Partitions of Random Graphs, Invertibility of Digraphs and the Erdős-Rothschild Problem

Emil Powierski

Merton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary Term 2025

Abstract

We begin by studying the number of partitions satisfying degree congruence conditions in random graphs. For $G = G_{n,1/2}$, the Erdős–Renyi random graph, let X_n be the random variable representing the number of distinct partitions of $V(G)$ into sets $A_1, \ldots, A_q$ so that the degree of each vertex in $G[A_i]$ is divisible by q for all $i \in [q]$. We prove that if $q \geqslant 3$ is odd then $X_n \xrightarrow{d} \text{Po}(1/q!)$, and if $q \geqslant 4$ is even then $X_n \xrightarrow{d} \text{Po}(2^q/q!)$. More generally, we show that the distribution is still asymptotically Poisson when we require all degrees in $G[A_i]$ to be congruent to x_i modulo q for each $i \in [q]$, where the residues x_i may be chosen freely. For $q = 2$, the distribution is not asymptotically Poisson, but it can be determined explicitly.

We next study a number of problems related to the invertibility of digraphs. For an oriented graph D and a set $X \subseteq V(D)$, the *inversion of X in D* is the digraph obtained by reversing the orientations of the edges of D with both endpoints in X. The *inversion number of D*, $\text{inv}(D)$, is the minimum number of inversions which can be applied in turn to D to produce an acyclic digraph. Answering a recent question of Bang-Jensen, da Silva, and Havet we show that, for each $k \in \mathbb{N}$ and tournament T, the problem of deciding whether $\text{inv}(T) \leqslant k$ is solvable in time $O_k(|V(T)|^2)$, which is tight for all k. In particular, the problem is fixed-parameter tractable when parameterised by k. On the other hand, we build on their work to prove their conjecture that for $k \geqslant 1$ the problem of deciding whether a general oriented graph D has $\text{inv}(D) \leqslant k$ is NP-complete. We also construct oriented graphs with inversion number equal to twice their cycle transversal number, confirming another conjecture of Bang-Jensen, da Silva, and Havet, and we provide a counterexample to their conjecture concerning the inversion number of so-called 'dijoin' digraphs while proving that it holds in certain cases. Furthermore, we asymptotically solve the natural extremal question in this setting, improving on previous bounds of Belkhechine, Bouaziz, Boudabbous, and Pouzet to show that the maximum inversion number of an n-vertex tournament is $(1 + o(1))n$.

Finally, we consider a generalised version of the Erdős-Rothschild problem from 1974. The original problem asks for the maximum number of s-edge-colourings in an n-vertex graph which avoid a monochromatic copy of K_k, given positive integers n, s, k. In this thesis, we systematically study the generalisation of this problem to a given forbidden family of colourings of K_k. This problem typically exhibits a dichotomy whereby for some values of s, the extremal graph is the 'trivial' one, namely the Turán graph on $k-1$ parts, with no copies of K_k; while for others, this graph is no longer extremal and determining the extremal graph becomes much harder. We generalise a framework developed for the monochromatic Erdős-Rothschild problem to the general setting and work in this framework to obtain our main results, which concern two specific forbidden families: triangles with exactly two colours, and improperly coloured cliques. We essentially solve these problems fully for all integers $s \geqslant 2$ and large n. In both cases we obtain an infinite family of structures which are extremal for some s, which are the first results of this kind. A consequence of our results is that for every non-monochromatic colour pattern, every extremal graph is complete partite. This last part extends work of Hoppen, Lefmann and Schmidt and of Benevides, Hoppen and Sampaio.

Acknowledgements

First, thank you to Alex for your advice and guidance, for proposing engaging problems and for bringing me into such a welcoming, supportive, and collaborative research group. I am also grateful for your patience during challenging moments.

Thank you to Katherine for introducing me to a research project I thoroughly enjoyed, for your patience and understanding, and for always being available with thoughtful advice on other matters. Thank you to Mathias for your guidance during my undergraduate studies and for introducing me to the world of extremal and probabilistic combinatorics.

I would also like to thank all my collaborators – it has been a privilege to work with you. I am especially grateful to Jane and Michael for making the many hours we spent thinking and writing maths together enjoyable, and maybe for a few other things. Thank you to everyone in Alex's research group and the wider combinatorics group in Oxford for creating such an inspiring and friendly environment. Thank you for the best roti Wednesdays, the (night) walks, being a reliable K_k at conferences, and for delivering me to the foot of two of the three peaks. These – and many further memorable moments – more than compensated for my failure in breaking the cake traditions.

Finally, thank you to my family and my friends in Oxford and beyond – for the fun times, for your steady support, for being on the phone, for listening or distracting me when it mattered most.

Statement of Originality

The material in this thesis has not been submitted for any other degree or qualification at any university.

Chapter 2 is based on joint work with Paul Balister, Alex Scott, and Jane Tan [8]. Chapter 3 is based on joint work with Noga Alon, Michael Savery, Alex Scott, and Elizabeth Wilmer [3]. Chapter 4 is based on joint work with Pranshu Gupta, Yani Pehova, and Katherine Staden [45].

Contents

List of Figures

Chapter 1

Introduction

This thesis focuses on a range of problems in combinatorics, with a particular emphasis on extremal and probabilistic graph theory. The introductory chapter provides an overview of the research areas explored and introduces our main results in detail. Each of the subsequent chapters delves into a specific research topic, providing a concise summary of our contributions and focusing on their proofs. We begin in Chapter 2 by deriving results on the distribution of the number of partitions of random graphs that satisfy certain degree congruence conditions. In Chapter 3, we address a number of problems concerning the 'inversion number' of digraphs, delivering structural characterisations, extremal bounds and complexity results. Finally, in Chapter 4 we turn to a generalisation of the Erdős-Rothschild problem, a classical problem in extremal graph theory centred around maximising the number of edge-colourings of a graph that avoid specified forbidden coloured subgraphs.

1.1 Counting partitions of $G_{n,1/2}$ with degree congruence conditions

A folklore result of Gallai (see [70], Exercise 5.17) states that every graph G has a vertex partition into two parts V_1 and V_2 (which may be empty) so that all degrees in the induced subgraphs $G[V_1]$ and $G[V_2]$ are even. An easy corollary of this is that there also exists a vertex partition into two parts for which the degrees in $G[V_1]$ are all odd whilst the degrees in $G[V_2]$ are all even. Another corollary in the same vein is a solution to a well-known riddle: given any graph with lights turned on at each vertex and buttons corresponding to each vertex that toggle the status (light on/off) of a vertex together with its neighbourhood, there is a sequence of button-pushes that turns all of the lights off.

Given that Gallai's theorem guarantees the existence of even/even and even/odd partitions into two parts, one line of research that has arisen is to investigate partitions where each part induces a subgraph with all degrees odd. By considering any graph with an odd number of vertices, it is clear that it is not always possible to find a partition into two parts satisfying this condition. However, improving on earlier work of Caro [25] and Scott [80], Ferber and Krivelevich [40] proved that any graph without isolated vertices contains a linearly sized induced subgraph with all degrees odd.

When we allow more than two parts, it is still trivially necessary that G must have an even number of vertices in each part and hence in each component. The following result due to Scott [81] states that this is in fact sufficient. We shall call a subgraph H of a graph G *odd* if $d_H(v)$ is odd for all $v \in V(H)$, and similarly *even* if all degrees are even.

Theorem 1.1 ([81]). *The vertices of a graph G can be partitioned into sets $A_1, \dots, A_k$ for some k such that $G[A_i]$ is an odd subgraph for all i if and only if every component of G has even order.*

More generally, we can consider congruence conditions modulo q for $q \geqslant 2$. Caro, Krasikov and Roditty [26] asked whether there exists a number $k = k(q)$ such that every graph G can be partitioned into k vertex-disjoint classes $A_1, \dots, A_k$ in such a way that all degrees in all induced subgraphs $G[A_i]$ are divisible by q. While this deterministic problem is very much open, more can be said in the context of random graphs.

Henceforth let $G_{n,1/2}$ be the standard binomial random graph with vertex set $[n]$ and edge probability $1/2$. Scott [81] formulated a random version of Caro, Krasikov and Roditty's problem in which we ask for partitions satisfying fixed degree residue conditions to exist for almost every $G_{n,1/2}$ (i.e. with probability tending to 1 as $n \to \infty$).

Problem 1.2 ([81]). *Let $q \geqslant 2$ and $0 \leqslant x < q$ be integers. Does there exist a number $k = k(q, x)$ such that almost every $G_{n,1/2}$ (with n even for simplicity) can be partitioned into k vertex-disjoint classes $A_1, \dots, A_k$ in such a way that in all the induced subgraphs $G[A_i]$ all degrees are $x \bmod q$?*

It was shown in [81] that three parts suffice for $q = 2$ and $x = 1$. Ferber, Hardiman and Krivelevich [38, 39] solved Problem 1.2 for general q, showing that $k(q, x) = q + 1$ is sufficient for all x.

Theorem 1.3 (Ferber, Hardiman and Krivelevich [39])**.** *For all $q \in \mathbb{N}$ and $0 \leqslant x < q$, almost every $G_{n,1/2}$ has a vertex partition into $q + 1$ sets $A_1, \ldots, A_{q+1}$ such that in each induced subgraph $G[A_i]$ all degrees are $x \bmod q$.*

It is not hard to show (by a first moment argument) that Theorem 1.3 does not hold for partitions into $q - 1$ parts. However, a natural question asked in [38] is whether the theorem still holds with q parts instead of $q + 1$. In Chapter 2, we provide a negative answer and moreover determine the asymptotic distribution of the number of 'good' partitions into q parts. The situation differs between the cases $q > 2$ and $q = 2$. Beginning with the former, for $G = G_{n,1/2}$, let X_n be the random variable representing the number of partitions of $V(G)$ into disjoint sets $A_1, \ldots, A_q$ so that the degree of each vertex in $G[V_i]$ is divisible by q for all $i \in [q]$, i.e.

$$X_n = \left| \left\{ \{A_1, \ldots, A_q\} : V(G) = \coprod_{i=1}^{q} A_i \text{ and } q \mid d_{G[A_i]}(v) \text{ for all } i \in [q],\, v \in A_i \right\} \right|.$$

Although we number the parts, we only consider partitions up to permutation of the parts. Our first result gives the distribution of X_n as $n \to \infty$.

Theorem 1.4. *If $q \geqslant 3$ is odd, then $X_n \xrightarrow{d} \mathrm{Po}(1/q!)$. If $q \geqslant 4$ is even, then $X_n \xrightarrow{d} \mathrm{Po}(2^q/q!)$.*

With some minor modifications to the proof of Theorem 1.4, we can obtain the following stronger statement which allows for greater flexibility in fixing congruence conditions. Specifically, for non-negative integers $a_0, \ldots, a_{q-1}$ with $\sum_{x=0}^{q-1} a_x = q$, we now let $X_n = X_n^{(a_0,\ldots,a_{q-1})}$ be the number of q-tuples $(\{A_{0,1}, \ldots, A_{0,a_0}\}, \{A_{1,1}, \ldots, A_{1,a_1}\}, \ldots)$, where each entry is an (unordered, possibly empty) set of parts and each part satisfies the degree condition given by its first index, i.e.

$$\begin{aligned} X_n = \Big| \Big\{ (\{A_{0,1}, \ldots, A_{0,a_0}\}, \{A_{1,1}, \ldots, A_{1,a_1}\}, \ldots \{A_{q-1,1}, \ldots, A_{q-1,a_{q-1}}\}) : \\ V(G) = \coprod_{(x,y)\in T} A_{x,y} \text{ and } d_{G[A_{x,y}]}(v) \equiv x \text{ for all } (x,y) \in T,\, v \in A_{x,y} \Big\} \Big|, \end{aligned}$$

where $T = \{(x, y) : x = 0, \ldots, q-1, y = 1, \ldots, a_x\}$ and the congruence is mod q. This definition ensures that (as before) two partitions that can be transformed into each other by reordering parts with the same degree condition are considered the same and only counted once. Setting $a_0 = q$ and $a_i = 0$ for $i \neq 0$ we would get the random variable we considered in Theorem 1.4.

Theorem 1.5. *Let $G = G_{n,1/2}$ and let $X_n = X_n^{(a_0,\dots,a_{q-1})}$ be as defined above for any non-negative integers $a_0, \dots, a_{q-1}$ with $\sum_{x=0}^{q-1} a_x = q$. For q even, write $c = \sum_{x=0}^{q/2-1} a_{2x}$ for the number of parts where the degree condition is even.*

(1) If $q \geqslant 3$ is odd, then $X_n \xrightarrow{d} \mathrm{Po}(1/\prod a_x!)$.

(2) If $q \geqslant 4$ is even and $c > 0$, then $X_n \xrightarrow{d} \mathrm{Po}(2^c/\prod a_x!)$.

(3) If $q \geqslant 4$ is even and $c = 0$, then $X_n \xrightarrow{d} \mathrm{Po}(2/\prod a_x!)$ as n runs over even *integers.*

Note that if q is even and $c = 0$ then all degrees in each $G[A_{x,y}]$ must be odd, so each $|A_{x,y}|$ is even and hence $n = \sum |A_{x,y}|$ must also be even.

For a cleaner presentation, Theorem 1.4 and Theorem 1.5 feature the random graph with edge probability $p = \frac{1}{2}$, although our proofs actually extend to any $p = p(n)$ with $C\frac{\log n}{n} \leqslant p \leqslant 1 - C\frac{\log n}{n}$ for some large constant C.

The case where $q = 2$ is really exceptional, which is especially intriguing given that the following question was one of the starting points in the random setting.

Problem 1.6 ([81])**.** *For n even, what is the probability that $G_{n,1/2}$ can be partitioned into two sets, each inducing odd subgraphs?*

It is known [81] that the answer is at least $1/2 + o(1)$. We show that the answer to Problem 1.6 is $2/3 + o(1)$. In fact, we provide the full distributions of $X_n^{(2,0)}$, $X_n^{(1,1)}$ and $X_n^{(0,2)}$, recalling that these are the number of bipartitions of $G_{n,1/2}$ inducing even/even, even/odd and odd/odd partitions respectively. The exact distributions are given in Chapter 2, and lead to the following asymptotic distributions. As usual, products over empty index sets, or with a lower limit exceeding the upper limit, are defined to be 1.

Theorem 1.7. *Let $G = G_{n,1/2}$. Then $X_n^{(2,0)} \xrightarrow{d} X$ and $X_n^{(1,1)} \xrightarrow{d} X$ where*

$$\mathbb{P}(X = 2^k) = c \prod_{i=1}^{k} (2^i - 1)^{-1}$$

for $k \in \mathbb{N} \cup \{0\}$ with constant $c = \prod_{i=0}^{\infty}(1 - 2^{-2i-1}) = (\sum_{j=1}^{\infty} \prod_{i=1}^{j}(2^i - 1)^{-1})^{-1}$, and $\mathbb{P}(X = x) = 0$ if $x \neq 2^k$ for any $k \in \mathbb{N} \cup \{0\}$. Furthermore, $X_n^{(0,2)} \xrightarrow{d} Z$ where n only *runs over even integers and*

$$\mathbb{P}(Z = 2^k) = c2^{-k} \prod_{i=1}^{k} (2^i - 1)^{-1}$$

for $k \in \mathbb{N} \cup \{0\}$, $\mathbb{P}(Z = 0) = \frac{1}{3}$ *and* $\mathbb{P}(Z = x) = 0$ *if* $x \neq 0$ *and* $x \neq 2^k$ *for any* $k \in \mathbb{N} \cup \{0\}$.

Chapter 2 is based on joint work with Paul Balister, Alex Scott, and Jane Tan [8].

1.2 Invertibility of digraphs and tournaments

In Chapter 3, we consider directed graphs without loops, digons, or parallel edges, for which we use the terms *digraph* and *oriented graph* interchangeably. For such a digraph $D = (V, E)$ and a set $X \subseteq V$, the *inversion of* X *in* D is the digraph obtained from D by reversing the direction of the edges with both endpoints in X; we say that we *invert* X *in* D. Given a family of sets $X_1, \ldots, X_k \subseteq V$, we can invert X_1 in D, then X_2 in the resulting digraph, and so on, noting that the final digraph produced by these inversions is independent of the order in which we perform them. If inverting $X_1, \ldots, X_k$ in turn transforms D into an acyclic digraph, then we say that these sets form a *decycling family* of D. We will refer to a set $X \subseteq V$ which forms a decycling family by itself as a *decycling set*. The *inversion number of* D, denoted $\mathrm{inv}(D)$, is defined to be the minimum size of a decycling family of D, and for $k \in \mathbb{N}_0$ we say that D is *k-invertible* if $\mathrm{inv}(D) \leqslant k$.

The study of inversions began in Houmem Belkhechine's PhD thesis [16] and continued in [17, 18, 79], in which many foundational results were established. The work presented in Chapter 3 is inspired by a paper of Bang-Jensen, da Silva, and Havet [11] which studied a wide range of questions about invertibility, with an emphasis on those of an algorithmic or extremal nature. They also posed a host of interesting conjectures and problems, some of which we answer in this thesis.

1.2.1 The inversion number of k-joins

The cornerstone of many of the conjectures made by Bang-Jensen, da Silva, and Havet in [11] is the following 'dijoin conjecture'. For oriented graphs L and R, the *dijoin* $L \to R$ *from* L *to* R is the oriented graph consisting of vertex-disjoint copies of L and R, with an edge $\overrightarrow{uv}$ for all $u \in V(L)$ and $v \in V(R)$.

Conjecture 1.8 ([11])**.** *For oriented graphs L and R we have* $\mathrm{inv}(L \to R) = \mathrm{inv}(L) + \mathrm{inv}(R)$.

Noting that the left-hand side is certainly at most the right-hand side for all L and R and that the conjecture holds trivially if $\mathrm{inv}(L) = 0$ or $\mathrm{inv}(R) = 0$, Bang-Jensen, da Silva, and Havet showed it to be true when $\mathrm{inv}(L) + \mathrm{inv}(R) \leqslant 3$, and when $\mathrm{inv}(L) = \mathrm{inv}(R) = 2$ and L and R are both strongly connected.[1] They also observed (see our Section 3.2) that the conjecture is equivalent to its restriction to tournaments. We disprove Conjecture 1.8 by exhibiting a tournament R which satisfies $\mathrm{inv}(R) = \mathrm{inv}(\overrightarrow{C_3} \to R)$, where $\overrightarrow{C_3}$ is the directed cycle on three vertices.

Theorem 1.9. *There exists a tournament R with* $\mathrm{inv}(R) = \mathrm{inv}(\overrightarrow{C_3} \to R) = 3$.

Just before making the material in Chapter 3 available online, we learned of independent simultaneous work by Aubian, Havet, Hörsch, Klingelhoefer, Nisse, Rambaud, and Vermande [6] who proved a stronger version of Theorem 1.9. Specifically, they showed that for any odd $\ell \geqslant 3$, there is a tournament L with $\mathrm{inv}(L) = \ell$ such that for any R with $\mathrm{inv}(R) \geqslant 1$ the pair L, R is a counterexample to the dijoin conjecture.

While Theorem 1.9 shows that the dijoin conjecture is false in general, we prove it in the case where $\mathrm{inv}(L) = \mathrm{inv}(R) = 2$.

Theorem 1.10. *If L and R are digraphs with* $\mathrm{inv}(L) = \mathrm{inv}(R) = 2$, *then* $\mathrm{inv}(L \to R) = 4$.

The proof of Theorem 1.10 relies on the strongly connected case and our next result, which concerns the following generalisation of dijoins to arbitrarily many digraphs. For $k \in \mathbb{N}$ the *k-join* of digraphs $D_1, \dots, D_k$, written $[D_1, \dots, D_k]$, is the digraph consisting of vertex-disjoint copies of $D_1, \dots, D_k$ with an additional edge $\overrightarrow{uv}$ whenever $u \in V(D_i), v \in V(D_j)$ for $i < j$. We write $[D]_k = [D, \dots, D]$ for the k-join of k copies of the same oriented graph D. The following result can be viewed as a k-join analogue of the dijoin conjecture holding under certain conditions. It generalises a theorem of Pouzet, Kaddour, and Thatte [79] which states that $\mathrm{inv}([\overrightarrow{C_3}]_k) = k$ for all k.

Theorem 1.11. *Let $k \in \mathbb{N}$ and let $D_1, \dots, D_k$ be oriented graphs. Assume that* $\mathrm{inv}(D_i) \leqslant 2$ *for all i, with equality for at most one i. Then*

$$\mathrm{inv}([D_1, \dots, D_k]) = \sum_{i=1}^{k} \mathrm{inv}(D_i). \tag{1.2.1}$$

[1]The case where $\mathrm{inv}(L) = 2$ and $\mathrm{inv}(R) = 1$ is not explicitly mentioned in [11], but follows easily from the case where $\mathrm{inv}(L) = 1$ and $\mathrm{inv}(R) = 2$ by inverting $V(L \to R)$.

Note that again the left-hand side is at most the right-hand side in (1.2.1) for any digraphs $D_1, \dots, D_k$. We will use Theorem 1.11 to confirm another conjecture from [11] which was made based on the dijoin conjecture (see Theorem 1.13 below). Theorem 1.11 and, in turn, Theorem 1.10 follow from a characterisation of the decycling families of size k of arbitrary k-joins of oriented graphs with inversion number 1. We will need some further terminology to state this result: for a digraph D, sets $X_1, \dots, X_k \subseteq V(D)$, and a vertex $v \in V(D)$, we define the *characteristic vector* of v in $X_1, \dots, X_k$ to be $(I_{\{v \in X_i\}} : i \in [k]) \in \mathbb{F}_2^k$, where $I_{\{v \in X_i\}}$ is the indicator function of the event $v \in X_i$. For vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}_2^k$ we write $\mathbf{u} \cdot \mathbf{v}$ for the usual scalar product of $\mathbf{u}$ and $\mathbf{v}$ over $\mathbb{F}_2$. This is not a genuine inner product, but we say nevertheless that a collection $\mathbf{u}_1, \dots, \mathbf{u}_\ell \in \mathbb{F}_2^k$ is *orthonormal* if $\mathbf{u}_i \cdot \mathbf{u}_i = 1$ for all i and $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ for all $i \neq j$. Finally, we refer to the canonical copy of D_i in $D = [D_1, \dots, D_k]$ as the *ith factor* of D. We are now ready to state our characterisation theorem, the case $k = 2$ of which was shown by Bang-Jensen, da Silva, and Havet [11]. Its proof is based on an approach used by Pouzet, Kaddour, and Thatte [79].

Theorem 1.12. *Let $D_1, \dots, D_k$ be oriented graphs with $\mathrm{inv}(D_i) = 1$ for all i and let $\widehat{D} = [D_1, \dots, D_k]$ be their k-join. Then sets $X_1, \dots, X_k \subseteq V(\widehat{D})$ form a decycling family of $\widehat{D}$ if and only if there are orthonormal vectors $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{F}_2^k$ and for each i a decycling set $V_i \subseteq V(D_i)$ of the ith factor of $\widehat{D}$ such that, for each i, the vertices in V_i have characteristic vector $\mathbf{u}_i$ (in $X_1, \dots, X_k$), and all other vertices have characteristic vector $\mathbf{0}$ (in $X_1, \dots, X_k$).*

In particular, any acyclic digraph obtained from $\widehat{D}$ by k inversions can also be obtained by inverting a decycling set for each factor in turn.

1.2.2 Computational complexity

One focus of Bang-Jensen, da Silva, and Havet's paper [11] was on the computational complexity of deciding whether an oriented graph is k-invertible. More formally, they considered, for fixed $k \in \mathbb{N}_0$, the problem of k-INVERSION:

> INPUT: an oriented graph D.
> PROBLEM: $\mathrm{inv}(D) \leqslant k$?

A first observation is that 0-INVERSION is equivalent to checking whether a digraph D is acyclic, which is well known to be possible in time $O(|V(D)|^2)$ (see [33, p. 612]), so we need only consider $k \geqslant 1$.

Bang-Jensen, da Silva, and Havet [11] showed that 1-Inversion is NP-complete using a reduction from Monotone 3-in-1 SAT. Then, using the special cases of the dijoin conjecture proved in that paper, they observed that for a digraph D we have $\text{inv}(D \to D) = 2$ if and only if $\text{inv}(D) = 1$, from which it follows that 2-Inversion is also NP-complete. They conjectured that NP-completeness extends to k-Inversion for all $k \geqslant 3$, noting that this would follow from a similar argument if the dijoin conjecture were true. Of course, the full dijoin conjecture is not required, and indeed it is easy to see that Theorem 1.11 is enough: it implies that $\text{inv}(D) = 1$ if and only if $\text{inv}([D]_k) = k$, which reduces 1-Inversion to k-Inversion and hence shows the following.

Theorem 1.13. *k-Inversion is NP-complete for all $k \in \mathbb{N}$.*

Bang-Jensen, da Silva, and Havet also considered the computational complexity of the same problem when the input is restricted to tournaments. For fixed $k \in \mathbb{N}$ the problem of k-Tournament-Inversion is:

> Input: a tournament T.
> Problem: $\text{inv}(T) \leqslant k$?

One way of analysing the complexity of this problem is to use k-inversion-critical tournaments: we say that a tournament T is a *k-inversion-critical tournament* if $\text{inv}(T) = k$ but $\text{inv}(T - \{v\}) < k$ for all $v \in V(T)$, and denote by $\mathcal{IC}_k$ the set of k-inversion-critical tournaments. It is not difficult to see that a tournament has inversion number at most k if and only if it contains no element of $\mathcal{IC}_{k+1} \cup \mathcal{IC}_{k+2}$ as a subtournament. Indeed, for any digraph D and vertex $v \in V(D)$ with out-neighbourhood $A \subseteq V(D)$, adding A and $A \cup \{v\}$ to a decycling family of $D - \{v\}$ gives a decycling family of D. We deduce that $\text{inv}(D) \leqslant \text{inv}(D - \{v\}) + 2$. Hence, if $\text{inv}(T) > k$, then by arbitrarily deleting vertices from T one by one, we can obtain a subtournament T' of T with $\text{inv}(T') \in \{k+1, k+2\}$. This T' contains a member of $\mathcal{IC}_{k+1} \cup \mathcal{IC}_{k+2}$ as a subtournament.

Belkhechine, Bouaziz, Boudabbous, and Pouzet [18] showed that $\mathcal{IC}_k$ is finite for all $k \in \mathbb{N}$. Writing m_k for the maximum number of vertices of an element of $\mathcal{IC}_k$, it follows that k-Tournament-Inversion can be solved in time $O(|V(T)|^{\max(m_{k+1}, m_{k+2})})$. Thus, in particular, k-Tournament-Inversion can be solved in polynomial time for any fixed k. Plainly $\mathcal{IC}_1 = \{\overrightarrow{C_3}\}$, so $m_1 = 3$, and $\mathcal{IC}_2$ was explicitly described in [79], giving $m_2 = 6$. However, no upper bound on m_k is known for $k \geqslant 3$, so for no

$k \geqslant 1$ does the above give a concrete polynomial bound on the complexity of k-Tournament-Inversion. Note also that this approach does not identify a decycling family of size k given a k-invertible tournament, it can only confirm the existence of one.

Bang-Jensen, da Silva, and Havet [11] used an alternative approach to show that 1-Tournament-Inversion can be solved in time $O(|V(T)|^3)$ while 2-Tournament-Inversion can be solved in time $O(|V(T)|^6)$. The idea behind their algorithm for 1-Tournament-Inversion is to check whether the tournament contains a vertex which can be made into a source, and for 2-Tournament-Inversion they check whether it contains a pair of vertices which can be made into a source and a sink respectively. They went on to ask for the least real numbers r_k such that k-Tournament-Inversion can be solved in time $O(|V(T)|^{r_k})$. We answer this question by showing that, perhaps surprisingly, for each fixed $k \in \mathbb{N}$ there is an algorithm solving k-Tournament-Inversion in time $O(|V(T)|^2)$. In the language of complexity theory, this means that the likely NP-hard problem of determining whether $\text{inv}(T) \leqslant k$ for inputs k and T (see Conjecture 3.11) is fixed-parameter tractable when parameterised by k.[2]

Theorem 1.14. *For fixed $k \in \mathbb{N}$, k-Tournament-Inversion can be solved for n-vertex tournaments in time $O(n^2)$. Moreover, if the input tournament is k-invertible, then our algorithm finds a decycling family of size at most k.*

Note that the exponent of n in this running time is optimal, since any algorithm solving k-Tournament-Inversion needs to inspect the orientation of every edge in the input tournament. However, the implied constant in the running time of our algorithm is doubly exponential in k, so it is unlikely to be of practical use for large k.

1.2.3 Relation to other parameters

Bang-Jensen, da Silva, and Havet [11] also considered the relationship between the inversion number and other digraph parameters. Two well studied parameters of particular interest are the cycle transversal number and the cycle edge-transversal number, defined as follows. A *cycle transversal* (or *feedback vertex set*) in a digraph D is a set of vertices of D whose removal from D leaves an acyclic digraph and the *cycle transversal number* of D, denoted $\tau(D)$, is the minimum size of a cycle transversal in D. Analogously, a *cycle edge-transversal* (or *feedback arc set*) in D is a set of edges of D whose removal leaves an acyclic digraph and the *cycle edge-transversal number*

[2]See [36] for the definition of fixed-parameter tractability and an exposition of the surrounding theory.

of D, $\tau'(D)$, is the minimum size of a cycle edge-transversal in D. Note that the inequality $\tau(D) \leqslant \tau'(D)$ always holds, since a set containing an endpoint of each edge in a cycle edge-transversal of D forms a cycle transversal of D.

Bang-Jensen, da Silva, and Havet [11] made the following observations concerning the relationships between $\mathrm{inv}(D)$, $\tau(D)$, and $\tau'(D)$ for a digraph D. Firstly, we have $\mathrm{inv}(D) \leqslant \tau'(D)$. This follows from the fact that if $F \subseteq E(D)$ is a cycle edge-transversal of D, then since $(V(D), E(D) \setminus F)$ is acyclic, there is a labelling $v_1, \ldots, v_n$ of $V(D)$ such that $\overrightarrow{v_j v_i} \notin E(D) \setminus F$ if $i < j$. Applying the family of inversions $(\{v_i, v_j\} \colon i < j, \overrightarrow{v_j v_i} \in F)$ transforms D into an acyclic digraph and hence $\mathrm{inv}(D) \leqslant \tau'(D)$ as claimed. They also observed that this inequality is tight for all values of $\tau'(D)$ as exhibited by $[\overrightarrow{C_3}]_k$, which clearly has cycle edge-transversal number k, and as mentioned above was shown in [79] to have inversion number k.

Turning to $\tau(D)$, the inequality $\mathrm{inv}(D) \leqslant 2\tau(D)$ was obtained in [11] as follows. After observing that $\tau(D) = 0$ implies $\mathrm{inv}(D) = 0$, we may assume that $\tau(D) \geqslant 1$. Let $S \subseteq V(D)$ be a cycle transversal in D of size $\tau(D)$ and pick $v \in S$. Then observe that $D - \{v\}$ has cycle transversal number $\tau(D) - 1$, with $S \setminus \{v\}$ a cycle transversal. Moreover, as noted in Section 1.2 we have $\mathrm{inv}(D) \leqslant \mathrm{inv}(D - \{v\}) + 2$, from which it follows by induction that $\mathrm{inv}(D) \leqslant 2\tau(D)$.

Bang-Jensen, da Silva, and Havet conjectured that this inequality is tight for all values of $\tau(D)$. Indeed, they considered the graph V_5 obtained by adding a vertex v and edges $\overrightarrow{v1}, \overrightarrow{2v}, \overrightarrow{v3}, \overrightarrow{4v}$ to the (transitive) tournament on vertex set $\{1, 2, 3, 4\}$ with edges $\overrightarrow{ij}$ for $i < j$, which can easily be seen to have $\tau(V_5) = 1$ and $\mathrm{inv}(V_5) = 2$. They noted that if the dijoin conjecture holds, then $\tau([V_5]_k) = k$ and $\mathrm{inv}([V_5]_k) = 2k$ for all k (in fact, since V_5 is strongly connected, the case $k = 2$ follows from the special cases for which they proved the dijoin conjecture). We construct digraphs with a similar character to V_5 which confirm their conjecture.

Theorem 1.15. *For all $k \in \mathbb{N}$ there exists an oriented graph D with* $\mathrm{inv}(D) = 2\tau(D) = 2k$.

1.2.4 The extremal problem

Finally, we consider $\mathrm{inv}(n)$, defined for each $n \in \mathbb{N}$ as the maximum inversion number of an oriented graph (or, equivalently, a tournament) on n vertices. Belkhechine, Bouaziz, Boudabbous, and Pouzet [18] were the first to study this parameter, obtaining bounds of the form

$$\frac{n}{2} - \log_2(n) + O(1) \leqslant \mathrm{inv}(n) \leqslant n + O(1).$$

Their lower and upper bounds follow from counting and inductive arguments respectively (see Section 3.6 for details), and they conjectured that $\mathrm{inv}(n) \geqslant \lfloor \frac{n-1}{2} \rfloor$ for all n. Bounds of the form above previously remained the best known, with Bang-Jensen, da Silva, and Havet [11] noting that the $O(1)$ term in the upper bound can be improved very slightly.

