Judicious partitions of directed graphs

Choongbum Lee [*] Po-Shen Loh [†] Benny Sudakov [‡]

Abstract

The area of *judicious partitioning* considers the general family of partitioning problems in which one seeks to optimize several parameters simultaneously, and these problems have been widely studied in various combinatorial contexts. In this paper, we study essentially the most fundamental judicious partitioning problem for directed graphs, which naturally extends the classical Max Cut problem to this setting: we seek bipartitions in which many edges cross in each direction. It is easy to see that a minimum outdegree condition is required in order for the problem to be nontrivial, and we prove that every directed graph with m edges and minimum outdegree at least two admits a bipartition in which at least $(\frac{1}{6}+o(1))m$ edges cross in each direction. We also prove that if the minimum outdegree is at least three, then the constant can be increased to $\frac{1}{5}$. If the minimum outdegree tends to infinity with n, then the constant increases to $\frac{1}{4}$. All of these constants are best-possible, and provide asymptotic answers to a question of Alex Scott.

1 Introduction

Partitioning problems have a long history in mathematics and theoretical computer science. One famous example is Max Cut, which seeks a bipartition of a given graph which maximizes the number of edges which cross between the two sides. This is a fundamental problem, and has been the subject of much investigation (see, e.g., [11, 12, 14, 22] and their references). Computing the exact solution can be quite difficult, since the Max Cut problem is known to be NP-complete. Still, it is possible to obtain some estimates on the size of the Max Cut in terms of the number of edges of the graph. A folklore bound (which comes from a simple and efficient algorithm) asserts that every graph with m edges has

$$\text{Max Cut} \geq \frac{m}{2}.$$

This immediately gives a 0.5-approximation algorithm, because no cut can have size greater than the total number of edges m. The current best known approximation ratio of 0.87856 is given by the celebrated algorithm of Goemans and Williamson [12], which is based on a ingenious application of semi-definite programming. From a purely combinatorial perspective, it is of interest to determine

[*]Department of Mathematics, MIT, Cambridge, MA 02139. E-mail: cb_lee@math.mit.edu.

[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213. E-mail: ploh@cmu.edu.
Research supported by NSF grant DMS-1201380, an NSA Young Investigators Grant and a USA-Israel BSF Grant.

[‡]Department of Mathematics, UCLA, Los Angeles, CA 90095. Email: bsudakov@math.ucla.edu. Research supported in part by NSF grant DMS-1101185, by AFOSR MURI grant FA9550-10-1-0569 and by a USA-Israel BSF grant.

best-possible bounds for parameters of optimal partitions. Edwards [8] improved on the folklore bound and proved that

$$\text{Max Cut} \geq \left\lceil \frac{m}{2} + \sqrt{\frac{m}{8} + \frac{1}{64}} - \frac{1}{8} \right\rceil,$$

which is tight, e.g., for complete graphs.

Empowered by the growth of probabilistic techniques, a new class of *judicious* partitioning results has emerged. In these problems, one simultaneously optimizes several properties, in contrast to the classical problems such as Max Cut where one attempts to optimize a single parameter. A classic result in this area is a theorem of Bollobás and Scott [4] which asserts that every m-edge graph has a bipartition $V = V_1 \cup V_2$ of its vertex set in which

$$e(V_1, V_2) \geq \left\lceil \frac{m}{2} + \sqrt{\frac{m}{8} + \frac{1}{64}} - \frac{1}{8} \right\rceil$$

and

$$\max\{e(V_1), e(V_2)\} \leq \frac{m}{4} + \sqrt{\frac{m}{32} + \frac{1}{256}} - \frac{1}{16}.$$

Note that their result simultaneously optimizes three parameters: the number of edges across the partition (matching the Edwards bound), and the number of edges inside each V_i. We direct the interested reader to any of [1, 2, 3, 5, 7, 13, 16, 17, 18, 19] (by no means a comprehensive list), or to either of the surveys [6, 21] for more background on the judicious partitioning literature.

In this paper we study essentially the most fundamental judicious partitioning problem for directed graphs. A directed graph is a pair (V, E) where V is a set of vertices, and E is a set of distinct edges $\overrightarrow{uv}$, where $u \neq v$. We disallow loops and multiple edges, but do allow both $\overrightarrow{uv}$ and $\overrightarrow{vu}$ to be present. In this context, any cut $V = V_1 \cup V_2$ is most naturally associated with two parameters: the number of edges from V_1 to V_2, and the number of edges from V_2 to V_1. Therefore, in contrast to undirected graphs, where Max Cut only needs to optimize a single parameter (the total number of crossing edges), in directed graphs one can measure the size of a cut in each direction. Therefore, the judicious analogue of the Max Cut problem would seek a bipartition which had many edges crossing in both directions.

Although it is easy to guarantee a partition with at least $1/4$ of the edges in a single direction, one immediately notices that the problem as stated above has the following issue. If the digraph is a star with all edges oriented from a central vertex, then regardless of the bipartition, one direction would always have zero edges. This is similar to the issue which arose in the judicious bisection problem in graphs (see [17]), and in both cases, it can be resolved by imposing a minimum-degree condition. The following natural question appears in the survey of Scott [21].

Problem. *Let d be a positive integer. What is the maximum constant c_d such that every m-edge directed graph of minimum outdegree at least d admits a bipartition $V = V_1 \cup V_2$ of its vertex set in which*

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \geq c_d \cdot m?$$

For $d = 1$, consider the graph $K_{1,n-1}$ and add a single edge inside the part of size $n-1$. This graph can be oriented so that the minimum outdegree is 1 and $\min\{e(V_1, V_2), e(V_2, V_1)\} \leq 1$ for every partition $V = V_1 \cup V_2$. This is because we have an cyclically oriented triangle, with lots of edges all pointing in to one of the vertices of that triangle. Then all of those edges will only contribute to $e(V_1, V_2)$ or $e(V_2, V_1)$, depending on whether the apex is in V_2 or V_1, respectively. Altogether, the other edges will only contribute a total of at most one edge back in the other direction. Hence we see that $c_1 = 0$.

For $d \geq 2$, first orient the edges of the complete graph K_{2d-1} along an Eulerian circuit. In this way we obtain a directed graph with $2d-1$ vertices, and all outdegrees equal to $d-1$. Moreover, in every bipartition of its vertex set, the number of edges crossing in each direction is exactly the same (this is easily seen by following the Eulerian circuit). Hence in every bipartition of its vertex set, the maximum number of edges in any direction is at most $\frac{d(d-1)}{2}$. Now consider the directed graph where we take k vertex disjoint copies of K_{2d-1} oriented as above, and a single vertex disjoint copy of K_{2d+1} oriented in a similar manner. Fix a vertex v_0 of K_{2d+1}, and add edges so that all the vertices belonging to the copies of K_{2d-1} are in-neighbors of v_0. This graph has minimum outdegree d, and its number of edges is

$$\begin{aligned} m &= k(d-1)(2d-1) + d(2d+1) + k(2d-1) \\ &= kd(2d-1) + d(2d+1). \end{aligned}$$

Moreover, for every partition $V = V_1 \cup V_2$ of its vertex set with $v_0 \in V_1$, we have

$$e(V_1, V_2) \leq k\frac{d(d-1)}{2} + \frac{d(d+1)}{2} = \frac{d-1}{2(2d-1)}m + \frac{d^2}{2d-1}.$$

Hence this graph shows that $c_d \leq \frac{d-1}{2(2d-1)}$. Our main theorem asserts that for $d = 2, 3$ this bound is asymptotically best possible.

Theorem 1.1. *For $d = 2, 3$, every directed graph of minimum outdegree at least d admits a bipartition $V = V_1 \cup V_2$ of its vertex set for which*

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \left(\frac{d-1}{2(2d-1)} + o(1)\right) m.$$

Thus $c_2 = \frac{1}{6} + o(1)$ and $c_3 = \frac{1}{5} + o(1)$.

Based on the constructions above and Theorem 1.1 we make the following conjecture.

Conjecture 1.2. *Let d be an integer satisfying $d \geq 4$. Every directed graph of minimum outdegree at least d admits a bipartition $V = V_1 \cup V_2$ of its vertex set for which*

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \left(\frac{d-1}{2(2d-1)} + o(1)\right) m.$$

This paper is organized as follows. In Sections 2 and 3, we prove the core results which drive the proof of our main theorem. We prove the $d = 2$ case of the main theorem in Section 4, and the $d = 3$ case in Section 5. The final section contains some concluding remarks, with a discussion of the obstacles that remain in the cases $d \geq 4$.

Notation. Graphs $G = (V, E)$ and directed graphs $D = (V, E)$ are given by pairs of vertex sets and edge sets. All of our objects will have no loops (endpoints of edges are distinct), and no multiple edges (edges are all distinct), although directed graphs are permitted to have antiparallel pairs $\overrightarrow{uv}$,

$\overrightarrow{vu}$. A directed graph is connected if the underlying undirected graph is connected. For an undirected graph $G = (V, E)$ and two vertex subsets X and Y, we let $e(X, Y) = |\{xy : x \in X, y \in Y, xy \in E\}|$. For a directed graph $D = (V, E)$ and a vertex v, let $d^-(v)$ and $d^+(v)$ be the number of v's in-neighbors and out neighbors, respectively, and let $d(v) = d^-(v) + d^+(v)$ be the total degree of v. Note that $d(v)$ can potentially be as high as $2(n-1)$ because edges in both directions are permitted between each pair. For a vertex subset A, let $e^-(A) = |\{\overrightarrow{xa} : a \in A, x \notin A, \overrightarrow{xa} \in E(D)\}|$ and $e^+(A) = |\{\overrightarrow{ax} : a \in A, x \notin A, \overrightarrow{ax} \in E(D)\}|$. Let $e(A) = e^-(A) + e^+(A)$. For two vertex subsets X and Y in a digraph, let $e(X, Y) = |\{\overrightarrow{xy} : x \in X, y \in Y, \overrightarrow{xy} \in E\}|$. For a vertex set A, we let $D[A]$ denote the induced subgraph of D on A. Since the majority of our results are asymptotic in nature, we will implicitly ignore rounding effects whenever these effects are of smaller order than our error terms. For two functions $f(n)$ and $g(n)$, we write $f(n) = o(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$. We often use subscripts such as $\varepsilon_{3.1}$ to indicate that ε is the constant coming from Theorem/Corollary/Lemma 3.1.

