component  $C$  such that for every arc  $(u, v)$  on the directed cycle  $D_\theta[C]$ , the vertices  $u$  and  $v$  have a common neighbour in  $G_\theta^{\text{frac}}$ . By Lemma 3.4,  $G_\theta^{\text{frac}}$  is a forest and therefore, by Lemma 6.2, there exists a common neighbour  $u$  of all vertices of  $C$  in  $G_\theta^{\text{frac}}$ . This implies that  $u \notin C$ . Moreover, by Lemma 3.5,  $u$  cannot be contained in another terminal component. Recall that since  $m < 2$ ,  $u$  has at most one out-arc in  $D_\theta$ , and if it had zero out-arcs it would be a singleton terminal component, which we just argued is not the case. Therefore,  $u$  has exactly one out-arc in  $D_\theta$ , and thus  $H$  contains a directed path from  $u$  to a terminal component  $C'$  of  $D_\theta$ .

First, suppose that  $C' \neq C$ . Let  $(v, w) \in A(D_\theta[C])$  be any arc and  $\theta'$  the  $m$ -allocation obtained from  $\theta$  by shifting along the triangle  $u, v, w$ . By Observation 3.2,  $D_{\theta'}$  is obtained from  $D_\theta$  by removing the arc  $(v, w)$  and adding one (or both) of the arcs  $(v, u), (u, w)$ . But as  $u$  has an out-arc in  $D_\theta$  which remains in  $D_{\theta'}$ , it cannot be that  $(u, w)$  gets added, since this would contradict  $m < 2$  and the fact that  $\theta'$  remains an  $m$ -allocation. Hence,  $D_{\theta'} = D_\theta - (v, w) + (v, u)$ . Note that by Lemma 3.1(b) and using that  $v$  has no out-arcs in  $D_\theta - (v, w)$ , the terminal components of  $D_\theta - (v, w)$  are the singleton  $\{v\}$  and the remaining terminal components of  $D_\theta$  besides  $C$ . But then, Lemma 3.1(a) implies that  $D_{\theta'} = D_\theta - (v, w) + (v, u)$  has strictly fewer terminal components than  $D_\theta$ , since  $(v, u)$  together with the path in  $H$  from  $u$  to  $C'$  forms a directed path from  $v$  to another terminal component of  $D_\theta - (v, w)$ . This contradicts the optimality of  $\theta$ .

Therefore, we may assume that  $C' = C$ . Thus,  $H$  contains a directed path from  $u$  to  $C$ . Let  $(v, w) \in A(D_\theta[C])$  be the directed edge in  $C$  such that  $w$  is the first vertex in  $C$  reached on this path. Let  $\theta'$  be the  $m$ -allocation obtained from  $\theta$  by shifting along the triangle  $u, v, w$ . As before, we can use Observation 3.2 and the fact that  $u$  has an out-arc in  $D_\theta$  that remains in  $D_{\theta'}$  (recall that  $u$  is connected only by fractional edges to the vertices in  $C$ , and hence cannot have an out-arc to one of  $v, w$ ) to conclude that  $D_{\theta'} = D_\theta - (v, w) + (v, u)$ . As above, by Lemma 3.1(b), the terminal components of  $D_\theta - (v, w)$  are the same as the ones of  $D_\theta$ , except that  $C$  is replaced by the new singleton terminal component  $\{v\}$ . Then, it follows by Lemma 3.1(e) and the optimality of  $\theta$  that  $D_{\theta'}$  has the same terminal components as  $D_\theta$  besides  $C$ , which is replaced by a terminal component of  $D_{\theta'}$  containing  $\{u, v\}$ . Let us denote the vertices of this component by  $A$ . Every vertex reachable from  $v$  in  $D_{\theta'}$  is contained in  $A$ . By our choice of  $w$ , we see that  $w$  is reachable from  $u$  and hence from  $v$  (via the arc  $(v, u)$ ) in  $D_{\theta'}$ . Therefore all vertices in  $C$  are reachable, since  $C - (v, w)$  is a path from  $w$  to  $v$  passing through all vertices of  $C$ . Therefore,  $C \subsetneq A$ , as  $u \in A \setminus C$ . Thus, there are strictly more vertices contained in terminal components of  $D_{\theta'}$  than in terminal components of  $D_\theta$ . Since the number of terminal components of  $D_\theta$  and  $D_{\theta'}$  is identical and since the number of integral edges of  $\theta$  and  $\theta'$  is also identical, this is a contradiction to the optimality of  $\theta$ , completing the proof.  $\square$

As argued above, Claim 6.1 finishes the last case, which concludes the proof of Theorem 1.7.  $\square$

## 7. CONCLUDING REMARKS

The results of this paper confirm the Kohayakawa–Kreuter conjecture, Conjecture 1.3, for all  $r$ -tuples of graphs. However, there remain a few natural directions for future work to explore.