Using a random construction, we show that $\mathrm{inv}(n) = (1 + o(1))n$.

Theorem 1.16. *For sufficiently large n we have*

$$\mathrm{inv}(n) \geqslant n - \sqrt{2n \log_2(n)}.$$

Moreover, a uniformly random labelled n-vertex tournament has at least this inversion number with probability tending to 1.

We will also show that $\mathrm{inv}(n) \leqslant n - \log_2(n + 1)$. Similar upper and lower bounds on $\mathrm{inv}(n)$ were also obtained in the simultaneous work by Aubian, Havet, Hörsch, Klingelhoefer, Nisse, Rambaud, and Vermande [6].

Chapter 3 is based on joint work with Noga Alon, Michael Savery, Alex Scott, and Elizabeth Wilmer [3].

1.3 The generalised Erdős-Rothschild problem with a focus on dichromatic triangles

Chapter 4 concerns an extremal problem on edge-colourings of graphs with certain forbidden structures: given a collection $\mathcal{X}$ of s-edge-coloured cliques on k vertices, which n-vertex graph has the maximum number of colourings which do not contain any elements of $\mathcal{X}$? More formally, let $k \geqslant 3$ be a clique size, $s \geqslant 2$ a number of colours, and let $\mathcal{X}$ be a family of s-edge-colourings of K_k. Given a graph G, we say that an s-edge-colouring of G is $\mathcal{X}$-*free* (or *valid* if $\mathcal{X}$ is clear from the context) if it avoids all coloured copies of K_k which lie in $\mathcal{X}$, and define $F(G; \mathcal{X})$ to be the number of $\mathcal{X}$-free s-edge-colourings of G. The goal is to determine

$$F(n; \mathcal{X}) := \max_{|V(G)|=n} F(G; \mathcal{X}),$$

as well as the set of $\mathcal{X}$-*extremal* graphs G, i.e. those satisfying $F(G;\mathcal{X}) = F(n;\mathcal{X})$. By considering all colourings of the Turán graph $T_{k-1}(n)$, which is the n-vertex complete $(k-1)$-partite graph with part sizes as equal as possible, we can see that

$$s^{t_{k-1}(n)} \leqslant F(n;\mathcal{X}) \leqslant s^{\binom{n}{2}}, \tag{1.3.1}$$

where $t_{k-1}(n) = (1 - \frac{1}{k-1} + o(1))\binom{n}{2}$ is the number of edges in $T_{k-1}(n)$. We refer to $s^{t_{k-1}(n)}$ as the *trivial lower bound* for $F(n;\mathcal{X})$ and to $T_{k-1}(n)$ as the *trivial example*. The inequalities in (1.3.1) give the order of magnitude of $\log F(n;\mathcal{X})$, so an asymptotic solution to the problem above would be to determine the limit

$$\lim_{n\to\infty} \frac{\log F(n;\mathcal{X})}{\binom{n}{2}},$$

which was shown to exist in [2] for forbidden monochromatic cliques using an entropy inequality of Shearer; the same proof applies for general families $\mathcal{X}$ (see [9, Lemma 7]).

1.3.1 The monochromatic Erdős-Rothschild problem

The problem of determining $F(n;\mathcal{X})$ was first posed by Erdős and Rothschild in 1974 [37], in the case when $\mathcal{X}$ is the family $\mathcal{X}^{(1)}_{k,s}$ of the s monochromatic colourings of K_k. They conjectured that the trivial lower bound is tight for $(k,s) = (3,2)$. This was verified by Yuster [85] for all $n \geqslant 6$ and by Alon, Balogh, Keevash and Sudakov [2], the latter of whom showed that in fact for both $s = 2,3$ and all $k \geqslant 3$ (and large n), the trivial lower bound is tight, but for $s \geqslant 4$ it is very far from tight. It was shown in [2, Theorem 1.2] that for $s = 4$, the complete 4-partite graph $T_4(n)$ has exponentially more colourings free of monochromatic triangles than $T_2(n)$. The construction can naturally be extended to $s \geqslant 4$ colours and inserted in place of three vertex classes of $T_{k-1}(n)$ to show that $F(n;\mathcal{X}^{(1)}_{k,s})$ exceeds $s^{t_{k-1}(n)}$ by an exponential factor. The authors also showed [2, Theorem 1.3] that $\lim_{n\to\infty} \log F(n;\mathcal{X}^{(1)}_{k,s})/\binom{n}{2} \to \left(1 - \frac{1}{k-1}\right)\log(s)$ as $\max\{k,s\} \to \infty$.

For $s \geqslant 4$, the only cases where $F(n;\mathcal{X}^{(1)}_{k,s})$ has been determined for large n are (k,s) equal to $(3,4)$, $(3,5)$ (asymptotically), $(3,6)$, $(3,7)$ and $(4,4)$ [22, 75, 77]. There is no conjecture for general (k,s). In all known exact results, there is a unique $\mathcal{X}^{(1)}_{k,s}$-extremal graph which is a Turán graph, with more than $k-1$ parts for $s \geqslant 4$. For example, the extremal graph for $(4,4)$ is $T_9(n)$. However, for $(3,5)$ there is a large family of graphs which are *almost* extremal, though we do not know which of these are extremal; nevertheless, these graphs are all complete multipartite. It is not known whether for all pairs (k,s) every extremal graph is complete multipartite.

1.3.2 The generalised Erdős-Rothschild problem

In Chapter 4 we study the *generalised Erdős-Rothschild problem* of determining $F(n; \mathcal{X})$ and the $\mathcal{X}$-extremal graphs for a general family $\mathcal{X}$ of forbidden s-edge-colourings of K_k. We develop a general proof framework for the generalised Erdős-Rothschild problem, which constitutes a powerful reduction to an optimisation problem, and give two applications to forbidden 2-edge-coloured triangles and to counting colourings of graphs in which every k-clique is properly coloured. Existing results on the generalised Erdős-Rothschild problem, which we will shortly survey, are sporadic in nature and any sort of comprehensive solution has so far been out of reach.

Balogh [9] was the first to extend the Erdős-Rothschild problem to forbidden non-monochromatic cliques. He considered the variant where $\mathcal{X}$ contains a single colouring $\chi : E(K_k) \to [2]$ and showed that, unless χ is monochromatic, the trivial lower bound is tight. Subsequent results in the literature consider predominantly what we call *symmetric* families $\mathcal{X}$ which are invariant under permutations of the colours. Equivalently, $\mathcal{X}$ is symmetric if there is a family $\mathcal{P}$ of partitions of $E(K_k)$ such that $\mathcal{X} = \bigcup_{P\in\mathcal{P}} \mathcal{X}_P$ where each $\mathcal{X}_P$ contains all colourings that can be obtained by injectively assigning colours to the classes of P (an example is given in Figure 1.3.1). We refer to such partitions as *colour patterns*. When $\mathcal{P}$ contains a single colour pattern P, we write (P, s) as shorthand for $\mathcal{X}_P$ and say that an s-edge-colouring of a graph G is *P-free* if it is $\mathcal{X}_P$-free. From now on, let $K_k^{(\ell)}$ denote the family of all patterns on K_k with ℓ classes. With this notation, the monochromatic pattern $\mathcal{X}_{k,s}^{(1)}$ is $(K_k^{(1)}, s)$.

The generalised Erdős-Rothschild problem for symmetric $\mathcal{X}$ has been studied in [2, 9, 10, 13, 14, 19, 53–55, 57–59, 74, 75, 77]; known results and the corresponding $\mathcal{X}_{\mathcal{P}}$-extremal graphs are summarised in Table 1.1. Observe that if the number $s \geqslant 2$ of colours is less than the number of parts in P, then the unique (P, s)-extremal graph is the complete graph since every s-edge-colouring of every graph is P-free. Benevides, Hoppen and Sampaio gave a simple proof [19, Theorem 1.1], using Zykov symmetrisation, to show that for every number $s \geqslant 2$ of colours and colour pattern P with at most s parts, there exists a complete partite extremal graph. It was much less clear whether there can be other extremal graphs, but the authors of [19] proved that for all but very few exceptional colour patterns, *every* extremal graph is complete partite.

Theorem 1.17 ([2, 9, 19, 22, 75])**.** *Let $s \geqslant 2$ and $k \geqslant 3$ and let P be a colour pattern of K_k. Suppose that none of the following hold.*

1. $P = K_k^{(1)}$ *is the monochromatic pattern and* $s \notin \{2, 3, 4, 6, 7\}$*;*

2. $k = 3$*,* $s \geqslant 4$ *and* $P = K_3^{(2)}$ *is the (unique) pattern on* K_3 *with* 2 *colour classes.*

Then for sufficiently large n*, every* n*-vertex* (P, s)*-extremal graph is complete partite.*

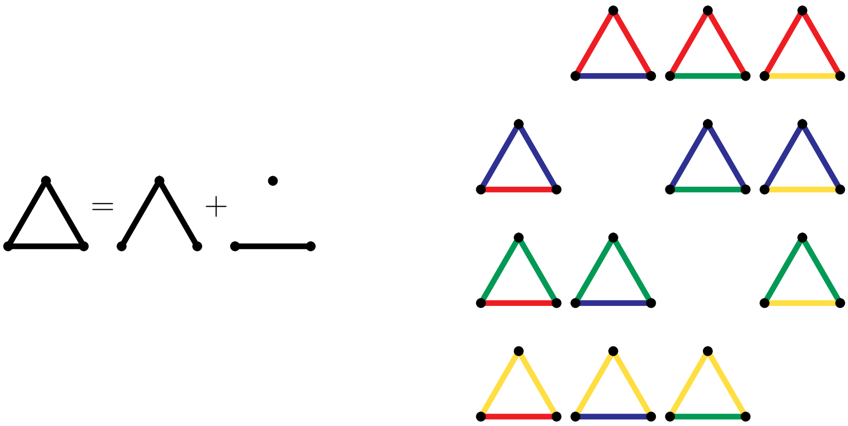


Figure 1.3.1: The pattern $K_3^{(2)}$ and the family $(K_3^{(2)}, 4)$.

Note that when the number of parts in P is larger than s, the problem is degenerate since every s-edge-colouring of any graph is valid and therefore (the complete multipartite graph) K_n is the unique (P, s)-extremal graph on n vertices. Otherwise, the proof in [19] works for all non-monochromatic patterns apart from the exception $K_3^{(2)}$ given in the theorem and a certain pattern on K_4 with two colours. It was also shown in [19] that $T_{k-1}(n)$ is the unique (P, s)-extremal graph when $s = 3$ and P is a colour pattern on K_k with two colour classes such that the graph formed by one colour class has Ramsey number at most k. In particular, this gives $T_2(n)$ as the unique extremal graph for the pattern $K_3^{(2)}$ and $s = 3$ colours. The other cases covered by Theorem 1.17 follow from [2, 22, 75] (monochromatic) and [9] (two colours).

As mentioned in Section 1.3.1, there are a handful of other sporadic (monochromatic) cases for which the conclusion of Theorem 1.17 holds. Pikhurko and Staden [75] proved that extremal graphs for the monochromatic case are complete partite, provided the solutions of a certain finite optimisation problem have a particular 'extension property'. As part of our proof framework we show that extremal graphs for the generalised Erdős-Rothschild problem are complete partite and described by solutions to the corresponding optimisation problem, provided $\mathcal{X}$ satisfies a strong analogue of the 'extension property'. The optimisation problem will later be formally stated in Section 4.2 and our main results linking it to the generalised Erdős-Rothschild problem are stated in Section 4.3.

The generalised Erdős-Rothschild problem for patterns on K_3 is the most widely studied so far. A $K_3^{(3)}$-free colouring of a graph (that is, rainbow triangles are forbidden) is also known as a *Gallai colouring*, a term introduced by Gyárfás and Simonyi [46] following an earlier work of Gallai [42] on comparability graphs. Balogh and Li [10] and Hoppen, Lefmann and Odermann [57] solved the forbidden rainbow triangle problem completely for large n. Their results imply that for $s = 3$, the complete graph K_n is the unique extremal graph, but for $s \geqslant 4$, $T_2(n)$ is the unique extremal graph. Hoppen, Lefmann and Schmidt [54, 59] studied the *dichromatic triangle* pattern $K_3^{(2)}$, which is the focus of our first main result. We discuss the dichromatic triangle pattern in more detail in Section 1.3.3. In terms of rainbow patterns, we proposed in [45] to study the next open case $K_4^{(6)}$ for any number $s \geqslant 6$ of colours. It had been shown in [55] that for $s \geqslant 5434$ the extremal graph is $T_3(n)$. We conjectured, by comparing the number of 5-colourings of K_n and the number of s-colourings of $T_3(n)$, that the extremal graph is $T_3(n)$ when $s \geqslant 12$, and K_n for $s \leqslant 11$. The first part of this conjecture concerning $s \geqslant 12$ was confirmed very recently by Hàn, Hoppen, Müller, and Schmidt [47].

The generalised Erdős-Rothschild problem has been studied for families of colourings that combine multiple patterns (these results are included in Table 1.1). Recall that $K_k^{(\ell)}$ denotes the family of patterns on K_k with exactly ℓ colour classes. Analogously, let $K_k^{(\leqslant \ell)}$ and $K_k^{(\geqslant \ell)}$ be the families of patterns with at most ℓ and at least ℓ classes. Hoppen, Lefmann and Nolibos [55] showed that for integers $k \geqslant 4$ and $2 \leqslant \ell \leqslant \binom{k}{2}$ and $s > s_0(k, \ell)$, the unique $(K_k^{(\geqslant \ell)}, s)$-extremal graph on n vertices is the trivial example $T_{k-1}(n)$, for large n. In particular, for sufficiently large s and every symmetric family containing the rainbow pattern, the trivial lower bound is tight. On the other hand, as noted in [14], a construction by Hoppen, Lefmann and Odermann [57, Remark 4.2] implies that for any symmetric family not containing the rainbow pattern, $T_{k-1}(n)$ is not extremal for $s \geqslant s_2(k) = \binom{k+1}{2}^{k^2}$ and sufficiently large n. For the family $K_k^{(\leqslant \binom{k}{2}-1)}$ of all non-rainbow colourings, their $s_2(k)$ was complemented by Bastos, Hoppen, Lefmann, Oertel and Schmidt [14] showing that whenever $s \leqslant s_1(k) \approx (k/2)^{k/2}$, the trivial example $T_{k-1}(n)$ is the unique $(K_k^{(\leqslant \binom{k}{2}-1)}, s)$-extremal graph for sufficiently large n.

This tells us exactly for which symmetric families the trivial bound is tight for large s. We may conjecture that for any symmetric family $\mathcal{X}$ containing the rainbow pattern, there is a threshold $s^* = s^*(\mathcal{X})$ such that the trivial bound is tight if and only if $s > s^*$. Among families not containing the rainbow pattern, we think that some (such as $K_k^{(\leqslant \binom{k}{2}-1)}$) might exhibit the opposite behaviour: the trivial bound is tight if and only if $s < s^*$ for some fixed s^*. However, for less restrictive families of this type (such as $K_k^{(\binom{k}{2}-1)}$), it seems that K_n might again be extremal for small s; the trivial bound could then be tight for somewhat larger s, before ceasing to be tight for sufficiently large s.

Results for the more interesting 'non-trivial' range are particularly scarce. Our two main results show that this dichotomy holds for $K_3^{(2)}$ and when $\mathcal{X}$ is the family of all improper colourings of K_k. For both problems, solutions in the non-trivial range exhibit an infinite sequence of phase transitions, giving an infinite family of extremal graphs. This is the first such result in the literature.

Perhaps surprisingly, the requirement that n is large cannot be removed in general. Indeed, Theorem 1.17 does not hold for small n. In [2] the authors showed that for $n \leqslant s^{\frac{k-2}{2}}$ the complete graph K_n has more s-edge-colourings free of monochromatic K_k than $T_{k-1}(n)$. In [48] this bound on n was shown to be of the correct order for $s = 2, 3$.

The Erdős-Rothschild problem has been studied in many other contexts. The results of [74–76] about the monochromatic pattern apply in the more general setting where the size of the forbidden monochromatic cliques depends on the colour. Patterns on graphs other than the clique K_k have also been studied, e.g. for bipartite graphs including stars and matchings [31, 52, 53, 57, 58], and for hypergraphs [32, 65–67]. There are many variants of the problem for different discrete structures, that is, determining the maximum number of colourings of a discrete structure in which certain local substructures are forbidden. Some examples include sum-free sets [29, 68], set systems [30, 51], linear vector spaces [30, 56], and partial orders [34].

A notable variant is another question of Erdős [37] posed in the same paper as the Erdős-Rothschild problem: given an oriented graph F, what is the maximum number of F-free orientations of an n-vertex graph G? This was solved for F a tournament by Alon and Yuster in [4]: this is closely related to the 2-colour monochromatic Erdős-Rothschild problem, and the unique extremal graph is the trivial example $T_{v(F)-1}(n)$. Recently, Bucić, Janzer and Sudakov [23] resolved all remaining cases in an asymptotic sense and fully resolved for large n the case when F is an oriented odd cycle, extending the work of Araújo, Botler and Mota [5].

Forbidden pattern(s) on K_k, $k \geqslant 3$	s	Extremal graph on n vertices, n large	Reference
$K_k^{(1)}$	$2, 3$	$T_{k-1}(n)$	[2]
	$\geqslant 4$	non-trivial	
$K_3^{(1)}$	4	$T_4(n)$	[77]
	5	asymptotically optimal family of complete partite graphs with 4, 6 or 8 parts	[22]
	6	$T_8(n)$	
	7	$T_8(n)$	[75]
$K_4^{(1)}$	4	$T_9(n)$	[77]
$K_3^{(2)}$	$\leqslant 26$	$T_2(n)$	[54, 59]
	$\geqslant 27$	$T_r(n)$ for some $r \in R_2(s)$	Theorem 1.21
any family $\subseteq K_k^{(2)}$	2	$T_{k-1}(n)$	[9]
any $P \in K_k^{(2)}$ s.t. some class J has $R(J, J) \leqslant k$	3	$T_{k-1}(n)$	[19]
$K_3^{(3)}$	3	K_n	[10, 12]
	$\geqslant 4$	$T_2(n)$	[10, 57]
$K_4^{(6)}$	$\geqslant 12$	$T_3(n)$	[47]
any $P \in K_k^{(\geqslant 2)}$	$\geqslant 2$	complete multipartite	[9, 19], Corollary 1.23
$K_k^{(\geqslant \ell)}$	$\geqslant s_0(k, \ell)$	$T_{k-1}(n)$	[55]
$K_k^{(\leqslant \binom{k}{2}-1)}$	$\leqslant s_1(k)$	$T_{k-1}(n)$	[57]
	$\geqslant s_2(k)$	non-trivial	[14]
all improper P	$\leqslant s(k)$	$T_{k-1}(n)$	Theorem 1.24
	$> s(k)$	$T_r(n)$ for some $r \in R_2(s)$	

Table 1.1: All results on the generalised Erdős-Rothschild problem for symmetric families. Results where the extremal graph is non-trivial are highlighted.

1.3.3 Forbidding dichromatic triangles

There are exactly three patterns on K_3, namely the *monochromatic pattern* $K_3^{(1)}$, the *dichromatic pattern* $K_3^{(2)}$ (shown in Figure 1.3.1), and the *rainbow pattern* $K_3^{(3)}$. The forbidden monochromatic triangle problem is notoriously difficult and is solved only for $2 \leqslant s \leqslant 7$ (asymptotically only for $s = 5$) [22, 75, 77]. We refer the reader to [75] for more details, where it was speculated that for general s, based on the existing results, extremal graphs could be related to Hadamard matrices. Recall that colourings avoiding the rainbow triangle pattern $K_3^{(3)}$ are maximised by K_n for $s = 3$ colours and $T_2(n)$ for $s \geqslant 4$ colours [10, 57].

The first results on the dichromatic pattern were obtained by Hoppen and Lefmann [54] who showed that for $2 \leqslant s \leqslant 12$ the trivial lower bound is tight and $T_2(n)$ is the unique extremal graph. The same authors together with Schmidt [59] showed that the result extends to $s \leqslant 26$, but not beyond, and conjectured that the extremal graph for $s = 27$ is $T_4(n)$.

Theorem 1.18 ([54, 59])**.** *For every integer $2 \leqslant s \leqslant 26$ and n sufficiently large,*

$$F(n; (K_3^{(2)}, s)) = s^{t_2(n)}$$

and the unique n-vertex graph with the maximum number of $K_3^{(2)}$-free s-edge-colourings is $T_2(n)$. For every integer $s \geqslant 27$, we have $F(n; (K_3^{(2)}, s)) > s^{t_2(n)}$.

The latter part of Theorem 1.18 can be seen by considering the following construction as described in [59] and shown in Figure 1.3.2. Let $s = 27$, let $V_1, \ldots, V_4$ be the vertex classes of $T_4(n)$ and let M_1, M_2, M_3 be the unique decomposition of K_4 into three perfect matchings. Split the set of 27 colours into three sets C_1, C_2, C_3 of size 9. For each edge ij in each matching M_ℓ, colour the edges in $V_i \times V_j$ arbitrarily with colours from C_ℓ. This generates $9^{t_4(n)}$ valid colourings of $T_4(n)$, which after taking all colour permutations yields at least $(1 - o(1))\binom{27}{9,9,9}9^{t_4(n)} > 27^{t_2(n)}$ colourings. The authors of [59] conjectured that $T_4(n)$ is in fact the unique extremal graph for $s = 27$.

Our first main result is an exact solution of the generalised Erdős-Rothschild problem for the dichromatic triangle pattern for all values of s, which as a corollary confirms the above conjecture and resolves the non-monochromatic cases excluded from Theorem 1.17. We show that the above construction on $T_4(n)$ is optimal for $27 \leqslant s \leqslant 496$, followed by analogous constructions on $T_6(n)$, $T_8(n)$, etc. We now describe these constructions before we introduce relevant notation and state our main result for general s.

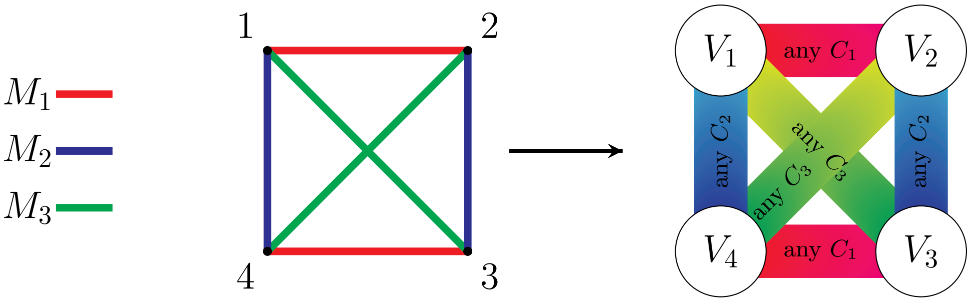


Figure 1.3.2: A lower bound on the number of $K_3^{(2)}$-free s-edge-colourings of $T_4(n)$ is given by splitting $[s] = C_1 \cup C_2 \cup C_3$ and colouring as shown.

Construction 1.19. Given s colours and an integer r, consider a copy of $T_r(n)$ with vertex classes $V_1, \dots, V_r$ and the set $\mathcal{M}$ of maximum matchings of K_r. Note that $\mathcal{M}$ will consist of perfect matchings if r is even and matchings of size $(r-1)/2$ if r is odd. Partition the set of colours $[s]$ into classes $\{C_M\}_{M\in\mathcal{M}}$. For each $ij \in E(K_r)$, colour the edges in $V_i \times V_j$ with colours from $\bigcup_{M\ni ij} C_M$.

Every colouring of $T_r(n)$ generated in this way is free of dichromatic triangles as the only monochromatic cherries (paths on three vertices) have both end vertices in the same vertex class, and thus cannot be completed to a triangle. This yields

$$\prod_{ij\in E(K_r)} \left(\sum_{M\ni ij} |C_M|\right)^{|V_i||V_j|} \tag{1.3.2}$$

colourings free of dichromatic triangles, which is maximised if the sums $\sum_{M\ni ij} |C_M|$ are all as equal as possible; by double counting this means that for every $i \in V(K_r)$ the sets of colours $\{\cup_{M\ni ij} C_M\}_{j\in[r]\setminus\{i\}}$ between V_i and other V_j form an equipartition of $[s]$, so (1.3.2) simplifies to

$$\prod_{i\in[r]} \left[\left\lfloor \frac{s}{r-1} \right\rfloor^{r-1-a} \left\lceil \frac{s}{r-1} \right\rceil^{a} \right]^{\frac{1}{2}(\frac{n}{r})^2 + O(n)},$$

where a is the remainder of s when divided by $r-1$.

We are able to show that for each $s \geqslant 2$, there exists a set $R_2(s)$ consisting of one integer or two consecutive even integers, such that Construction 1.19 is optimal only for $r \in R_2(s)$. We define a few key quantities that capture $R_2(s)$ and the number of colourings generated by Construction 1.19. We write log for the natural logarithm.

s	$R_2(s)$
$[2, 26]$	$\{2\}$
27	$\{2, 4\}$
$[28, 496]$	$\{4\}$
$[497, 5856]$	$\{6\}$
$[5857, 59470]$	$\{8\}$
$[59471, 559116]$	$\{10\}$
$[559117, 5015852]$	$\{12\}$
$[5015853, 10^7]$	$\{14\}$

Table 1.2: $R_2(s)$ for $s \leqslant 10^7$ generated by the script `optr.py`.

Definition 1.20 ($g_s(r)$, $g(s)$, $R_2(s)$, $r_2(s)$)**.** Let $s \geqslant 2$ be the number of colours, and let $2 \leqslant r < s$ be an integer. Let z and $a \in \{0, \ldots, r-2\}$ be the quotient and remainder of s when divided by $(r-1)$; that is, $z = \left\lfloor \frac{s}{r-1} \right\rfloor$ and $a = s - (r-1)z$. Define

$$g_s(r) := \frac{r-1-a}{r} \log(z) + \frac{a}{r} \log(z+1),$$

and

$$g(s) := \max_{\substack{r \in 2\mathbb{N} \\ r < s}} g_s(r), \quad R_2(s) := \{r \in 2\mathbb{N} : g_s(r) = g(s)\}.$$

If $s < e^2$ we let $r_2(s) := 2$; otherwise let $r_2(s)$ be the largest even integer r such that $(r-1)e^r \leqslant s$.

With this notation, Construction 1.19 will yield $O_{r,s}(1) \cdot e^{\frac{r}{r-1} g_s(r) t_r(n)}$ colourings free of dichromatic triangles. (Note that this simplifies to $O_{r,s}(1) \cdot e^{g_s(r)n^2/2}$ when $r|n$; otherwise these quantities differ by an exponential factor in n.) By definition of $R_2(s)$, taking any $r \in R_2(s)$ yields $O_{r,s}(1) \cdot e^{\frac{r}{r-1} g(s) t_r(n)}$ colourings, which we will prove is the maximum for $s \geqslant 2$, and $T_r(n)$ is the extremal graph. A priori, there is no reason why $R_2(s)$ should be easy to describe, or finite, but we are able to show that $R_2(s) \subseteq \{r_2(s), r_2(s)+2\}$. A list of the resulting values of $R_2(s)$ for $s \leqslant 10^7$ is shown in Table 1.2. We are now ready to state our first main result.

Theorem 1.21. *For every integer $s \geqslant 2$ we have $R_2(s) \subseteq \{r_2(s), r_2(s)+2\}$, so in particular $g(s) = \max\{g_s(r_2(s)), g_s(r_2(s)+2)\}$. For sufficiently large n, every n-vertex $(K_3^{(2)}, s)$-extremal graph is $T_r(n)$ for some $r \in R_2(s)$ and*

$$F(n; (K_3^{(2)}, s)) = (C + o(1)) \cdot e^{\frac{r}{r-1} g(s) t_r(n)},$$

where C depends only on s, r and on $n \pmod r$.
Moreover, for all $s \in [2, 10^7] \setminus \{27\}$, the unique extremal graph is $T_r(n)$ where r is the

unique value in $R_2(s)$ shown in Table 1.2, and the unique extremal graph for $s = 27$ is $T_4(n)$.

Theorem 1.21 (together with Lemma 4.22) implies that

$$\lim_{n\to\infty} \frac{\log F(n;(K_3^{(2)}, s))}{\binom{n}{2}} = g(s) \sim W(s/e)$$

where W is the Lambert W-function, defined to be the inverse of $f(x) = xe^x$, which is closely related to the definition of $r_2(s)$.

Even though $g_{27}(2) = g_{27}(4)$, as described above, $T_4(n)$ has more valid colourings than $T_2(n)$, by the multiplicative constant $(1 + o(1))\binom{27}{9,9,9} > 10^{12}$. Thus Theorem 1.21 implies Theorem 1.18 from [54, 59] and confirms the conjecture of Hoppen, Lefmann and Schmidt [59] for the case $s = 27$. We conjecture that $R_2(s)$ is a singleton unless $s = 27$ but are unable to prove this.

Prior to this work, the only pattern which had been solved for all s was the rainbow triangle pattern, for which the extremal graph is one of two, depending on s (described above). By combining Theorem 1.21 with further analysis of the set $R_2(s)$, we show that there is an infinite family of graphs which are extremal for some s, in the following strong sense:

Corollary 1.22. *For any $r \in 2\mathbb{N}$, there are integers $s^-(r) \leqslant s^+(r)$ such that whenever n is sufficiently large, the Turán graph $T_r(n)$ is uniquely $(K_3^{(2)}, s)$-extremal for all $s^-(r) \leqslant s \leqslant s^+(r)$. Moreover, $s^-(r) \in ((r-3)e^{r-2}, (r-1)e^r)$ and there is at most one value of s between $s^+(r)$ and $s^-(r+2)$. If such a value of s exists, for this s every n-vertex $(K_3^{(2)}, s)$-extremal graph lies in $\{T_r(n), T_{r+2}(n)\}$.*

Finally, we immediately obtain the following corollary by combining Theorem 1.21 with Theorem 1.17.

Corollary 1.23. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let P be a non-monochromatic colour pattern of K_k. Then every (P, s)-extremal graph is complete partite.*

1.3.4 Forbidding improperly coloured cliques

Closely related to the dichromatic triangle problem is the following generalised Erdős-Rothschild question on *proper k-clique colourings*. Among all graphs G on n vertices, what is the maximum number of s-edge-colourings of G in which every copy of K_k is properly coloured? Here, an edge-colouring is *proper* if every colour class is a

k	3	4	5	6	7	8	9
$s(k)$	26	3124	531440	$\approx 1 \cdot 10^8$	$\approx 2.6 \cdot 10^{10}$	$\approx 7.5 \cdot 10^{12}$	$\approx 2.5 \cdot 10^{15}$

Table 1.3: Phase transitions in Theorem 1.24 for small values of k generated by the script `improper_patterns.py` attached as an ancillary file.

matching. Equivalently, a proper k-clique colouring is one where two adjacent edges of the same colour must not be contained in any copy of K_k. Proper edge-colourings and particularly the chromatic index $\chi'(G)$, the smallest number of colours needed for a proper edge-colouring of G, have been studied extensively. In e.g. [64, 69, 71] the authors consider the related Linial-Wilf problem of finding graphs with a fixed number of vertices and edges having the maximum number of (globally) proper colourings.

Using the language we introduced so far, determining the graphs with the maximum number of proper k-clique colourings corresponds to finding $F(n; \mathcal{X}^{\wedge}_{k,s})$ and the $\mathcal{X}^{\wedge}_{k,s}$-extremal graphs for $\mathcal{X}^{\wedge}_{k,s} = \{\text{all improper } s\text{-edge-colourings of } K_k\}$. In particular, when $k = 3$, this is equivalent to forbidding both the patterns $K_3^{(2)}$ and $K_3^{(1)}$. Note that for any fixed graph the sets of $\mathcal{X}^{\wedge}_{k,s}$-free colourings form an upward chain as we increase k. As a second main result concerning the Erdős-Rothschild problem, we give the maximum number of proper k-clique colourings for all k and s and classify all extremal graphs. Note that a proper edge-colouring of K_k has at least $k-1$ colours so we only consider the problem in this range.