2 Basic probabilistic approach

A simple, yet powerful, method of obtaining an effective partition is to apply randomness, by independently placing each vertex to each side with some specified probability. Even though this method is not powerful enough to immediately solve our main problem, it serves as a useful starting point, and in fact provides a sufficiently good partition for some range of the parameter space. In this section, we develop this idea in a slightly more general form, keeping in mind later applications. The following lemma estimates the number of edges across a random partition using the first and second moment methods.

Lemma 2.1. *Let $D = (V, E)$ be a directed graph with m edges. Let $0 \le p \le 1$ be a real number. Suppose that we are also given a subset $A \subset V$, with partition $A = A_1 \cup A_2$. Let $B = V \setminus A$ and consider a random bipartition $B = B_1 \cup B_2$ obtained by independently placing each vertex of B in B_1 with probability p, and in B_2 with probability $1-p$. Let $V_1 = A_1 \cup B_1$ and $V_2 = A_2 \cup B_2$. Then*

$$\begin{aligned} \mathbb{E}[e(V_1, V_2)] &= e(A_1, A_2) + (1-p)\cdot e^+(A_1) + p \cdot e^-(A_2) + p(1-p)\cdot e(B) \quad \text{and} \\ Var[e(V_1, V_2)] &< 2m \cdot \max_{v \in B} d(v). \end{aligned}$$

Proof. For each edge $e = vw$ of the directed graph D, let $\mathbf{1}_e$ be the indicator random variable of the event that the edge e becomes an edge from V_1 to V_2. We have

$$e(V_1, V_2) = \sum_e \mathbf{1}_e.$$

Note that

$$\mathbb{E}[\mathbf{1}_e] = \begin{cases} 1 & \text{if } v \in A_1, w \in A_2, \\ 1-p & \text{if } v \in A_1, w \in B, \\ p & \text{if } v \in B, w \in A_2, \\ p(1-p) & \text{if } v \in B, w \in B, \\ 0 & \text{otherwise.} \end{cases}$$

The claim on the expected value of $e(V_1, V_2)$ immediately follows from linearity of expectation.

To estimate the variance of $e(V_1, V_2)$, it suffices to focus on the edges $e = vw$ for which ($v \in A_1, w \in B$), ($v \in B, w \in A_2$), or ($v, w \in B$), as all other edges have constant contribution towards $e(V_1, V_2)$. Let $E_{1,2}$ be the set of such edges. We have

$$\mathrm{Var}\left[\sum_e \mathbf{1}_e\right] = \mathrm{Var}\left[\sum_{e \in E_{1,2}} \mathbf{1}_e\right] = \sum_{e \in E_{1,2}} \mathrm{Var}\left[\mathbf{1}_e\right] + \sum_{e,e' \in E_{1,2}, e \neq e'} \mathrm{Cov}[\mathbf{1}_e, \mathbf{1}_{e'}].$$

For $e \in E_{1,2}$, we have $\mathrm{Var}\left[\mathbf{1}_e\right] \leq \mathbb{E}[\mathbf{1}_e] \leq 1$. For the second term, we have $\mathrm{Cov}[\mathbf{1}_e, \mathbf{1}_{e'}] = 0$ if e and e' do not share a vertex. If $e, e' \in E_{1,2}$ go between the same pair of endpoints, but in opposite directions, then they can never simultaneously contribute to $e(V_1, V_2)$, and hence $\mathrm{Cov}[\mathbf{1}_e, \mathbf{1}_{e'}] \leq 0$. Furthermore, if $e, e' \in E_{1,2}$ share an endpoint in A but have distinct endpoints in B, then

$$\mathrm{Cov}[\mathbf{1}_e, \mathbf{1}_{e'}] = \mathbb{E}[\mathbf{1}_e \mathbf{1}_{e'}] - \mathbb{E}[\mathbf{1}_e]\mathbb{E}[\mathbf{1}_{e'}] = 0.$$

Hence the only positive contributions to $\mathrm{Cov}[\mathbf{1}_e, \mathbf{1}_{e'}]$ come when e and e' share a vertex in B. Since $\mathrm{Cov}[\mathbf{1}_e, \mathbf{1}_{e'}] \leq \mathbb{E}[\mathbf{1}_e \mathbf{1}_{e'}] \leq 1$, we have

$$\begin{aligned}
\sum_{e,e' \in E_{1,2}, e \neq e'} \mathrm{Cov}[\mathbf{1}_e, \mathbf{1}_{e'}] &\leq \sum_{v \in B} d(v)(d(v) - 1) \\
&\leq \left(\sum_{v \in B} d(v)\right)\left(\max_{v \in B} d(v) - 1\right) \\
&\leq 2m \left(\max_{v \in B} d(v) - 1\right).
\end{aligned}$$

Thus

$$\mathrm{Var}\left[\sum_{e \in E_{1,2}} \mathbf{1}_e\right] \leq \left(\sum_{e \in E_{1,2}} 1\right) + 2m\left(\max_{v \in B} d(v) - 1\right) < 2m \cdot \max_{v \in B} d(v).$$

□

This implies the following lemma.

Lemma 2.2. *Let $D = (V, E)$ be a given directed graph with m edges. Let p be a real satisfying $p \in [0, 1]$, and ε be a positive real. Suppose that a subset $A \subset V$ and its partition $A = A_1 \cup A_2$ are given, and let $B = V \setminus A$. Further suppose that $\max_{v \in B} d(v) \leq \frac{\varepsilon^2}{4} m$. Then there exists a partition $V_1 \cup V_2$ for which*

$$\begin{aligned}
&e(V_1, V_2) \geq e(A_1, A_2) + (1-p) \cdot e^+(A_1) + p \cdot e^-(A_2) + p(1-p) \cdot e(B) - \varepsilon m \quad \textit{and} \\
&e(V_2, V_1) \geq e(A_2, A_1) + p \cdot e^+(A_2) + (1-p) \cdot e^-(A_1) + p(1-p) \cdot e(B) - \varepsilon m.
\end{aligned}$$

Proof. Let $V_1 \cup V_2$ be the partition obtained by placing each vertex in B independently in V_1 or V_2, with probability p and $1 - p$, respectively. Let

$$m_{1,2} = e(A_1, A_2) + (1-p) \cdot e^+(A_1) + p \cdot e^-(A_2) + p(1-p) \cdot e(B),$$

and

$$m_{2,1} = e(A_2, A_1) + p \cdot e^+(A_2) + (1-p) \cdot e^-(A_1) + p(1-p) \cdot e(B).$$

By Lemma 2.1 and Chebyshev's inequality,

$$\mathbf{P}\Big(e(V_1, V_2) \geq m_{1,2} - \varepsilon m\Big) \leq \frac{\mathrm{Var}\left[e(V_1, V_2)\right]}{\varepsilon^2 m^2} < \frac{2m \cdot \max_{v \in B} d(v)}{\varepsilon^2 m^2} \leq \frac{(\varepsilon^2/2)m^2}{\varepsilon^2 m^2} = \frac{1}{2}.$$

Similarly, we have

$$\mathbf{P}\Big(e(V_1, V_2) \geq m_{2,1} - \varepsilon m\Big) < \frac{1}{2}.$$

Hence there exists a partition $V = V_1 \cup V_2$ for which $e(V_1, V_2) \geq m_{1,2} - \varepsilon m$ and $e(V_2, V_1) \geq m_{2,1} - \varepsilon m$ both hold. $\square$

The next statement is an immediate corollary of the lemma.

Proposition 2.3. For $\varepsilon > 0$, let $D = (V, E)$ be an n-vertex directed graph with m edges, such that all degrees are at most $\frac{\varepsilon^2}{4}m$, or $m \geq 8\varepsilon^{-2}n$. Then there exists a partition $V_1 \cup V_2$ for which both $e(V_1, V_2)$ and $e(V_2, V_1)$ are at least $\left(\frac{1}{4} - \varepsilon\right)m$.

Note that $m \geq 8\varepsilon^{-2}n$ implies that the maximum degree is at most $\frac{\varepsilon^2}{4}m$, since all degrees of a directed graph are at most $2n$. Hence Proposition 2.3 indeed is an immediate corollary of the lemma.

3 Large bipartition

As noticed in [10, 20], the results that Edwards proved in [9] implicitly imply that connected graphs with n vertices and m edges admit a bipartition of size at least

$$\frac{m}{2} + \frac{n-1}{4}.$$

In fact, for even integers n we have $\lceil \frac{m}{2} + \frac{n-1}{4} \rceil \geq \frac{m}{2} + \frac{n}{4}$, and thus the above bound implies that a graph with τ odd components admit a bipartition of size at least

$$\frac{m}{2} + \frac{n-\tau}{4}.$$

A bisection of a graph is a bipartition of its vertex set in which the number of vertices in the two parts differ by at most one. In [17], we extended the bound above to bisections and proved that every graph with n vertices, m edges, τ odd components, and maximum degree Δ admits a bisection of size at least

$$\frac{m}{2} + \frac{n - \max\{\tau, \Delta - 1\}}{4}.$$

We then developed a randomized algorithm which asymptotically achieves the bound above (and some other estimates as well), based on the proof of this theorem. This algorithm turned out to be a powerful new tool in obtaining a judicious bisection result. In this paper, we adjust the randomized algorithm for directed graphs. The following theorem is one of the main tools of this paper.