The first question we wish to highlight is a conjecture of Kuperwasser, Samotij, and Wigderson [19, Conjecture 1.5], which strengthens the conclusion of Theorem 1.6 by demanding that  $F$  is a forest, rather than a pseudoforest.

**Conjecture 7.1** (Kuperwasser–Samotij–Wigderson). *Let  $m > 0$  be a real number. Every graph  $G$  with  $m(G) \leq m$  contains a forest  $F \subseteq G$  such that  $m_2(G - F) \leq m$ .*

Note that this conjecture is in fact a common strengthening of both Theorem 1.6 and Theorem 1.7, since  $m_{\frac{4}{3}}(G - F) < \max\{m, m_2(G - F)\}$  assuming  $m > \frac{3}{2}$ . This is an interesting

and natural graph decomposition in its own right, but would also have further implications for the study of Ramsey properties of random graphs. Namely, as proved in [19, Theorem 1.10], Conjecture 7.1 would imply a natural generalization of Conjecture 1.3 where one wishes to color  $G_{n,p}$  and avoid not just a single graph  $H_i$  in color  $i$ , but some finite family  $\{H_i^{(1)}, \dots, H_i^{(t)}\}$ . It seems very possible that our techniques, potentially coupled with a more careful choice of  $\theta$  and  $F$ , could resolve Conjecture 7.1. We note that Conjecture 7.1 was proved in [19] under the extra assumption that  $m$  is an integer, using matroid-theoretic techniques; however, those techniques do not seem to work in the general case.

Another natural question that deserves more study concerns the Ramsey properties of random hypergraphs. In this setting, even the symmetric case (i.e. the generalization of Theorem 1.1 to hypergraphs) is not well-understood. The analogous 1-statement to that in Theorem 1.1 is known to be true [3, 7]. However, it was shown in [9] that in uniformity 4 and above, the corresponding 0-statement is not true in general; there exist hypergraphs whose symmetric Ramsey threshold is controlled by an intermediate “reason” between the “local” and “global” reasons discussed in the introduction. Despite this, Bowtell, Hancock, and Hyde [1] proved an analogue of Theorem 1.4 in the hypergraph setting, so in some sense all that remains to understand is when the hypergraph analogue of Theorem 1.5 holds.

#### ACKNOWLEDGEMENTS

We are indebted to Benjamin Moore for many insightful discussions on this topic, and for carefully reading an earlier draft of this paper. We would also like to thank the anonymous referees for many helpful comments and corrections.

#### REFERENCES

- [1] C. Bowtell, R. Hancock, and J. Hyde, Proof of the Kohayakawa–Kreuter conjecture for the majority of cases, 2023. Preprint available at arXiv:2307.16760.
- [2] D. Conlon, J. Fox, and B. Sudakov, Recent developments in graph Ramsey theory, in *Surveys in combinatorics 2015, London Math. Soc. Lecture Note Ser.*, vol. 424, Cambridge Univ. Press, Cambridge, 2015, 49–118.
- [3] D. Conlon and W. T. Gowers, Combinatorial theorems in sparse random sets, *Ann. of Math. (2)* **184** (2016), 367–454.
- [4] G. Fan, Y. Li, N. Song, and D. Yang, Decomposing a graph into pseudoforests with one having bounded degree, *J. Combin. Theory Ser. B* **115** (2015), 72–95.
- [5] A. Frank and A. Gyárfás, How to orient the edges of a graph?, in *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, Vol. I, *Colloq. Math. Soc. János Bolyai*, vol. 18, North-Holland, Amsterdam-New York, 1978, 353–364.
- [6] P. Frankl and V. Rödl, Large triangle-free subgraphs in graphs without  $K_4$ , *Graphs Combin.* **2** (1986), 135–144.
- [7] E. Friedgut, V. Rödl, and M. Schacht, Ramsey properties of random discrete structures, *Random Structures Algorithms* **37** (2010), 407–436.
- [8] L. Grout and B. Moore, The pseudoforest analogue for the strong nine dragon tree conjecture is true, *J. Combin. Theory Ser. B* **145** (2020), 433–449.
- [9] L. Gugelmann, R. Nenadov, Y. Person, N. Škorić, A. Steger, and H. Thomas, Symmetric and asymmetric Ramsey properties in random hypergraphs, *Forum Math. Sigma* **5** (2017), Paper No. e28, 47pp.
- [10] B. Hajek, Performance of global load balancing by local adjustment, *IEEE Trans. Inform. Theory* **36** (1990), 1398–1414.
- [11] S. L. Hakimi, On the degrees of the vertices of a directed graph, *J. Franklin Inst.* **279** (1965), 290–308.
- [12] R. Hancock, K. Staden, and A. Treglown, Independent sets in hypergraphs and Ramsey properties of graphs and the integers, *SIAM J. Discrete Math.* **33** (2019), 153–188.