Theorem 1.24. *For every $k \geqslant 3$ there exists an integer $s(k)$ such that for every $s \geqslant k-1$ and sufficiently large n every n-vertex $\mathcal{X}^{\wedge}_{k,s}$-extremal graph is either*

- *$T_{k-1}(n)$ if $s \leqslant s(k)$, in which case $F(n; \mathcal{X}^{\wedge}_{k,s}) = s^{t_{k-1}(n)}$; or*
- *$T_r(n)$ for some $r \in R_2(s)$ if $s > s(k)$, in which case $F(n; \mathcal{X}^{\wedge}_{k,s}) = (C + o(1)) \cdot e^{\frac{r}{r-1} g(s) t_r(n)}$ where C is a constant depending only on s, r and on $n \pmod r$.*

Some precise values and estimates of $s(k)$ for small k are shown in Table 1.3. Theorem 1.24 is almost a corollary of Theorem 1.21, which explains the similarity in the attained values of $F(n; \mathcal{X})$. We prove Theorem 1.24 similarly using the same reduction to an optimisation problem as promised for the proof of Theorem 1.21, and it turns out that both forbidden families reduce to optimising a version of Construction 1.19. Indeed, note that in Construction 1.19 every clique is properly coloured.

Chapter 4 is based on joint work with Pranshu Gupta, Yani Pehova, and Katherine Staden [45].

Chapter 2

Counting partitions of $G_{n,1/2}$ with degree congruence conditions

Recall from Section 1.1 that for $G = G_{n,1/2}$, we first defined X_n to be the random variable representing the number of partitions of $V(G)$ into disjoint sets $A_1, \ldots, A_q$ so that the degree of each vertex in $G[A_i]$ is divisible by q for all $i \in [q]$, i.e.

$$X_n = \left| \left\{ \{A_1, \ldots, A_q\} : V(G) = \coprod_{i=1}^{q} A_i \text{ and } q \mid d_{G[A_i]}(v) \text{ for all } i \in [q],\ v \in A_i \right\} \right|.$$

The first main result of this chapter determines the distribution of X_n as $n \to \infty$.

Theorem 1.4. *If $q \geqslant 3$ is odd, then $X_n \xrightarrow{d} \mathrm{Po}(1/q!)$. If $q \geqslant 4$ is even, then $X_n \xrightarrow{d} \mathrm{Po}(2^q/q!)$.*

We next restate our more general theorem, which accommodates broader congruence conditions. For non-negative integers $a_0, \ldots, a_{q-1}$ with $\sum_{x=0}^{q-1} a_x = q$, we now let $X_n = X_n^{(a_0,\ldots,a_{q-1})}$ be the number of q-tuples $(\{A_{0,1}, \ldots, A_{0,a_0}\}, \{A_{1,1}, \ldots, A_{1,a_1}\}, \ldots)$, where each entry is an (unordered, possibly empty) set of parts and each part satisfies the degree condition given by its first index, i.e.

$$\begin{aligned} X_n = \Big| \Big\{ (\{A_{0,1}, \ldots, A_{0,a_0}\}, \{A_{1,1}, \ldots, A_{1,a_1}\}, \ldots \{A_{q-1,1}, \ldots, A_{q-1,a_{q-1}}\}) : \\ V(G) = \coprod_{(x,y)\in T} A_{x,y} \text{ and } d_{G[A_{x,y}]}(v) \equiv x \text{ for all } (x,y) \in T,\ v \in A_{x,y} \Big\} \Big|, \end{aligned}$$

where $T = \{(x,y) : x = 0, \ldots, q-1, y = 1, \ldots, a_x\}$ and the congruence is mod q.

Theorem 1.5. *Let $G = G_{n,1/2}$ and let $X_n = X_n^{(a_0,\ldots,a_{q-1})}$ be as defined above for any non-negative integers $a_0, \ldots, a_{q-1}$ with $\sum_{x=0}^{q-1} a_x = q$. For q even, write $c = \sum_{x=0}^{q/2-1} a_{2x}$ for the number of parts where the degree condition is even.*

(1) If $q \geqslant 3$ is odd, then $X_n \xrightarrow{d} \mathrm{Po}(1/\prod a_x!)$.

(2) If $q \geqslant 4$ is even and $c > 0$, then $X_n \xrightarrow{d} \mathrm{Po}(2^c/\prod a_x!)$.

(3) If $q \geqslant 4$ is even and $c = 0$, then $X_n \xrightarrow{d} \mathrm{Po}(2/\prod a_x!)$ as n runs over even *integers.*

The proofs of the two above theorems form the main part of this chapter and will be presented in Section 2.1 (proof of Theorem 1.4), Section 2.2 (proof of Theorem 1.5), and Section 2.3 (proof of algebraic lemmas used in the previous sections).

The case $q = 2$, which is strikingly different from the above, will be dealt with in Section 2.4, where we obtain the full distributions of $X_n^{(2,0)}$, $X_n^{(1,1)}$, and $X_n^{(0,2)}$. These lead to the following asymptotic distributions.

Theorem 1.7. *Let $G = G_{n,1/2}$. Then $X_n^{(2,0)} \xrightarrow{d} X$ and $X_n^{(1,1)} \xrightarrow{d} X$ where*

$$\mathbb{P}(X = 2^k) = c \prod_{i=1}^{k} (2^i - 1)^{-1}$$

for $k \in \mathbb{N} \cup \{0\}$ with constant $c = \prod_{i=0}^{\infty}(1 - 2^{-2i-1}) = (\sum_{j=1}^{\infty} \prod_{i=1}^{j}(2^i - 1)^{-1})^{-1}$, and $\mathbb{P}(X = x) = 0$ if $x \neq 2^k$ for any $k \in \mathbb{N} \cup \{0\}$. Furthermore, $X_n^{(0,2)} \xrightarrow{d} Z$ where n only runs over even integers and

$$\mathbb{P}(Z = 2^k) = c 2^{-k} \prod_{i=1}^{k} (2^i - 1)^{-1}$$

for $k \in \mathbb{N} \cup \{0\}$, $\mathbb{P}(Z = 0) = \frac{1}{3}$ and $\mathbb{P}(Z = x) = 0$ if $x \neq 0$ and $x \neq 2^k$ for any $k \in \mathbb{N} \cup \{0\}$.

2.1 Proof of Theorem 1.4

For some $n, q \in \mathbb{N}$, consider a graph $G = ([n], E)$ and a partition $A = \{A_i\}_{i \in [q]}$ of $[n]$ into q parts. Define the *subgraph induced by A* to be the disjoint union of the subgraphs induced by each part, that is, the spanning subgraph

$$G[A] := G[A_1] \coprod \cdots \coprod G[A_q].$$

Equivalently, $G[A]$ can be obtained from G by removing all edges for which the endvertices lie in different parts of A. We shall call a partition A *good* if the degree of every vertex in $G[A]$ is divisible by q.

With $G = G_{n,1/2}$, let $X = X_n$ be the random variable representing the number of good partitions. To show that X is asymptotically Poisson distributed, we use the method of moments. Let $\mathbb{E}((X)_k) := \mathbb{E}(X(X-1)\dots(X-k+1))$ denote the falling factorial expectation of X.

Theorem 2.1 (see e.g. [21], Theorem 1.22)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of non-negative integer-valued random variables, and $\lambda \geqslant 0$. If*

$$\lim_{n\to\infty} \mathbb{E}((X_n)_k) = \lambda^k$$

for $k = 0, 1, \dots$, then X_n converges in distribution to $\mathrm{Po}(\lambda)$ as $n \to \infty$.

In view of Theorem 2.1, the proof of Theorem 1.4 entails counting the expected number of ordered k-tuples $\mathbf{A} = (A^{(1)}, \dots, A^{(k)})$ of *distinct* partitions $A^{(j)}$ such that each $A^{(j)}$ is good. We use discrete Fourier analysis to write a deterministic expression for the expectation, which we then compute via combinatorial and algebraic means. It turns out that this count mostly comprises k-tuples which intersect 'generically', or are 'independent' in some sense. To describe this situation, we introduce some notation that will be used throughout this section as well as Section 2.2.

For a particular $\mathbf{A}$ let the *coordinates* of a vertex $v \in [n]$ be the k-tuple $\mathbf{c}(v) = (c_1, \dots, c_k)$ such that $v \in (A^{(j)})_{c_j}$ for all j. Then for vertices $u, v \in [n]$, let $I_{u,v} := \{j \in [k] : c(u)_j = c(v)_j\}$. That is, $I_{u,v}$ corresponds to the set of partitions in our k-tuple in which u and v lie in the same part. Given $\mathbf{c}, \mathbf{c}' \in [q]^k$, we will similarly let $I_{\mathbf{c},\mathbf{c}'} = \{j \in [k] : c_j = c'_j\}$. It is quite possible for many vertices to share the same coordinates. Indeed, given $\mathbf{c} \in [q]^k$ we define the *box* $V_{\mathbf{c}} = \{v \in [n] : \mathbf{c}(v) = \mathbf{c}\}$ to be the set of vertices with coordinates $\mathbf{c}$. Equivalently,

$$V_{\mathbf{c}} = \bigcap_{j\in[k]} (A^{(j)})_{c_j}.$$

We will show that the k-tuples of partitions for which all of the associated boxes are reasonably large (and hence the partitions will be essentially independent) contribute $q!^{-k} + o(1)$ to the expectation $\mathbb{E}((X)_k)$ when $q \geqslant 3$ is odd and $2^q q!^{-k} + o(1)$ when $q \geqslant 4$ is even, whilst the remaining configurations contribute $o(1)$ as $n \to \infty$. Since we are just concerned with asymptotics, all statements made throughout should be interpreted with the implicit assumption that n is sufficiently large.

To start the proof, let $q \geqslant 3$ and fix a k-tuple of partitions of $[n]$, say $\mathbf{A} = (A^{(1)}, A^{(2)}, \dots, A^{(k)})$. Our goal is to determine the probability over choice of $G_{n,1/2}$ that all partitions $A^{(j)}$ in this tuple are good. Let $\zeta_q = e^{2\pi i/q}$ ($i = \sqrt{-1}$ in this instance only) and $\mu_q = \{1, \zeta_q, \ldots, \zeta_q^{q-1}\}$ be the qth roots of unity. We denote by $\boldsymbol{\zeta} = (\zeta_{v,j})$ an assignment of roots of unity to vertices for each partition, consisting of a root $\zeta_{v,j} \in \mu_q$ corresponding to each vertex v and $j \in [k]$. Given some $\boldsymbol{\zeta}$, we denote the k-tuple of roots for a fixed vertex v by $\boldsymbol{\zeta}_v = (\zeta_{v,1}, \ldots, \zeta_{v,k})$. Let R be the set of all possible assignments $\boldsymbol{\zeta}$, so that $|R| = q^{kn}$. For each $d \in \mathbb{Z}$ we have

$$\frac{1}{q} \sum_{\zeta \in \mu_q} \zeta^d = \begin{cases} 1, & \text{if } q \mid d; \\ 0, & \text{otherwise.} \end{cases} \tag{2.1.1}$$

Thinking of d as the degree of a vertex v in the spanning subgraph $G_j := G[A^{(j)}]$ induced by the partition $A^{(j)}$, the assigned roots will be used to determine whether v satisfies the required degree condition. Writing $\mathbb{1}\{\ldots\}$ for the indicator function of an event and $\mathbb{E}$ for the expectation over choices of the random graph $G_{n,1/2}$, it follows from (2.1.1) that

$$\begin{aligned} \mathbb{P}(\text{all } A^{(j)} \text{ are good}) &= \mathbb{E}\, \mathbb{1}\{d_{G_j}(v) \equiv 0 \bmod q \text{ for all } v \text{ and } j\} \\ &= \mathbb{E} \prod_{j \in [k]} \prod_{v \in [n]} \frac{1}{q} \sum_{\zeta \in \mu_q} \zeta^{d_{G_j}(v)} \\ &= \mathbb{E} \frac{1}{q^{kn}} \sum_{\boldsymbol{\zeta} \in R} \prod_{j \in [k]} \prod_{v \in [n]} \zeta_{v,j}^{d_{G_j}(v)} \\ &= \mathbb{E} \frac{1}{q^{kn}} \sum_{\boldsymbol{\zeta} \in R} \prod_{j \in [k]} \prod_{vw \in E(G_j)} \zeta_{v,j}\zeta_{w,j} \\ &= \frac{1}{q^{kn}} \sum_{\boldsymbol{\zeta} \in R} \mathbb{E} \prod_{vw \in E(G)} \prod_{j \in I_{v,w}} \zeta_{v,j}\zeta_{w,j} \\ &= \frac{1}{q^{kn}} \sum_{\boldsymbol{\zeta} \in R} \prod_{\{v,w\}} \frac{1}{2}\Big(1 + \prod_{j \in I_{v,w}} \zeta_{v,j}\zeta_{w,j}\Big) \end{aligned}$$

where the last equality holds because $G = G_{n,1/2}$, and the outer product is taken over all 2-element subsets of $[n]$. We can then write

$$\mathbb{E}((X)_k) = \frac{1}{q^{kn}} \sum_{\mathbf{A}} \sum_{\boldsymbol{\zeta} \in R} \prod_{\{v,w\}} \frac{1}{2}\Big(1 + \prod_{j \in I_{v,w}} \zeta_{v,j}\zeta_{w,j}\Big) \tag{2.1.2}$$

with the first sum taken over all k-tuples of (not necessarily good) distinct partitions. The expression in (2.1.2) is deterministic in the sense that it no longer involves the random graph, and we are instead left to work with *configurations*, which are choices $(\mathbf{A}, \boldsymbol{\zeta})$ of a k-tuple of partitions and assignment of roots. To dismiss the possibility that the contributions from different configurations, which in general are complex numbers, may cancel each other out, we will work with the expression

$$\frac{1}{q^{kn}} \sum_{\mathbf{A}} \sum_{\boldsymbol{\zeta} \in R} \prod_{\{v,w\}} \left| \frac{1}{2}\Big(1 + \prod_{j \in I_{v,w}} \zeta_{v,j}\zeta_{w,j}\Big) \right|, \tag{2.1.3}$$

where we take the modulus of the individual contributions. For the computation, we will group together configurations that share certain characteristics and bound and compare the contributions of those groups to (2.1.2) and (2.1.3). For a given configuration, observe that the contribution is small unless almost all pairs v, w satisfy $\prod_{j \in I_{v,w}} \zeta_{v,j}\zeta_{w,j} = 1$. Specifically, say that distinct vertices v and w are in *conflict* if $\prod_{j \in I_{v,w}} \zeta_{v,j}\zeta_{w,j} \neq 1$. For every conflicted pair,

$$\left| \frac{1}{2}\Big(1 + \prod_{j \in I_{v,w}} \zeta_{v,j}\zeta_{w,j}\Big) \right| \leqslant \frac{1}{2}\big|1 + \zeta_q\big| = \cos(\pi/q) \leqslant e^{-1/q^2}. \tag{2.1.4}$$

We will in fact see that the total contribution of all configurations with at least one conflict is $o(1)$. This will show that (2.1.2) and (2.1.3) converge to the same limit, so that we may consider contributions to (2.1.3) rather than directly to $\mathbb{E}((X)_k)$.

Fix $K > kq^2 \log q$. Since there are at most q^{kn} choices of k-tuples $\mathbf{A}$ and the same number of possible assignments $\boldsymbol{\zeta}$, the total contribution to (2.1.3) from configurations with more than Kn conflicted pairs is at most

$$q^{-kn} \cdot q^{2kn} \cdot \left(e^{-1/q^2}\right)^{Kn} = e^{-(K - kq^2 \log q)n/q^2} = o(1) \tag{2.1.5}$$

as $n \to \infty$.

Now fix $C > 2q^2$, and call a vertex *bad* if it is involved in more than $C \log n$ conflicted pairs. A vertex that is not bad is *good*. From (2.1.5), we may assume that there are at most Kn conflicted pairs and hence, allowing for the possibility that bad vertices may be in conflict with each other, at most $2Kn/(C \log n) = o(n)$ bad vertices.

Since there are q^k boxes and q^k choices of $\boldsymbol{\zeta}_v$ for each vertex v, there must be a particular box $V^\star = V_{\mathbf{c}}$ and vertex $v^\star \in V^\star$ such that $\boldsymbol{\zeta}_v = \boldsymbol{\zeta}_{v^\star}$ for at least $q^{-2k}n$ vertices $v \in V^\star$. Such a vertex (in $V^\star$ and assigned the same k-tuple of roots as $v^\star$) will be called a *most common* vertex.

Lemma 2.2. *Let* $(\mathbf{A}, \boldsymbol{\zeta})$ *be a configuration with* $o(n)$ *bad vertices. Then all most common vertices are good. Moreover, each most common vertex is not conflicted with any other good vertices, including other most common vertices.*

Proof. It is enough to observe that since all most common vertices v have the same coordinates and $\boldsymbol{\zeta}_v$, they must all be in conflict with the same set of vertices. As there are only $o(n)$ bad vertices, some most common vertices are good, meaning they are in conflict with at most $C \log n$ vertices. Similarly, for the second statement, if any good vertex is in conflict with a most common vertex then it must be in conflict with all most common vertices, but this exceeds the allowed number $C \log n$ of conflicts for a good vertex. □

We will show that bad vertices may be replaced by most common vertices at a small cost in contribution to (2.1.3). This will allow us to assume that there are no bad vertices in the remainder of the argument.

Lemma 2.3. *The total contribution to* (2.1.3) *from configurations with bad vertices is* $o(1)$ *times the contribution from configurations where all vertices are good.*

Proof. For a k-tuple of partitions $\mathbf{A}$ and assignment $\boldsymbol{\zeta}$, let t be the number of bad vertices. We modify the configuration by replacing all the bad vertices by duplicates of the most common vertices. That is remove all t of them and add t vertices to the box $V^\star$, each assigned $\boldsymbol{\zeta}^\star$. Note that this produces a configuration in which all vertices are good. Indeed, removing vertices cannot increase the number of conflicts, and the most common vertices are not conflicted with any good vertex by Lemma 2.2.

The preceding construction sends at most $\binom{n}{t}(q^k)^t(q^k)^t \leqslant (nq^{2k})^t$ configurations that have t bad vertices to a single configuration without bad vertices. To see this, reversing the process allows at most $\binom{n}{t}$ choices for the bad vertices, q^k choices each for which box they were in and q^k choices each for their original ζ values. On the other hand, we have removed at least $t(C \log n)/2$ conflicts as after this transformation the 'new' most common vertices do not participate in any conflicts by Lemma 2.2. Using the bound from (2.1.4) and the choice of $C > 2q^2$, the sum of contributions to (2.1.3) from all configurations with bad vertices that are transformed into a particular configuration with all vertices good divided by the contribution from that particular configuration is a factor of at most

$$\sum_{t>0} (nq^{2k})^t e^{-(t(C \log n)/2)/q^2} = \sum_{t>0} \left(n^{1-C/2q^2} q^{2k}\right)^t = o(1)$$

as $n \to \infty$. This is true for all configurations without bad vertices, so the lemma follows. □

Henceforth, we assume that there are no bad vertices. This has some useful consequences for boxes $V_{\mathbf{c}}$ with $|V_{\mathbf{c}}| > 2C\log n + 2$. We call such boxes *large*, and the remaining non-empty boxes *small*. Given any vertex $v \in [n]$ and $I \subseteq [k]$ define $\zeta_{v,I} := \prod_{j\in I} \zeta_{v,j}$.

Lemma 2.4. *Let $(\mathbf{A}, \boldsymbol{\zeta})$ be a configuration without bad vertices. Take any pair of boxes $V_{\mathbf{c}}$ and $V_{\mathbf{c}'}$ (possibly the same box), with the former being large, and let $I = I_{\mathbf{c},\mathbf{c}'}$. Then we have $\zeta_{u,I} = \zeta_{v,I}$ for all vertices $u, v \in V_{\mathbf{c}'}$. If $V_{\mathbf{c}}$ and $V_{\mathbf{c}'}$ are both large boxes, then $\zeta_{v,I} = \zeta_{w,I}^{-1}$ for all $v \in V_{\mathbf{c}'}$ and $w \in V_{\mathbf{c}}$. In particular, there are no conflicted pairs within or between large boxes.*

Proof. If $v \in V_{\mathbf{c}'}$ and $w \in V_{\mathbf{c}}$ then v and w are in conflict iff $\zeta_{v,I} \neq \zeta_{w,I}^{-1}$, where $I = I_{\mathbf{c},\mathbf{c}'}$. Since the $v \in V_{\mathbf{c}'}$ are good we deduce that the $\zeta_{v,I}$ must all be equal to more than $|V_{\mathbf{c}}| - 1 - C\log n > |V_{\mathbf{c}}|/2$ values of $\zeta_{w,I}^{-1}$. (The -1 is due to the fact that v lies in $V_{\mathbf{c}}$ when $\mathbf{c} = \mathbf{c}'$.) This proves the first statement. Assuming that $V_{\mathbf{c}}$ and $V_{\mathbf{c}'}$ are both large, applying the first result twice shows that there is a common $\zeta_{v,I}$ for all $v \in V_{\mathbf{c}'}$ and also a common $\zeta_{w,I}$ for all $w \in V_{\mathbf{c}}$, and moreover that $\zeta_{v,I} = \zeta_{w,I}^{-1}$ must be true for all $v \in V_{\mathbf{c}'}$, $w \in V_{\mathbf{c}}$. The final statement follows immediately, noting that we may take $\mathbf{c} = \mathbf{c}'$. □

In light of the first statement in Lemma 2.4, one can define a common value $\zeta_{\mathbf{c}',I}$ equal to $\zeta_{v,I}$ for all $v \in V_{\mathbf{c}'}$ provided there is some large $V_{\mathbf{c}}$ with $I = I_{\mathbf{c},\mathbf{c}'}$.

We now begin to evaluate (2.1.3) by grouping together k-tuples in the first sum depending on how their component partitions intersect. For a k-tuple of partitions $\mathbf{A}$, define $L = L_{\mathbf{A}} := \{\mathbf{c} \in [q]^k : V_{\mathbf{c}} \text{ is large}\}$. Suppose that $L = [q]^k$, meaning all boxes are large. In this case we note that for each $\mathbf{c}$, any $j \in [k]$ and any $v \in V_{\mathbf{c}}$, we can find vertices u, w that satisfy

$$I_{v,u} = I_{v,w} = I_{u,w} = \{j\}.$$

To see this, take for instance $u \in V_{\mathbf{c}'}$ and $w \in V_{\mathbf{c}''}$ where $c''_j = c'_j = c_j$, but c''_i, c'_i, c_i are distinct for all $i \neq j$. Such coordinates exist since $q \geqslant 3$. Now by Lemma 2.4 there are no conflicted vertices between large boxes, so $I_{v,u} = \{j\}$ implies that $\zeta_{v,j} = \zeta_{u,j}^{-1}$ and similarly for the other two pairs of vertices. This gives $\zeta_{v,j} = \zeta_{u,j}^{-1} = \zeta_{w,j} = \zeta_{v,j}^{-1}$, so $\zeta_{v,j} \in \{\pm 1\}$. Moreover, for any u with $\mathbf{c}(u)_j = \mathbf{c}(v)_j$ we can find a w with $I_{v,w} = I_{w,u} = \{j\}$ by picking values to ensure that each $\mathbf{c}(w)_i \neq \mathbf{c}(u)_i, \mathbf{c}(v)_i$, $i \neq j$, and $\mathbf{c}(w)_j = \mathbf{c}(u)_j = \mathbf{c}(v)_j$. Then $\zeta_{v,j} = \zeta_{w,j}^{-1} = \zeta_{u,j}$. Hence $\zeta_{v,j}$ depends only on the value of $\mathbf{c}(v)_j$. Thus, we can write $\zeta_{v,j} = \zeta_{\mathbf{c}(v)_j,j}$ for some choice of $\zeta_{i,j} \in \{\pm 1\}$ with

$i \in [q]$ and $j \in [k]$. Conversely any such choice gives rise to no conflicts. Indeed, for $I = I_{v,w}$ we have

$$\zeta_{v,I}\zeta_{w,I} = \prod_{j\in I} \zeta_{\mathbf{c}(v)_j,j}\zeta_{\mathbf{c}(w)_j,j} = \prod_{j\in I} \zeta^2_{\mathbf{c}(v)_j,j} = 1.$$

Thus there are precisely 2^{kq} choices of ζ values giving rise to no bad vertices when q is even, and only one (all $\zeta_{v,j} = 1$) when q is odd as then $-1 \notin \mu_q$.

Now there are q^n *ordered* partitions (allowing empty parts). An unordered partition without empty parts corresponds to exactly $q!$ ordered partitions and so, as there are only $O((q-1)^n)$ partitions with empty parts, we have $(1 + o(1))q^n/q!$ unordered partitions, whether or not we allow empty parts.

Also, it is easy to see that only $o(q^{kn})$ k-tuples of partitions have $L \neq [q]^k$. Indeed, we may choose a k-tuple of *not necessarily distinct* ordered partitions uniformly at random by independently including each vertex in any box with probability q^{-k}. The probability that a fixed box is not large is $o(1)$. The assertion then follows by taking a union bound over all boxes and noting that the number of not necessarily distinct ordered partitions gives an upper bound on the number of tuples of distinct unordered partitions.

The number of k-tuples with $L = [q]^k$ is then $(q^n/q!)_k(1 + o(1))$ and hence the total contribution to (2.1.3) from these k-tuples is

$$q^{-kn} \cdot (q^n/q!)_k(1 + o(1)) = (1/q!)^k + o(1) \tag{2.1.6}$$

for q odd and

$$q^{-kn} \cdot 2^{kq} \cdot (q^n/q!)_k(1 + o(1)) = (2^q/q!)^k + o(1) \tag{2.1.7}$$

when q is even.

We now introduce a special family of subsets of $[q]^k$ that plays a key role in our analysis of the remaining k-tuples. We say that a subset $B \subseteq [q]^k$ is a *combinatorial subspace* if there are a positive integer r, indices $i_1, \ldots, i_k \in [r]$ and permutations $\phi_1, \ldots, \phi_k$ of $[q]$ such that

$$B = \{(\phi_1(x_{i_1}), \ldots, \phi_k(x_{i_k})) \in [q]^k : x_1, \ldots, x_r \in [q]\}. \tag{2.1.8}$$

In other words, up to permutations, each coordinate follows one of the variables x_i, but different coordinates may follow the same variable. Equivalently, it is a non-empty set of points $(c_1, \dots, c_k) \in [q]^k$ that can be expressed as the intersection of some number of constraints of the form $c_i = \phi_{ij}(c_j)$ where the ϕ_{ij} are permutations of $[q]$.

Combinatorial subspaces capture the situation where two partitions in the tuple are the same, accounting for relabelling of parts. That is, if $L \neq [q]^k$ but forms a combinatorial subspace, and there are no small boxes, then two of the partitions in $\mathbf{A}$ must be the same. Indeed, in the notation of (2.1.8), under these assumptions there must exist distinct $j, j' \in [k]$ such that $i_j = i_{j'}$. Then for $\phi = \phi_{j'} \circ \phi_j^{-1}$ we have $A_i^{(j)} = A_{\phi(i)}^{(j')}$ for all $i \in [q]$. To see this, letting $v \in A_i^{(j)}$ we get $v \in V_c$ with $c \in L$ as there are no small boxes. It follows that $c_{j'} = \phi(c_j) = \phi(i)$, i.e. $v \in A_{\phi(i)}^{(j')}$ as desired.

However, such k-tuples where two partitions are the same do not occur in (2.1.2). The k-tuples yet to be considered therefore fall into one of two classes: those for which L is not a combinatorial subspace and there are no small boxes, and those for which there are small boxes. It suffices to prove that the contribution to (2.1.3) is $o(1)$ in both cases. Then, configurations with at least one conflict contribute $o(1)$ in total to (2.1.3), and (2.1.2) and (2.1.3) converge to the same limit. This, together with (2.1.6) and (2.1.7), would allow us to conclude that $\mathbb{E}((X)_k) \to (1/q!)^k$ for $q \geqslant 3$ odd and $\mathbb{E}((X)_k) \to (2^q/q!)^k$ for $q \geqslant 4$ even as $n \to \infty$, whence Theorem 2.1 completes the proof.

To handle the remaining cases, we introduce some notation. Given a set $B \subseteq [q]^k$ (which we will always choose to be either L or the set $D := \{\mathbf{c} \in [q]^k : V_\mathbf{c} \neq \emptyset\}$ of all non-empty boxes), to each large box $V_\mathbf{c}$, $\mathbf{c} \in L \subseteq B$, we associate the matrix $M^{(\mathbf{c},B)} = \left(M_{j,\mathbf{c}'}^{(\mathbf{c},B)}\right)_{j\in[k],\mathbf{c}'\in B}$ where $M_{j,\mathbf{c}'}^{(\mathbf{c},B)} = \mathbb{1}\{c_j = c'_j\}$. In other words $M^{(\mathbf{c},B)}$ is a $k \times |B|$ matrix and each column is a 0-1 vector of length k corresponding to some $\mathbf{c}' \in B$ such that there is a 1 in row j if $\mathbf{c}$ and $\mathbf{c}'$ agree in the jth component, and 0 otherwise.

The columns of $M^{(\mathbf{c},B)}$ may be viewed as elements of $(\mathbb{Z}/q\mathbb{Z})^k$. Let $\langle M_{\text{col}}^{(\mathbf{c},B)} \rangle$ be the subgroup of $(\mathbb{Z}/q\mathbb{Z})^k$ generated by the columns of $M^{(\mathbf{c},B)}$ and define

$$N_{\mathbf{c},B} = |(\mathbb{Z}/q\mathbb{Z})^k / \langle M_{\text{col}}^{(\mathbf{c},B)} \rangle|$$

to be the size of the quotient group. This quantity is useful due to the following two lemmas which tie combinatorial subspaces to the present algebraic setup.

Lemma 2.5. *If there is at least one solution $\boldsymbol{a} \in (\mathbb{Z}/q\mathbb{Z})^k$ to the congruence*

$$\boldsymbol{a} M^{(\mathbf{c},B)} \equiv \boldsymbol{b} \bmod q \tag{2.1.9}$$

for fixed $\boldsymbol{b} \in (\mathbb{Z}/q\mathbb{Z})^{|B|}$, then the total number of solutions is given by $N_{\mathbf{c},B}$.

Lemma 2.6. *Let $q \geqslant 3$. For all $L \subseteq [q]^k$ we have*

$$\sum_{\mathbf{c} \in L} N_{\mathbf{c},L} \leqslant q^k. \tag{2.1.10}$$

Equality holds if and only if L is a combinatorial subspace.

The proofs of both Lemma 2.5 and Lemma 2.6 are deferred to Section 2.3. We use them in the following lemma, in which we bound the contributions of two particular configuration types. These will allow us to complete the main computation; of the two classes of remaining configurations identified earlier, the first where L is not a combinatorial subspace and there are no small boxes will be covered directly by (i) below, whilst the second in which there are small boxes can be reduced to the special case of (ii). Crucially, the configurations considered are instances for which $\sum_{\mathbf{c} \in L} N_{\mathbf{c},D} < q^k$.

Lemma 2.7. *(i) The total contribution to (2.1.3) from configurations without conflicts, without small boxes, and where L is not a combinatorial subspace is at most $2^{q^k} q^{q^{2k}} (1 - q^{-k})^n$.*

(ii) The total contribution to (2.1.3) from configurations without conflicts and with exactly one small box which has size 1 is at most $2^{q^k} q^{q^{2k}} n(1 - q^{-k})^{n-1}$.

Proof. Given a particular $\mathbf{A}$ we determine the contribution from all configurations $(\mathbf{A}, \boldsymbol{\zeta})$ of the types described above. Since these configurations do not admit conflicts involving large boxes, for all large boxes $V_{\mathbf{c}}$ and non-empty boxes $V_{\mathbf{c}'}$ we can define a common value $\zeta_{\mathbf{c},I_{\mathbf{c},\mathbf{c}'}}$ equal to $\zeta_{v,I_{\mathbf{c},\mathbf{c}'}}$ for all $v \in V_{\mathbf{c}}$. Now also fix a choice of these $\zeta_{\mathbf{c},I_{\mathbf{c},\mathbf{c}'}}$ and let $\zeta_{\mathbf{c},I_{\mathbf{c},\mathbf{c}'}} = \zeta_q^{b_{\mathbf{c},\mathbf{c}'}}$ for all $\mathbf{c} \in L, \mathbf{c}' \in D$. We determine the number of assignments $\boldsymbol{\zeta}$ that comply with these values of $\zeta_{\mathbf{c},I_{\mathbf{c},\mathbf{c}'}}$. Fixing a large box $V_{\mathbf{c}}$ and writing $\zeta_{v,j} = \zeta_q^{a_{v,j}}$ for $v \in V_{\mathbf{c}}$ leads to the constraints

$$\sum_{j \in I_{\mathbf{c},\mathbf{c}'}} a_{v,j} \equiv b_{\mathbf{c},\mathbf{c}'} \bmod q,$$

one for every $\mathbf{c}' \in D$. These can also be formulated as

$$\boldsymbol{a} M^{(\mathbf{c},D)} \equiv \boldsymbol{b} \bmod q,$$

where $\boldsymbol{a} = (a_{v,j})_{j \in [k]} \in (\mathbb{Z}/q\mathbb{Z})^k$ and $\boldsymbol{b} = (b_{\mathbf{c},\mathbf{c}'})_{\mathbf{c}' \in [D]} \in (\mathbb{Z}/q\mathbb{Z})^{|D|}$. Then, with $\boldsymbol{b}$ fixed, Lemma 2.5 tells us that the number of solutions $\boldsymbol{a}$ to this system is at most $N_{\mathbf{c},D}$. Thus, there are at most $N_{\mathbf{c},D}^{|V_{\mathbf{c}}|}$ choices of $\boldsymbol{\zeta}$ for vertices in $V_{\mathbf{c}}$. By repeating the argument for each box and noting that there are at most $q^{q^{2k}}$ choices of $(b_{\mathbf{c},\mathbf{c}'})_{\mathbf{c},\mathbf{c}'}$, we find that for our fixed $\mathbf{A}$ the number of assignments $\boldsymbol{\zeta}$ for which there are no conflicts involving vertices in large boxes is bounded by

$$q^{q^{2k}} \prod_{\mathbf{c} \in L} N_{\mathbf{c},D}^{|V_\mathbf{c}|}.$$

It follows that the configurations for a fixed L (and coordinates $\mathbf{c}^s$ for the small box in case (ii)) and fixed box sizes $|V_\mathbf{c}|$ for $\mathbf{c} \in L$ contribute at most

$$\frac{1}{q^{kn}} \frac{n!}{\prod_{\mathbf{c} \in L} |V_\mathbf{c}|!} \cdot q^{q^{2k}} \prod_{\mathbf{c} \in L} N_{\mathbf{c},D}^{|V_\mathbf{c}|}$$

to the expectation in (2.1.3). Here, the second factor is the multinomial coefficient representing the choices of $\mathbf{A}$ that lead to the box sizes that were fixed before. Note that this is also valid for case (ii) since $|V_{c^s}| = 1$.