Theorem 3.1. *Given any real constants $C, \varepsilon > 0$, there exist $\gamma, n_0 > 0$ for which the following holds. Let $D = (V, E)$ be a given directed graph with $n \geq n_0$ vertices and at most Cn edges, and let $A \subset V$ be a set of at most γn vertices which have already been partitioned into $A_1 \cup A_2$. Let $B = V \setminus A$, and suppose that every vertex in B has degree at most γn (with respect to the full D). Let τ be the* *number of odd components in $D[B]$. Then, there is a bipartition $V = V_1 \cup V_2$ with $A_1 \subset V_1$ and $A_2 \subset V_2$, such that both*

$$e(V_1, V_2) \geq e(A_1, A_2) + \frac{e(A_1, B) + e(B, A_2)}{2} + \frac{e(B)}{4} + \frac{n-\tau}{8} - \varepsilon n$$
$$e(V_2, V_1) \geq e(A_2, A_1) + \frac{e(B, A_1) + e(A_2, B)}{2} + \frac{e(B)}{4} + \frac{n-\tau}{8} - \varepsilon n\,.$$

Informally, Theorem 3.1 asserts that if the number of edges and the maximum degree satisfy certain conditions, then we can in fact obtain an additive term of $\frac{n-\tau}{8}$ over the expected number of edges in a purely random bipartition. Consider the directed graphs given in Section 1 which achieve the upper bound of Theorem 1.1. In the notation of Theorem 3.1, A is the set whose only element is the vertex of degree $n-1$, and B is the set of other vertices. Note that the induced subgraph on B consists of components of odd size. These graphs are designed to maximize τ, and hence these graphs will turn out to be the graphs which give the worst bound in Theorem 3.1. The proof of this theorem is somewhat involved, although it is similar to the that of the corresponding theorem in [17]. The rest of this section is devoted to its proof.

3.1 Decomposing the graph

We start with a technical lemma which will provide structural information about the underlying undirected (simple) graph obtained by ignoring edge orientations and removing redundant parallel edges when edges in both directions appear between pairs of vertices. A *star* is a bipartite graph on n vertices consisting of a unique vertex of degree $n-1$, and $n-1$ other vertices of degree one. We refer to the unique vertex of degree $n-1$ as the *apex*. The following lemma decomposes an undirected graph (with no loops or multiple edges) into induced stars plus some leftover vertices.

Lemma 3.2. *Let ε and C be arbitrary positive reals. Let G be an undirected graph with n vertices, $m \leq Cn$ edges, maximum degree Δ, and τ odd components. Then there exists a partition $V = T_1 \cup T_2 \cup \cdots \cup T_s \cup U$ of its vertex set such that*

(i) *each T_i induces a star, and $2 \leq |T_i| \leq \Delta + 1$,*

(ii) *all but at most one non-apex vertex in each T_i has degree (in the full graph) at most $\frac{2C}{\varepsilon}$, and*

(iii) *U is an independent set of order $|U| \leq \tau + \varepsilon n$.*

The lemma above is implicitly proved in [17]. A similar lemma also appears in the paper of Erdős, Gyárfas, and Kohayakawa [10], but their bound is in terms of the number of connected components, not the number of odd components. In order to prove this lemma, we first take a maximum matching. Afterwards, for the leftover vertices which are not covered by the matching, we attempt to find an edge in the matching with which the vertex will create an induced star. By systematically assigning each leftover vertex in this way, we will eventually obtain the partition described in Lemma 3.2. In order to provide the full details for this argument, it is convenient to introduce the following concept.

Definition 3.3. *Let $\{e_1, \ldots, e_s\}$ be the edges of a maximum matching in a graph $G = (V, E)$, and let W be the set of vertices not in the matching. With respect to this fixed matching, say that a vertex v in a matching edge e_i is a* free neighbor *of a vertex $w \in W$ if w is adjacent to v, but w is not adjacent to the other endpoint of e_i. In this case, we also say that e_i is a* free neighbor *of w. Call a vertex $w \in W$ a* free vertex *if it has at least one free neighbor.*

A *tight component* is a connected component T such that for every $v \in T$, the subgraph induced by $T \setminus \{v\}$ contains a perfect matching, and every perfect matching of $T \setminus \{v\}$ has the property that no edge of the perfect matching has exactly one endpoint adjacent to v. Note that a tight component is necessarily an odd component. The following lemma delineates the relationship between non-free vertices and tight components.

Lemma 3.4. *Let $\{e_1, \ldots, e_s\}$ be the edges of a maximum matching in an undirected graph $G = (V, E)$, and let W be the set of vertices not in the matching. Further assume that among all matchings of maximum size, we have chosen one which maximizes the number of free vertices in W. Then, every tight component contains a distinct non-free vertex of W, and all non-free W-vertices are covered in this way (there is a bijective correspondence).*

Proof. The matching $\{e_1, \ldots, e_s\}$ must be maximal within each connected component. One basic property of a tight component is that it contains an almost-perfect matching which misses only one vertex. Consequently, by maximality, in every tight component T, $\{e_1, \ldots, e_s\}$ must miss exactly one vertex $w \in W$. Furthermore, the second property of a tight component is that w must have either 0 or 2 neighbors in each edge e_i in T (and w must have 0 neighbors in each edge e_j not in T, since T is the connected component containing w). Therefore, the unique vertex w is in fact a non-free W-vertex contained in T.

The remainder of the proof concentrates on the more substantial part of the claim, which is that each non-free W-vertex is contained in some tight component. Consider such a vertex w, and let T be a maximal set of vertices which (i) contains w, (ii) induces a connected graph which is a tight component, and (iii) does not cut any e_i. Since the set $\{w\}$ satisfies (i)–(iii), our optimum is taken over a non-empty set, and so T exists.

If T is already disconnected from the rest of the graph, then we are done. So, consider a vertex $v \notin T$ which has a neighbor $v' \in T$. If $v \in W$, then we can modify our matching by taking the edge vv', and changing the matching within T by using property (ii) to generate a new matching of $T \setminus \{v'\}$. This will not affect the matching outside of $T \cup \{v\}$, because property (iii) insulates the adjustments within T from the rest of the matching outside. We would then obtain a matching with one more edge, contradicting maximality. Therefore, all vertices $v \notin T$ which have neighbors in T also satisfy $v \notin W$.

Let us then consider a vertex $v_1 \notin T \cup W$ with a neighbor $v' \in T$. We now know that v_1 must be covered by a matching edge; let v_2 be the other endpoint of that edge. By (iii), we also have $v_2 \notin T$. Note that v_2 cannot have a neighbor $w' \in W \setminus T$, or else we could improve our matching by replacing v_1v_2 with $w'v_2$ and v_1v', and then using (ii) to take a perfect matching of $T \setminus \{v'\}$.

Our next claim is that v_2 must be adjacent to v' as well. Indeed, assume for contradiction that this is not the case. Then, consider modifying our matching by replacing v_1v_2 with the edge v_1v' and changing the matching within T by using (ii) to generate a new matching of $T \setminus \{v'\}$. As before, (iii) ensures that the result is still a matching. This time, the new matching has the same size as the original one, but with more free W-vertices (note that the vertex v_2 replaced the vertex w in the set W). To see this, observe that v_2 is now unmatched and free because it is adjacent to v_1 but not v'. Previously, the only W-vertex inside T was our original w, which we assumed to be non-free in the first place. Also, no other vertices outside of T changed from being free to non-free, because we already showed that no W-vertices outside of T were adjacent to $T \cup \{v_2\}$, and so any vertex that was free by virtue of its adjacency with v_1 but not v_2 is still free because it is not adjacent to v' either. This contradiction to maximality establishes that v_2 must be adjacent to v'

We now have v_1, v_2, and v' all adjacent to each other, and no vertices of $W \setminus T$ are adjacent to $T \cup \{v_1, v_2\}$. Our argument also shows that for any $v'' \in T$ which is adjacent to one of v_1 or v_2, it also must be adjacent to the other. Our final objective is to show that $T' = T \cup \{v_1, v_2\}$ also satisfies (i)–(iii), which would contradict the maximality of T. Properties (i) and (iii) are immediate, so it remains to verify the conditions of a tight component. Since T is tight and v_1v_2 is an edge, $T' \setminus \{u\}$ has a perfect matching for any $u \in T$. The tightness of T and the pairwise adjacency of v_1, v_2, and v' also produce this conclusion if $u \in \{v_1, v_2\}$. It remains to show that for any $u \in T'$ and any perfect matching of $T' \setminus \{u\}$, u has either 0 or 2 vertices in every matching edge. But if this were not the case, then we could replace the matching within T' with the violating matching of $T' \setminus \{u\}$. The two matchings would have the same size, but u would become a free vertex. No other vertex of W is adjacent to T' by our observation above, so the number of free vertices would increase, contradicting the maximality of our initial matching. Therefore, T' induces a tight component, contradicting the maximality of T. We conclude that T must have been disconnected from the rest of the graph, as required. $\square$

We are now ready to prove Lemma 3.2.

Proof of Lemma 3.2. Start by taking a maximum matching $\{e_1, \ldots, e_s\}$ which secondarily maximizes the number of free vertices in $W = V \setminus \{e_1, \ldots, e_s\} = \{w_1, \ldots, w_r\}$, so that we can apply Lemma 3.4. By maximality, W is an independent set. Let $U \subset W$ be the set of vertices which are either not free, or have degree at least $\frac{2C}{\varepsilon}$. Since all tight components have odd order, by Lemma 3.4, there are at most τ non-free vertices. On the other hand, since there are at most Cn edges total, there are at most εn vertices which have degree at least $\frac{2C}{\varepsilon}$. Hence $|U| \leq \tau + \varepsilon n$, giving (iii).