- [13] J. Hyde, Towards the 0-statement of the Kohayakawa-Kreuter conjecture, *Combin. Probab. Comput.* **32** (2023), 225–268.
- [14] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.
- [15] H. Jiang and D. Yang, Decomposing a graph into forests: the nine dragon tree conjecture is true, *Combinatorica* **37** (2017), 1125–1137.
- [16] Y. Kohayakawa and B. Kreuter, Threshold functions for asymmetric Ramsey properties involving cycles, *Random Structures Algorithms* **11** (1997), 245–276.
- [17] Y. Kohayakawa, M. Schacht, and R. Spöhel, Upper bounds on probability thresholds for asymmetric Ramsey properties, *Random Structures Algorithms* **44** (2014), 1–28.
- [18] E. Kuperwasser and W. Samotij, The list-Ramsey threshold for families of graphs, 2023. Preprint available at arXiv:2305.19964.
- [19] E. Kuperwasser, W. Samotij, and Y. Wigderson, On the Kohayakawa–Kreuter conjecture. Preprint available at arXiv:2307.16611.
- [20] A. Liebenau, L. Mattos, W. Mendonça, and J. Skokan, Asymmetric Ramsey properties of random graphs involving cliques and cycles, *Random Structures Algorithms* **62** (2023), 1035–1055.
- [21] T. Łuczak, A. Ruciński, and B. Voigt, Ramsey properties of random graphs, *J. Combin. Theory Ser. B* **56** (1992), 55–68.
- [22] M. Marciniszyn, J. Skokan, R. Spöhel, and A. Steger, Asymmetric Ramsey properties of random graphs involving cliques, *Random Structures Algorithms* **34** (2009), 419–453.
- [23] S. Mies and B. Moore, The strong nine dragon tree conjecture is true for  $d \leq k + 1$ , *Combinatorica* **43** (2023), 1215–1239.
- [24] S. Mies, B. Moore, and E. Smith-Roberge, Beyond the pseudoforest strong nine dragon tree theorem. Preprint available at arXiv:2310.00931.
- [25] F. Mousset, R. Nenadov, and W. Samotij, Towards the Kohayakawa–Kreuter conjecture on asymmetric Ramsey properties, *Combin. Probab. Comput.* **29** (2020), 943–955.
- [26] W. Nadara and M. Smulewicz, Decreasing the maximum average degree by deleting an independent set or a  $d$ -degenerate subgraph, *Electron. J. Combin.* **29** (2022), Paper No. 1.49, 15pp.
- [27] C. S. J. A. Nash-Williams, Decomposition of finite graphs into forests, *J. London Math. Soc.* **39** (1964), 12.
- [28] R. Nenadov and A. Steger, A short proof of the random ramsey theorem, *Combinatorics, Probability and Computing* **25** (2016), 130–144.
- [29] F. P. Ramsey, On a problem of formal logic, *Proc. London Math. Soc. (2)* **30** (1929), 264–286.
- [30] V. Rödl and A. Ruciński, Lower bounds on probability thresholds for Ramsey properties, in *Combinatorics, Paul Erdős is eighty, Vol. 1*, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 1993, 317–346.
- [31] V. Rödl and A. Ruciński, Random graphs with monochromatic triangles in every edge coloring, *Random Structures Algorithms* **5** (1994), 253–270.
- [32] V. Rödl and A. Ruciński, Threshold functions for Ramsey properties, *J. Amer. Math. Soc.* **8** (1995), 917–942.

DEPARTMENT OF COMPUTER SCIENCE, INSTITUTE OF THEORETICAL COMPUTER SCIENCE, ETH ZÜRICH,  
SWITZERLAND

*Email address:* `micha.christoph@inf.ethz.ch`

DEPARTMENT OF COMPUTER SCIENCE, INSTITUTE OF THEORETICAL COMPUTER SCIENCE, ETH ZÜRICH,  
SWITZERLAND

*Email address:* `anders.martinsson@inf.ethz.ch`

DEPARTMENT OF COMPUTER SCIENCE, INSTITUTE OF THEORETICAL COMPUTER SCIENCE, ETH ZÜRICH,  
SWITZERLAND

*Email address:* `raphaelmario.steiner@inf.ethz.ch`

INSTITUTE FOR THEORETICAL STUDIES, ETH ZÜRICH, SWITZERLAND

*Email address:* `yuval.wigderson@eth-its.ethz.ch`