In case (i), since L is not a combinatorial subspace we obtain from Lemma 2.6 that $\sum_{\mathbf{c} \in D} N_{\mathbf{c},D} = \sum_{\mathbf{c} \in L} N_{\mathbf{c},L} < q^k$. Taking a sum over possible box sizes and applying the multinomial theorem, we see that the contribution to (2.1.3) from configurations of type (i) with just a fixed L is at most

$$\begin{aligned} \sum_{(|V_\mathbf{c}|:\mathbf{c} \in L)} \frac{n!}{\prod_{\mathbf{c} \in L} |V_\mathbf{c}|!} \cdot q^{-kn} q^{q^{2k}} \prod_{\mathbf{c} \in L} N_{\mathbf{c},D}^{|V_\mathbf{c}|} &= q^{q^{2k}} q^{-kn} \Big(\sum_{c \in L} N_{\mathbf{c},D} \Big)^n \\ &\leqslant q^{q^{2k}} q^{-kn} (q^k - 1)^n \\ &= q^{q^{2k}} (1 - q^{-k})^n. \end{aligned} \tag{2.1.11}$$

This establishes statement (i) as there are at most 2^{q^k} choices of L that do not form a combinatorial subspace (since these are subsets of $[q]^k$).

We claim that the strict inequality $\sum_{\mathbf{c} \in L} N_{\mathbf{c},D} < q^k$ also holds in case (ii). This is immediate from Lemma 2.6 if L is not a combinatorial subspace as clearly $N_{\mathbf{c},D} \leqslant N_{\mathbf{c},L}$. If L is a combinatorial subspace, then Lemma 2.6 only tells us that $\sum_{\mathbf{c} \in L} N_{\mathbf{c},L} = q^k$. However, adding a column corresponding to the small box $V_{\mathbf{c}^s}$ will necessarily decrease $N_{\mathbf{c},D}$ for some $\mathbf{c} \in L$, i.e. $N_{\mathbf{c},D} < N_{\mathbf{c},L}$. Indeed, when L is of the form (2.1.8) there must be some j_1, j_2 with $i_{j_1} = i_{j_2}$ but $\phi_{j_1}^{-1}(c_{j_1}^s) \neq \phi_{j_2}^{-1}(c_{j_2}^s)$. But then adding a column corresponding to $\mathbf{c}^s$ to $M^{(\mathbf{c},L)}$ for any $\mathbf{c}$ with $c_{j_1} = c_{j_1}^s$ and hence $c_{j_2} \neq c_{j_2}^s$ will ensure that $|\langle M_{\mathrm{col}}^{(\mathbf{c},D)} \rangle| > |\langle M_{\mathrm{col}}^{(\mathbf{c},L)} \rangle|$, and hence $N_{\mathbf{c},D} < N_{\mathbf{c},L}$. Thus, we again get $\sum_{\mathbf{c} \in D} N_{\mathbf{c},D} < \sum_{\mathbf{c} \in L} N_{\mathbf{c},L} \leqslant q^k$.

Now using an analogous calculation to case (i), the configurations of type (ii) with fixed L and $\mathbf{c}^s$ contribute at most

$$\begin{aligned}\sum_{(|V_{\mathbf{c}}|:\mathbf{c}\in L)} \frac{n!}{\prod_{\mathbf{c}\in L}|V_{\mathbf{c}}|!}\cdot q^{-kn}q^{q^{2k}}\prod_{\mathbf{c}\in L}N_{\mathbf{c},D}^{|V_{\mathbf{c}}|} &= q^{q^{2k}}q^{-kn}n\Big(\sum_{c\in L}N_{\mathbf{c},D}\Big)^{n-1}\\ &\leqslant q^{q^{2k}}q^{-kn}n(q^k-1)^{n-1}\\ &= q^{q^{2k}}q^{-k}n(1-q^{-k})^{n-1}. \qquad (2.1.12)\end{aligned}$$

The second result follows as there are at most q^k choices of $\mathbf{c}^s$ and 2^{q^k} choices of L. □

Lemma 2.7(i) directly yields that k-tuples of partitions for which L is not a combinatorial subspace and there are no small boxes, contribute $o(1)$ to (2.1.3). Indeed, by Lemma 2.3 we need only consider configurations without bad vertices and by Lemma 2.4 these do not admit any conflicts. These together provide the hypothesis of Lemma 2.7(i).

The remaining k-tuples of partitions are those for which there is at least one small box. Again by Lemma 2.3 and Lemma 2.4, we may assume that in all such configurations all vertices $v \in V_{\mathbf{c}'}$ are good and have the same value of $\zeta_{v,I_{\mathbf{c},\mathbf{c}'}}$ for all large boxes $V_{\mathbf{c}}$.

Fix a small box $V_{\mathbf{c}^s} \neq \emptyset$ and an arbitrary vertex $v^\star \in V_{\mathbf{c}^s}$. Remove all vertices in all small boxes except $v^\star$ and all vertices from large boxes that are in conflict with $v^\star$. Then replace each removed vertex by a duplicate most common vertex. Note that the resulting configurations do not have any conflicts and hence, by Lemma 2.7(ii), contribute at most $2^{q^k}q^{q^{2k}}n(1-q^{-k})^{n-1}$ to (2.1.3). The total number of vertices replaced by most common vertices is $t \leqslant q^k(2C\log n + 2)$ as small boxes contain at most $2C\log n + 2$ vertices and $v^\star$ was not in conflict with more than $C\log n$ vertices in any large box. As previously noted, this construction sends at most $\sum_t (nq^{2k})^t$ configurations to a single resulting configuration. Therefore, the total contribution to (2.1.3) from configurations with small boxes is at most

$$\sum_{t\leqslant q^k(2C\log n+2)} (nq^{2k})^t \cdot 2^{q^k}q^{q^{2k}}n(1-q^{-k})^{n-1} \leqslant e^{O((\log n)^2)-q^{-k}n} = o(1).$$

This completes the proof of Theorem 1.4.

2.2 Proof of Theorem 1.5

As the proof of Theorem 1.5 is very similar to that of Theorem 1.4, we will only describe the points at which the calculation deviates.

Fix non-negative integers $a_0, \ldots, a_{q-1}$ that satisfy $\sum_{x=0}^{q-1} a_x = q$. Recall that T is the index set $\{(x, y) : x = 0, \ldots, q-1, y = 1, \ldots, a_x\}$. We shall say that a vertex partition A into sets $A_{x,y}$ with $(x, y) \in T$ is *good* if

$$d_{G[A_{x,y}]}(v) \equiv x \bmod q$$

for all $(x, y) \in T$ and $v \in A_{x,y}$. We again begin by determining the probability that each partition in a fixed k-tuple $\mathbf{A} = (A^{(1)}, A^{(2)}, \ldots, A^{(k)})$ is good. Assign roots of unity $\zeta_{v,j}$ as before and let R again denote the set of all possible assignments of roots to vertices for each partition. For each vertex v we have

$$\frac{1}{q} \sum_{\zeta \in \mu_q} \zeta^{d_{G_j}(v) - x_{v,j}} = \begin{cases} 1 & \text{if } d_{G_j}(v) \equiv x_{v,j} \bmod q \\ 0 & \text{otherwise} \end{cases} \tag{2.2.1}$$

where $x_{v,j}$ represents the congruence class corresponding to the part of $A^{(j)}$ to which the vertex v belongs (i.e. $v \in (A^{(j)})_{x_{v,j},y}$ for some y). Letting Y be the event that all $A^{(j)}$ are good, then

$$\begin{aligned}
\mathbb{P}(Y) &= \mathbb{E} \frac{1}{q^{kn}} \sum_{\boldsymbol{\zeta} \in R} \prod_{j \in [k]} \prod_{v \in [n]} \zeta_{v,j}^{d_{G_j}(v) - x_{v,j}} \\
&= \frac{1}{q^{kn}} \sum_{\boldsymbol{\zeta} \in R} \left(\prod_{j \in [k]} \prod_{v \in [n]} \zeta_{v,j}^{-x_{v,j}} \right) \left(\mathbb{E} \prod_{j \in [k]} \prod_{vw \in E(G_j)} \zeta_{v,j} \zeta_{w,j} \right) \\
&= \frac{1}{q^{kn}} \sum_{\boldsymbol{\zeta} \in R} \left(\prod_{j \in [k]} \prod_{v \in [n]} \zeta_{v,j}^{-x_{v,j}} \right) \left(\mathbb{E} \prod_{vw \in E(G)} \prod_{j \in I_{v,w}} \zeta_{v,j} \zeta_{w,j} \right) \\
&= \frac{1}{q^{kn}} \sum_{\boldsymbol{\zeta} \in R} \left(\prod_{j \in [k]} \prod_{v \in [n]} \zeta_{v,j}^{-x_{v,j}} \right) \left(\prod_{\{v,w\}} \frac{1}{2} \Big(1 + \prod_{j \in I_{v,w}} \zeta_{v,j} \zeta_{w,j} \Big) \right).
\end{aligned}$$

Hence, we can write

$$\mathbb{E}((X_n)_k) = \frac{1}{q^{kn}} \sum_{\mathbf{A}} \sum_{\boldsymbol{\zeta} \in R} \left(\prod_{j \in [k]} \prod_{v \in [n]} \zeta_{v,j}^{-x_{v,j}} \right) \left(\prod_{\{v,w\}} \frac{1}{2} \Big(1 + \prod_{j \in I_{v,w}} \zeta_{v,j} \zeta_{w,j} \Big) \right)$$

with the first sum taken over all k-tuples of distinct partitions. Following the arguments of the proof of Theorem 1.4 we can show that summing over all configurations $(\mathbf{A}, \boldsymbol{\zeta})$ with at least one conflict, or with $L_{\mathbf{A}} \neq [q]^k$,

$$\frac{1}{q^{kn}} \sum_{(\mathbf{A}, \boldsymbol{\zeta})} \left| \prod_{\{v,w\}} \frac{1}{2} \Big(1 + \prod_{j \in I_{v,w}} \zeta_{v,j} \zeta_{w,j} \Big) \right| = o(1).$$

Furthermore we have $\left|\prod_{j\in[k]}\prod_{v\in[n]}\zeta_{v,j}^{-x_{v,j}}\right| = 1$, giving that the contribution from these configurations is indeed still $o(1)$.

For configurations without conflicts and with all boxes large, we may observe that $\prod_{\{v,w\}}\frac{1}{2}\left(1+\prod_{j\in I_{v,w}}\zeta_{v,j}\zeta_{w,j}\right)$ is always 1 and recall that for a fixed $\mathbf{A}$ the value of $\zeta_{v,j}$ depends only on the part of the jth partition to which v belongs. As $\zeta_{v,j} = \zeta_{t,j} \in \{\pm 1\}$ when $v \in (A^{(j)})_t$, this allows us to write

$$\prod_{j\in[k]}\prod_{v\in[n]}\zeta_{v,j}^{-x_{v,j}} = \prod_{t=(x,y)\in T}\prod_{j\in[k]}\zeta_{t,j}^{-x|(A^{(j)})_t|}.$$

If q is odd then all $\zeta_{t,j} = 1$ for all t and j, and this factor is 1. If q is even then we must sum over choices of $\zeta_{t,j} \in \{\pm 1\}$. We find that there is no contribution from k-tuples of partitions $\mathbf{A}$ such that there is some $t = (x, y) \in T$ and $j \in [k]$ with $x|(A^{(j)})_t|$ odd, as then

$$\sum_{(\zeta_{t,j})}\left(\prod_{t=(x,y)\in T}\prod_{j\in[k]}\zeta_{t,j}^{-x|(A^{(j)})_t|}\right) = 0.$$

Hence only tuples $\mathbf{A}$ where $x|(A^{(j)})_t|$ is even for all t and j contribute, in which case we get a factor of 2^{kq} when we sum over the choices of the $\zeta_{t,j}$. Thus for q even and recalling that $c = \sum a_{2x}$ is the number of parts with even degree conditions, we shall insist that the $q - c$ parts $A_{x,y}$ where x is odd are of even size in all partitions.

To count the number of such partitions, it is convenient to work through *ordered* partitions. For this, let us view the congruence conditions as a list consisting of a_x entries equal to x for each $x = 0, \dots q-1$, and ordered so that all $q - c$ odd values appear before the c even ones. So, let P be the number of ordered partitions of $[n]$ with the additional condition that when q is even the first $q - c$ parts have even size. Allowing for some of the parts to be empty, we note that P is also the number of functions $[n] \to [q]$ where the preimages of $1, \dots, q-c$ are of even size. The number of functions mapping n_i elements to i for $i = 1, \dots, q$ is precisely the coefficient of $z_1^{n_1}z_2^{n_2}\dots$ in the multinomial expansion of $(z_1 + z_2 + \dots + z_q)^n$. Thus, the number for which $n_1, \dots, n_{q-c}$ are even can be obtained by averaging this expression over all choices of $z_i = \pm 1$ when $i \leqslant q - c$ while fixing $z_i = 1$ for $i > q - c$, i.e.

$$\frac{1}{2^{q-c}}\sum_{\varepsilon_i\in\{\pm 1\}}(\varepsilon_1 + \dots + \varepsilon_{q-c} + 1 + \dots + 1)^n = \frac{1}{2^{q-c}}\sum_{i=0}^{q-c}\binom{q-c}{i}(q-2i)^n.$$

When $c > 0$, this is $2^{-(q-c)}(q^n + O((q-2)^n))$ as $n \to \infty$. However, when $c = 0$ this becomes $2^{-q}(q^n + (-q)^n + O((q-2)^n))$. Thus the number of ordered partitions with the first $q-c$ parts of even size is $(1+o(1))2^{-(q-c)}q^n$ if $c>0$, and $(1+o(1))2^{-(q-1)}q^n$ if $c=0$ and n is even (and in fact precisely zero if $c=0$ and n is odd). To summarise, we have

$$P = \begin{cases} q^n, & q \text{ odd}, \\ (1+o(1))2^c \cdot q^n/2^q, & q \text{ even}, \ c>0, \\ (1+o(1))2 \cdot q^n/2^q, & q \text{ even}, \ c=0, \ n \text{ even}. \end{cases}$$

It remains to adjust our counts for unordered partitions. Taking k-tuples of ordered partitions, it is possible to have a degenerate case in which two partitions $A^{(i)}$ and $A^{(j)}$ are the same unordered partition. Now either $x_1 = x_2$ whenever $(A^{(i)})_{x_1,y_1} = (A^{(j)})_{x_2,y_2}$ meaning that corresponding parts have the same congruence condition, or else they are incompatible. For the former case, by ordering the elements of $\{A_{x,1}, \dots, A_{x,a_x}\}$ for each $x \in \{0, \dots, q-1\}$, each unordered good partition without empty parts gives rise to $F := \prod_{x=0}^{q-1} a_x!$ ordered partitions with identical degree conditions. All of these count as a single partition when calculating the random variable X_n.

The latter case occurs when there are $(A^{(i)})_{x_1,y_1} = (A^{(j)})_{x_2,y_2}$ for which $x_1 \neq x_2$. However, it is then impossible for the vertices of $G[(A^{(i)})_{x_1,y_1}] = G[(A^{(j)})_{x_2,y_2}]$ to satisfy both degree conditions simultaneously so such k-tuples do not contribute to $\mathbb{E}((X_n)_k)$. Thus, again noting that only a $o(1)$ proportion of the partitions counted by P have empty parts, there are $((1+o(1))P/F)_k$ tuples $\mathbf{A}$ of partitions where any part with an odd degree condition is even when q is even. As $o(q^{kn})$ of these tuples have small boxes, we find that there are indeed $((1+o(1))P/F)_k$ tuples contributing 1 each when q is odd and 2^{qk} each when q is even. Thus for $\mathbb{E}((X_n)_k)$ we get $q^{-kn}(P/F)_k(1+o(1))$ when q is odd, and $q^{-kn}(P/F)_k 2^{qk}(1+o(1))$ when q is even. The result now follows by application of Theorem 2.1.

2.3 Proofs of algebraic lemmas

In this section we prove the deferred algebraic lemmas used in the proofs of our main theorems.

The following general statement implies Lemma 2.5. The proof is standard, but we include it for completeness.

Lemma 2.8. *Let M be a $k \times \ell$ matrix with entries in $\mathbb{Z}/q\mathbb{Z}$ and $\boldsymbol{b} \in (\mathbb{Z}/q\mathbb{Z})^\ell$. If there is at least one solution $\boldsymbol{a} \in (\mathbb{Z}/q\mathbb{Z})^k$ to the congruence*

$$\boldsymbol{a}M \equiv \boldsymbol{b} \bmod q \tag{2.3.1}$$

then the total number of solutions is given by

$$N = \left| {(\mathbb{Z}/q\mathbb{Z})^k} \big/ {\langle M_{\mathrm{col}} \rangle} \right| = \frac{q^k}{|\langle M_{\mathrm{col}} \rangle|}.$$

Proof. Define a group homomorphism

$$\begin{aligned} M_{\mathrm{Grp}} \colon (\mathbb{Z}/q\mathbb{Z})^k &\to (\mathbb{Z}/q\mathbb{Z})^\ell \\ \boldsymbol{a} &\mapsto \boldsymbol{a}M \end{aligned}$$

where elements $\boldsymbol{a}$ are viewed as row vectors. The number of solutions to (2.3.1) is the size of the preimage of $\boldsymbol{b}$, which is either empty or a coset of the kernel. That is, if there exists a solution, the number of solutions is $|\ker M_{\mathrm{Grp}}|$. Consider the natural bijection $\phi \colon (\mathbb{Z}/q\mathbb{Z})^k \to \mathrm{Hom}((\mathbb{Z}/q\mathbb{Z})^k, \mathbb{Z}/q\mathbb{Z})$ given by sending $\boldsymbol{a} \mapsto (\boldsymbol{v} \mapsto \boldsymbol{a}\boldsymbol{v}^T)$. We have $\boldsymbol{a}M = \mathbf{0}$ if and only if $\boldsymbol{a}M\boldsymbol{v}^T = \mathbf{0}$ for all $\boldsymbol{v} \in (\mathbb{Z}/q\mathbb{Z})^\ell$ viewed as row vectors, or equivalently $\boldsymbol{a}\boldsymbol{b} = 0$ for all $\boldsymbol{b} \in \langle M_{\mathrm{col}} \rangle = \{M\boldsymbol{v}^T : \boldsymbol{v} \in (\mathbb{Z}/q\mathbb{Z})^\ell\}$. Hence ϕ induces a bijection between elements of $\ker M_{\mathrm{Grp}}$ and morphisms in $\mathrm{Hom}((\mathbb{Z}/q\mathbb{Z})^k, \mathbb{Z}/q\mathbb{Z})$ that annihilate $\langle M_{\mathrm{col}} \rangle$. Such morphisms form a subgroup that is naturally isomorphic to $\mathrm{Hom}((\mathbb{Z}/q\mathbb{Z})^k/\langle M_{\mathrm{col}} \rangle, \mathbb{Z}/q\mathbb{Z})$, which by duality for finite abelian groups is isomorphic to $(\mathbb{Z}/q\mathbb{Z})^k/\langle M_{\mathrm{col}} \rangle$. □

For Lemma 2.6, we recall the surrounding setup from Section 2.1. Since the matrix $M^{(c,B)}$ only appears with $B = L$ in this statement, let us slightly simplify our notation here: given any subset $L \subseteq [q]^k$, associate each $\mathbf{c} \in L$ with the matrix

$$M^{(\mathbf{c})} = \left(M^{(\mathbf{c})}_{j,\mathbf{c}'}\right)_{j \in [k], \mathbf{c}' \in L} \tag{2.3.2}$$

and let

$$N_{\mathbf{c}} = \left| {(\mathbb{Z}/q\mathbb{Z})^k} \big/ {\langle M^{(\mathbf{c})}_{\mathrm{col}} \rangle} \right| = \frac{q^k}{|\langle M^{(\mathbf{c})}_{\mathrm{col}} \rangle|}.$$

Also recall that $q \geqslant 3$.

Proof of Lemma 2.6. We use induction on k. The case $k = 1$ is clear as the $\mathbf{c}$ column of $M^{(\mathbf{c})}$ is (1) and so generates the full group $\mathbb{Z}/q\mathbb{Z}$. Thus

$$\sum_{c \in L} N_c = \sum_{c \in L} 1 \leqslant q,$$

with equality if and only if $L = [q]$.

Now assume that $k > 1$. Define the *k-compression* $\Gamma_k = \Gamma \colon [q]^k \to [q]^{k-1}$ for $\mathbf{c} = (c_1, \ldots, c_k)$ by $\Gamma(\mathbf{c}) = (c_1, \ldots, c_{k-1})$, and let $\Gamma(L) = \{\Gamma(\mathbf{c}) \colon \mathbf{c} \in L\}$. For

$\Gamma(\mathbf{c}) \in [q]^{k-1}$ we define $M^{(\Gamma(\mathbf{c}))}$ and $N_{\Gamma(\mathbf{c})}$ as before, but in $k-1$ dimensions and with respect to $\Gamma(L)$. We note that the matrix $M^{(\Gamma(\mathbf{c}))}$ is obtained from $M^{(\mathbf{c})}$ by deleting the kth row and then possibly deleting some duplicate columns, so $N_{\mathbf{c}} = t \cdot N_{\Gamma(\mathbf{c})}$ for some $t \mid q$. Indeed, this t is the smallest positive integer such that $te_k \in \langle M^{(\mathbf{c})}_{\text{col}} \rangle$ where $e_k = (0, \ldots, 0, 1)^T$. To see this, note that the kernel of the surjective group homomorphism $\langle M^{(\mathbf{c})}_{\text{col}} \rangle \to \langle M^{(\Gamma(\mathbf{c}))}_{\text{col}} \rangle$ which forgets the last coordinate is precisely $\langle te_k \rangle$, so $N_{\mathbf{c}}/N_{\Gamma(\mathbf{c})} = q/(|\langle M^{(\mathbf{c})}_{\text{col}} \rangle|/|\langle M^{(\Gamma(\mathbf{c}))}_{\text{col}} \rangle|) = q/|\langle te_k \rangle| = t$.

We have by induction that

$$\sum_{\mathbf{a} \in \Gamma(L)} N_{\mathbf{a}} \leqslant q^{k-1}, \tag{2.3.3}$$

with equality if and only if $\Gamma(L)$ is a combinatorial subspace.

Fix some $\mathbf{a} \in \Gamma(L)$ and first suppose that there are $\mathbf{c}_1, \mathbf{c}_2 \in L$ with $\mathbf{c}_1 \neq \mathbf{c}_2$ but $\Gamma(\mathbf{c}_1) = \mathbf{a} = \Gamma(\mathbf{c}_2)$. These conditions imply that $\mathbf{c}_1$ and $\mathbf{c}_2$ differ only in the kth coordinate, so subtracting the columns corresponding to $\mathbf{c}_1$ and $\mathbf{c}_2$ shows that e_k is in $\langle M^{(\mathbf{c}_1)}_{\text{col}} \rangle$ and $\langle M^{(\mathbf{c}_2)}_{\text{col}} \rangle$ and hence $N_{\mathbf{c}_1} = N_{\mathbf{c}_2} = N_{\mathbf{a}}$.

If on the other hand $\Gamma^{-1}(\mathbf{a}) \cap L = \{\mathbf{c}\}$ for $\mathbf{a} \in \Gamma(L)$, then $N_{\mathbf{c}} = tN_{\mathbf{a}} \leqslant qN_{\mathbf{a}}$. Thus we see that

$$\sum_{\mathbf{c} : \Gamma(\mathbf{c}) = \mathbf{a}} N_{\mathbf{c}} \leqslant q \cdot N_{\mathbf{a}}, \tag{2.3.4}$$

with equality when $|\{\mathbf{c} \in L : \Gamma(\mathbf{c}) = \mathbf{a}\}| = q$ or when $|\{\mathbf{c} \in L : \Gamma(\mathbf{c}) = \mathbf{a}\}| = 1$ and $te_k \notin \langle M^{(\mathbf{c})}_{\text{col}} \rangle$ for all $t < q$ where $\mathbf{c}$ is the unique preimage of $\mathbf{a}$. Hence we deduce that

$$\sum_{\mathbf{c} \in L} N_{\mathbf{c}} \leqslant q \sum_{\mathbf{a} \in \Gamma(L)} N_{\mathbf{a}} \leqslant q \cdot q^{k-1} = q^k. \tag{2.3.5}$$

as required.

Now for equality we must have $|\{\mathbf{c} \in L : \Gamma(\mathbf{c}) = \mathbf{a}\}| \in \{1, q\}$ for every $\mathbf{a} \in \Gamma(L)$. However, if there exists any such $\mathbf{a}$ with $|\{\mathbf{c} \in L : \Gamma(\mathbf{c}) = \mathbf{a}\}| = q$ then e_k lies in the column span of every $M^{(\mathbf{c})}$. Indeed, we need only subtract the columns corresponding to $\mathbf{c}'$ and $\mathbf{c}''$, where $\Gamma(\mathbf{c}') = \Gamma(\mathbf{c}'') = \mathbf{a}$, $c'_k = c_k$ and $c''_k \neq c_k$. Such $\mathbf{c}'$ and $\mathbf{c}''$ exist as all $\mathbf{c}$ with $\Gamma(\mathbf{c}) = \mathbf{a}$ lie in L. Hence, for equality to hold in (2.3.5) we would now need $|\{\mathbf{c} \in L : \Gamma(\mathbf{c}) = \mathbf{a}\}| = q$ for all $\mathbf{a} \in \Gamma(L)$.

Thus we are reduced to two cases. The first is that $L = \Gamma(L) \times [q]$. In this case $N_{\mathbf{c}} = N_{\Gamma(\mathbf{c})}$ for all $\mathbf{c} \in L$ and equality in (2.3.5) occurs if and only if it does in (2.3.3), which in turn happens if and only if $\Gamma(L)$ is a combinatorial subspace. But $\Gamma(L)$ is a combinatorial subspace if and only if $L = \Gamma(L) \times [q]$ is a combinatorial subspace.

The second case is that $|\{\mathbf{c} \in L : \Gamma(\mathbf{c}) = \mathbf{a}\}| = 1$ for all $\mathbf{a} \in \Gamma(L)$. Again, for equality we must have $\Gamma(L)$ equal to a combinatorial subspace, and moreover e_k is not in the subgroup $\langle M_{\text{col}}^{(\mathbf{c})} \rangle$ for any $\mathbf{c} \in L$. Writing $\Gamma(L)$ as in (2.1.8), we may assume all the ϕ_j are equal to the identity as that just corresponds to permuting the values in $[q]$ in the jth coordinate and this does not affect the matrix. We may also suppose that all the i_j are distinct, since if any component of the combinatorial subspace is bound to another, say $x_{i_a} = x_{i_b}$, this would mean that rows a and b are identical in every matrix $M^{(\mathbf{c})}$. It is clear that deleting duplicate rows does not affect $|\langle M_{\text{col}}^{(\mathbf{c})} \rangle|$. Hence for the a-compression Γ_a obtained by removing coordinate a, we have $qN_{\Gamma_a(\mathbf{c})} = N_{\mathbf{c}}$ for all $\mathbf{c}$ where $N_{\Gamma_a(\mathbf{c})}$ is defined with respect to $\Gamma_a(L)$ in $k-1$ dimensions. Now we only have equality if $\Gamma_a(L)$ is a combinatorial subspace. But that implies that L is also a combinatorial subspace as, given a combinatorial subspace representation of $\Gamma_a(L)$, we can bind coordinate a to b.

Thus we are reduced to the case $\Gamma(L) = [q]^{k-1}$. Define $L(i) = \{\Gamma(\mathbf{c}) : \mathbf{c} \in L,\, c_k = i\}$ to be the layer consisting of elements of $\Gamma(L) = [q]^{k-1}$ whose corresponding element $\mathbf{c} \in L$ has kth coordinate equal to i. As there are only q possible values of i, there must exist an i for which $|L(i)| \geqslant q^{k-2}$.

Fix such an i and for $\mathbf{a}_1 = \Gamma(\mathbf{c}_1), \mathbf{a}_2 = \Gamma(\mathbf{c}_2) \in L(i)$ let $S_{12} = \{j \in [k-1] : (\mathbf{a}_1)_j \neq (\mathbf{a}_2)_j\}$. We note that $(\mathbb{1}\{j \in S_{12}^c\})_j^T$ is equal to the column corresponding to $\mathbf{c}_2$ in $M^{(\mathbf{c}_1)}$. Also, if we have some $\mathbf{a}_3 = \Gamma(\mathbf{c}_3) \in L(i')$ for $i' \neq i$ and $S_{13} = \{j : (\mathbf{a}_1)_j \neq (\mathbf{a}_3)_j\}$ is equal to S_{12}, then subtracting columns $\mathbf{c}_2$ and $\mathbf{c}_3$ in $M^{(\mathbf{c}_1)}$ gives e_k which is a contradiction. Hence, every $\mathbf{a}_3$ with $S_{12} = S_{13}$ must lie in $L(i)$.

Now fix $\mathbf{a}_1, \mathbf{a}_2 \in L(i)$ so as to maximise $|S_{12}|$. We claim that $\mathbf{a}_3 \in L(i)$ if and only if $S_{13} \subseteq S_{12}$. First suppose there exists some $\mathbf{a}_3 \in L(i)$ with S_{13} not a subset of S_{12}. Then take some $\mathbf{a}_4 \in \Gamma(L) = [q]^{k-1}$ for which $(\mathbf{a}_4)_j = (\mathbf{a}_3)_j$ for $j \notin S_{12} \cap S_{13}$, and in addition so that $(\mathbf{a}_4)_j$ is distinct from $(\mathbf{a}_1)_j$ and $(\mathbf{a}_2)_j$ for all $j \in S_{12} \cap S_{13}$. Such a $\mathbf{a}_4$ can be found since $q \geqslant 3$, meaning we have room to pick component values satisfying these conditions. With this choice of $\mathbf{a}_4$, observe that $S_{14} = S_{13}$ and hence $\mathbf{a}_4 \in L(i)$. However $S_{24} = S_{12} \cup S_{13}$ then has more elements than S_{12}, contradicting the choice of $\mathbf{a}_1$ and $\mathbf{a}_2$.

Conversely, suppose $S_{13} \subseteq S_{12}$. Pick $\mathbf{a}_4$ such that $(\mathbf{a}_4)_j = (\mathbf{a}_2)_j$ if $j \notin S_{13}$, and $(\mathbf{a}_4)_j \neq (\mathbf{a}_1)_j, (\mathbf{a}_3)_j$ when $j \in S_{13}$. Then $S_{14} = S_{12}$, so $\mathbf{a}_4 \in L(i)$. Also $S_{34} = S_{14}$, so (basing our arguments at $\mathbf{a}_4$ and noting that $\mathbf{a}_1, \mathbf{a}_4 \in L(i)$) we see that $\mathbf{a}_3 \in L(i)$.

Now given that $\mathbf{a}_3 \in L(i)$ if and only if $S_{13} \subseteq S_{12}$, as $|L(i)| \geqslant q^{k-2}$ we must have $|S_{12}| = k-2$ or $k-1$. If $|S_{12}| = k-1$, then $L(i) = \Gamma(L)$ and so every other $L(j)$ is empty. In this case for any $\mathbf{c} \in L$ we can find $\mathbf{c}' \in L$ that differs from $\mathbf{c}$ in all but the kth coordinate. This means that e_k is a column of every $M^{(\mathbf{c})}$ and we do not have equality.