We now construct the induced stars. Let T_i be the union of the set of vertices of e_i and the set of vertices $w \in W \setminus U$ for which i is the minimum index where e_i is a free neighbor of w. This is a partition $V = T_1 \cup T_2 \cup \cdots \cup T_s \cup U$ because each vertex in $W \setminus U$ has at least one free neighbor by construction. Since our matching is maximal, there cannot be any $e_i = vv'$ such that vw and $v'w'$ are both edges to distinct vertices $w, w' \in W$. So, if two vertices $w, w' \in W$ each have a free neighbor in an e_i, then their free neighbor is the same vertex. This, together with the fact that W is an independent set, implies that each T_i induces a star, giving (i). Finally, all vertices in each T_i outside of e_i have degree at most $\frac{2C}{\varepsilon}$, and thus (ii) holds. $\square$

3.2 Randomized algorithm

In this subsection, we use the following martingale concentration result (essentially the Hoeffding-Azuma inequality) to control the performance of the randomized partitioning algorithm at the heart of Theorem 3.1.

Theorem 3.5. *(Corollary 2.27 in [15].) Given real numbers $\lambda, C_1, \ldots, C_n > 0$, let $f : \{0,1\}^n \to \mathbb{R}$ be a function satisfying the following Lipschitz condition: whenever two vectors $z, z' \in \{0,1\}^n$ differ only in the i-th coordinate (for any i), we always have $|f(z) - f(z')| \leq C_i$. Suppose $X_1, X_2, \ldots, X_n$ are independent random variables, each taking values in $\{0,1\}$. Then, the random variable $Y = f(X_1, \ldots, X_n)$ satisfies*

$$\mathbf{P}\left(|Y - \mathbb{E}[Y]| \geq \lambda\right) \leq 2 \exp\left\{-\frac{\lambda^2}{2\sum C_i^2}\right\}.$$

We are now ready to prove Theorem 3.1, which will be the core result for our main theorem.

Proof of Theorem 3.1. Without loss of generality, assume that $C > 1$, $\epsilon < 1$, and n is sufficiently large. Apply Lemma 3.2 to the underlying undirected (simple) graph induced by B, and obtain a partition $B = T_1 \cup \cdots \cup T_s \cup U$. Note that since we are assured that the full digraph D contains at most Cn edges, we can actually establish in part (ii) of Lemma 3.2 that the degree bound of $\frac{2C}{\varepsilon}$ holds with respect to total degrees (in- plus outdegrees) in the full digraph D, not just in the undirected simple graph on B. (When constructing $U \subset W$ in the proof of that lemma, one may absorb all vertices with degree greater than $\frac{2C}{\varepsilon}$ with respect to the whole graph, not just in the underlying undirected graph on B.) So, we may assume that each T_i has at most one non-apex vertex with full D-degree greater than $\frac{2C}{\varepsilon}$, and we still have $|U| \le \tau + \varepsilon n$. Let v_i denote the apex vertex of tree T_i, arbitrarily distinguishing an apex if T_i has only two vertices. We now randomly construct a bipartition $V = V_1 \cup V_2$ by placing each A_i in V_i, and partitioning each T_i by independently placing each apex v_i on a uniformly random side, and then placing the rest of $T_i \setminus \{v_i\}$ on the other side. Each remaining vertex (from the set U) is independently placed on a uniformly random side.

Define the random variables $Y_1 = e(V_1, V_2)$ and $Y_2 = e(V_2, V_1)$. For an edge $e = \overrightarrow{vw}$ of the digraph, let $\mathbf{1}_e$ be the indicator random variable of the event that $v \in V_1$ and $w \in V_2$. Thus $Y_1 = \sum_e \mathbf{1}_e$ and

$$\mathbb{E}[Y_1] = \sum_e \mathbb{E}[\mathbf{1}_e].$$

We have $\mathbb{E}[\mathbf{1}_e] = 1$ if $v \in A_1$ and $w \in A_2$, and $\mathbb{E}[\mathbf{1}_e] = \frac{1}{2}$ if either $v \in A_1$ and $w \in B$, or $v \in B$ and $w \in A_2$. For edges in $D[B]$, the gain comes from edges e in the digraph which correspond to edges in the stars T_i in the underlying undirected graph on B: there, we have $\mathbb{E}[\mathbf{1}_e] = \frac{1}{2}$, while all other edges in $D[B]$ give the regular $\mathbb{E}[\mathbf{1}_e] = \frac{1}{4}$. Note that the total number of edges in the stars induced by the sets T_i is at least $\frac{|B|-|U|}{2} \ge \frac{(n-\gamma n)-(\tau+\varepsilon n)}{2}$. Therefore,

$$\mathbb{E}[Y_1] \ge e(A_1, A_2) + \frac{e(A_1, B) + e(B, A_2)}{2} + \frac{e(B)}{4} + \frac{1}{4} \cdot \frac{(n-\gamma n) - (\tau + \varepsilon n)}{2}. \tag{3.1}$$

Similarly, we have

$$\mathbb{E}[Y_2] \ge e(A_2, A_1) + \frac{e(B, A_1) + e(A_2, B)}{2} + \frac{e(B)}{4} + \frac{n-\tau}{8} - \frac{(\varepsilon+\gamma)n}{8}.$$

For each $1 \le i \le s$, let C_i be the sum of the degrees of all vertices in T_i. Clearly, flipping the assignment of v_i cannot affect Y_1 by more than C_i. Also, flipping the assignment of any $w \in U$ cannot change Y_1 by more than the degree $d(w)$ of w. Therefore, if we define

$$L = \sum_{i=1}^{s} \left(\sum_{u \in T_i} d(u) \right)^2 + \sum_{w \in U} d(w)^2 ,$$

the Hoeffding-Azuma inequality (Theorem 3.5) gives

$$\mathbf{P}\left(Y_1 \le \mathbb{E}[Y_1] - \frac{\varepsilon n}{2}\right) \le 2 \exp\left\{ -\frac{\varepsilon^2 n^2}{8L} \right\} . \tag{3.2}$$

Let us now control L. Each T_i induces a star with apex $v_i \in e_i$. Let u_i be the unique non-apex vertex with degree greater than $\frac{2C}{\varepsilon}$ (if no such vertex exists, then let u_i be an arbitrary non-apex vertex).

Since every vertex in T_i other than v_i and u_i has degree at most $\frac{2C}{\varepsilon}$ in the full D, we see that

$$\sum_{u \in T_i} d(u) \leq d(v_i) + d(u_i) + (d(v_i) - 1) \cdot \frac{2C}{\varepsilon} \leq (d(u_i) + d(v_i)) \frac{4C}{\varepsilon},$$

and hence

$$\begin{aligned} L &\leq \frac{16C^2}{\varepsilon^2} \sum_{i=1}^{s} (d(u_i) + d(v_i))^2 + \sum_{w \in U} d(w)^2 \\ &\leq \frac{32C^2}{\varepsilon^2} \sum_{i=1}^{s} (d(u_i)^2 + d(v_i)^2) + \sum_{w \in U} d(w)^2 \\ &\leq \frac{32C^2}{\varepsilon^2} \sum_{v \in B} d(v)^2 \leq \frac{32C^2}{\varepsilon^2} (\gamma n) \sum_{v \in B} d(v) \\ &\leq \frac{32C^2}{\varepsilon^2} (\gamma n)(2Cn), \end{aligned}$$

where we used that $d(v) \leq \gamma n$ for all $v \in B$, and the degree sum of D is at most $2Cn$. Therefore, we choose $\gamma = \frac{\varepsilon^4}{1024C^3}$, so that $L \leq \frac{\varepsilon^2 n^2}{16}$. Substituting this into (3.2), we conclude that

$$\mathbf{P}\left(Y_1 \leq \mathbb{E}[Y_1] - \frac{\varepsilon n}{2}\right) \leq 2e^{-2} < \frac{1}{2},$$

as desired. By symmetry, we have $\mathbf{P}(Y_2 \leq \mathbb{E}[Y_2] - \frac{\varepsilon n}{2}) < \frac{1}{2}$ as well. Hence, there is a partition $V = V_1 \cup V_2$ with the properties that $Y_1 \geq \mathbb{E}[Y_1] - \frac{\varepsilon n}{2}$ and $Y_2 \geq \mathbb{E}[Y_2] - \frac{\varepsilon n}{2}$, which is a desired partition. $\square$

4 Minimum outdegree two

In this section we prove the $d = 2$ case of Theorem 1.1. Suppose that $D = (V, E)$ is a given directed graph of minimum outdegree at least 2 with n vertices and m edges. Our goal is to find a partition $V = V_1 \cup V_2$ for which $\min\{e(V_1, V_2), e(V_2, V_1)\} \geq (\frac{1}{6} - \varepsilon)m$. Throughout the proof we tacitly assume that the number of vertices n is large enough.

Suppose that $m \geq 1152n$. Then by applying Proposition 2.3 with $\varepsilon = \frac{1}{12}$, we obtain a partition $V = V_1 \cup V_2$ for which

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \frac{m}{4} - \frac{m}{12} = \frac{m}{6}.$$

Hence it suffices to consider the case when $m < 1152n$. Since the minimum outdegree is at least two, we see that $2n \leq m < 1152n$.