The other possibility is that $|S_{12}| = k-2$, in which case S_{12} omits exactly one coordinate. Without loss of generality, say it is coordinate 1. But then $L(i) = \{\mathbf{a} \in [q]^{k-1} : (\mathbf{a})_1 = x_i\}$ for some $x_i \in [q]$. Now the remaining $L(j)$ have $(q-1)q^{k-2}$ elements in total, so there is some $j \neq i$ with $|L(j)| \geqslant q^{k-2}$. By the above argument this is again a co-dimension 1 subspace given by an equation of the form $(\mathbf{a})_\ell = x_j$. But $L(j) \cap L(i) = \emptyset$, so $\ell = 1$ and $x_j \neq x_i$.

Continuing in this manner, we see that each $L(j) = \{\mathbf{a} : (\mathbf{a})_1 = x_j\}$ for distinct elements $x_1, \ldots, x_q \in [q]$. Writing $\phi_1(j) = x_j$, $\phi_i(j) = j$ for all $i > 1$, $i_j = j$ for $j < k$ and $i_k = 1$, and $r = k-1$, we obtain a form satisfying (2.1.8) so L is in fact a combinatorial subspace. Moreover, rows 1 and k of each $M^{(\mathbf{c})}$ are identical so we have $N_{\mathbf{c}} \geqslant q$ for all $\mathbf{c} \in L$. Hence equality in (2.1.10) must hold. □

2.4 When $q = 2$

Observe that we have crucially used the assumption that $q > 2$ in both the proof of Lemma 2.6 and the main body of the proof of Theorem 1.4. In fact, one can show directly that it is not possible to have a Poisson distribution when $q = 2$. In this section, we determine the distribution of the number of partitions into two parts both inducing even graphs, both inducing odd graphs, and inducing one odd and one even graph.

It is convenient to work in the setting of uniform random symmetric matrices over $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$, that is, symmetric matrices with entries on the diagonal and upper triangle chosen independently randomly to be 0 or 1 each with probability 1/2. Henceforth, we will just call these *random symmetric matrices*. Let $\mathbf{1} = \mathbf{1}_n$ denote the $n \times 1$ vector of all 1s, $\mathbf{0} = \mathbf{0}_n$ denote the $n \times 1$ vector of all 0s, and $I = I_n$ denote the $n \times n$ identity matrix. We shall call a vector $\mathbf{v}$ even if $\mathbf{1}^T\mathbf{v} = 0$, and odd otherwise. Then say that a matrix A is even if $\mathbf{1}^T A = \mathbf{0}^T$, and odd if $\mathbf{1}^T A = \mathbf{1}^T$. In other words, a matrix is even if all column sums are even, and similarly for odd.

To bring this back to our partitions, let G be a graph on n vertices with degree sequence $d_1, \ldots, d_n$, and let A be its adjacency matrix. Then A is a random symmetric matrix with a fixed diagonal of 0s for $G = G_{n,1/2}$. Let D be the $n \times n$ diagonal matrix with entries on the diagonal being d_i mod 2. Any ordered partition of the vertex set into $q = 2$ parts can be represented by a 0-1 column vector of length n, where all of the positions with 0 entries form one part, whilst the 1 entries form the other.

For each degree parity condition on the two parts (even/even, even/odd, odd/odd), we shall set up a nonhomogeneous linear system $M\mathbf{v} = \mathbf{b}$ where M is a random symmetric matrix over $\mathbb{F}_2$ that is either even or odd, and the solutions are precisely good (ordered) partitions. This immediately implies that the number of solutions, if there are any, must be a power of 2 and hence already precludes the possibility of a Poisson distribution. To obtain the distribution, it suffices to know the probability that the system is inconsistent (this can only occur in the odd/odd case), together with the rank distribution of random symmetric matrices. The latter can be deduced from the following counts. The interested reader may also find Lemma 3.8 useful, where we revisit these probabilities and give a simple proof of a rough upper bound on the probability that the rank of these matrices is at most r.

Theorem 2.9 (MacWilliams [72])**.** *For $0 \leqslant r \leqslant n$, the number of symmetric $n \times n$ matrices over $\mathbb{F}_2$ with rank r is*

$$\prod_{i=1}^{\lfloor r/2 \rfloor} \frac{2^{2i}}{2^{2i}-1} \cdot \prod_{i=0}^{r-1} (2^{n-i}-1).$$

Corollary 2.10. *The probability of a uniformly random even symmetric $n \times n$ matrix having rank r is*

$$2^{-\binom{n}{2}} \cdot \prod_{i=1}^{\lfloor r/2 \rfloor} \frac{2^{2i}}{2^{2i}-1} \cdot \prod_{i=0}^{r-1} (2^{n-i-1}-1)$$

if $0 \leqslant r \leqslant n-1$, and 0 otherwise.

Proof. In an even symmetric matrix M, the entries in the last row and column are completely determined by the minor $(M)_{n,n}$. Explicitly, we can uniquely recover M by adding an nth row given by the sum of the existing rows in the $(n-1) \times (n-1)$ matrix, and then the nth column given by the sum of columns in the $n \times (n-1)$ matrix. As this preserves rank, it follows that the rank distribution of a random even symmetric $n \times n$ matrix is the same as that of a random symmetric $(n-1) \times (n-1)$ matrix with no parity conditions. This shows M cannot have full rank, and that the probability of having rank $r \leqslant n-1$ is equal to the probability that a (fully) random symmetric $(n-1) \times (n-1)$ has rank r. The required probability now follows directly from Theorem 2.9, noting that there are $2^{\binom{n}{2}}$ many $(n-1)\times(n-1)$ symmetric matrices. $\square$

We now address each of the three possible parity conditions in turn.

Theorem 2.11. *Let $G = G_{n,1/2}$, and let $X_n = X_n^{(2,0)}$ be the number of partitions of $V(G)$ into two sets both inducing even subgraphs. Then*

$$\mathbb{P}(X_n = 2^k) = 2^{-\binom{n}{2}} \cdot \prod_{i=1}^{\lfloor (n-1-k)/2 \rfloor} (1-2^{-2i})^{-1} \cdot \prod_{i=k+1}^{n-1} (2^i - 1)$$

for $0 \leqslant k \leqslant n-1$, and if $x \neq 2^k$ for any such k then $\mathbb{P}(X_n = x) = 0$.

Proof. We use the notation defined at the start of the section. Suppose we have an ordered partition of $V(G)$ into two parts, represented by a 0-1 column vector $\mathbf{v}$. The entries of $A\mathbf{v}$ represent degrees into the 1-part of the partition, and entries of $D(\mathbf{1}-\mathbf{v})$ are total degrees (i.e. degrees in G) for vertices of the 0-part and are 0 for vertices of the 1-part. Hence if $\mathbf{v}$ corresponds to a vertex partition into two parts such that both induce even subgraphs, then it is a solution to

$$A\mathbf{v} = D(\mathbf{1}-\mathbf{v}) \iff (A+D)\mathbf{v} = D\mathbf{1},$$

and conversely. The solutions $\mathbf{v}$ to this linear equation form an affine subspace of $\mathbb{F}_2$ and Gallai's theorem ensures that there is at least one solution $\mathbf{v}$ to this linear equation. Hence, the probability that there are 2^k *unordered* partitions is equal to the probability that $A+D$ has rank $r = n-1-k$. Noting that $A+D$ is a uniformly random even symmetric matrix, the even condition being guaranteed by definition of D, the result now follows from Corollary 2.10. $\square$

Theorem 2.12. *Let $G = G_{n,1/2}$, and let $Y_n = X_n^{(1,1)}$ be the number of partitions of $V(G)$ into two sets, one inducing an even subgraph and the other inducing an odd subgraph. If n is odd, then*

$$\mathbb{P}(Y_n = 2^k) = 2^{-\binom{n}{2}} \cdot \prod_{i=1}^{\lfloor (n-1-k)/2 \rfloor} (1-2^{-2i})^{-1} \cdot \prod_{i=k+1}^{n-1} (2^i - 1)$$

for $0 \leqslant k \leqslant n-1$, and $\mathbb{P}(Y_n = x) = 0$ when $x \neq 2^k$ for any such k. If n is even, then $\mathbb{P}(Y_n = x) = \mathbb{P}(Y_{n-1} = x)$.

Proof. It suffices to count ordered partitions for which the 1-part is odd and the 0-part is even since every unordered partition satisfying the degree conditions has a unique ordering for which this is true. The partitions we wish to count are then precisely the solutions to

$$A\mathbf{v} = D(\mathbf{1} - \mathbf{v}) + \mathbf{v} \iff (A + D + I)\mathbf{v} = D\mathbf{1}.$$

By the corollary to Gallai's theorem, this system has at least one solution. Thus, to determine the distribution of Y_n it suffices to know the rank distribution of $O = A + D + I$.

Let n be odd. Since $A + D$ is a uniform random even symmetric matrix, both $A + D + I$ and $A + D + \mathbf{1}\mathbf{1}^T$ are uniform random odd symmetric matrices and hence have the same rank distribution as each other. We will work with the latter. Note that $\mathrm{rk}(A + D + \mathbf{1}\mathbf{1}^T) \leqslant \mathrm{rk}(A + D) + \mathrm{rk}(\mathbf{1}\mathbf{1}^T) = \mathrm{rk}(A + D) + 1$. The column space of $(A+D+\mathbf{1}\mathbf{1}^T)$ contains $(A+D+\mathbf{1}\mathbf{1}^T)\mathbf{1} = \mathbf{1}$ and thus also contains the column space of $A+D$ as, for any vector $\mathbf{v}$, we have $(A+D+\mathbf{1}\mathbf{1}^T)\mathbf{v} \in \{(A+D)\mathbf{v}, (A+D)\mathbf{v}+\mathbf{1}\}$. The odd vector $\mathbf{1}$ is not in the column space of $A+D$, giving $\mathrm{rk}(A+D) < \mathrm{rk}(A+D+\mathbf{1}\mathbf{1}^T)$ and hence $\mathrm{rk}(A + D + \mathbf{1}\mathbf{1}^T) = \mathrm{rk}(A + D) + 1$.

Then for $1 \leqslant r \leqslant n$ we have

$$\begin{aligned}
\mathbb{P}(Y_n = 2^{n-r}) &= \mathbb{P}(\mathrm{rk}(A + D + I) = r) \\
&= \mathbb{P}(\mathrm{rk}(A + D + \mathbf{1}\mathbf{1}^T) = r) \\
&= \mathbb{P}(\mathrm{rk}(A + D) = r - 1)
\end{aligned}$$

which is given by Corollary 2.10.

For n even, the odd random symmetric matrices $O = A + D + I$ and $O + \mathbf{1}\mathbf{1}^T$ have the same rank. To see this, it is enough to note that

$$\{O\mathbf{v}, O(\mathbf{v} + \mathbf{1})\} = \{(O + \mathbf{1}\mathbf{1}^T)\mathbf{v}, (O + \mathbf{1}\mathbf{1}^T)(\mathbf{v} + \mathbf{1})\}$$

for all $\mathbf{v}$, which is easily verified by considering the parity of the number of 1s in $\mathbf{v}$. Thus, we may assume that O is an odd random symmetric matrix with a 1 in the top-left entry as this still has the same rank distribution.

Note that since O is odd, it has $\mathbf{1}^T$ in its row space. We now add $\mathbf{1}^T$ to all rows of O which have a 1 in the first entry, except the first row. The resulting matrix O' still has both row and column sums odd as we added an even vector to an even number of rows. In particular, $\mathbf{1}^T$ is still in the row space and we thus did not change the row space, i.e. $\text{rk}(O') = \text{rk}(O)$. In the same way, we can add $\mathbf{1}$ to all columns of O' which have a 1 in the first entry, except the first column to obtain an odd symmetric matrix O'' which has the same column space as O'. In particular, $\text{rk}(O'') = \text{rk}(O') = \text{rk}(O)$ and the first row and column of O'' is e_1. The remaining $(n-1) \times (n-1)$ matrix $O_{<11>}$ is an odd (uniform) random symmetric matrix since, given the first row of O, we can reconstruct O from $O_{<11>}$. As $\text{rk}(O_{<11>}) = \text{rk}(O) - 1$, the distribution of nullities, and hence number of solutions $\mathbf{v}$, is the same for n as for $n-1$ when n is even.

□

Theorem 2.13. *Let $G = G_{n,1/2}$ with n even, and let $Z_n = X_n^{(0,2)}$ be the number of partitions of $V(G)$ into two sets both inducing odd subgraphs. Then*

$$\mathbb{P}(Z_n = 2^k) = \frac{2^{n-k-1} - 1}{2^{\binom{n}{2}}(2^{n-1} - 1)} \cdot \prod_{i=1}^{\lfloor (n-1-k)/2 \rfloor} (1 - 2^{-2i})^{-1} \cdot \prod_{i=k+1}^{n-1} (2^i - 1)$$

for $0 \leqslant k \leqslant n-1$. We have $\mathbb{P}(Z_n = x) = 0$ if $x \neq 0$ and $x \neq 2^k$ for any $0 \leqslant k \leqslant n-1$, while $\mathbb{P}(Z_n = 0) \neq 0$ can be obtained from the probability of the complement.

Proof. Fix an even n. Ordered partitions for which both parts induce odd subgraphs are given by the solutions of

$$A\mathbf{v} = D(\mathbf{1} - \mathbf{v}) + \mathbf{1} \iff (A + D)\mathbf{v} = D\mathbf{1} + \mathbf{1},$$

where $A + D$ is a random even symmetric matrix. Unlike the other degree conditions, it is possible for this system to be inconsistent, meaning G has no such partitions. We can write $\mathbb{P}(Z_n = 2^k)$ as the probability that there is at least one solution and the rank of $A + D$ is $r = n - k - 1$. The probability for the latter is given by Corollary 2.10, so it remains to show that the probability that there is at least one solution conditioned on the rank of $A + D$ being r is $(2^r - 1)/(2^{n-1} - 1)$.

Let $E = A + D$. By Gallai's theorem we already know that $D\mathbf{1}$ is in the column space of E, so we equivalently determine the probability that $\mathbf{1}$ is also in the column space of E which is a random even symmetric matrix with given rank r. Let H be the set of (not necessarily symmetric) odd invertible matrices M (note that being odd here only requires the condition $\mathbf{1}^T M = \mathbf{1}^T$ on the column sums, the row sums can have either parity.)

We claim that H is a group. It is clear that H contains the identity and is closed under multiplication. Since H is finite and its elements are invertible, these must all have finite order, so it follows that H is also closed under inverses. Now let $\mathcal{E}$ be the set of all even symmetric matrices. Then H acts on $\mathcal{E}$ by the group action which sends $E \mapsto MEM^T$, noting that $\mathbf{1}^T MEM^T = \mathbf{1}^T EM^T = 0$ so MEM^T is even.

Given the basis $\{\mathbf{e}_1 + \mathbf{e}_i \colon i > 1\}$ for the subspace of even vectors, and any set of $n-1$ linearly independent even vectors $\mathbf{v}_2, \ldots, \mathbf{v}_n$, there exists $M \in H$ such that $M(\mathbf{e}_1 + \mathbf{e}_i) = \mathbf{v}_i$ for all $i = 2, \ldots n$. Indeed, take M to be the matrix with first column $\mathbf{e}_1$ and ith column $\mathbf{e}_1 + \mathbf{v}_i$ for $i > 1$. As H is a group it follows that we can map any r-dimensional subspace of even vectors to any other r-dimensional subspace of even vectors by a suitable choice of M. Thus, for any fixed E of rank r, the column spaces of the matrices MEM^T hit every r-dimensional subspace of even vectors with the same frequency. Thus, the proportion of times $\mathbf{1}$ lies in the column space is just the proportion of r-dimensional subspaces of even vectors that contain $\mathbf{1}$, which by symmetry is $(2^r - 1)/(2^{n-1} - 1)$. Picking M and E uniformly independently randomly with $\mathrm{rk}(E) = r$ means that MEM^T is a uniform random even symmetric matrix with rank r, so the probability that $\mathbf{1}$ is in the column space of E conditioned on the rank is as claimed. □

Finally, we deduce the asymptotic results stated in the introduction.

Proof of Theorem 1.7. For $k \in \mathbb{N} \cup \{0\}$ and n even we have

$$\begin{aligned}
\mathbb{P}(Z_n = 2^k) &= \frac{2^{n-k-1} - 1}{2^{\binom{n}{2}}(2^{n-1} - 1)} \cdot \prod_{i=1}^{\lfloor (n-1-k)/2 \rfloor} (1 - 2^{-2i})^{-1} \cdot \prod_{i=k+1}^{n-1} (2^i - 1) \\
&= \frac{2^{n-k-1} - 1}{2^{n-1} - 1} \prod_{i=1}^{\lfloor (n-1-k)/2 \rfloor} (1 - 2^{-2i})^{-1} \cdot \prod_{i=1}^{n-1} (1 - 2^{-i}) \prod_{i=1}^{k} (2^i - 1)^{-1} \\
&\to 2^{-k} \prod_{i=0}^{\infty} (1 - 2^{-2i-1}) \prod_{i=1}^{k} (2^i - 1)^{-1} = \mathbb{P}(Z = 2^k) \quad \text{as } n \to \infty.
\end{aligned}$$

Analogously, for $X_n = X_n^{(2,0)}, Y_n = X_n^{(1,1)}$ and any $k \in \mathbb{N}$ we find that

$$\lim_{n\to\infty} \mathbb{P}(X_n = 2^k) = \lim_{n\to\infty} \mathbb{P}(Y_n = 2^k) = c \prod_{i=1}^{k} (2^i - 1)^{-1},$$

where $c = \prod_{i=0}^{\infty} (1 - 2^{-2i-1})$ as defined in the statement. It remains to argue that $\mathbb{P}(Z_n = 0) \to \frac{1}{3}$. Note that $\mathbb{P}(Z_n = 2^k) \leqslant 2^{-k+1}$ for all $k, n \in \mathbb{N}$, so $\mathbb{P}(Z_n = 0)$ converges to $1 - \sum_{k=0}^{\infty} \mathbb{P}(Z = 2^k)$. Let

$$f(x) = c \sum_{k=0}^{\infty} x^k \prod_{i=1}^{k} (2^i - 1)^{-1}$$

be the generating function of the limiting distribution of $\tilde{X}_n$ where $2^{\tilde{X}_n}$ is the number X_n of even/even partitions. Then $f(1) = 1$ and

$$\begin{aligned} f(2x) &= c\sum_{k=0}^{\infty} x^k(1 + 2^k - 1)\prod_{i=1}^{k}(2^i - 1)^{-1} \\ &= c\sum_{k=0}^{\infty} x^k\Big(\prod_{i=1}^{k}(2^i - 1)^{-1} + \prod_{i=1}^{k-1}(2^i - 1)^{-1}\Big) = (1 + x)f(x). \end{aligned}$$

Furthermore,

$$g(x) := f(x/2) = c\sum_{k=0}^{\infty} 2^{-k}x^k\prod_{i=1}^{k}(2^i - 1)^{-1} = \sum_{k=0}^{\infty} x^k\mathbb{P}(Z = 2^k),$$

so $1 - \lim_{n\to\infty}\mathbb{P}(Z_n = 0) = g(1) = f(1/2)$. Using $f(1) = 1$ and $f(2x) = (1 + x)f(x)$ we get at $x = 1/2$ that $1 = (3/2)f(1/2)$, so $f(1/2) = 2/3$, hence $\lim_{n\to\infty}\mathbb{P}(Z_n = 0) = 1/3$. $\square$

Chapter 3

Invertibility of digraphs and tournaments

We remind the reader that for a digraph $D = (V, E)$ and a set $X \subseteq V$, the *inversion of* X *in* D is the digraph obtained from D by reversing the direction of the edges with both endpoints in X. A sequence of sets $X_1, \dots, X_k$ is a *decycling family* of D if inverting these sets in turn transforms D into an acyclic digraph. The *inversion number of* D, denoted $\mathrm{inv}(D)$, is defined to be the minimum size of a decycling family of D, and for $k \in \mathbb{N}_0$ we say that D is k*-invertible* if $\mathrm{inv}(D) \leqslant k$.

Before turning to proofs, we briefly summarise the main results we establish in this chapter. We disprove the conjecture of Bang-Jensen, da Silva, and Havet [11] that for any oriented graphs L, R we have $\mathrm{inv}(L \to R) = \mathrm{inv}(L) + \mathrm{inv}(R)$, where the dijoin $L \to R$ is obtained from disjoint copies of L and R by adding all edges from L to R.

Theorem 1.9. *There exists a tournament* R *with* $\mathrm{inv}(R) = \mathrm{inv}(\overrightarrow{C_3} \to R) = 3$.

While Theorem 1.9 shows that the dijoin conjecture is false in general, we prove it in the case where $\mathrm{inv}(L) = \mathrm{inv}(R) = 2$.

Theorem 1.10. *If* L *and* R *are digraphs with* $\mathrm{inv}(L) = \mathrm{inv}(R) = 2$*, then* $\mathrm{inv}(L \to R) = 4$.

The proof of Theorem 1.10 relies on the strongly connected case covered in [11] and our result for k-joins. For $k \in \mathbb{N}$, the k*-join* of digraphs $D_1, \dots, D_k$, written $[D_1, \dots, D_k]$, is the digraph consisting of vertex-disjoint copies of $D_1, \dots, D_k$ with additional edges from D_i to D_j whenever $i < j$.

Theorem 1.11. *Let $k \in \mathbb{N}$ and let $D_1, \ldots, D_k$ be oriented graphs. Assume that* $\text{inv}(D_i) \leqslant 2$ *for all i, with equality for at most one i. Then*

$$\text{inv}([D_1, \ldots, D_k]) = \sum_{i=1}^{k} \text{inv}(D_i). \tag{1.2.1}$$

In the introduction, we also stated the characterisation theorem (Theorem 1.12) which characterises the decycling families of size k of arbitrary k-joins of oriented graphs with inversion number 1. In Section 3.3, we will recall the statement of this theorem, prove it and deduce Theorem 1.11.

Theorem 1.11 also implies NP-completeness of the problem of k-Inversion, which asks, for fixed $k \in \mathbb{N}_0$:

Input: an oriented graph D.
Problem: $\text{inv}(D) \leqslant k$?

Theorem 1.13. *k-Inversion is NP-complete for all $k \in \mathbb{N}$.*

In stark contrast, we show that the problem of k-Tournament-Inversion which restricts the inputs to tournaments can be solved in quadratic time.

Theorem 1.14. *For fixed $k \in \mathbb{N}$, k-Tournament-Inversion can be solved for n-vertex tournaments in time $O(n^2)$. Moreover, if the input tournament is k-invertible, then our algorithm finds a decycling family of size at most k.*

A *cycle transversal* in a digraph D is a set of vertices of D whose removal from D leaves an acyclic digraph and the *cycle transversal number* of D, denoted $\tau(D)$, is the minimum size of a cycle transversal in D. In the introduction, we recalled the bound $\text{inv}(D) \leqslant 2\tau(D)$, first proved in [11], and in this chapter we will establish that it is tight for all values of $\tau(D)$.

Theorem 1.15. *For all $k \in \mathbb{N}$ there exists an oriented graph D with* $\text{inv}(D) = 2\tau(D) = 2k$.

Finally, we determine the asymptotic behaviour of the extremal parameter $\text{inv}(n)$, defined for each $n \in \mathbb{N}$ as the maximum inversion number of an oriented graph on n vertices.

Theorem 1.16. *For sufficiently large n we have*

$$\text{inv}(n) \geqslant n - \sqrt{2n \log_2(n)}.$$

Moreover, a uniformly random labelled n-vertex tournament has at least this inversion number with probability tending to 1.

We also show that $\mathrm{inv}(n) \leqslant n - \log_2(n+1)$.

The remainder of this chapter is organised as follows. In Section 3.1 we introduce some further notation, definitions, and preliminary observations which will be useful in the rest of the chapter. In the very short Section 3.2 we prove Theorem 1.9, constructing a counterexample to the dijoin conjecture. Our results on the inversion number of k-joins, Theorem 1.11 and Theorem 1.12, are proved in Section 3.3, along with Theorem 1.10 and Theorem 1.13. Section 3.4 concerns the complexity of k-TOURNAMENT-INVERSION and contains the proof of Theorem 1.14. We give the proof of Theorem 1.15 in Section 3.5. In Section 3.6 we discuss the existing bounds on $\mathrm{inv}(n)$ before proving Theorem 1.16 and the improved upper bound. Finally, in Section 3.7 we restate some conjectures and questions which remain open and pose some new ones of our own.

3.1 Preliminaries

In this section we detail some of the definitions, observations, and notation to be used in the rest of the chapter. As noted above, all digraphs will be oriented graphs, that is, loopless directed graphs with at most one edge between each pair of vertices. An *acyclic digraph* is a digraph with no directed cycles. In the case where the digraph is a tournament, we use the term *transitive* instead of acyclic. Note that for each $n \in \mathbb{N}$ there is a unique unlabelled transitive tournament on n vertices. To a transitive tournament T we associate the total order $<$ on $V(T)$ where $u < v$ for all $u, v \in V(T)$ such that $\overrightarrow{uv} \in E(T)$. We write $[n]$ for the set $\{1, 2, \ldots, n\}$. For a digraph D and a set $S \subseteq V(D)$ we write $D - S$ for the digraph produced by deleting the vertices in S from D. We now give the following key definitions.

Definition 3.1. Recall that for a digraph D, sets $X_1, \ldots, X_k \subseteq V(D)$, and a vertex $v \in V(D)$, the *characteristic vector* of v in $X_1, \ldots, X_k$ is $(I_{\{v \in X_i\}} : i \in [k]) \in \mathbb{F}_2^k$, where $I_{\{v \in X_i\}}$ is the indicator function of the event $v \in X_i$. Define an equivalence relation $\sim$ on $V(D)$ by setting $u \sim v$ if u and v have the same characteristic vector in $X_1, \ldots, X_k$. The *atoms* of $X_1, \ldots, X_k$ in D are the equivalence classes of this relation.

Note that, equivalently, the atoms of $X_1, \ldots, X_k$ in D are the atoms of the set algebra on $V(D)$ generated by $X_1, \ldots, X_k$, and that there are at most 2^k atoms for given D and $X_1, \ldots, X_k$. The next observation will be useful throughout the chapter.

Observation 3.2. Let D be a digraph and suppose that $u, v \in V(D)$ are joined by an edge in D. Let $X_1, \dots, X_k \subseteq V(D)$. Write $\mathbf{u}, \mathbf{v} \in \mathbb{F}_2^k$ for the characteristic vectors of u and v in $X_1, \dots, X_k$ respectively. Then the edge between u and v undergoes a net change in orientation when $X_1, \dots, X_k$ are inverted in D if and only if $\mathbf{u} \cdot \mathbf{v} = 1$.

This follows from the fact that $\mathbf{u} \cdot \mathbf{v}$ is the parity of the number of $X_1, \dots, X_k$ which contain both u and v. An obvious implication of Observation 3.2 is that given D and $X_1, \dots, X_k$, for every pair of (not necessarily distinct) atoms A and B, either all edges $\{ab \colon a \in A, b \in B\}$ undergo a net orientation change when $X_1, \dots, X_k$ are inverted, or none of them do. In particular, for every vertex v and atom A, either all edges $\{va \colon a \in A\}$ change orientation or none of them do.

Finally, we note some simple observations which will be used freely in what follows.

(i) If D' is a subdigraph of an oriented graph D, then $\mathrm{inv}(D') \leqslant \mathrm{inv}(D)$.

(ii) For every oriented graph D and every non-negative integer $k \leqslant \mathrm{inv}(D)$, there exists a spanning subdigraph of D with inversion number k.

(iii) If $X_1, \dots, X_k$ is a decycling family of an oriented graph D, then D can be extended to a tournament T for which $X_1, \dots, X_k$ is still a decycling family. In particular $\mathrm{inv}(T) = \mathrm{inv}(D)$.

For (ii), delete edges of D one by one, noting that the inversion number drops by at most 1 at each step. For (iii), invert the decycling family in D, extend the resulting acyclic digraph to a transitive tournament, then invert the decycling family again.

3.2 A counterexample to the dijoin conjecture

In this short section we give a counterexample to the dijoin conjecture of Bang-Jensen, da Silva, and Havet [11], that is, the conjecture that $\mathrm{inv}(L \to R) = \mathrm{inv}(L) + \mathrm{inv}(R)$ for all oriented graphs L and R. As noted in the introduction, this conjecture is equivalent to its restriction to tournaments. Indeed, suppose that L and R are digraphs with $\mathrm{inv}(L \to R) < \mathrm{inv}(L) + \mathrm{inv}(R)$. Extend $L \to R$ to a tournament of the same inversion number and observe that this tournament is $L' \to R'$ for some tournaments L' and R' extending L and R respectively. These clearly satisfy $\mathrm{inv}(L') \geqslant \mathrm{inv}(L)$ and $\mathrm{inv}(R') \geqslant \mathrm{inv}(R)$, so we have tournaments L' and R' with $\mathrm{inv}(L' \to R') < \mathrm{inv}(L) + \mathrm{inv}(R) \leqslant \mathrm{inv}(L') + \mathrm{inv}(R')$.

Proof of Theorem 1.9. Let L be a copy of $\overrightarrow{C_3}$. Suppose that R is a tournament with $\mathrm{inv}(R) = 3$ for which there exist disjoint $A, B, C \subseteq V(R)$ such that $A \cup B$, $A \cup C$ and

$B \cup C$ form a decycling family of R. Then for distinct vertices $u, v \in V(L)$ the sets $A \cup B \cup \{u, v\}$, $A \cup C \cup \{u, v\}$ and $B \cup C \cup \{u, v\}$ form a decycling family of $L \to R$, demonstrating that

$$\text{inv}(L \to R) = 3 < 4 = \text{inv}(L) + \text{inv}(R).$$

One way to construct such an R is as follows: let R be a tournament with vertex set [9], let $A = \{1, 3\}$, $B = \{4, 6\}$, and $C = \{7, 9\}$, and let the edge ij be directed backwards (that is, from j to i when $i < j$) if and only if i and j are both in $A \cup B \cup C$, but not both in A, B, or C. It is clear that inverting $A \cup B$, $A \cup C$ and $B \cup C$ transforms R into a transitive tournament, and a computer check shows that $\text{inv}(R) = 3$, as required. □

3.3 Decycling families of k-joins

In this section we prove Theorem 1.12, which we first restate here. For vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}_2^k$ we write $\mathbf{u} \cdot \mathbf{v}$ for the usual scalar product of $\mathbf{u}$ and $\mathbf{v}$ over $\mathbb{F}_2$. This is not a genuine inner product, but we say nevertheless that a collection $\mathbf{u}_1, \dots, \mathbf{u}_\ell \in \mathbb{F}_2^k$ is *orthonormal* if $\mathbf{u}_i \cdot \mathbf{u}_i = 1$ for all i and $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ for all $i \neq j$. Finally, we refer to the canonical copy of D_i in $D = [D_1, \dots, D_k]$ as the *ith factor* of D.

Theorem 1.12. *Let $D_1, \dots, D_k$ be oriented graphs with $\text{inv}(D_i) = 1$ for all i and let $\widehat{D} = [D_1, \dots, D_k]$ be their k-join. Then sets $X_1, \dots, X_k \subseteq V(\widehat{D})$ form a decycling family of $\widehat{D}$ if and only if there are orthonormal vectors $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{F}_2^k$ and for each i a decycling set $V_i \subseteq V(D_i)$ of the ith factor of $\widehat{D}$ such that, for each i, the vertices in V_i have characteristic vector $\mathbf{u}_i$ (in $X_1, \dots, X_k$), and all other vertices have characteristic vector $\mathbf{0}$ (in $X_1, \dots, X_k$).*

In particular, any acyclic digraph obtained from $\widehat{D}$ by k inversions can also be obtained by inverting a decycling set for each factor in turn.

We will then deduce Theorem 1.11 from this characterisation theorem, and use Theorem 1.11 to obtain Theorem 1.10 and Theorem 1.13. The bulk of the work in our proof of Theorem 1.12 is put towards proving Lemma 3.3, which deals with the case $\widehat{D} = [\overrightarrow{C_3}]_k$.

Lemma 3.3. *Let $k \in \mathbb{N}$, let $\widehat{D} = [\overrightarrow{C_3}]_k$, and let $X_1, \dots, X_k \subseteq V(\widehat{D})$ be a decycling family of $\widehat{D}$. Then there exist orthonormal vectors $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{F}_2^k$ such that in the* *ith factor of* $[\overrightarrow{C_3}]_k$*, one vertex has characteristic vector* $\mathbf{0}$ *and the other two have characteristic vector* $\mathbf{u}_i$*.*

We will use the setup that Pouzet, Kaddour, and Thatte [79] introduced in their proof that $\mathrm{inv}(\widehat{D}) = k$. The first part of our argument is essentially a reformulation of theirs, but we include it for completeness and to build intuition.