Let A be the set of *large* vertices, which are defined to be the vertices with total degree at least $n^{3/4}$, and let $B = V \setminus A$. Note that

$$|A| \cdot n^{3/4} \leq 2m < 2304n,$$

from which it follows that $|A| \leq 2304n^{1/4} \leq \varepsilon n$, and $e(A) \leq 2304^2 n^{1/2} \leq \frac{\varepsilon m}{2}$. For sake of simplicity we remove all the edges within A, and update m to be the new total number of edges in the digraph. In terms of this new m, it suffices to obtain a partition $V = V_1 \cup V_2$ for which

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \frac{m}{6} - \frac{\varepsilon m}{2}.$$

Indeed, this will be a desired partition even after recovering the removed edges, since at most $\frac{\varepsilon m}{2}$ edges were removed from within A. So, for the remainder of this section, we focus on the case when A induces no edges, and we let $m_A = e(A,B)$ and $m_B = e(B)$. Note that now $m = m_A + m_B$.

4.1 Partition of large vertices

Given a partition $A = A_1 \cup A_2$, define

$$\Theta = \big(e^+(A_1) + e^-(A_2)\big) - \big(e^-(A_1) + e^+(A_2)\big). \tag{4.1}$$

We call Θ the *gap* of the partition. Consider the following greedy algorithm to partition A. Since we now assume that A induces no edges, each vertex $v \in A$ contributes one of $\pm(d^+(v) - d^-(v))$ to the expression in (4.1), and thus each contribution is bounded in magnitude by n. Process the vertices of A in an arbitrary sequential order $v_1, v_2, \ldots$, and when assigning v_i to a side, choose the side which makes the sign of v_i's contribution to (4.1) opposite to the sign of the cumulative contribution of all previous $v_1, \ldots, v_{i-1}$ thus far. Since each contribution is bounded in magnitude by n, the final gap Θ of this greedy partition will also be bounded in magnitude by n. Now, let $A_1 \cup A_2$ be a partition of A which minimizes $|\Theta|$, and without loss of generality, assume that $\Theta \geq 0$. The greedy partition provides the upper bound $\Theta \leq n$.

Since $\max_{v \in B} d(v) \leq n^{3/4} \leq \frac{\varepsilon^2}{16} m$, by Lemma 2.2 with $p = \frac{1}{2}$ and $\varepsilon_{2.2} = \frac{\varepsilon}{2}$, there exists a partition $V = V_1 \cup V_2$ for which

$$\begin{aligned}
\min\{e(V_1,V_2), e(V_1,V_2)\} &\geq \frac{1}{2}\min\{e^+(A_1) + e^-(A_2), e^-(A_1) + e^+(A_2)\} + \frac{1}{4}e(B) - \frac{\varepsilon}{2}m \\
&= \frac{1}{2} \cdot \frac{m_A - \Theta}{2} + \frac{1}{4}m_B - \frac{\varepsilon}{2}m = \frac{m - \Theta}{4} - \frac{\varepsilon}{2}m.
\end{aligned}$$

Hence if $\Theta \leq \frac{m}{3}$, then we obtain a desired partition. Thus we assume for the remainder that

$$\Theta > \frac{m}{3}. \tag{4.2}$$

For a vertex $v \in A$, we let the *in-surplus* of v be $s^-(v) = d^-(v) - d^+(v)$, and the *out-surplus* of v be $s^+(v) = d^+(v) - d^-(v)$. Let the *surplus* of v be $s(v) = \max\{s^+(v), s^-(v)\}$. Note that the in-surplus and out-surplus differ only in their sign, and the surplus is equal to their magnitude. Call a vertex $v \in A$ a *forward vertex* if either $v \in A_1$ and $s^+(v) > 0$, or $v \in A_2$ and $s^-(v) > 0$. Similarly, call a vertex $v \in A$ a *backward vertex* if either $v \in A_1$ and $s^-(v) > 0$, or $v \in A_2$ and $s^+(v) > 0$. Observe that Θ is the difference between the sum of surpluses of forward vertices and the sum of surpluses of backward vertices. Let the *forward edges* be the edges out of A_1, and the edges in to A_2. Similarly, let the *backward edges* be the edges in to A_1, and the edges out of A_2. Let m_A^f and m_A^b be the numbers of forward and backward edges, respectively.

4.2 Structure of large vertices

Call a vertex *huge* if $s(v) \geq \Theta$. If there are no huge vertices, then the greedy algorithm of the previous section will immediately give a partition of the large vertices which has gap smaller than Θ, contradicting the minimality of Θ. Hence there exists at least one huge vertex. Suppose that the vertex v_0 of largest surplus has surplus Δ. By our analysis of the greedy partition of A, we must have $\Theta \leq \Delta$. Yet if the sum of the surpluses of the remaining vertices of A is at least $\Delta + \Theta$, then the total number of edges is at least

$$m \geq \Delta + (\Delta + \Theta) \geq 3\Theta > m$$

by (4.2), and this is a contradiction. Hence the sum g of the surpluses of the remaining vertices of A is strictly less than $\Delta + \Theta$. Consider the partition of A which puts v_0 in A_1, and places all other vertices of A such that their surplus contributes oppositely to the surplus of v_0. The gap of the resulting partition would have magnitude $|\Delta - g|$, and therefore, the above observation, together with the minimality of Θ, implies that $g \leq \Delta - \Theta$. Yet our minimal partition achieves a gap of exactly Θ, and therefore it must have a single forward vertex of surplus Δ, and all the other large vertices of positive surpluses must be backward vertices with surpluses summing to exactly $\Delta - \Theta$.

Note that the edges contributing to $\sum_{v \in A} (d(v) - s(v))$ come in pairs of in-edges and out-edges. Call these the *buffer edges*, and let $2b = \sum_{v \in A}(d(v) - s(v))$. The observation above implies that $m_A^f = \Delta + b$ and $m_A^b = \Delta - \Theta + b$. Moreover, since the graph has minimum outdegree at least two, and there are at least b buffer edges directed out of A, it also implies that the total number of edges in D is at least

$$m \geq b + 2|B| \geq b + 2n - 2\varepsilon n. \tag{4.3}$$

Note that we in fact have

$$m \geq b + 2|B| + \sum_{v \in A} \max\{s^+(v), 0\},$$

and thus the first inequality in (4.3) is tight only if all vertices in A have in-surplus.

4.3 Obtaining a large partition

By the given condition, we know that all the vertices in B have at least two out-edges incident to them. At most one out-edge of each vertex can be incident to v_0 (the vertex of largest surplus), and there are at most $\Delta - \Theta + b$ edges directed in to A which are not incident to v_0. Therefore, the induced digraph on B has at most $\Delta - \Theta + b$ isolated vertices. Since all the other odd components of $D[B]$ have size at least three, this implies that the number of odd components is at most

$$\tau \leq (\Delta - \Theta + b) + \frac{|B| - (\Delta - \Theta + b)}{3} \leq \frac{n + 2(\Delta - \Theta + b)}{3}.$$

Let γ be the constant from Theorem 3.1 where $C = 1152$ and $\varepsilon_{3.1} = \frac{\varepsilon}{4}$. Since $|A| \leq 2304 n^{1/4} \leq \gamma n$ and the maximum degree of vertices in B is at most $n^{3/4} \leq \gamma n$, by Theorem 3.1, we obtain a partition $V = V_1 \cup V_2$ for which

$$\begin{aligned}
\min\{e(V_1, V_2), e(V_2, V_1)\} &\geq \frac{m_A - \Theta}{4} + \frac{m_B}{4} + \frac{n - \tau}{8} - \frac{\varepsilon}{4}n \\
&\geq \frac{m}{4} - \frac{\Theta}{4} + \frac{n - (\Delta - \Theta + b)}{12} - \frac{\varepsilon}{4}n \\
&= \frac{m}{6} + \frac{m + n - \Delta - 2\Theta - b}{12} - \frac{\varepsilon}{4}n
\end{aligned}$$

By (4.3) and $\Theta \leq \Delta \leq n$, we see that

$$m + n - \Delta - 2\Theta - b \geq (b + 2n - 2\varepsilon n) + n - 3n - b = -2\varepsilon n,$$

from which our conclusion follows. Note that the inequality has the least amount of slackness when all large vertices have in-surplus; see the remark following (4.3). This observation fits well with the construction given in Section 1, where there is a single large vertex having huge in-surplus. This completes the proof of the minimum outdegree $d = 2$ case of Theorem 1.1.

5 Minimum outdegree three

The $d = 3$ case of Theorem 1.1 is more complicated, and we provide its proof in this section. Suppose that $D = (V, E)$ is a given directed graph of minimum outdegree at least 3 with n vertices and m edges. Our goal is to find a partition $V = V_1 \cup V_2$ for which $\min\{e(V_1, V_2), e(V_2, V_1)\} \geq (\frac{1}{5} - \varepsilon)m$. Throughout the proof we tacitly assume that the number of vertices n is large enough.

Suppose that $m \geq 3200n$. Then by applying Proposition 2.3 with $\varepsilon = \frac{1}{20}$, we obtain a partition $V = V_1 \cup V_2$ for which

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \frac{m}{4} - \frac{m}{20} = \frac{m}{5}.$$

Hence it suffices to consider the case $m < 3200n$. Since the minimum degree of D is at least three, we see that $3n \leq m < 3200n$.

Let A be the set of *large* vertices, which are defined as vertices that have total degree at least $n^{3/4}$, and let $B = V \setminus A$. Note that

$$|A| \cdot n^{3/4} \leq 2m < 6400n,$$

from which it follows that $|A| \leq 6400n^{1/4} \leq \varepsilon n$ and $e(A) \leq (6400)^2 n^{1/2} \leq \frac{\varepsilon m}{2}$. By the same argument which we used at the beginning of Section 4, it suffices to consider the case when A induces no edges, and seek a partition $V = V_1 \cup V_2$ for which

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \frac{m}{5} - \frac{\varepsilon m}{2}.$$

Let $m_A = e(A, B)$ and $m_B = e(B)$, and note that $m = m_A + m_B$.

5.1 Partition of large vertices

This section follows essentially the same argument as Section 4.1, but the proofs deviate in the next section. As before, if we define the gap Θ of a partition $A = A_1 \cup A_2$ to be

$$\Theta = \big(e^+(A_1) + e^-(A_2)\big) - \big(e^-(A_1) + e^+(A_2)\big),$$

we may take a partition which minimizes the magnitude of the gap, with $0 \leq \Theta \leq n$. (The upper bound is provided by the greedy partition.)

Since $\max_{v \in B} d(v) \leq n^{3/4} \leq \frac{\varepsilon^2}{16} m$, by Lemma 2.2 with $p = \frac{1}{2}$ and $\varepsilon_{2.2} = \frac{\varepsilon}{2}$, there exists a partition $V = V_1 \cup V_2$ such that

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \frac{m - \Theta}{4} - \frac{\varepsilon m}{2} = \frac{m - \Theta}{4} - \frac{\varepsilon m}{2},$$

and hence if $\Theta \leq \frac{m}{5}$, then this partition already gives a desired partition. Thus we may assume that

$$\Theta > \frac{m}{5} \geq \frac{3n}{5}. \tag{5.1}$$

For the remainder of our proof, we will re-use the terms *in-surplus*, *out-surplus*, *surplus*, *forward vertex*, *backward vertex*, *forward edge*, and *backward edge* as originally defined in Section 4.1. Let m_A^f and m_A^b be the numbers of forward and backward edges, respectively.

5.2 Structure of large vertices

Call a vertex $v \in A$ a *huge vertex* if $s(v) \geq \Theta$. Such vertices exist for the same reason as in the $d = 2$ case.

Lemma 5.1. *There exist at least one and at most four huge vertices, and the sum of the surpluses of the rest of the large vertices is at most $n - \Theta$.*

Proof. We have $\Theta > \frac{m}{5}$ by (5.1), so it immediately follows that there are at most four huge vertices. By minimality of the gap, if we switch the side of a forward vertex v, then we obtain a partition whose gap is $\Theta - 2s(v)$. Since we started with a partition which minimized the absolute value of the gap, we must have $|\Theta - 2s(v)| \geq \Theta$. Since the surplus $s(v)$ is always nonnegative, this forces $\Theta - 2s(v) \leq -\Theta$, or equivalently, $s(v) \geq \Theta$. Thus, all forward vertices have surplus at least Θ, and are actually huge. Moreover, since we chose the partition with $\Theta > 0$, there are more forward edges than backward edges, and we see that there exists at least one forward vertex.

Now pick an arbitrary forward vertex v of surplus $\Delta \geq \Theta$. If we move v to the other side, then the number of forward edges becomes $m_A^f - \Delta$ and the number of backward edges becomes $m_A^b + \Delta$. We must have

$$m_A^f - \Delta \leq m_A^b < m_A^f \leq m_A^b + \Delta$$

as otherwise it contradicts the minimality of the gap. By the observation above, all the large vertices which are not huge are backward vertices. If the sum of the surpluses of these vertices is greater than $\Delta - \Theta$, then since each of these vertices has surplus less than Θ, by choosing one such vertex at a time and switching its sides, we will eventually reach a partition in which the number of forward edges is greater than $m_A^f - \Delta + (\Delta - \Theta) = m_A^f - \Theta = m_A^b$ and less than $m_A^f - \Delta + \Delta = m_A^f$. This again contradicts the minimality of the gap. Therefore the sum of the surpluses of the large vertices that are not huge is at most $\Delta - \Theta \leq n - \Theta$. □

To reduce the number of cases, our next step is to further restrict the number of huge vertices.

Lemma 5.2. *The number of huge vertices is either one or three.*

Proof. By Lemma 5.1, we already know that the number of huge vertices is between one and four inclusive. Let g be the sum of the surpluses of the large vertices that are not huge. By Lemma 5.1, we have $g \leq n - \Theta$. Suppose that there are two huge vertices v_1 and v_2, which have surpluses Δ_1 and Δ_2, respectively, where $\Delta_1 \geq \Delta_2$. Re-partition A so that v_1 is the only forward vertex, and all the other vertices are backward vertices. Then the gap of this partition of A is $\Delta_1 - (\Delta_2 + g)$. However, $\Delta_1 - (\Delta_2 + g) \leq n - (\Theta + 0)$ and $(\Delta_2 + g) - \Delta_1 \leq g \leq n - \Theta$, and thus the magnitude of the gap of the new partition is at most $n - \Theta$, which is less than Θ by (5.1). This is a contradiction.

Now suppose that there are four huge vertices v_1, v_2, v_3, v_4, which have surpluses Δ_1, Δ_2, Δ_3, and Δ_4, respectively, where $\Delta_1 \geq \Delta_2 \geq \Delta_3 \geq \Delta_4$. Re-partition A so that v_1 and v_3 are forward vertices, and all other vertices are backward vertices. Then the gap of this partition is $(\Delta_1+\Delta_3)-(\Delta_2+\Delta_4+g)$. Since

$$(\Delta_1 + \Delta_3) - (\Delta_2 + \Delta_4 + g) \leq \Delta_1 - \Delta_4 \leq n - \Theta < \Theta$$

and

$$(\Delta_2 + \Delta_4 + g) - (\Delta_1 + \Delta_3) \leq g \leq n - \Theta < \Theta,$$

this again gives a contradiction. Hence we either have one or three huge vertices. □

5.3 One huge vertex

This section completes the proof of our main theorem for $d = 3$ when there is only one huge vertex. The final case with three huge vertices is finished in the next section. So, for this section, let v_0 be the huge vertex and let $\Delta = s(v_0)$. Since the sum of surpluses of all the other large vertices is at most $n - \Theta$ (by Lemma 5.1) and $n - \Theta < \Theta \le \Delta$ (by (5.1) and the greedy partition), we see that v_0 is the unique forward vertex, and all the other large vertices are backward vertices.

Recall that the edges contributing to $\sum_{v \in A} (d(v) - s(v))$ come in pairs of in-edges and out-edges. Call these the *buffer edges*, and let $2b = \sum_{v \in A}(d(v) - s(v))$. The observation above implies that $m_A^f = \Delta + b$ and $m_A^b = \Delta - \Theta + b$. Moreover, since the graph has minimum outdegree at least three, and there are at least b buffer edges directed in to B, it also implies that the number of edges in D is at least

$$m \ge b + 3|B| \ge b + 3n - 3\varepsilon n. \tag{5.2}$$

In the $d = 2$ case, at this point we applied Theorem 3.1, with a simple bound on the number of odd components in $D[B]$, which provided a bound on the number of tight components. There, it was sufficient to control the number of odd components with only one vertex, which was easy because 1-vertex components have extremely simple structure (they consist of a single vertex, inducing no edges). For the $d = 3$ case, it turns out that we must also control the number of 3-vertex tight components. This would still be particularly easy in the case of oriented graphs, where each pair of vertices spans at most one edge, but an additional twist is required to handle the general case of directed graphs, where edges can go in both directions between the same pair of vertices.

Nevertheless, oriented graphs are still the extremal case, because there is an additional way to gain from pairs of edges running in opposite directions (which we call pairs of *antiparallel* edges). We strengthen Theorem 3.1 to take advantage of this phenomenon. Recall that 3-vertex tight components are undirected graphs with the property that for every one of the 3 vertices, the remaining two vertices form an edge, and the first vertex is either adjacent to both or none of the other two. A moment's inspection reveals that 3-vertex tight components are undirected K_3's. This is particularly useful.

Lemma 5.3. *Given any real constants $C, \varepsilon > 0$, there exist $\gamma, n_0 > 0$ for which the following holds. Let $D = (V, E)$ be a given directed graph with $n \ge n_0$ vertices and at most Cn edges, and let $A \subset V$ be a set of at most γn vertices which have already been partitioned into $A_1 \cup A_2$. Let $B = V \setminus A$, and suppose that every vertex in B has degree at most γn (with respect to the full D). Let τ' be the number of **tight** components in the underlying undirected (simple) graph induced by B, **not counting the 3-vertex components which contain edges that lift to antiparallel pairs in D**. Then, there is a bipartition $V = V_1 \cup V_2$ with $A_1 \subset V_1$ and $A_2 \subset V_2$, such that both*

$$e(V_1, V_2) \ge e(A_1, A_2) + \frac{e(A_1, B) + e(B, A_2)}{2} + \frac{e(B)}{4} + \frac{n - \tau'}{8} - \varepsilon n$$
$$e(V_2, V_1) \ge e(A_2, A_1) + \frac{e(B, A_1) + e(A_2, B)}{2} + \frac{e(B)}{4} + \frac{n - \tau'}{8} - \varepsilon n\,.$$

Remark. The only differences between this statement and Theorem 3.1 are indicated in bold. Importantly, the bounds in the conclusion are the same.

Proof. Let τ be the number of tight components in the underlying undirected (simple) graph induced by B, and let σ be the number of 3-vertex tight components which contain an antiparallel pair when lifted to D. The first step of the proof of Theorem 3.1 was to apply Lemma 3.2 to partition $B = T_1 \cup \cdots \cup T_s \cup U$. By the proof of that lemma, each 3-vertex tight component contributes exactly one star T_i, with its original e_i coming from two of the vertices, and the third vertex contributing a non-free vertex to U. (Here, we use the fact that 3-vertex tight components are K_3's, so that we are assured that the third vertex is always non-free.) Using this structural fact again, observe that actually, no matter which edge of the K_3 we use for our e_i to seed the T_i, the third vertex will always be non-free. In particular, if the 3-vertex tight component contains an antiparallel pair, then we may select the corresponding edge in the underlying undirected (simple) graph as the e_i, and the total number of non-free edges will remain the same as before. This is important because in the proof of Lemma 3.2, it is essential that we start with a maximal matching, which secondarily maximizes the number of free vertices.