Proof of Lemma 3.3. Let T be the transitive tournament obtained by inverting the sets $X_1, \dots, X_k$ in $\widehat{D}$, and let $<$ be the total order on $V(\widehat{D})$ associated to T. Note that for all i, after inverting $X_1, \dots, X_k$ the ith factor has one vertex that has out-edges to the other two vertices in the factor and exactly one of these edges has undergone a net reversal. Thus we can label the vertices in the ith factor as u_i, v_i, w_i where $\overrightarrow{u_iv_iw_i}$ is a directed 3-cycle in $\widehat{D}$, and the edge between u_i and w_i undergoes a net reversal under $X_1, \dots, X_k$ while the edge between u_i and v_i does not. In particular, we will use throughout that $u_i < v_i, w_i$ and that, by Observation 3.2, $\mathbf{u}_i \cdot \mathbf{v}_i = 0$ and $\mathbf{u}_i \cdot \mathbf{w}_i = 1$ where $\mathbf{u}_i, \mathbf{v}_i, \mathbf{w}_i \in \mathbb{F}_2^k$ are the respective characteristic vectors of u_i, v_i, w_i in $X_1, \dots, X_k$. We have the following claim, originally proved in [79].

Claim 3.4 ([79])**.** *The vectors* $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{F}_2^k$ *are linearly independent.*

Proof. The statement is equivalent to the claim that for all non-empty $I \subseteq [k]$ we have $\sum_{i \in I} \mathbf{u}_i \neq \mathbf{0}$. Fix such an I and note that it is sufficient to show that there exists some $\mathbf{x} \in \mathbb{F}_2^k$ such that $(\sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{x} \neq 0$. Let $m \in I$ be such that $u_i < u_m$ for all $i \in I \setminus \{m\}$. Note that $u_m < v_m, w_m$, so by the transitivity of T we have $u_i < v_m, w_m$ for all $i \in I$. It is straightforward to deduce from this that for all $i \in I \setminus \{m\}$, the orientations of the edges u_iv_m and u_iw_m are either both unchanged after $X_1, \dots, X_k$ are inverted, or both reversed. By Observation 3.2, in other words we have $\mathbf{u}_i \cdot \mathbf{v}_m = \mathbf{u}_i \cdot \mathbf{w}_m$ for all $i \in I \setminus \{m\}$. On the other hand we have $\mathbf{u}_m \cdot \mathbf{v}_m = 0$ while $\mathbf{u}_m \cdot \mathbf{w}_m = 1$, so it follows by linearity of the dot product that $(\sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{v}_m \neq (\sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{w}_m$. One of these two dot products is thus non-zero, and we deduce that $\sum_{i \in I} \mathbf{u}_i \neq \mathbf{0}$, as required. □

We now build on Claim 3.4 as follows.

Claim 3.5. *Let* $\ell \in [k]$ *and suppose that the vectors* $\mathbf{u}_i, \mathbf{v}_i, \mathbf{w}_i$ *for* $\ell \leqslant i \leqslant k$ *all lie in a subspace* V *of* $\mathbb{F}_2^k$ *of dimension* $k - \ell + 1$*. Then* $\mathbf{u}_\ell, \dots, \mathbf{u}_k$ *are orthonormal, and for all* $\ell \leqslant i \leqslant k$ *we have* $\mathbf{u}_i = \mathbf{w}_i$ *and* $\mathbf{v}_i = \mathbf{0}$*.*

Proof. We will prove the claim by reverse induction on ℓ. In the $\ell = k$ case the claim follows easily from the fact that $\mathbf{u}_k \cdot \mathbf{w}_k = 1$ while $\mathbf{u}_k \cdot \mathbf{v}_k = 0$. Thus, let $\ell \leqslant k-1$ and write $[\ell, k]$ for $\{\ell, \ell+1, \ldots, k\}$. Let z be the $<$-minimal vertex among $v_\ell, \ldots, v_k, w_\ell, \ldots, w_k$. Write $\mathbf{z} \in V \subseteq \mathbb{F}_2^k$ for the characteristic vector of z in $X_1, \ldots, X_k$ and let $t \in [\ell, k]$ be such that $z \in \{v_t, w_t\}$. By Claim 3.4, the vectors $\mathbf{u}_\ell, \ldots, \mathbf{u}_k$ form a basis of V so there exists $I \subseteq [\ell, k]$ such that $\mathbf{z} + \sum_{i \in I} \mathbf{u}_i = \mathbf{0}$.

First suppose that $I \notin \{\emptyset, \{t\}\}$ and let $m \in I$ be such that $u_i < u_m$ for all $i \in I \setminus \{m\}$. If $m \neq t$, then we have $z < v_m, w_m$, so $\mathbf{z} \cdot \mathbf{v}_m = \mathbf{z} \cdot \mathbf{w}_m$ by Observation 3.2. As in the proof of Claim 3.4, we have $\mathbf{u}_i \cdot \mathbf{v}_m = \mathbf{u}_i \cdot \mathbf{w}_m$ for all $i \in I \setminus \{m\}$, but $\mathbf{u}_m \cdot \mathbf{v}_m \neq \mathbf{u}_m \cdot \mathbf{w}_m$, so $(\mathbf{z} + \sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{v}_m \neq (\mathbf{z} + \sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{w}_m$, and hence $\mathbf{z} + \sum_{i \in I} \mathbf{u}_i \neq \mathbf{0}$. If $m = t$, then let $j \in I \setminus \{t\}$ and note that $z < v_j, w_j$ by the minimality of z. Consequently, $\mathbf{z} \cdot \mathbf{v}_j = \mathbf{z} \cdot \mathbf{w}_j$. Moreover, since $u_m = u_t < z$, we have $u_m < v_j, w_j$. From this it follows that $u_i < v_j, w_j$ for all $i \in I$. Thus, $\mathbf{u}_i \cdot \mathbf{v}_j = \mathbf{u}_i \cdot \mathbf{w}_j$ for all $i \in I \setminus \{j\}$, while $\mathbf{u}_j \cdot \mathbf{v}_j \neq \mathbf{u}_j \cdot \mathbf{w}_j$. Hence, similarly to above, we have $(\mathbf{z} + \sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{v}_j \neq (\mathbf{z} + \sum_{i \in I} \mathbf{u}_i) \cdot \mathbf{w}_j$, so $\mathbf{z} + \sum_{i \in I} \mathbf{u}_i \neq \mathbf{0}$.

The remaining cases are $I = \emptyset$ and $I = \{t\}$, so we have $\mathbf{z} \in \{\mathbf{0}, \mathbf{u}_t\}$. Suppose that $\mathbf{z} = \mathbf{u}_t$. If $z = v_t$, then we have $\mathbf{v_t} = \mathbf{z} = \mathbf{u_t}$, so $\mathbf{v_t} \cdot \mathbf{w_t} = \mathbf{u_t} \cdot \mathbf{w_t} = 1$, i.e. the edge between v_t and w_t undergoes a net reversal under $X_1, \ldots, X_k$. This would imply that $w_t < v_t = z$, which contradicts the minimality of z. Similarly, if $z = w_t$, then since the edge between u_t and v_t is not inverted, neither is the edge between w_t and v_t, so $v_t < w_t = z$, another contradiction. Therefore $\mathbf{z} = \mathbf{0}$. This means no edges incident to z are reversed when $X_1, \ldots, X_k$ are inverted so by the minimality of z we have $z = v_\ell$.

We have shown that $\mathbf{v}_\ell = \mathbf{0}$, so the only vertex among the u_i, v_i, w_i with $i \geqslant \ell$ which precedes v_ℓ in $<$ is u_ℓ. It follows that u_ℓ is the least element among the u_i, v_i, w_i with $i \geqslant \ell$. Hence, by Observation 3.2 we have $\mathbf{u}_\ell \cdot \mathbf{u}_i = \mathbf{u}_\ell \cdot \mathbf{v}_i = \mathbf{u}_\ell \cdot \mathbf{w}_i = 0$ for all $i \geqslant \ell + 1$, so if V' is the subspace of V spanned by the $\mathbf{u}_i, \mathbf{v}_i, \mathbf{w}_i$ with $i \geqslant \ell + 1$, then $\mathbf{u}_\ell \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in V'$. We have $\mathbf{u}_\ell \cdot \mathbf{w}_\ell = 1$, so V' is a proper subspace of V, but $\mathbf{u}_{\ell+1}, \ldots, \mathbf{u}_k \in V'$ are linearly independent, so we deduce that V' has dimension $k - \ell$. Therefore by the induction hypothesis $\mathbf{u}_{\ell+1}, \ldots, \mathbf{u}_k$ are orthonormal, and we have $\mathbf{u}_i = \mathbf{w}_i$ and $\mathbf{v}_i = \mathbf{0}$ for $i \geqslant \ell + 1$.

To complete the induction step it remains to show that $\mathbf{u}_\ell = \mathbf{w}_\ell$ and $\mathbf{u}_\ell \cdot \mathbf{u}_\ell = 1$. The latter follows from the fact that $\mathbf{u}_\ell, \ldots, \mathbf{u}_k$ is a basis for V with $\mathbf{u}_\ell \cdot \mathbf{u}_i = 0$ for all $i \geqslant \ell + 1$, but $\mathbf{w}_\ell \in V$ has $\mathbf{u}_\ell \cdot \mathbf{w}_\ell = 1$. For the former, note that $\mathbf{w}_\ell = \sum_{i \in I} \mathbf{u}_i$ for some $I \subseteq [\ell, k]$ and by the established properties of the $\mathbf{u}_i$ this set I contains exactly those i for which $\mathbf{u}_i \cdot \mathbf{w}_\ell = 1$. Thus, we certainly have $\ell \in I$. Suppose that $\mathbf{u}_i \cdot \mathbf{w}_\ell = 1$ for some $i \geqslant \ell + 1$. Since $\mathbf{w}_i = \mathbf{u}_i$ and $\mathbf{v}_i = \mathbf{0}$, by Observation 3.2 we find that the cycle $\overrightarrow{w_\ell v_i w_i}$ appears in T, which is a contradiction. Hence $I = \{\ell\}$ and $\mathbf{u}_\ell = \mathbf{w}_\ell$, as required. □

The lemma now follows from the $\ell = 1$ case of Claim 3.5. □

We will now deduce Theorem 1.12 from the lemma. In the proof, we will use the easy fact that every family of orthonormal vectors in $\mathbb{F}_2^k$ is linearly independent.

Proof of Theorem 1.12. The sufficiency of the given conditions for $X_1, \dots, X_k$ to be a decycling family of $\widehat{D}$ is straightforward to verify using Observation 3.2. This observation also allows the 'in particular' part of the theorem statement to be easily deduced from the preceding part. It remains to prove that the given conditions are necessary.

Given a decycling family $X_1, \dots, X_k$ of $\widehat{D}$, extend $\widehat{D}$ to a tournament T for which $X_1, \dots, X_k$ is still a decycling family. For each i, let T_i be the subtournament of T induced on the vertex set of the ith factor of $\widehat{D}$. Since D_i contains a directed cycle, so does T_i, and hence the latter contains a copy of $\overrightarrow{C_3}$. We can thus find a copy of $[\overrightarrow{C_3}]_k$ in T whose ith factor is contained in T_i. It follows by Lemma 3.3 that there are orthonormal vectors $\mathbf{u}_1, \dots, \mathbf{u}_k \in \mathbb{F}_2^k$ and for each i a triangle $\overrightarrow{u_i v_i w_i}$ in T_i such that u_i and w_i have characteristic vector $\mathbf{u}_i$ and v_i has characteristic vector $\mathbf{0}$ in $X_1, \dots, X_k$.

We next show that for all i, all vertices in T_i have characteristic vector either $\mathbf{u}_i$ or $\mathbf{0}$ in $X_1, \dots, X_k$. Let $z \in V(T_i)$ and let $\mathbf{z}$ be its characteristic vector. Since $\mathbf{u}_1, \dots, \mathbf{u}_k$ form a basis of $\mathbb{F}_2^k$, there exists $J \subseteq [k]$ such that $\mathbf{z} = \sum_{j \in J} \mathbf{u}_j$. If there exists $\ell \in J \setminus \{i\}$, then $\mathbf{z} \cdot \mathbf{u}_\ell = \mathbf{u}_\ell \cdot \mathbf{u}_\ell = 1$ and hence the directions of the edge between z and u_ℓ and the edge between z and w_ℓ are reversed under $X_1, \dots, X_k$. If $\ell < i$, then the cycle $\overrightarrow{u_\ell v_\ell z}$ appears in T and if $i < \ell$, then the cycle $\overrightarrow{z v_\ell w_\ell}$ appears in T. We have a contradiction in both cases, so $J = \emptyset$ or $J = \{i\}$ as desired.

We have shown that all vertices in the ith factor of $\widehat{D}$ have characteristic vector either $\mathbf{u}_i$ or $\mathbf{0}$ in $X_1, \dots, X_k$. The effect on this copy of D_i of inverting these sets in $\widehat{D}$ is therefore the same as inverting the set of vertices with characteristic vector $\mathbf{u}_i$, which we call V_i. The latter is therefore a decycling set for the ith factor of $\widehat{D}$, that is, the family that contains just V_i is decycling. This completes the proof of the theorem. □

Theorem 1.11 now follows easily.

Proof of Theorem 1.11. It is clear that the left-hand side of equation (1.2.1) is at most the right-hand side. For the reverse inequality, let $\widehat{D} = [D_1, \dots, D_k]$ and note that we may assume that none of the D_i have inversion number 0. Indeed, if $\text{inv}(D_i) = 0$ for some $i \geqslant 2$, then view $\widehat{D}$ as the $(k-1)$-join $[D_1, \dots, D_{i-2}, D_{i-1} \to D_i, D_{i+1}, \dots, D_k]$ and, since $\text{inv}(D_{i-1} \to D_i) = \text{inv}(D_{i-1})$, the result follows by induction on k. The case where $i = 1$ can be handled similarly.

Thus, consider the case where $\text{inv}(D_i) = 1$ for all i and suppose for a contradiction that $X_1, \dots, X_k$ is a decycling family of $\widehat{D}$ with $X_k = \emptyset$. By Theorem 1.12 there exist k orthonormal, and hence linearly independent, vectors in $\mathbb{F}_2^k$ each of which occurs as the characteristic vector of some vertex of $\widehat{D}$ in $X_1, \dots, X_k$. This contradicts the fact that all such vectors have a 0 in their final coordinate. Hence, in this case, $\text{inv}(\widehat{D}) = k$.

It remains to check the case where $\text{inv}(D_j) = 2$ for some j and $\text{inv}(D_i) = 1$ for all $i \neq j$. Start by letting D'_j be a spanning subdigraph of the jth factor of $\widehat{D}$ with $\text{inv}(D'_j) = 1$, then define $\widehat{D}'$ to be the digraph obtained by replacing the jth factor of $\widehat{D}$ by D'_j. Assume for a contradiction that $X_1, \dots, X_k$ is a decycling family of $\widehat{D}$, in which case it is also a decycling family of $\widehat{D}'$. Theorem 1.12 thus yields a vector $\mathbf{u}_j \in \mathbb{F}_2^k$ with $\mathbf{u}_j \cdot \mathbf{u}_j = 1$ such that all the vertices in the jth factor of $\widehat{D}'$ (and hence also the jth factor of $\widehat{D}$) have characteristic vector either $\mathbf{0}$ or $\mathbf{u}_j$ in $X_1, \dots, X_k$. Inverting $X_1, \dots, X_k$ in $\widehat{D}$ therefore has the same effect on its jth factor as inverting the set of vertices with characteristic vector $\mathbf{u}_j$. It follows that this set of vertices is a decycling set for D_j, contradicting $\text{inv}(D_j) = 2$. □

As mentioned in the introduction, it follows from Theorem 1.11 that for any digraph D we have $\text{inv}(D) = 1$ if and only if $\text{inv}([D]_k) = k$, which in turn implies Theorem 1.13 (which states that k-Inversion is NP-complete for all $k \in \mathbb{N}$). Indeed, Theorem 1.11 directly gives $\text{inv}([D]_k) = k$ in the case $\text{inv}(D) = 1$, and if $\text{inv}(D) = 0$ then clearly $\text{inv}([D]_k) = 0$. If $\text{inv}(D) > 1$, then there are subdigraphs D' and D'' of D with $\text{inv}(D') = 1$ and $\text{inv}(D'') = 2$. The k-join $D'' \to [D']_{k-1}$, which has inversion number $k+1$ by Theorem 1.11, is a subdigraph of $[D]_k$ and thus $\text{inv}([D]_k) \geqslant k+1$ as required.

Finally, we deduce Theorem 1.10 (which states that $\text{inv}(L \to R) = 4$ for all digraphs L and R with inversion number 2) from Theorem 1.11. We will use the fact, shown in [11], that if L and R are strongly connected digraphs with $\text{inv}(L), \text{inv}(R) \geqslant 2$, then $\text{inv}(L \to R) \geqslant 4$.

Proof of Theorem 1.10. Let L and R be digraphs with $\text{inv}(L) = \text{inv}(R) = 2$. It is immediate that $\text{inv}(L \to R) \leqslant 4$, so it is sufficient to prove the lower bound. For this, extend $L \to R$ to a tournament T of the same inversion number and let the tournaments to which L and R are extended be L' and R' respectively. Note that $\text{inv}(L'), \text{inv}(R') \geqslant 2$ and T is $L' \to R'$.

Every tournament can be written as the k-join of its strongly connected components, so let L' be $[L_1, \dots, L_{k_1}]$ and R' be $[R_1, \dots, R_{k_2}]$ for some $k_1, k_2 \in \mathbb{N}$ and strongly connected tournaments $L_1, \dots, L_{k_1}, R_1, \dots, R_{k_2}$. Since $\text{inv}(L') \geqslant 2$, either there is some L_i with $\text{inv}(L_i) \geqslant 2$, or there are $i < j$ such that $\text{inv}(L_i) = \text{inv}(L_j) = 1$. An analogous condition holds for R'. If there are i and j such that $\text{inv}(L_i), \text{inv}(R_j) \geqslant 2$, then since T contains $L_i \to R_j$, we have $\text{inv}(T) \geqslant \text{inv}(L_i \to R_j) \geqslant 4$ by the above result of [11]. Otherwise, either there exist $i < j$ such that $\text{inv}(L_i) = \text{inv}(L_j) = 1$, in which case $\text{inv}(T) \geqslant \text{inv}([L_i, L_j, R]) = 4$ by Theorem 1.11, or there exist $i < j$ with $\text{inv}(R_i) = \text{inv}(R_j) = 1$, in which case it follows similarly that $\text{inv}(T) \geqslant 4$. □

3.4 Complexity of k-Tournament-Inversion

In this section we prove Theorem 1.14 by constructing, for each fixed $k \in \mathbb{N}$, an algorithm solving k-Tournament-Inversion in time $O(|V(T)|^2)$. Our proof uses a technique known as iterative compression; see [36] for a description of this method and other applications of it. The most involved part of our proof concerns the 'compression step' of the algorithm. This step is handled by the following lemma, which roughly says that for constant k, given an n-vertex tournament T_0 and a decycling family of T_0 of constant size, in time linear in n we can find a decycling family of T_0 of size k if one exists. Throughout this section, we represent a total order $<$ on a finite set $S = \{s_1, \dots, s_m\}$ by the tuple $(s_1, \dots, s_m)$ where $s_1 < \cdots < s_m$.

Lemma 3.6. *Fix $k, s \in \mathbb{N}$. There is an algorithm which solves the following problem for n-vertex tournaments in time $O(n)$:*

<u>INPUTS</u>:
- *a tournament T_0;*
- *a decycling family $X_1, \ldots, X_s$ of T_0 (transforming T_0 into T, say);*
- *the order on $V(T_0)$ associated to T.*

<u>OUTPUTS</u>:
EITHER
- *that T_0 is not k-invertible;*

OR
- *a decycling family $Y_1, \ldots, Y_k$ of T_0 (transforming T_0 into T', say);*
- *the order on $V(T_0)$ associated to T'.*

We now use iterative compression to prove Theorem 1.14 before returning to Lemma 3.6.

Proof of Theorem 1.14. Fix $k \geqslant 1$. We will induct on n to define an algorithm solving the following problem for n-vertex tournaments in time $C_k \cdot n^2$ for some constant C_k:

<u>INPUT</u>:
- a tournament T_0.

<u>OUPUTS</u>:
EITHER
- that T_0 is not k-invertible;

OR
- a decycling family $Y_1, \ldots, Y_k$ of T_0 (transforming T_0 into T, say);
- the order on $V(T_0)$ associated to T.

In particular, this algorithm solves k-TOURNAMENT-INVERSION.

Fix $n \geqslant 2$ and assume that we have defined such an algorithm for all smaller tournaments. Let T_0 be an n-vertex tournament and pick some $v \in V(T_0)$. Applying the induction hypothesis, in time $C_k \cdot (n-1)^2$ we either find that $T_0 - \{v\}$ is not k-invertible or we obtain a decycling family $X_1, \ldots, X_k$ of $T_0 - \{v\}$ and the order on $V(T_0) \setminus \{v\}$ associated to the transitive tournament obtained by inverting these sets in T_0. In the former case, it follows that T_0 is also not k-invertible and we can output that fact. In the latter case, let A be the out-neighbourhood of v in T_0, and define $X_{k+1} = A \cup \{v\}$ and $X_{k+2} = A$. Then $X_1, \ldots, X_{k+2}$ is a decycling family of T_0, and we can obtain the order associated to the resulting transitive tournament by adding v to the previous order as the maximal element. By Lemma 3.6 we can now, in linear time, either find that T_0 is not k-invertible or obtain a decycling family $Y_1, \ldots, Y_k$ of T_0 of size k and the order associated to the transitive tournament obtained by inverting these sets in T_0. As required, this algorithm runs in time $C_k \cdot (n-1)^2 + O(n)$, which is at most $C_k \cdot n^2$ if C_k is large enough. □

It is left to prove Lemma 3.6. To this end, we describe an algorithm which explores what happens if, starting from T, we invert $X_1, \dots, X_s$ and k further sets $Y_1, \dots, Y_k$ to obtain a tournament T_Y, where $Y = (Y_1, \dots, Y_k)$. Since T_Y is the tournament obtained by inverting $Y_1, \dots, Y_k$ in T_0, these k sets are a decycling family of T_0 if and only if T_Y is transitive. If we were to examine each possibility individually there would be too many for this exploration process to be tractable. However, the fact that we are starting from a transitive tournament T makes it possible to identify cycles in the final tournament T_Y without fully specifying the sets $Y_1, \dots, Y_k$. This means there are far fewer cases to consider, indeed few enough that the exploration process is linear in n for fixed k and s.

Proof of Lemma 3.6. Fix $k, s \in \mathbb{N}$ and let T_0, $X_1, \dots, X_s$, and T be as in the statement of the lemma. Let $n = |V(T_0)|$ and label the vertices of T_0 as $u_1, \dots, u_n$ in T-increasing order. With notation as above, we wish to investigate for which Y the tournament T_Y is transitive. For each Y we write $\mathbf{u}_i \in \mathbb{F}_2^{s+k}$ for the characteristic vector of u_i in $X_1, \dots, X_s, Y_1, \dots, Y_k$ (suppressing the dependence on Y in the notation) and then let $\mathbf{u} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$. There is a bijective correspondence between Y and $\mathbf{u}$ and it will be more convenient to work with the latter, so let $T_{\mathbf{u}} = T_Y$ and write $\mathcal{U}$ for the set of all possible $\mathbf{u}$. Our first aim is to determine in linear time whether there exists $\mathbf{u} \in \mathcal{U}$ such that $T_{\mathbf{u}}$ is transitive, and to identify such a $\mathbf{u}$ if so.

The tournament $T_{\mathbf{u}}$ is transitive exactly when it contains no cyclic triples. It is straightforward to use Observation 3.2 to show that this is equivalent to the condition that there are no $a < b < c$ in $[n]$ such that $\mathbf{u}_a \cdot \mathbf{u}_b = \mathbf{u}_b \cdot \mathbf{u}_c$ but $\mathbf{u}_a \cdot \mathbf{u}_b \neq \mathbf{u}_a \cdot \mathbf{u}_c$. We describe the triple $(\mathbf{u}_a, \mathbf{u}_b, \mathbf{u}_c)$ as *bad* if this occurs. Thus, $T_{\mathbf{u}}$ is transitive if and only if $B(\mathbf{u}) = \{(\mathbf{u}_a, \mathbf{u}_b, \mathbf{u}_c) \colon a < b < c\}$ contains no bad triples, and T_0 is k-invertible if and only if $\mathcal{B} = \{B(\mathbf{u}) \colon \mathbf{u} \in \mathcal{U}\}$ contains a set which is free of bad triples. Our algorithm will construct this set $\mathcal{B}$ and check whether any of its elements are free of bad triples. If one of these sets *is* free of bad triples, then we need to be able to output a corresponding decycling family of T_0, so for each $B \in \mathcal{B}$ we will also record some $\mathbf{u} \in \mathcal{U}$ for which $B = B(\mathbf{u})$.

We will now explain how the above can be achieved in linear time. First note that we may assume that $n \geqslant 4$. Let $\mathcal{U}'$ be the set of all possible vectors $\mathbf{u}' = (\mathbf{u_1}, \dots, \mathbf{u}_{n-1})$ of characteristic vectors of $u_1, \dots, u_{n-1}$ in $X_1, \dots, X_s, Y_1, \dots, Y_k$. For $\mathbf{u}' \in \mathcal{U}'$, let

$B'(\mathbf{u}') = \{(\mathbf{u}_a, \mathbf{u}_b, \mathbf{u}_c) \colon 1 \leqslant a < b < c \leqslant n-1\}$ and let $\mathcal{B}' = \{B'(\mathbf{u}') \colon \mathbf{u}' \in \mathcal{U}'\}$. We may assume inductively that there is a constant C depending only on k and s such that in time $C \cdot (n-1)$ we can construct $\mathcal{B}'$ and associate to each $B' \in \mathcal{B}'$ some $\mathbf{u}' \in \mathcal{U}'$ such that $B' = B'(\mathbf{u}')$. For the induction step, we need to show that we can use this to obtain in time C the set $\mathcal{B}$ and for each $B \in \mathcal{B}$ some $\mathbf{u} \in \mathcal{U}$ such that $B = B(\mathbf{u})$.

The key observation is that there are only 2^{s+k} possible characteristic vectors for each of $u_1, \ldots, u_n$, so the number of triples of characteristic vectors is at most $2^{3(s+k)}$ and the sizes of $\mathcal{B}$ and $\mathcal{B}'$ are at most $2^{2^{3(s+k)}}$. In particular, there are only boundedly many pairs $(B', \mathbf{u}_n)$ where $B' \in \mathcal{B}'$ and $\mathbf{u}_n$ is a possible characteristic vector for u_n. For each such pair, we can construct in bounded time the set $S(B', \mathbf{u}_n)$ consisting of all triples in B', and all triples of the form $(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}_n)$ for $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \in B'$ and $1 \leqslant i < j \leqslant 3$. It is not hard to see that $\mathcal{B}$ equals the set of all sets $S(B', \mathbf{u}_n)$ and that each $S(B', \mathbf{u}_n)$ can be associated with the $\mathbf{u} \in \mathcal{U}$ formed by appending $\mathbf{u}_n$ to the $\mathbf{u}' \in \mathcal{U}'$ associated with B'. Indeed, given B' and $\mathbf{u}_n$ and defining $\mathbf{u}$ as in the previous sentence, since $n \geqslant 4$, we have $S(B', \mathbf{u}_n) = B(\mathbf{u})$. For the other direction, given $\mathbf{u} = (\mathbf{u_1}, \ldots, \mathbf{u}_n) \in \mathcal{U}$ and letting $\mathbf{u}' = (\mathbf{u}_1, \ldots, \mathbf{u}_{n-1})$, we have $B(\mathbf{u}) = S(B'(\mathbf{u}'), \mathbf{u}_n)$.

We can construct this set in bounded time and then forget about all but one of the elements of $\mathcal{U}$ associated to each $B \in \mathcal{B}$. Thus, in constant time we have obtained $\mathcal{B}$ and for each $B \in \mathcal{B}$ some $\mathbf{u} \in \mathcal{U}$ such that $B = B(\mathbf{u})$, and the induction continues.

Once we have constructed $\mathcal{B}$ in linear time, since it has bounded size we can check whether any of its members is free of bad triples in bounded time. If not, then T_0 is not k-invertible. If so, then pick $B \in \mathcal{B}$ with no bad triples and use the $\mathbf{u}$ associated to it to construct a decycling family $Y_1, \ldots, Y_k$ of T_0. Let T' be the transitive tournament obtained by inverting these sets in T_0.

It remains to show that we can obtain the order on $V(T_0)$ associated to T' in linear time. Inverting the sets $X_1, \ldots, X_s, Y_1, \ldots, Y_k$ transforms T into T', and we have the characteristic vector of each vertex in these sets as well as the order on the vertices associated to T. We can therefore in linear time obtain the atoms of these $s + k$ inversions and for each atom A the restriction to A of the order associated to T. By reversing the order on each atom whenever the edges within it undergo a net reversal under the inversions, we obtain the order on that atom associated to T'. The T'-minimal vertex is now the minimal vertex of one of the atoms under their current orderings. There are at most 2^{s+k} atoms so we can identify the T'-minimal vertex in constant time. After deleting this vertex from its atom, the second smallest vertex according to T' is one of the new minimal vertices of the atoms so can be found in constant time again. Continuing in this way we can obtain the full ordering in linear time, as required. □

Note that the implicit constant in the running time given by this proof is doubly exponential in $s + k$.

3.5 Cycle transversals

In this section we will prove Theorem 1.15, constructing for each $k \in \mathbb{N}$ a digraph D with $\tau(D) = k$ and $\text{inv}(D) = 2k$. We will use the so-called Eventown theorem, proved independently by Berlekamp [20] and Graver [44].

Theorem 3.7 (Eventown [20], [44])**.** *Let $n \in \mathbb{N}$ and let $\mathcal{F} \subseteq \mathcal{P}([n])$ be a family of subsets of $[n]$ such that $|F_1 \cap F_2|$ is even for all $F_1, F_2 \in \mathcal{F}$. Then $|\mathcal{F}| \leqslant 2^{\lfloor n/2 \rfloor}$.*

For a digraph D and vertices $u, v, w \in V(D)$, we will say that u and v *differ* on w if either $\overrightarrow{uw}, \overrightarrow{wv} \in E(D)$ or $\overrightarrow{vw}, \overrightarrow{wu} \in E(D)$. We are now ready to prove the theorem.

Proof of Theorem 1.15. Fix $k \in \mathbb{N}$ and let $n \in \mathbb{N}$ be large and divisible by 2^k. We will define a digraph D on vertex set $\{u_0, \dots, u_{k-1}, v_0, \dots, v_{n-1}\}$ and then show that it satisfies the conditions of the theorem. Start by including all directed edges $\overrightarrow{v_i v_j}$ for $i < j$, so that the subdigraph of D induced on $\{v_0, \dots, v_{n-1}\}$ is a transitive tournament. For $i \in \{0, \dots, k-1\}$ and $j \in \{0, \dots, n-1\}$, add the edge $\overrightarrow{u_i v_j}$ if in the binary expansion of j, the digit in the 2^i place is a 0, and add the edge $\overrightarrow{v_j u_i}$ otherwise. For ease of exposition we will not include any edges among the u_i (though including any combination of such edges would still give a valid construction), so this completes the definition of D. As noted above, the removal of the vertices $u_0, \dots, u_{k-1}$ from D leaves an acyclic digraph, so $\tau(D) \leqslant k$.

It remains to show that $\text{inv}(D) \geqslant 2k$, as then $\text{inv}(D) = 2\tau(D) = 2k$ follows from $\text{inv}(D) \leqslant 2\tau(D)$. Suppose for a contradiction that $X_1, \dots, X_{2k-1} \subseteq V(D)$ form a decycling family of D and let D' be the acyclic digraph obtained by inverting these sets in D. Consider the characteristic vectors of $v_0, \dots, v_{n-1}$ in $X_1, \dots, X_{2k-1}$, which we will denote by $\mathbf{v}_0, \dots, \mathbf{v}_{n-1} \in \mathbb{F}_2^{2k-1}$ respectively. Let $K = 2^k$. By the pigeonhole principle, if n is large enough, then there exist distinct $i, i' \in \{0, \dots, n/K - 1\}$ such that

$$(\mathbf{v}_{iK}, \mathbf{v}_{iK+1}, \dots, \mathbf{v}_{(i+1)K-1}) = (\mathbf{v}_{i'K}, \mathbf{v}_{i'K+1}, \dots, \mathbf{v}_{(i'+1)K-1}).$$

We may assume that $i = 0$ and $i' = 1$.