Therefore, we may assume that σ of the stars T_i contain at least one edge which lifts to an antiparallel pair in D. This implies that the total number of edges in the stars T_i is now σ more than that in Theorem 3.1. Continuing the proof along the lines of Theorem 3.1, observe that the gain of $+\frac{1}{4}$ comes from the edges of D which correspond to edges of the stars T_i. Therefore, we may improve inequality (3.1) to

$$\begin{aligned}\mathbb{E}[Y_1] &\geq e(A_1, A_2) + \frac{e(A_1, B) + e(B, A_2)}{2} + \frac{e(B)}{4} + \frac{1}{4}\left[\frac{(n - \gamma n) - (\tau + \varepsilon n)}{2} + \sigma\right] \\ &\geq e(A_1, A_2) + \frac{e(A_1, B) + e(B, A_2)}{2} + \frac{e(B)}{4} + \frac{n - \tau'}{8} - \frac{(\varepsilon + \gamma)n}{8},\end{aligned}$$

because $\tau' = \tau - \sigma$. At this point, we have reached the same formula as in the proof of Theorem 3.1, except that τ has been fully replaced with τ'. Therefore, the rest of the proof completes in the same way as before. □

In order to use Lemma 5.3, we must now control τ'. We do this with a similar argument to what was used in Section 4.3.

Lemma 5.4. *The number of tight components in the underlying undirected (simple) graph induced by B, not counting 3-vertex components which contain edges that lift to antiparallel pairs in D, satisfies:*

$$\tau' \leq \frac{n + 2(\Delta - \Theta + b)}{5}.$$

Proof. Let τ_1 be the number of isolated vertices, τ_3' be the number of tight components of order three, not counting those with contain antiparallel pairs, and τ_5 be the number of odd components of order at least five, each in the induced subgraph on B. Note that $\tau' \leq \tau_1 + \tau_3' + \tau_5$. By considering the number of vertices, we obtain the inequality

$$\tau_1 + 3\tau_3' + 5\tau_5 \leq n. \tag{5.3}$$

The vertices in B must have outdegree at least three in the whole graph. Each vertex has at most one edge incident to v_0, and there are at most $\Delta - \Theta + b$ edges from B to A which are not incident to v_0. Each isolated vertex in B uses at least two edges out of the $\Delta - \Theta + b$ edges. Similarly, since a 3-vertex component counted by τ_3' contains at most 3 edges (it cannot have antiparallel pairs), in order to obtain such a component, we must use at least three edges out of the $\Delta - \Theta + b$ edges, per component. Thus we obtain the inequality

$$2\tau_1 + 3\tau_3' \leq \Delta - \Theta + b.$$

By adding two times this inequality to (5.3), we obtain

$$5\tau + 4\tau_3' = 5\tau_1 + 9\tau_3' + 5\tau_5 \le n + 2(\Delta - \Theta + b).$$

Hence

$$\tau' \le \frac{n + 2(\Delta - \Theta + b)}{5}.$$

$\square$

Let γ be the constant from Theorem 3.1, where $C = 3200$ and $\varepsilon_{3.1} = \frac{\varepsilon}{4}$. Since $m \le 3200n$, $|A| \le 6400\varepsilon n^{1/4} \le \gamma n$, and $\max_{v \in B} d(v) \le n^{3/4} \le \gamma n$, by Theorem 3.1 and Lemma 5.4, we obtain a bipartition $V = V_1 \cup V_2$ for which

$$\begin{aligned}
\min\{e(V_1,V_2), e(V_2,V_1)\} &\ge \frac{1}{2}\min\{m_A^f, m_A^b\} + \frac{1}{4}m_B + \frac{n-\tau'}{8} - \frac{\varepsilon}{4}n. \\
&\ge \frac{1}{4}(m - \Theta) + \frac{n}{8} - \frac{n + 2(\Delta - \Theta + b)}{40} - \frac{\varepsilon}{4}n. \\
&= \frac{1}{4}m - \frac{1}{5}\Theta + \frac{1}{10}n - \frac{1}{20}\Delta - \frac{1}{20}b - \frac{\varepsilon}{4}n.
\end{aligned}$$

Thus it suffices to prove that the right hand side of above is at least $\frac{m}{5} - \frac{\varepsilon n}{2}$, or equivalently that

$$\begin{aligned}
&\left(\frac{1}{4}m - \frac{1}{5}\Theta + \frac{1}{10}n - \frac{1}{20}\Delta - \frac{1}{20}b - \frac{\varepsilon}{4}n\right) - \left(\frac{m}{5} - \frac{\varepsilon}{2}n\right) \\
&= \frac{m}{20} - \frac{1}{5}\Theta + \frac{1}{10}n - \frac{1}{20}\Delta - \frac{1}{20}b + \frac{\varepsilon}{4}n
\end{aligned}$$

is at least zero. Recall that by (5.2), we have $m \ge b + 3n - 3\varepsilon n$. By substituting this bound on m in the equation above, we get

$$\begin{aligned}
&\frac{(b + 3n - 3\varepsilon n)}{20} - \frac{1}{5}\Theta + \frac{1}{10}n - \frac{1}{20}\Delta - \frac{1}{20}b + \frac{\varepsilon}{4}n \\
&= \frac{1}{4}n - \frac{1}{5}\Theta - \frac{1}{20}\Delta + \frac{\varepsilon}{10}n.
\end{aligned}$$

Since $n \ge \Delta \ge \Theta$, the right hand side is indeed at least zero, and this proves the theorem when there is one huge vertex.

5.4 Three huge vertices

Let v_1, v_2, v_3 be the three huge vertices, and let $\Delta_1, \Delta_2, \Delta_3$ be their respective surpluses so that $\Delta_1 \ge \Delta_2 \ge \Delta_3$. Let g be the sum of surpluses of the large vertices which are not huge. By Lemma 5.1, we know that $g \le n - \Theta$. Recall that the edges contributing to $\sum_{v \in A}(d(v) - s(v))$ come in pairs of in-edges and out-edges. Call these the *buffer edges*, and let $2b = \sum_{v \in A}(d(v) - s(v))$. Note that

$$m = (\Delta_1 + \Delta_2 + \Delta_3 + g + 2b) + m_B. \tag{5.4}$$

We re-partition A as follows. First place the three vertices v_1, v_2, v_3 so that v_1 is a forward vertex and v_2, v_3 are backward vertices. Depending on the range of parameters, we will choose where to place the large vertices that are not huge. Let $m_A^f = \Delta_1 + b + X$ and $m_A^b = \Delta_2 + \Delta_3 + b + Y$, where

$X+Y=g$. We will either use the partition that gives $(X,Y)=(g,0)$, or the partition that gives $(X,Y)=(0,g)$. Note that such partitions always exist.

If v_1 has positive out-surplus, then let $p=\frac{2}{5}$, and if v_1 has positive in-surplus, then let $p=\frac{3}{5}$. By Lemma 2.2 with such choice of p and $\varepsilon_{2.2}=\frac{\varepsilon}{2}$, we obtain a bipartition of V in which

$$e(V_1,V_2)\geq(1-p)e^+(A_1)+p\cdot e^-(A_2)+p(1-p)e(B)-\frac{\varepsilon m}{2}.$$

Also, note that $e^+(A_1)+e^-(A_2)=m_A^f$, and $\{p,1-p\}=\left\{\frac{2}{5},\frac{3}{5}\right\}$, so $(1-p)e^+(A_1)+p\cdot e^-(A_2)$ has the form $\frac{3}{5}Z+\frac{2}{5}(m_A^f-Z)$ for some Z. By how we placed the vertex v_1, we always have $Z\geq s(v_1)$, and therefore

$$(1-p)e^+(A_1)+p\cdot e^-(A_2)\geq\frac{3}{5}s(v_1)+\frac{2}{5}(m_A^f-s(v_1))=\frac{3}{5}\Delta_1+\frac{2}{5}(b+X).$$

Hence

$$e(V_1,V_2)\geq\frac{3}{5}\Delta_1+\frac{2}{5}(b+X)+\frac{6}{25}m_B-\frac{\varepsilon m}{2},$$

and for

$$m_{1,2}=\frac{3}{5}\Delta_1+\frac{2}{5}(b+X)+\frac{6}{25}m_B,$$

it suffices to prove that $m_{1,2}\geq\frac{m}{5}$. For $e(V_2,V_1)$, we simply use the observation that $\min\{p,1-p\}=\frac{2}{5}$ together with $e^+(A_2)+e^-(A_1)=m_A^b$, and therefore Lemma 2.2 gives

$$e(V_2,V_1)\geq\frac{2}{5}m_A^b+\frac{6}{25}m_B-\frac{\varepsilon m}{2}.$$

Hence for

$$m_{2,1}=\frac{2}{5}m_A^b+\frac{6}{25}m_B=\frac{2}{5}(\Delta_2+\Delta_3+b+Y)+\frac{6}{25}m_B,$$

it suffices to prove that $m_{2,1}\geq\frac{m}{5}$.

Thus our goal is to show that $m_{1,2}-\frac{m}{5}$ and $m_{2,1}-\frac{m}{5}$ are both non-negative. By (5.4) asserting $m=\Delta_1+\Delta_2+\Delta_3+g+2b+m_B$, we have

$$m_{1,2}-\frac{m}{5}=\frac{2\Delta_1-\Delta_2-\Delta_3}{5}+\frac{2X-g}{5}+\frac{1}{25}m_B$$

and

$$m_{2,1}-\frac{m}{5}=\frac{\Delta_2+\Delta_3-\Delta_1}{5}+\frac{2Y-g}{5}+\frac{1}{25}m_B.$$

Two cases complete the rest of this section.

Case 1. $2\Delta_1-\Delta_2-\Delta_3-g>0$.

We partition A so that $(X,Y)=(0,g)$. The condition in this case immediately implies that $m_{1,2}-\frac{m}{5}\geq 0$, and thus it suffices to show that $m_{2,1}-\frac{m}{5}\geq 0$. Note that since $\Delta_2,\Delta_3\geq\Theta$ and $\Delta_1\leq n$, we have

$$\begin{aligned}m_{2,1}-\frac{m}{5}&=\frac{\Delta_2+\Delta_3-\Delta_1}{5}+\frac{2Y-g}{5}+\frac{1}{25}m_B\\&\geq\frac{2\Theta-n}{5}+\frac{g}{5}+\frac{1}{25}m_B.\end{aligned}$$

By (5.1) asserting $\Theta > \frac{m}{5} \geq \frac{3n}{5}$, we have $m_{2,1} - \frac{m}{5} \geq 0$. Hence we obtain a desired partition.

Case 2. $2\Delta_1 - \Delta_2 - \Delta_3 - g \leq 0$.

We partition A so that $(X, Y) = (g, 0)$. We have

$$m_{1,2} - \frac{m}{5} = \frac{2\Delta_1 - \Delta_2 - \Delta_3}{5} + \frac{g}{5} + \frac{1}{25} m_B,$$

and this is non-negative since $\Delta_1 \geq \Delta_2 \geq \Delta_3$. On the other hand,

$$m_{2,1} - \frac{m}{5} = \frac{\Delta_2 + \Delta_3 - \Delta_1 - g}{5} + \frac{1}{25} m_B,$$

and since $\Delta_1 \leq \frac{\Delta_2 + \Delta_3 + g}{2}$ by the condition in this case, we have

$$m_{2,1} - \frac{m}{5} \geq \frac{\Delta_2 + \Delta_3}{10} - \frac{3g}{10} + \frac{1}{25} m_B.$$

By (5.1) asserting $\Theta > \frac{3n}{5}$, we have $\Delta_2 + \Delta_3 \geq 2\Theta > \frac{6n}{5}$ and $3g < 3(n - \Theta) < \frac{6n}{5}$. Hence

$$m_{2,1} - \frac{m}{5} \geq 0,$$

and we obtain a desired partition. This concludes the proof.

6 Concluding remarks

The structure of the proof for $d = 2, 3$ can be described as follows. First, we identify the vertices A which have large total degree, and consider an optimal partition of these vertices. We then further identify the "huge" vertices, which are vertices whose surplus is at least as large as the gap of the partition. It turns out that there can only be a small number of huge vertices. Finally, we partition the set $B = V \setminus A$ depending on the structure of the huge vertices. For this, we used two different probabilistic approaches. One was through the estimate on the number of odd components (Theorem 3.1), and another was through making a random unbalanced partition of B (Lemma 2.2). However, both methods turn out to be too limited in strength to cover the cases $d \geq 4$.

To see why we needed both probabilistic techniques, consider the orientation of $K_{3,n-3}$ where all edges are oriented from the part of size $n-3$ to the part of size 3. This digraph essentially has minimum outdegree 3, with only three vertices in violation, and a constant-size addition would give it that property without affecting its asymptotic partition performance. So, we would expect there to be a partition for which $\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \frac{m}{5} + o(m)$. Note that the set A of large vertices would then be the three vertices of degree $n-3$, and the remainder B would be the $n-3$ vertices of degree 3. If we try to use only Theorem 3.1, the resulting bipartition will nearly be a bisection (a bipartition into two equal size parts), since that method distributes vertices into the two sides with equal probability. Yet if we only consider bisections of this graph, then in every bisection $V = V_1 \cup V_2$, we have

$$\min\{e(V_1, V_2), e(V_2, V_1)\} \leq \left(\frac{1}{2} + o(1)\right) n = \left(\frac{1}{6} + o(1)\right) m,$$

which is already too small.

A different example shows that even an unbalanced straightforward random partition of B (Lemma 2.2) is insufficient. Indeed, add a 3-out-regular graph inside the larger part of the bipartite graph above, so that $m = 6(n-3)$. By merely taking a random partition of B, and assuming that V_1 contains one vertex of degree $n-3$ and V_2 contains two vertices of degree $n-3$, we obtain a partition for which

$$\begin{aligned} e(V_1, V_2) &\approx 2np + p(1-p) \cdot 3n \\ e(V_2, V_1) &\approx n(1-p) + p(1-p) \cdot 3n. \end{aligned}$$

In order to maximize $\min\{e(V_1, V_2), e(V_2, V_1)\}$, we take $p = \frac{1}{3}$, and obtain $\min\{e(V_1, V_2), e(V_2, V_1)\} \approx \frac{4}{3}n \approx \frac{2}{9}m$. Even though this graph does not quite have minimum outdegree six, only three vertices are deficient, and so if Conjecture 1.2 is true, we expect there to be a bipartition for which $\min\{e(V_1, V_2), e(V_2, V_1)\} \geq \frac{5}{22}m + o(m)$. Hence Lemma 2.2 is also too weak on its own.

Therefore in order to proceed further under the same framework, we must combine the two ideas. A naive combination will fail for the following reason. Consider the orientation of $K_{5,n-5}$ where out of the 5 vertices on one side, one vertex v_1 has outdegree $n-5$ and the other four vertices v_2, v_3, v_4, v_5 have indegree $n-5$. The set A of large vertices is precisely $\{v_1, \ldots, v_5\}$. Suppose that in the optimal partition of A has $v_2 \in A_1$ and $v_1, v_3, v_4, v_5 \in A_2$. (It is possible to slightly modify the graph to ensure that this is the unique optimal partition of A.) No matter how we complete this partition into a partition of the whole vertex set, we have

$$e(V_2, V_1) = n - 5 = \frac{m}{5}.$$

Since the minimum outdegree is essentially 4, we need the factor $\frac{3}{14}$ to prove Conjecture 1.2, and thus we fall short. Hence our example shows that in some cases we must start with a sub-optimal partition of A. Indeed, we used this idea in our proof for the case $d = 3$, but in a brute force, ad-hoc manner. It would be interesting to find a systematic way to combine all of these ideas to resolve the general case $d \geq 4$.

References

[1] N. Alon, B. Bollobas, M. Krivelevich and B. Sudakov, Maximum cuts and judicious partitions in graphs without short cycles, *J. Combinatorial Theory Ser. B 88* (2003), 329–346.

[2] N. Alon, M. Krivelevich and B. Sudakov, MaxCut in H-free graphs, *Combinatorics, Probability and Computing* 14 (2005), 629–647.

[3] B. Bollobás, B. Reed, and A. Thomason An extremal function for the achromatic number, **Graph Structure Theory**, Eds. N. Robertson and P. Seymour, American Mathematical Society, Providence, Rhode Island (1993), 161–165.

[4] B. Bollobás and A. Scott, Exact bounds for judicious partitions of graphs, *Combinatorica* 19 (1999), 473–486.

[5] B. Bollobás and A. Scott, Judicious partitions of 3-uniform hypergraphs, *Eur. J. Combin.* 21 (2000), 289–300.

[6] B. Bollobás and A. Scott, Problems and results on judicious partitions, *Random Structures and Algorithms* 21 (2002), 414–430.

[7] B. Bollobás and A. Scott, Max k-cut and judicious k-partitions, *Discrete Mathematics* 310 (2010), 2126–2139.

[8] C. Edwards, Some extremal properties of bipartite subgraphs, *Canad. J. Math.* 25 (1973), 475–485.

[9] C. Edwards, An improved lower bound for the number of edges in a largest bipartite subgraph, *Proc. 2nd Czech Symp. Graph Theory, Prague* (1975), 167–181.

[10] P. Erdős, A. Gyárfás, and Y. Kohayakawa, The size of the largest bipartite subgraphs, *Discrete Math.* 177 (1997), 267–271.

[11] A. Frieze and M. Jerrum, Improved approximation algorithms for MAX k-CUT and MAX BISECTION, *Algorithmica* 18 (1997), 61–77.

[12] M. Goemans and D. Williamson, 0.878 approximation algorithms for MAX CUT and MAX 2-SAT, *Proc. 26th ACM Symp. Theory Comput.* (1994), 422–431; Updated as: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *J. ACM* 42 (1995), 1115–1145.

[13] J. Haslegrave, The Bollobás-Thomason conjecture for 3-uniform hypergraphs, *Combinatorica* 32 (2012), 451–471.

[14] J. Håstad, Some optimal inapproximability results, *J. ACM* 48 (2001), 798–859.

[15] S. Janson, T. Łuczak, and A. Ruciński, **Random Graphs**, Wiley, New York (2000)

[16] D. Kühn and D. Osthus, Maximizing several cuts simultaneously, *Combinatorics, Probability and Computing* 16 (2007), 277–283.

[17] C. Lee, P. Loh, and B. Sudakov, Bisections of graphs, *J. Comb. Theory, Ser. B*, to appear.

[18] J. Ma, P. Yan, and X. Yu, On several partition problems of Bollobás and Scott, *J. Comb. Theory, Ser. B* 100 (2010), 631–649.

[19] J. Ma and X. Yu, Partitioning 3-uniform hypergraphs, *J. Comb. Theory, Ser. B* 102 (2012), 212–232.

[20] S. Poljak and Zs. Tuza, Bipartite subgraphs of triangle-free graphs, *SIAM J. Discrete Math.* 7 (1994), 307–313.

[21] A. Scott, Judicious partitions and related problems, in *Surveys in combinatorics*, London Math. Soc. Lecture Note Ser., 327, Cambridge Univ. Press, Cambridge (2005), 95–117.

[22] L. Trevisan, G. Sorkin, M. Sudan, and D. Williamson, Gadgets, approximation, and linear programming, *SIAM J. Comput.* 29 (2000), 2074–2097.