We will show that $\mathbf{v}_0, \ldots, \mathbf{v}_{K-1}$ are pairwise distinct and that $\mathbf{v}_i \cdot \mathbf{v}_j$ is constant as $i, j \in \{0, \ldots, K-1\}$ vary. We claim that these conditions force a contradiction. Indeed, in the case where $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all i, j, we have that the $\mathbf{v}_i$ are indicator vectors of pairwise distinct subsets of $[2k-1]$ which each have even size, and each pair of which have even intersection. By Eventown, every such collection has at most $2^{(2k-1)/2} < 2^k = K$ members, giving the required contradiction. On the other hand, if $\mathbf{v}_i \cdot \mathbf{v}_j = 1$ for all i, j, then consider the 'complement' vectors $\mathbf{w}_0, \ldots, \mathbf{w}_{K-1}$, which have 1's where the $\mathbf{v}_i$ have 0's and 0's where the $\mathbf{v}_i$ have 1's. It is straightforward to use the fact that the vectors have odd length to show that these $\mathbf{w}_i$ are pairwise distinct and satisfy $\mathbf{w}_i \cdot \mathbf{w}_j = 0$ for all i, j, from which we can derive a contradiction as above.

We continue by showing that the vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{K-1}$ are pairwise distinct, which is equivalent to showing that each of $v_0, \ldots, v_{K-1}$ is in a different atom. Suppose for a contradiction that v_i and v_j are in the same atom for some $i < j$ in $\{0, \ldots, K-1\}$, and note that v_{K+i} and v_{K+j} are in this atom too by assumption. By the construction of D there is some $\ell \in \{0, \ldots, k-1\}$ such that v_i and v_j differ on u_ℓ in D. Since they are in the same atom as each other, they also differ on u_ℓ in D'. If $\overrightarrow{v_i u_\ell}, \overrightarrow{u_\ell v_j} \in E(D')$, then to avoid a cyclic triple in D' we have $\overrightarrow{v_i v_j} \in E(D')$. This means that the edges within v_i and v_j's atom have the same orientations in D as they do in D', so in particular we have $\overrightarrow{v_j v_{K+i}} \in E(D')$. Moreover v_i and v_{K+i} are in the same atom and are either both in-neighbours of u_ℓ in D or both out-neighbours, so since $\overrightarrow{v_i u_\ell} \in E(D')$ we also have $\overrightarrow{v_{K+i} u_\ell} \in E(D')$. Hence, the cycle $\overrightarrow{v_j v_{K+i} u_\ell}$ appears in D'. Similarly if $\overrightarrow{v_j u_\ell}, \overrightarrow{u_\ell v_i} \in E(D')$, then we have $\overrightarrow{v_j v_i} \in E(D')$. In this case the edges within the atom of v_i and v_j switch orientation between D and D', so the cycle $\overrightarrow{v_j u_\ell v_{K+i}}$ appears in D'. In both cases we have the desired contradiction, and we deduce that the vertices $v_0, \ldots, v_{K-1}$ are all in different atoms.

It remains to show that $\mathbf{v}_i \cdot \mathbf{v}_j$ is constant as $i, j \in \{0, \ldots, K-1\}$ vary. Suppose for a contradiction that this is not the case, then there exists $i \in \{0, \ldots, K-1\}$ such that $\mathbf{v}_i \cdot \mathbf{v}_j$ is not constant as $j \in \{0, \ldots, K-1\}$ varies. For such i we can pick $j \in \{0, \ldots, K-1\}$ such that $\mathbf{v}_i \cdot \mathbf{v}_i \neq \mathbf{v}_i \cdot \mathbf{v}_j$. Now if $\mathbf{v}_i \cdot \mathbf{v}_i = 0$, then $\mathbf{v}_i \cdot \mathbf{v}_j = 1$ so by Observation 3.2, D' contains the cycle $\overrightarrow{v_{K+i} v_j v_i}$ if $i < j$ or the cycle $\overrightarrow{v_{K+i} v_{K+j} v_i}$ if $i > j$. Similarly if $\mathbf{v}_i \cdot \mathbf{v}_i = 1$, then D' contains one of the cycles $\overrightarrow{v_i v_j v_{K+i}}$ and $\overrightarrow{v_i v_{K+j} v_{K+i}}$. We have a contradiction in all cases, so the value of $\mathbf{v}_i \cdot \mathbf{v}_j$ is constant as $i, j \in \{0, \ldots, K-1\}$ vary, as required. □

3.6 Bounds on inv(n)

3.6.1 Lower bounds

In this section we discuss the previous known lower bound on inv(n) and give the proof of Theorem 1.16. As noted in the introduction, Belkhechine, Bouaziz, Boudabbous, and Pouzet [18] used a counting argument to lower bound inv(n). They observed that since there are $n!$ labelled transitive tournaments on n vertices, there are at most $n! \cdot 2^{n(k-1)}$ labelled $(k-1)$-invertible tournaments on n vertices. There are a total of $2^{n(n-1)/2}$ labelled n-vertex tournaments, so for any k such that $2^{n(n-1)/2} > n! \cdot 2^{n(k-1)}$ we have $\mathrm{inv}(n) \geqslant k$. Taking logarithms base 2 and rearranging, this condition becomes[1] $k < (n-1)/2 - \log(n!)/n$, so we have

$$\mathrm{inv}(n) \geqslant \left\lfloor \frac{n-1}{2} - \frac{\log(n!)}{n} \right\rfloor \geqslant \left\lfloor \frac{n-1}{2} - \log(n) \right\rfloor,$$

where for the final inequality we used $n! \leqslant n^n$. Lower bounds on inv(n) of this form were the best known (disregarding very slight tightenings of the argument).

The proof of Theorem 1.16 uses the following lemma which gives a bound on the probability that a random symmetric binary matrix has at most a certain rank. This bound can again be deduced from the exact counts [72] which we met earlier in Theorem 2.9, but we also include a short proof for the simpler bound we use.

Lemma 3.8. *The probability that a uniformly random $n \times n$ symmetric matrix over $\mathbb{F}_2$ has rank at most $n - s$ (over $\mathbb{F}_2$) is at most $2^{s\log(n) - \binom{s}{2}}$.*

Proof. Construct the random matrix in n steps, in the ith step choosing the first i entries of the ith row of the matrix (and also, by symmetry, the ith column). For each $i \in [n]$, let M_i be the random symmetric $i \times i$ matrix obtained after step i.

Note that for each i the nullity increases by at most 1 between M_i and M_{i+1}. It follows that if the nullity of M_n is at least s, then for all $1 \leqslant j \leqslant s-1$ we can define k_j to be the smallest i such that the nullity of M_i is $j+1$, and we have $2 \leqslant k_1 < k_2 < \cdots < k_{s-1} \leqslant n$. For each j, the ranks of $M_{k_j - 1}$ and M_{k_j} are equal, so the first $k_j - 1$ entries of the k_jth row of M_{k_j} lie in the $(k_j - 1 - j)$-dimensional row space of M_{k_j-1}, which happens with probability 2^{-j}. There are $\binom{n}{s-1}$ ways to choose $k_1, \ldots, k_{s-1}$ as above, so the probability that M_n has rank at most $n-s$ is at most

$$\binom{n}{s-1} \prod_{j=1}^{s-1} 2^{-j} \leqslant 2^{s\log(n) - \binom{s}{2}},$$

[1]All logarithms in this chapter are taken base 2, which differs from our use in other chapters.

as required. ☐

We are now ready to prove Theorem 1.16.

Proof of Theorem 1.16. Let T be a uniformly random tournament on vertex set $[n]$ and let $M_T = (m_{ab})$ be the $n \times n$ matrix over $\mathbb{F}_2$ defined as follows. For $a < b$, let m_{ab} be 0 if $\overrightarrow{ab}$ is an edge of T and 1 otherwise, then define $m_{ba} = m_{ab}$, and finally choose each diagonal entry uniformly at random. Note that the $\binom{n}{2}$ entries of M_T above the diagonal determine T, and the other entries are defined such that M_T is a uniformly random symmetric binary matrix.

Let $s = \left\lfloor \sqrt{2n\log(n)} \right\rfloor$ and write $k = n - s$. Suppose that $\mathrm{inv}(T) \leqslant k$ and let $X_1, \ldots, X_k$ be a decycling family of T. For each X_i, let M_i be the $n \times n$ binary matrix whose (a, b) entry is 1 if and only if $a, b \in X_i$. Observe that, working over $\mathbb{F}_2$, we have $\mathrm{rk}(M_i) \leqslant 1$ for all i, and thus $\mathrm{rk}(\sum_i M_i) \leqslant k$. By construction, $M_T + \sum_i M_i$ is a matrix whose entries above the diagonal correspond to a transitive tournament on $[n]$ (its diagonal entries can be anything). Let $\mathcal{M}$ be the set of binary matrices corresponding in this manner to a transitive tournament on $[n]$, and note that $|\mathcal{M}| = n!2^n$.

Putting all of this together, we have that if $\mathrm{inv}(T) \leqslant k$, then there exists $M \in \mathcal{M}$ such that $\mathrm{rk}(M_T + M) \leqslant k$. For each fixed M, we have that $M_T + M$ is a uniformly random symmetric binary matrix and hence has rank at most k with probability at most $2^{s\log(n) - \binom{s}{2}}$ by Lemma 3.8. Taking a union bound over all $M \in \mathcal{M}$ we obtain

$$\mathbb{P}(\mathrm{inv}(T) \leqslant k) \leqslant n!2^n 2^{s\log(n) - \binom{s}{2}}.$$

Since $n! = O(\sqrt{n}(n/e)^n)$, the right-hand side is $O(2^{f(n)})$ where

$$\begin{aligned} f(n) &= \frac{\log(n)}{2} + n\log(n) - n\log(e) + n + s\log(n) - \binom{s}{2} \\ &= -n(\log(e) - 1) + o(n), \end{aligned}$$

and thus $\mathbb{P}(\mathrm{inv}(T) \leqslant k) \to 0$ as $n \to \infty$ as desired. ☐

3.6.2 Upper bounds

The only approach which has been used to prove upper bounds on $\mathrm{inv}(n)$, introduced in [18], is to 'solve' one vertex at a time, as follows. Given a tournament T, pick a vertex v and invert the set consisting of v and its out-neighbourhood. In the resulting tournament T_1, v is a sink. Using a further $\mathrm{inv}(n-1)$ inversions we can transform

$T_1 - \{v\}$ into a transitive tournament, so $\text{inv}(n) \leqslant \text{inv}(n-1)+1$ for all $n \geqslant 2$. The authors of [18] observed that $\text{inv}(4) = 1$, so $\text{inv}(n) \leqslant n-3$ for $n \geqslant 4$. For $n \geqslant 6$ this was improved by 1 in [11] using the fact that $\text{inv}(6) = 2$ (which they attribute to [17] and which we have verified by a computer check). We introduce a slightly different approach to prove the following.

Proposition 3.9. *For all $n \in \mathbb{N}$,*

$$\text{inv}(n) \leqslant \left\lfloor \frac{n-1}{2} \right\rfloor + \text{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil \right).$$

Proof. Let $n \in \mathbb{N}$ and let T be an n-vertex tournament. Pick $v \in V(T)$ and write A and B for the in- and out-neighbourhoods of v respectively. We may assume that $|A| \geqslant \lceil (n-1)/2 \rceil$ (the case where B is the larger of the two is similar). By 'solving' each vertex in B one after another, we can find at most $|B|$ inversions which transform T into a tournament T' such that the subtournament of T' induced on $B \cup \{v\}$ is transitive (with v as the minimal element) and every edge of T' between A and $B \cup \{v\}$ is oriented away from A. With a further $\text{inv}(A) \leqslant \text{inv}(|A|)$ inversions we can transform T' into a transitive tournament. Thus, $\text{inv}(T) \leqslant |B| + \text{inv}(|A|)$.

We have $\text{inv}(k) \leqslant \text{inv}(k-1)+1$ for all $k \in \mathbb{N}$ and we can apply this $|A| - \lceil (n-1)/2 \rceil$ times to obtain $\text{inv}(|A|) \leqslant \text{inv}(\lceil (n-1)/2 \rceil) + |A| - \lceil (n-1)/2 \rceil$. Using the fact that $|A| + |B| = n-1$, this yields

$$\text{inv}(T) \leqslant |B| + \text{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil \right) + |A| - \left\lceil \frac{n-1}{2} \right\rceil = \left\lfloor \frac{n-1}{2} \right\rfloor + \text{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil \right),$$

and the claim follows. □

We can use this result to improve (for large n) the upper bound on $\text{inv}(n)$.

Corollary 3.10. *For all $n \in \mathbb{N}$, $\text{inv}(n) \leqslant n - \log(n+1)$.*

Proof. We prove the statement by induction for $n \geqslant 0$, noting that the definition of inv extends naturally with $\text{inv}(0) = 0$. If $n \geqslant 1$ and the claim holds for all smaller values, then we have

$$\begin{aligned}
\text{inv}(n) &\leqslant \left\lfloor \frac{n-1}{2} \right\rfloor + \text{inv}\left(\left\lceil \frac{n-1}{2} \right\rceil \right), \\
&\leqslant \left\lfloor \frac{n-1}{2} \right\rfloor + \left\lceil \frac{n-1}{2} \right\rceil - \log\left(\left\lceil \frac{n-1}{2} \right\rceil + 1 \right), \\
&\leqslant n - 1 - \log\left(\frac{n+1}{2} \right), \\
&= n - \log(n+1).
\end{aligned}$$

□

3.7 Conclusion

In this chapter, we have answered several of the questions posed in [11]. We have shown that their 'dijoin conjecture', that $\mathrm{inv}(L \to R) = \mathrm{inv}(L) + \mathrm{inv}(R)$, is false in general, but have verified it in the case where $\mathrm{inv}(L) = \mathrm{inv}(R) = 2$ and have also shown that a k-join analogue holds under certain conditions. In addition, we have confirmed their related conjectures that k-Inversion is NP-complete for all $k \geqslant 1$, and that the inequality $\mathrm{inv}(D) \leqslant 2\tau(D)$ is tight. We have answered their question concerning the minimal r_k such that k-Tournament-Inversion can be solved in time $O(|V(T)|^{r_k})$, showing that $r_k = 2$ for all k, and have improved the lower bound on $\mathrm{inv}(n)$ to show that $\mathrm{inv}(n) = (1+o(1))n$. There are, however, still many interesting open problems in this area. Before discussing some of them, we will touch on two operations similar to inversion.

3.7.1 Similar operations

We first consider an operation on permutations which is used by molecular biologists as a model for genetic mutations, and could loosely be seen as a vertex analogue of inversions in tournaments. Given a permutation $\pi = (\pi_1\ \pi_2\ \ldots\ \pi_n)$ of $[n]$, for $1 \leqslant i < j \leqslant n$, the *reversal of the interval* $[i, j]$ is the permutation obtained by reversing the order of $\pi_i, \ldots, \pi_j$ in π. The *reversal distance*, $d(\pi)$, of a permutation π is the minimum number of reversals required to transform π into the identity permutation. For a survey of reversals and the reversal distance (and many other combinatorial models of genome rearrangements) see [41]. We highlight some results of particular relevance to the work covered in this chapter. With regards to computational complexity, Caprara [24] showed that the problem of Sorting by Reversals, that is, determining whether $d(\pi) \leqslant k$ for inputs of a permutation π and $k \in \mathbb{N}$, is NP-complete, while Hannenhalli and Pevzner [49, 50] showed that it is fixed-parameter tractable when parameterised by k. The natural extremal problem was solved by Bafna and Pevzner [7] who proved that for a permutation π of $[n]$, we have $d(\pi) \leqslant n-1$ with equality if and only if $\pi \in \{\gamma_n, \gamma_n^{-1}\}$ for an explicit γ_n.

Inversions in digraphs can also be thought of as generalisations of *edge reversals*, i.e. the operations which reverse the orientation of a single edge. It is not difficult to see (using an argument from Section 1.2.3) that the minimum number of such operations required to transform a digraph D into an acyclic digraph is equal to $\tau'(D)$, the cycle edge-transversal number of D. Determining this quantity is the famous feedback arc set problem, which has been widely studied (see [62] for an overview). In particular the problem of determining for inputs D and k whether $\tau'(D) \leqslant k$ was one of the first shown to be NP-complete [60] and it remains NP-complete when the input is restricted to tournaments [1, 27]. However, Chen, Liu, Lu, O'Sullivan, and Razgon [28] showed that this problem is again fixed-parameter tractable when parameterised by k. On the extremal side, it was shown by Spencer [82, 83] that the maximum cycle edge-transversal number of an n-vertex tournament is $\frac{1}{2}\binom{n}{2} - \Theta(n^{3/2})$ and that a random labelled n-vertex tournament has this cycle edge-transversal number with probability tending to 1. Bounds of this form remain the best known (see also [35, 78]).

3.7.2 Open problems

We have shown (in Theorem 1.14) that the problem which takes as inputs a tournament T and an integer $k \in \mathbb{N}$, and asks whether $\text{inv}(T) \leqslant k$, is fixed-parameter tractable when parameterised by k. In keeping with the pattern exhibited in the settings discussed in Section 3.7.1, Bang-Jensen, da Silva, and Havet [11] conjectured that the full problem is NP-complete.

Conjecture 3.11 ([11]). *The problem of deciding whether* $\text{inv}(T) \leqslant k$ *for inputs of* $k \in \mathbb{N}$ *and a tournament* T *is NP-complete.*

Note that Theorem 1.14 does not make any progress towards disproving this because the implied constant in the $O(n^2)$ running time is not polynomial in k. In fact, as noted above, the constant arising from our algorithm is doubly exponential in k. However, again in keeping with both settings discussed in Section 3.7.1 (and indeed many natural fixed-parameter tractable problems), we conjecture that this constant can be taken to be singly exponential in k, perhaps with a higher power of n.

Conjecture 3.12. *There exist constants* $c_1, c_2 > 0$ *such that* k-TOURNAMENT-INVERSION *can be solved in time* $O(2^{k^{c_1}}|V(T)|^{c_2})$ *for any* $k \in \mathbb{N}$.

As discussed in the introduction, the set $\mathcal{IC}_k$ of k-inversion-critical tournaments was shown to be finite for all k in [18]. They explicitly described $\mathcal{IC}_1$ and $\mathcal{IC}_2$, for the latter using results of Gallai [43] (see [73] for an English translation) and Latka [63], but for $k \geqslant 3$ very little is known about these sets. In particular, it would be interesting to determine m_k, the maximum number of vertices in a tournament in $\mathcal{IC}_k$, for $k \geqslant 3$.

Question 3.13 ([11]). *What is the value of* m_k *for* $k \geqslant 3$*?*

Finding the minimum possible size of a k-inversion-critical tournament is equivalent to the problem of determining $\mathrm{inv}(n)$. The best known bounds on $\mathrm{inv}(n)$ for large n are now

$$n - \sqrt{2n\log(n)} \leqslant \mathrm{inv}(n) \leqslant n - \log(n+1),$$

and it would be interesting to tighten these further.

Question 3.14. *What is the asymptotic behaviour of* $n - \mathrm{inv}(n)$*?*

In light of our improved lower bound on $\mathrm{inv}(n)$, the lack of an explicit construction for a tournament of large inversion number is even more apparent: no n-vertex construction with inversion number more than about $n/3$ (as given by the $(n/3)$-join $[\overrightarrow{C_3}]_{n/3}$) is known.

Problem 3.15. *Construct n-vertex tournaments with inversion number closer to* $\mathrm{inv}(n)$*.*

Belkhechine, Bouaziz, Boudabbous, and Pouzet ([17]; see [11]) defined for each $n \in \mathbb{N}$ a tournament Q_n on vertex set $[n]$ in which for $i < j$ the edge ij is oriented towards j, except if $j = i+1$, in which case it is oriented towards i, and conjectured that these graphs satisfy $\mathrm{inv}(Q_n) = \left\lfloor \frac{n-1}{2} \right\rfloor$.

Conjecture 3.16 ([17])**.** *For all $n \in \mathbb{N}$ we have* $\mathrm{inv}(Q_n) = \left\lfloor \frac{n-1}{2} \right\rfloor$*.*

The conjecture is known to hold for $n \leqslant 8$ [11, 18], and it is certainly true that $\mathrm{inv}(Q_n) \leqslant \left\lfloor \frac{n-1}{2} \right\rfloor$ for all n since the sets

$$\{2,3\}, \{4,5\}, \{6,7\}, \ldots, \{2\lfloor (n-1)/2 \rfloor, 2\lfloor (n-1)/2 \rfloor + 1\}$$

form a decycling family of Q_n.

Defining the *inversion distance*, $\mathrm{inv}(T, T')$, between two labelled tournaments T and T' on the same vertex set to be the minimum number of inversions required to transform T into T', we remark that the matrix rank techniques developed in Section 3.6.1 can be used to show that the maximum inversion distance between two n-vertex tournaments is exactly $n-1$. Moreover, combining these ideas with Lemma 3.8 gives an upper bound of $2^{\binom{n}{2}+n-\binom{s}{2}+s\log(n)}$ on the number of labelled tournaments within inversion distance $n-s$ of a given labelled tournament.

It is natural in this context to study the random walk $\mathcal{W}$ on the space of labelled tournaments on $[n]$ where each step in the walk consists of picking a uniform random subset of $[n]$ and inverting that set in the current tournament. In particular, we ask the following.

Question 3.17. *What is the mixing time of $\mathcal{W}$? Does it satisfy the cutoff phenomenon?*

Returning to the dijoin conjecture, Theorem 1.10 completes the work of Bang-Jensen, da Silva, and Havet in showing that the conjecture holds in the cases where $\text{inv}(L), \text{inv}(R) \leqslant 2$. We have also shown (Theorem 1.11) a k-join analogue of the dijoin conjecture for collections of 2-invertible digraphs $D_1, \ldots, D_k$ at most one of which has inversion number 2. We conjecture that this final condition can be removed.

Conjecture 3.18. *Let $k \in \mathbb{N}$ and let $D_1, \ldots, D_k$ be oriented graphs satisfying $\text{inv}(D_i) \leqslant 2$ for all i. Then*

$$\text{inv}([D_1, \ldots, D_k]) = \sum_{i=1}^{k} \text{inv}(D_i).$$

On the other hand, Theorem 1.9 gives a counterexample to the dijoin conjecture where $\text{inv}(L) = 1$ and $\text{inv}(R) = 3$. From this, we can obtain counterexamples with $\text{inv}(L) = k$ and $\text{inv}(R) = 3$ for any $k \in \mathbb{N}$: let $L = [\overrightarrow{C_3}]_k$ and let R be as in the proof of Theorem 1.9. The tournaments obtained from these by inverting the whole vertex set give counterexamples in which $\text{inv}(L) = 3$ and $\text{inv}(R) = k$. In our paper, we conjectured that here 3 can be replaced with any larger integer, or in other words that the only values of $\text{inv}(L)$ and $\text{inv}(R)$ for which the dijoin conjecture always holds are those where $\text{inv}(L), \text{inv}(R) \leqslant 2$ or where one of $\text{inv}(L)$ or $\text{inv}(R)$ is 0.

However, it has since been proven by Wang, Yang and Lu [84] that the dijoin conjecture holds when $\text{inv}(L) = 1$ and $\text{inv}(R)$ is even (or vice versa), thereby refuting part of our conjecture. Together with the counterexamples of Aubian, Havet, Hörsch, Klingelhoefer, Nisse, Rambaud, and Vermande [6] and our Theorem 1.10 covering the case where L and R both have inversion number two, the following cases are the only ones to remain open.

Question 3.19. *Let $\ell, r \geqslant 2$ be even integers with $\ell + r \geqslant 6$. Is it true that for all oriented graphs L, R with $\text{inv}(L) = \ell, \text{inv}(R) = r$ we have $\text{inv}(L \to R) = \text{inv}(L) + \text{inv}(R)$?*

A potentially closely related direction, suggested by Behague in her research statement [15] and also by McDiarmid, concerns how badly the dijoin conjecture can fail. Recall that the defect $\alpha(L, R) = \text{inv}(L) + \text{inv}(R) - \text{inv}(L \to R)$ is always non-negative; it can be zero and – as shown in the disproof of the conjecture – it can also be one. But can it be (much) larger?

Question 3.20. *Do there exist oriented graphs L and R with $\alpha(L, R) \geqslant 2$?*

Finally, we noted in Section 1.2.2 that $\mathrm{inv}(D) \leqslant \mathrm{inv}(D - \{v\}) + 2$ for all digraphs D and vertices $v \in V(D)$. It is certainly the case that this inequality is tight for some D and v. Indeed, a reformulation of Theorem 1.15 yields the stronger statement that for all $k \in \mathbb{N}$ there exists a digraph D and a set $S \subseteq V(D)$ with $|S| = k$ such that for all $T \subseteq S$ we have $\mathrm{inv}(D - T) = \mathrm{inv}(D) - 2|T|$. We conjecture, however, that the inequality $\mathrm{inv}(D) \leqslant \mathrm{inv}(D - \{v\}) + 2$ cannot be tight for all vertices v in a given digraph D.

Conjecture 3.21. *Let D be a digraph with at least one vertex. Then there exists $v \in V(D)$ such that* $\mathrm{inv}(D - \{v\}) \geqslant \mathrm{inv}(D) - 1$.

Chapter 4

A framework for the generalised Erdős-Rothschild problem and a resolution of the dichromatic triangle case

This chapter is devoted to the generalised Erdős-Rothschild problem, which we now revisit. Let $k \geqslant 3$ and $s \geqslant 2$, and let $\mathcal{X}$ be a family of s-edge-colourings of K_k. We say an s-edge-colouring of a graph G is $\mathcal{X}$*-free* if it avoids every coloured copy of K_k in $\mathcal{X}$. We write $F(G;\mathcal{X}) = |\{\mathcal{X}\text{-free } s\text{-edge-colourings of } G\}|$, and our goal is to determine $F(n;\mathcal{X}) = \max_{|V(G)|=n} F(G;\mathcal{X})$ as well as the $\mathcal{X}$-*extremal* graphs G, i.e. those satisfying $F(G;\mathcal{X}) = F(n;\mathcal{X})$.

The cornerstone result of this chapter is an exact solution of the generalised Erdős-Rothschild problem for the dichromatic triangle pattern $K_3^{(2)}$ for all values of s. Recall that we write $(K_3^{(2)}, s)$ for the family of s-edge-colourings of K_3 that use exactly two colours. In Construction 1.19, we constructed $(K_3^{(2)}, s)$-free s-edge-colourings of the r-partite Turán graph $T_r(n)$, where each colour class corresponds to a maximal matching of K_r, and the number of colours occurring between pairs of parts varies by at most one. We computed in the introduction that this construction gives rise to $O_{r,s}(1) \cdot e^{\frac{r}{r-1} g_s(r) t_r(n)}$ colourings free of dichromatic triangles, where $t_r(n)$ is the number of edges in $T_r(n)$ and $g_s(r)$ was defined accordingly (this definition will be recalled in Section 4.4.1). We then define $g(s) := \max_{r \in 2\mathbb{N}, r<s} g_s(r)$, $R_2(s) := \{r \in 2\mathbb{N} : g_s(r) = g(s)\}$ as well as the integer $r_2(s)$ which is the largest even integer r such that $(r-1)e^r \leqslant s$ when $s \geqslant e^2$ (and 2 otherwise). Armed with these definitions, we now restate our main theorem.

Theorem 1.21. *For every integer $s \geqslant 2$ we have $R_2(s) \subseteq \{r_2(s), r_2(s)+2\}$, so in particular $g(s) = \max\{g_s(r_2(s)), g_s(r_2(s)+2)\}$. For sufficiently large n, every n-vertex $(K_3^{(2)}, s)$-extremal graph is $T_r(n)$ for some $r \in R_2(s)$ and*

$$F(n; (K_3^{(2)}, s)) = (C + o(1)) \cdot e^{\frac{r}{r-1} g(s) t_r(n)},$$

where C depends only on s, r and on $n \pmod r$.
Moreover, for all $s \in [2, 10^7] \setminus \{27\}$, the unique extremal graph is $T_r(n)$ where r is the unique value in $R_2(s)$ shown in Table 1.2, and the unique extremal graph for $s = 27$ is $T_4(n)$.

With some further analysis we will deduce that there is an infinite family of graphs which are extremal for some s, in the following strong sense:

Corollary 1.22. *For any $r \in 2\mathbb{N}$, there are integers $s^-(r) \leqslant s^+(r)$ such that whenever n is sufficiently large, the Turán graph $T_r(n)$ is uniquely $(K_3^{(2)}, s)$-extremal for all $s^-(r) \leqslant s \leqslant s^+(r)$. Moreover, $s^-(r) \in ((r-3)e^{r-2}, (r-1)e^r)$ and there is at most one value of s between $s^+(r)$ and $s^-(r+2)$. If such a value of s exists, for this s every n-vertex $(K_3^{(2)}, s)$-extremal graph lies in $\{T_r(n), T_{r+2}(n)\}$.*

Recall that a *pattern* P is a partition of the edges of K_k and, given s, naturally induces a family of colourings (P, s) which includes all s-colourings where colour classes correspond to parts of P. Building on previous results, Theorem 1.21 also implies that for non-monochromatic patterns (i.e. patterns with more than one part) all extremal graphs are complete partite.

Corollary 1.23. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let P be a non-monochromatic colour pattern of K_k. Then every (P, s)-extremal graph is complete partite.*

In order to restate our related result on the case where improper colourings are forbidden, let $\mathcal{X}_{k,s}^{\wedge} = \{$all improper s-edge-colourings of $K_k\}$.

Theorem 1.24. *For every $k \geqslant 3$ there exists an integer $s(k)$ such that for every $s \geqslant k-1$ and sufficiently large n every n-vertex $\mathcal{X}_{k,s}^{\wedge}$-extremal graph is either*

- *$T_{k-1}(n)$ if $s \leqslant s(k)$, in which case $F(n; \mathcal{X}_{k,s}^{\wedge}) = s^{t_{k-1}(n)}$; or*
- *$T_r(n)$ for some $r \in R_2(s)$ if $s > s(k)$, in which case $F(n; \mathcal{X}_{k,s}^{\wedge}) = (C + o(1)) \cdot e^{\frac{r}{r-1} g(s) t_r(n)}$ where C is a constant depending only on s, r and on $n \pmod r$.*

4.1 Methods, chapter outline and notation

In the literature, results on the generalised Erdős-Rothschild problem are proved by considering a valid colouring of an n-vertex extremal graph and approximating it using a coloured version of Szemerédi's regularity lemma, or occasionally the hypergraph container method, e.g. [10, 22, 48]. Then, it suffices to solve a corresponding problem where edges may be given a set of colours, and every assignment from one of these sets is counted. In the regularity lemma approach, this is a combinatorial optimisation problem whose feasible solutions are vertex-weighted edge-coloured multigraphs which contain no forbidden patterns and whose size does not depend on n. This problem was explicitly stated for monochromatic patterns in [76] but has implicitly appeared in all earlier and most later works. Tho ugh this optimisation problem can be solved by brute force in time $O_{s,k}(1)$, it is very difficult to solve in practice. Most solutions have been obtained via a linear relaxation which only yields feasible solutions in a few cases, hence the sporadic results. A different optimisation problem arises from the container method (see [22]); again it appears difficult to solve in general.

As mentioned in Section 1.3.2, we prove our main results for forbidden dichromatic triangles and improperly coloured cliques by first deriving a proof framework for the generalised Erdős-Rothschild problem. Our framework is inspired by existing theory for monochromatic patterns developed in [74–76]. The gist of our general results is that the generalised Erdős-Rothschild problem reduces to the optimisation problem sketched above. In both cases, we are able to solve it for all s. For this, we do not use any linear programs, but mainly use analytic tools, as well as combinatorial considerations and some computer assistance for small s. We hope the ideas in the proof will be useful for solving the generalised Erdős-Rothschild problem for other colour patterns.

The first part of this chapter contains some general theory about forbidden colour patterns in cliques. In Section 4.2 we introduce the optimisation problem $Q(\mathcal{X})$ which underpins this chapter, and then in Section 4.3 state some general results which show the strong connection between the generalised Erdős-Rothschild problem and this optimisation problem. One of these is a new 'exact' result which we will use to prove the main results covered in this chapter. Section 4.4 concerns the dichromatic triangle problem, and in it we use the theory developed in the first part of the chapter and solve the appropriate instances of $Q(\mathcal{X})$ to prove our first main result, Theorem 1.21. Corollary 1.22 and Corollary 1.23 will follow immediately.

In Section 4.5 we apply the same methods to the problem of maximising the number of proper k-clique colourings and prove our second main result, Theorem 1.24. The proof of Theorem 1.24 is closely related to Theorem 1.21 and will use our tools for optimising a version of Construction 1.19 when $\mathcal{X} = (K_3^{(2)}, s)$. We prove the general results of Section 4.3 in Section 4.6. These proofs are adaptations of analogous results in [74–76], so we provide sketches only. Section 4.7 contains some concluding remarks on the generalised Erdős-Rothschild problem and further avenues for investigation.

In this chapter, log always denotes the natural logarithm $\log_e$ (this is in contrast to many papers on this topic, but useful here as we make extensive use of calculus). We also set $\log 0 := 0$ to ease notation, in particular, given a multiset $\mathcal{A}$ of sets A, we write $\sum_{A\in\mathcal{A}} \log(|A|)$ instead of $\sum_{A\in\mathcal{A}:A\neq\emptyset} \log(|A|)$. We write $x = a \pm b$ as short-hand for $x \in [a-b, a+b]$. The L^1-norm of a vector $\boldsymbol{x}$ is $\|\boldsymbol{x}\|_1 = \sum_i |x_i|$. We write $2\mathbb{N}$ to denote the set of positive even integers. We write $0 < a \ll b \ll c < 1$ to mean that we can choose the constants a, b, c from right to left. More precisely, there exist non-decreasing functions $f : (0,1] \to (0,1]$ and $g : (0,1] \to (0,1]$ such that for all $a \leqslant f(b)$ and $b \leqslant g(c)$ our calculations and arguments in our proofs are correct. Larger hierarchies are defined similarly. We suppress rounding where it does not affect the validity of the argument. Most of the specific notation will be introduced in the next section.

4.2 An optimisation problem

In [76], Pikhurko, Staden and Yilma showed that the Erdős-Rothschild problem is asymptotically solved by complete partite colouring constructions such as the one shown in Figure 1.3.2. That is, it is enough to maximise the number of colourings of a complete multipartite graph with parts $V_1, \ldots, V_r$ such that each edge in $V_i \times V_j$ is coloured arbitrarily from a list of colours $\phi(ij)$. This can be stated more precisely as an optimisation problem with parameters r, the sizes of the V_i, and ϕ. They also proved a stability result which states that any almost extremal graph for the Erdős-Rothschild problem resembles a solution to this optimisation problem. Pikhurko and Staden then showed that, under certain conditions on solutions to the optimisation problem, a stronger stability theorem holds [74], and also an exact result [75]. They solved a new case of the Erdős-Rothschild problem by solving the corresponding optimisation problem.

An analogous optimisation problem, Problem $Q(\mathcal{X})$, can be formulated for the generalised Erdős-Rothschild problem for a family $\mathcal{X}$ of forbidden colourings. As we mentioned before, we consider graphs with a partite structure, and count only the $\mathcal{X}$-free colourings such that the colours of edges are determined entirely by the parts they lie between. Thus these colourings are defined by a function ϕ which maps pairs of parts to sets of colours. This function ϕ must be such that every colouring generated by ϕ will be $\mathcal{X}$-free.

So suppose we have $r \in \mathbb{N}$, which will be the number of parts $V_1, \dots, V_r$, and a function $\phi : \binom{[r]}{2} \to 2^{[s]}$, which maps pairs of vertices to sets of colours. Given a colouring $\sigma : E(K_k) \to [s]$, we say that ϕ is *σ-free* if there is no injective map $\psi : [k] \to [r]$ with $\sigma(ij) \in \phi(\psi(i)\psi(j))$ for all $ij \in \binom{[k]}{2}$. We say that ϕ is *$\mathcal{X}$-free* if it is σ-free for every $\sigma \in \mathcal{X}$.

Before we state our optimisation problem, we define $\Phi_{\mathcal{X},t}(r)$ to be the set of $\mathcal{X}$-free colourings $\phi : \binom{[r]}{2} \to 2^{[s]}$ such that $|\phi(ij)| \geqslant t$ for all $i \neq j$. We call such ϕ *colour templates* and, given $c \in [s]$, also write $\phi^{-1}(c) := \{ij \in \binom{[r]}{2} : c \in \phi(ij)\}$ which can be considered as (the edge-set of) a graph on the vertex set $[r]$. We write $\phi_{ij} := |\phi(ij)|$ for brevity and often call this the *multiplicity* (of ij), which comes from thinking of ϕ as the multigraph made up of colour graphs $\phi^{-1}(1), \dots, \phi^{-1}(s)$. We also define Δ^r to be the set of non-negative vectors $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r)$ such that $\sum_{i \in [r]} \alpha_i = 1$. By convention, given $\boldsymbol{\alpha}_j \in \Delta^r$, we write $\boldsymbol{\alpha}_j = (\alpha_{j,1}, \dots, \alpha_{j,r})$.

Problem $Q_t(\mathcal{X})$.
Maximise
$$q(\phi, \boldsymbol{\alpha}) := 2 \sum_{ij \in \binom{[r]}{2}} \alpha_i \alpha_j \log \phi_{ij}$$
subject to $r \in \mathbb{N}$, $\phi \in \Phi_{\mathcal{X},t}(r)$ and $\boldsymbol{\alpha} \in \Delta^r$.

Denote by $\text{FEAS}_t(\mathcal{X})$ the set of feasible triples $(r, \phi, \boldsymbol{\alpha})$ for this problem, and by $\text{OPT}_t(\mathcal{X})$ the set of optimal triples. The maximum of the objective function we denote by $Q_t(\mathcal{X})$.

For ease of notation, we sometimes omit the subscript t when $t = 0$, for example, $Q(\mathcal{X}) := Q_0(\mathcal{X})$ and $\text{FEAS}(\mathcal{X}) := \text{FEAS}_0(\mathcal{X})$.

We now show how this optimisation problem gives lower bounds on $F(n; \mathcal{X})$. Let $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_0(\mathcal{X})$. Let G be an n-vertex graph with parts $V_1, \dots, V_r$, where $|V_i| \in \{\lfloor \alpha_i n \rfloor, \lceil \alpha_i n \rceil\}$, such that $G[V_i]$ has no edges for all i; if $\phi(ij) = \emptyset$ for some

$ij \in \binom{[r]}{2}$ then $G[V_i, V_j]$ has no edges, and otherwise it is a complete bipartite graph. Consider the set of all s-edge-colourings χ of G where, for every $ij \in \binom{[r]}{2}$ and every $xy \in E(G[V_i, V_j])$, we have $\chi(xy) \in \phi(ij)$. Every such χ is $\mathcal{X}$-free since ϕ is $\mathcal{X}$-free. Counting the number of such colourings, we see that

$$F(n; \mathcal{X}) \geqslant F(G; \mathcal{X}) \geqslant \prod_{ij \in \binom{[r]}{2}} \phi_{ij}^{|V_i||V_j|}.$$

Thus

$$\log F(n; \mathcal{X}) \geqslant \sum_{ij \in \binom{[r]}{2}} |V_i||V_j| \log \phi_{ij} = q(\phi, \boldsymbol{\alpha}) \binom{n}{2} + O(n) \tag{4.2.1}$$

and therefore we have the lower bound

$$F(n; \mathcal{X}) \geqslant e^{Q_0(\mathcal{X}) \binom{n}{2} + O(n)}. \tag{4.2.2}$$

Note that

$$\tfrac{k-2}{k-1} \log(s) \leqslant Q_0(\mathcal{X}) \leqslant \log(s),$$

which should be compared to our bounds in (1.3.1) for graphs. Here, the upper bound comes from $2 \sum \alpha_i \alpha_j = 1 - \sum \alpha_i^2 \leqslant 1$ and $\phi_{ij} \leqslant s$, while the lower bound – the trivial bound – comes from taking $r = k-1$, $\boldsymbol{\alpha}$ uniform and $\phi \equiv [s]$.

A central result of this chapter, presented in Section 4.3, states that, under certain conditions, the bound in (4.2.2) is tight, thus reducing finding $\log(F(n; \mathcal{X}))/\binom{n}{2}$ asymptotically to finding $Q_0(\mathcal{X})$.

The set $\text{OPT}_0(\mathcal{X})$ of solutions is rather degenerate, because we can extend every $(r, \phi, \boldsymbol{\alpha}) \in \text{OPT}_0(\mathcal{X})$ by adding a new part $r+1$ with $\alpha_{r+1} = 0$, or by splitting any part into two. It will be helpful to consider the set of solutions in which parts cannot be merged or deleted to obtain a smaller optimal solution. We define the set of *basic optimal solutions* $\text{OPT}^*(\mathcal{X})$ to be the set of $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}_2(\mathcal{X})$ such that $\alpha_i^* > 0$ for all $i \in [r^*]$. We now show that the set $\text{OPT}_0(\mathcal{X})$ contains a basic optimal solution. The following fact also justifies our omission of subscripts when referring to $Q(\mathcal{X})$.

Fact 4.1. Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a family of s-edge-colourings of K_k. If $\text{OPT}_0(\mathcal{X}) \neq \emptyset$, then $\text{OPT}^*(\mathcal{X}) \neq \emptyset$. Moreover, $Q_2(\mathcal{X}) = Q_1(\mathcal{X}) = Q_0(\mathcal{X})$ and $\text{OPT}_2(\mathcal{X}) \subseteq \text{OPT}_1(\mathcal{X}) \subseteq \text{OPT}_0(\mathcal{X})$.

Proof. Let $(r, \phi, \boldsymbol{\alpha}) \in \text{OPT}_0(\mathcal{X})$ be an optimal solution with minimum number of parts r. It holds that $\alpha_i > 0$ for all $i \in [r]$. Indeed, otherwise we would obtain another optimal solution with fewer parts. Let us, for the sake of contradiction, assume that $(r, \phi, \boldsymbol{\alpha}) \notin \text{OPT}_2(\mathcal{X})$, say, $\phi_{12} \leqslant 1$. For any ε with $-\alpha_2 \leqslant \varepsilon \leqslant \alpha_1$ define $\tilde{\alpha}_1 = \alpha_1 - \varepsilon$, $\tilde{\alpha}_2 = \alpha_2 + \varepsilon$, and $\tilde{\alpha}_i = \alpha_i$ for all other i. (This is a version of 'Zykov symmetrisation', see Section 4.6.) Then $q(\phi, \tilde{\boldsymbol{\alpha}}) = Q(\mathcal{X}) - \varepsilon \sum_{i=3}^{r} \alpha_i (\log \phi_{1i} - \log \phi_{2i})$ since the contribution $\alpha_1 \alpha_2 \log \phi_{12}$ is zero. By optimality of $(r, \phi, \boldsymbol{\alpha})$ we must have that $\sum_{i=3}^{r} \alpha_i (\log \phi_{1i} - \log \phi_{2i}) = 0$, for otherwise either a positive or negative ε would yield a better solution. Thus setting $\varepsilon = \alpha_1$ and removing part 1 (of size $\tilde{\alpha}_1$) yields a solution $(r-1, \phi|_{[2,r]}, \tilde{\boldsymbol{\alpha}}) \in \text{OPT}_0(\mathcal{X})$ with fewer parts, a contradiction to the minimality of r.

For the second part, as $\text{FEAS}_2(\mathcal{X}) \subseteq \text{FEAS}_1(\mathcal{X}) \subseteq \text{FEAS}_0(\mathcal{X})$ we have $Q_2(\mathcal{X}) \leqslant Q_1(\mathcal{X}) \leqslant Q_0(\mathcal{X})$. However, from our proof above, $q(\phi, \boldsymbol{\alpha}) = Q_0(\mathcal{X}) = Q_2(\mathcal{X})$, so we have equality throughout. Since the optimum is the same for $t = 0, 1, 2$, we also have the inclusion $\text{OPT}_2(\mathcal{X}) \subseteq \text{OPT}_1(\mathcal{X}) \subseteq \text{OPT}_0(\mathcal{X})$. □

Fact 4.2. Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a family of s-edge-colourings of K_k. Then for all $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$, the colour template ϕ^* is maximal. That is, if one obtains a new colour template ϕ by adding to ϕ^* a new colour in $[s]$ between any pair $ij \in \binom{[r^*]}{2}$, then ϕ is not $\mathcal{X}$-free.

Proof. Suppose that there is some $ij \in \binom{[r^*]}{2}$ and $c \in [s]$ such that ϕ as described above is $\mathcal{X}$-free. Then $(r^*, \phi, \boldsymbol{\alpha}^*) \in \text{FEAS}_2(\mathcal{X})$ and $q(\phi, \boldsymbol{\alpha}^*) - q(\phi^*, \boldsymbol{\alpha}^*) = \alpha_i^* \alpha_j^* (\log \phi_{ij} - \log \phi_{ij}^*) \geqslant \log(\frac{s}{s-1}) \alpha_i^* \alpha_j^* > 0$, a contradiction to the optimality of $(r^*, \phi^*, \boldsymbol{\alpha}^*)$. □

4.2.1 Contributions

Given $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}(\mathcal{X})$ and $i \in [r]$, we define

$$q_i(\phi, \boldsymbol{\alpha}) := \sum_{j \in [r] \setminus \{i\}} \alpha_j \log \phi_{ij}$$

and refer to q_i as the *contribution of part* i, as $q(\phi, \boldsymbol{\alpha}) = \sum_{i \in [r]} \alpha_i q_i(\phi, \boldsymbol{\alpha})$.

The following lemma, a version of [74, Proposition 2.1], is crucial in our proof of Theorem 1.21 as well as in the other general results of the next section. It states that every vertex in an optimal solution contributes optimally to q. Indeed, if two vertices had differing contributions, we can move weight from one to the other to increase q, contradicting optimality. This can be proved via the method of Lagrange multipliers.

Lemma 4.3. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a family of s-edge-colourings of K_k. Then for any $(r, \phi, \boldsymbol{\alpha}) \in \text{OPT}(\mathcal{X})$, we have*

$$q_i(\phi, \boldsymbol{\alpha}) = Q(\mathcal{X}) \quad \textit{whenever } \alpha_i > 0.$$

4.2.2 Properties of families and solutions

Now we introduce some properties of $\mathcal{X}$ – or, more specifically, the set $\textsc{opt}^*(\mathcal{X})$ of basic optimal solutions of Problem $Q_2(\mathcal{X})$ – which are required to state our general results. Given a description of $\textsc{opt}^*(\mathcal{X})$, these properties all tend to be easy to check.

Given $r \in \mathbb{N}$ and $\phi \in \Phi_{\mathcal{X},0}(r)$, we say that $i \in [r]$ is

- a *clone of* $j \in [r] \setminus \{i\}$ *(under* ϕ*)* if $\phi(i\ell) = \phi(j\ell)$ for all $\ell \in [r] \setminus \{i,j\}$ and $\phi_{ij} \leqslant 1$.
- a *strong clone of* j *(under* ϕ*)* if i is a clone of j and $\phi_{ij} = 0$.

For $(r, \phi, \boldsymbol{\alpha}) \in \textsc{feas}(\mathcal{X})$ and $\phi' \in \Phi_{\mathcal{X}}(r+1)$ such that $\phi'|_{\binom{[r]}{2}} = \phi$, we define

$$\mathrm{ext}(\phi', \boldsymbol{\alpha}) := q_{r+1}(\phi', (\alpha_1, \ldots, \alpha_r, 0)) = \sum_{i \in [r]} \alpha_i \log |\phi'(\{i, r+1\})|.$$

Definition 4.4 (Bounded, (strong) extension property, stable inside, hermetic)**.** Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a family of s-colourings of K_k. We say that $\mathcal{X}$

- is *bounded* if $\textsc{opt}^*(\mathcal{X})$ is non-empty and there is some $R > 0$ such that $r^* \leqslant R$ for all $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$,
- has the *(strong) extension property* if for any $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$ and $\phi \in \Phi_{\mathcal{X}}(r^*+1)$ with $\phi|_{\binom{[r^*]}{2}} = \phi^*$ and $\mathrm{ext}(\phi, \boldsymbol{\alpha}^*) = Q(\mathcal{X})$, there exists $j \in [r^*]$ such that r^*+1 is a (strong) clone of j under ϕ,
- is *stable inside* if for every $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$ the following holds: if we make a non-strong clone of a vertex $x \in [r^*]$ (i.e. define ϕ' on r^*+1 parts by setting $\phi'(\{r^*+1, i\}) = \phi(xi)$ for all $i \in [r^*] \setminus \{x\}$ and $\phi'(\{r^*+1, x\}) = \{c\}$ for some $c \in [s]$), then there is a forbidden configuration (i.e. $\phi' \notin \Phi_{\mathcal{X}}(r^*+1)$),
- is *hermetic* if, for all $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$, there is no $\phi \in \Phi_{\mathcal{X},1}(r^*+1)$ such that $\phi|_{\binom{[r^*]}{2}} = \phi^*$.

Observe that (i) if $\mathcal{X}$ is hermetic, then it is stable inside, since adding a non-strong clone to a basic optimal solution is a specific extension of that basic optimal solution where the new vertex sends at least one colour to all existing vertices; and (ii) if $\mathcal{X}$ has the extension property and is hermetic, then it has the strong extension property, since being hermetic implies that a non-strong clone cannot have optimal contribution.

4.2.3 Examples and specific cases

In this section we derive properties of $\text{OPT}^*(\mathcal{X})$ for some important patterns P and resulting symmetric families (P, s) of s-edge-colourings of K_k.

Forbidden monochromatic cliques: $P = K_k^{(1)}$

For every $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_1(K_k^{(1)}, s)$, we have $r < R_s(k)$, the s-colour Ramsey number of K_k. Thus $\text{FEAS}_1(K_k^{(1)}, s)$ is compact, and, since q is continuous, we have that $\text{OPT}_1(K_k^{(1)}, s) \neq \emptyset$. In particular, by Fact 4.1, $(K_k^{(1)}, s)$ is bounded. Fact 4.2 implies that for $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(K_k^{(1)}, s)$, for every colour $c \in [s]$, the graph $\phi^{*-1}(c)$ is maximally K_k-free. Every known exact result for the (monochromatic) Erdős-Rothschild problem can be deduced via auxiliary results from the corresponding optimisation problem; in all cases, every $\phi^{*-1}(c)$ is a $(k-1)$-partite Turán graph and $\boldsymbol{a}^*$ is uniform.

Forbidden dichromatic triangles: $P = K_3^{(2)}$

Lemma 4.5. *Let $s \geqslant 2$ be an integer. Then the following hold.*

(i) *For every $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_2(K_3^{(2)}, s)$, we have that $\phi^{-1}(c)$ is a matching for all $c \in [s]$. In particular, $\sum_{j \in [r] \setminus \{i\}} \phi_{ij} \leqslant s$ for all $i \in [r]$.*

(ii) *For every $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(K_3^{(2)}, s)$, we have that $\phi^{*-1}(c)$ is a maximal matching for all $c \in [s]$.*

(iii) *$(K_3^{(2)}, s)$ is bounded.*

Proof. We first prove (i). Let $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_2(K_3^{(2)}, s)$. If $\phi^{-1}(c)$ is not a matching for some $c \in [s]$, then there exist two adjacent edges, say hi and hj, such that $c \in \phi(hi)$ and $c \in \phi(hj)$. Since all edges receive at least two colours, the edge ij receives at least one colour different from c. The multicoloured graph induced on the vertices h, i, and j contradicts the $K_3^{(2)}$-free property of the function ϕ. The last part follows from the identity $\sum_{j \in [r] \setminus \{i\}} \phi_{ij} = \sum_{c \in [s]} d_{\phi^{-1}(c)}(i)$.

Part (ii) follows from part (i) and Fact 4.2.

For (iii), let $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_2(K_3^{(2)}, s)$. Similarly to Section 4.2.3, it suffices to show that there is some R such that $r < R$ for every $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_2(K_3^{(2)}, s)$. As noted, $\phi^{-1}(c)$ is a matching for each $c \in [s]$, therefore it contains at most $r/2$ edges. Since

$2\binom{r}{2}/\left(\frac{r}{2}\right) = 2(r-1)$ and together the matchings are required to cover the edges of K_r at least twice over, we have that $s \geqslant 2(r-1)$. In other words, $r \leqslant s/2+1$. $\square$

Forbidden 2-edge-colourings

In the case of $s = 2$, the fact that we may restrict to ϕ with $\phi_{ij} \geqslant 2$ for all ij (see Fact 4.1) means we can find $\text{OPT}^*(\mathcal{X})$ for all $\mathcal{X}$.

Lemma 4.6. *Let $k \geqslant 3$ be an integer and let $\mathcal{X}$ be a family of 2-edge-colourings of K_k. Then $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$ if and only if $r^* = k-1$, $\phi^*(ij) = [2]$ for all $ij \in \binom{[r^*]}{2}$, and $\boldsymbol{\alpha}^*$ is uniform.*

Proof. Take some $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$. Elements of $\text{OPT}^*(\mathcal{X})$ see at least two colours everywhere so $\phi^*(ij) = [2]$ for all ij. This contains any pattern on r^* vertices so we must have $r^* \leqslant k-1$. Then $q(\phi^*, \boldsymbol{\alpha}^*) = 2\log(2)\sum_{ij}\alpha_i^*\alpha_j^*$ which for fixed r^* is maximised by uniform $\boldsymbol{\alpha}^*$. Indeed, if some $\alpha_i^* > \alpha_j^*$, perturbing α_i^* down by a sufficiently small ε and α_j^* up by ε yields an increase proportional to $\varepsilon(\alpha_i^* - \alpha_j^* - \varepsilon) > 0$ in q, contradicting optimality. Therefore $\boldsymbol{\alpha}^*$ is uniform and $q(\phi^*, \boldsymbol{\alpha}^*) = \log(2)\frac{r^*-1}{r^*}$. This function is increasing in r^* so it is maximised by $r^* = k-1$. $\square$

Forbidden rainbow cliques: $P = K_k^{(\binom{k}{2})}$

Let $k \geqslant 3$ and $s \geqslant \binom{k}{2}$. Now, for $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_1(\mathcal{X})$, the number r of parts can be arbitrarily large since any solution with ϕ using less than $\binom{k}{2}$ colours is feasible. Note that this does not imply $(K_k^{(\binom{k}{2})}, s)$ is not bounded.

Lemma 4.7. *Let $s \geqslant 3$ be an integer. Let $\mathcal{X} = (K_3^{(3)}, s)$. If $s = 3$, then $\text{OPT}^*(\mathcal{X}) = \emptyset$, so in particular, $\mathcal{X}$ is not bounded. If $s \geqslant 4$, then $\text{OPT}^*(\mathcal{X})$ contains the unique element $(2, \phi^*, (\frac{1}{2}, \frac{1}{2}))$ where $\phi^* \equiv [s]$, so in particular, $\mathcal{X}$ is bounded.*

Proof. Consider the feasible solution $(r, \phi, \boldsymbol{\alpha})$ such that there is $A \subseteq [s]$ of size two for which $\phi(ij) = A$ for all pairs ij in $[r]$. By convexity, $\boldsymbol{\alpha}$ is uniform (otherwise $q(\phi, \boldsymbol{\alpha})$ can be increased), and we have $q(\phi, \boldsymbol{\alpha}) = (1 - \frac{1}{r})\log(2)$. However, we can find a better solution simply by adding one to r, so $(r, \phi, \boldsymbol{\alpha})$ is not optimal, and we obtain a sequence of feasible solutions whose q value tends to $\log(2)$ from below. Thus $Q^*(\mathcal{X}) \geqslant \log(2)$.

Suppose that $\text{OPT}^*(\mathcal{X}) \neq \emptyset$. Let $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$. By the above, at least three colours appear in ϕ^*.

Suppose first that they never appear on a single pair. Then there are distinct $h, i, j \in [r^*]$ such that, without loss of generality, $\phi^*(hi) = \{1, 2\}$ and $\phi^*(hj) = \{1, 3\}$. But then $\phi^*(ij) = \emptyset$, a contradiction.

Thus there is some ij with $|\phi^*(ij)| \geqslant 3$. If there is any other $h \in [r^*] \setminus \{i, j\}$, since there are at least two colours on hi and hj, we see there is a rainbow triangle on h, i, j, a contradiction. Thus $r = 2$ and, optimising, $\phi^*(12) = [s]$ and $\boldsymbol{\alpha}^*$ is uniform, with $Q^*(\mathcal{X}) = q(\phi^*, \boldsymbol{\alpha}^*) = \frac{1}{2}\log(s)$. So $\frac{1}{2}\log(s) \geqslant \log(2)$. This implies that $s \geqslant 4$. □

This example suggests that we should expand our set of feasible solutions to allow colours inside parts, that is, expand the domain of ϕ to both pairs and singletons. Then the above argument would show that every optimal solution for the rainbow triangle problem with 2 or 3 colours corresponds to a 2-edge-coloured clique, that is, a one-part solution where the part receives two colours only. This aligns with the results in [10], which show that the maximum number of 3-colour Gallai colourings is attained by K_n and is equal to $(3 + o(1))2^{\binom{n}{2}}$. Many of our general results can be formulated in this setting but we chose not to do this, as in our two applications, giving colours to singletons is not necessary. We do however discuss this extension in Section 4.7.

4.3 Results for general colour patterns

This section contains some general results for bounded families $\mathcal{X}$ of forbidden edge-colourings. Their culmination is Theorem 4.11, an 'exact' result for hermetic families with the extension property. We prove our main result, Theorem 1.21, by combining this with the solution to the optimisation problem for dichromatic triangles, which we obtain in Section 4.4.

4.3.1 Asymptotic upper bound

We saw in Section 4.2 that optimal solutions of $Q(\mathcal{X})$ give a lower bound for $F(n; \mathcal{X})$ (see (4.2.2)). Using Szemerédi's regularity lemma, one can prove that there is a matching upper bound.

Theorem 4.8. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a family of s-edge-colourings of K_k for which* $\text{OPT}(\mathcal{X}) \neq \emptyset$. *Then*

$$F(n; \mathcal{X}) = e^{Q(\mathcal{X})\binom{n}{2} + o(n^2)}.$$

This result and its proof has appeared for specific colour patterns many times in the literature, starting with [2] which considered the monochromatic pattern $K_k^{(1)}$ and $s \in \{2,3\}$ colours. A general monochromatic version appeared in [76]. The corresponding result for general colour patterns is no harder, so we don't prove it separately. A proof is obtained en route to proving Theorem 4.10 (see Section 4.6.3).

4.3.2 Stability for optimal solutions to $Q(\mathcal{X})$

The next key theorem is a generalisation of the main result of [74] which is for monochromatic patterns. It is a 'stability' theorem which states that every almost optimal feasible solution to Problem $Q(\mathcal{X})$ has a similar structure to a basic optimal solution.

Theorem 4.9. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a bounded family of s-edge-colourings of K_k which has the extension property. Let $\nu > 0$. Then there exists $\varepsilon > 0$ such that for every $(r, \phi, \boldsymbol{\alpha}) \in \textsc{feas}(\mathcal{X})$ with*

$$q(\phi, \boldsymbol{\alpha}) > Q(\mathcal{X}) - \varepsilon,$$

there exist $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$ and a partition $[r] = Y_0 \cup \ldots \cup Y_{r^*}$ such that, defining $\beta_i = \sum_{i' \in Y_i} \alpha_{i'}$ for all $i \in [r^*]$, the following hold.

(SO1) $\|\boldsymbol{\beta} - \boldsymbol{\alpha}^*\|_1 < \nu$. *(In particular, $\sum_{i' \in Y_0} \alpha_{i'} < \nu$.)*

(SO2) *For all $ij \in \binom{[r^*]}{2}$, $i' \in Y_i$ and $j' \in Y_j$, we have that $\phi(i'j') \subseteq \phi^*(ij)$.*

(SO3) *For all $i \in [r^*]$ there is a colour $c_i \in [s]$ such that for every $i'j' \in \binom{Y_i}{2}$, we have $\phi(i'j') \subseteq \{c_i\}$. If the pattern is stable inside, then for all $i \in [r^*]$ and every $i'j' \in \binom{Y_i}{2}$, we have $\phi(i'j') = \emptyset$.*

The proof follows [74] closely, so we provide a sketch in Section 4.6.2.

4.3.3 Stability for graphs

The next general result states that almost optimal graphs have a very similar structure to the blow-up of an optimal solution (and almost all valid colourings follow an optimal colour template). Again, a monochromatic version appears in [74] and the proof is very similar, so we sketch it in Section 4.6.3.

Given a graph G, disjoint $A, B \subseteq V(G)$ and $0 \leqslant d \leqslant 1$, we say that $G[A, B]$ is (δ, d)-*regular* if $d_G(A, B) := e_G(A, B)|A|^{-1}|B|^{-1} \in (d - \delta, d + \delta)$, and $|d_G(X, Y) - d_G(A, B)| <$

δ for all $X \subseteq A$, $Y \subseteq B$ with $|X|/|A|, |Y|/|B| \geqslant \delta$. We suppress the subscript G when it is clear from the context.

Theorem 4.10. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a bounded family of s-edge-colourings of K_k which has the extension property. Then for all $\delta > 0$ there exist $n_0 \in \mathbb{N}$ and $\varepsilon > 0$ such that the following holds. If G is a graph on $n \geqslant n_0$ vertices such that*

$$\frac{\log F(G;\mathcal{X})}{\binom{n}{2}} \geqslant Q(\mathcal{X}) - \varepsilon,$$

then for at least $(1 - e^{-\varepsilon n^2})F(G;\mathcal{X})$ valid colourings $\chi : E(G) \to [s]$ of G there exist $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$ and a partition $V_1 \cup \ldots \cup V_{r^*}$ of $V(G)$ such that:

(SG1) *For all $i \in [r^*]$ we have $||V_i| - \alpha_i^* n| < 1$.*

(SG2) *For all $ij \in \binom{[r^*]}{2}$ and $c \in \phi^*(ij)$, the graph $\chi^{-1}(c)[V_i, V_j]$ is $(\delta, |\phi^*(ij)|^{-1})$-regular.*

(SG3) *For each $i \in [r^*]$ there is a colour $c_i \in [s]$ with $\sum_{i \in [r^*]} |E(G[V_i]) \backslash \chi^{-1}(c_i)| \leqslant \delta n^2$. Moreover, if $\mathcal{X}$ is stable inside, then $\sum_{i \in [r^*]} e(G[V_i]) \leqslant \delta n^2$.*

Note that (SG2) implies that all but at most an $s\delta$ proportion of pairs in (V_i, V_j) are edges in G, and given a colour in $\phi^*(ij)$ by χ.

4.3.4 An exact result

Our final general result is 'exact', in the sense that it guarantees that every extremal graph is a complete partite graph with part sizes very close to some optimal solution, and moreover, almost all valid colourings follow an optimal colour template perfectly. One of the main results of [75] is a monochromatic version that assumes the strong extension property. Here we assume both the extension property and that $\mathcal{X}$ is hermetic. As observed, these together imply the strong extension property. Being hermetic is a very strong property that, we will show, is satisfied by the families of interest in this chapter, but we do not think it holds for any of the other families where the optimisation problem has been solved. Assuming that the family is hermetic leads to a much simpler proof. We discuss what might arise from dropping this assumption in Section 4.7.

Theorem 4.11. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a bounded, hermetic family of s-edge-colourings of K_k which has the extension property. Then, for all $\delta > 0$, there exists ε with $0 < \varepsilon < \delta$ such that whenever n is sufficiently large, the following hold for every $\mathcal{X}$-extremal graph G on n vertices: There is an integer r^* and $\boldsymbol{\alpha}^* \in \Delta^{r^*}$ such that*

(SE1) *G is a complete r^*-partite graph whose i-th part W_i has size $(\alpha_i^* \pm \delta)n$ for all $i \in [r^*]$, and*

(SE2) *for at least $(1-e^{-\varepsilon n})F(G;\mathcal{X})$ valid colourings χ of G there is ϕ^* such that*

(SE2.1) $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$.

(SE2.2) *$\chi^{-1}(c)[W_i, W_j]$ is $(\delta, |\phi^*(ij)|^{-1})$-regular for all $ij \in \binom{[r^*]}{2}$ and $c \in \phi^*(ij)$.*

(SE2.3) *χ is* perfect*: that is, for all $ij \in \binom{[r^*]}{2}$ and $x \in W_i$ and $y \in W_j$, we have $\chi(xy) \in \phi^*(ij)$.*

Part (SE2.3) is crucial in determining the extremal graph(s). This is because, to do so, we need to compare the number of valid colourings in very similar graphs, which all have r^* parts of roughly the same sizes. For this, we need to understand very accurately what most colourings look like.

The proof of Theorem 4.11 is given in Section 4.6.4.

4.4 Forbidding dichromatic triangles

In this section, we apply our general results connecting the generalised Erdős-Rothschild problem to the optimisation problem to study a specific forbidden pattern, the dichromatic triangle; that is, the triangle coloured by exactly two colours. Throughout this section we set $k=3$, the forbidden pattern $P = K_3^{(2)}$ is the unique partition of K_3 into two non-empty classes, and $\mathcal{X} = (P, s)$. Only s will vary. Thus we write $Q(s) := Q(\mathcal{X})$, $\text{OPT}(s) := \text{OPT}(\mathcal{X})$, $F(n;s) := F(n;\mathcal{X})$ and so on, and call an $\mathcal{X}$-extremal graph *s-extremal*. We will prove Theorem 1.21, which determines, for $s \geqslant 2$ and large n, the value of $F(n;s)$ up to a multiplicative error of $1+o(1)$, as well as the extremal graphs, which are indeed complete partite, as required to prove Corollary 1.23. In fact, we show that there are at most two extremal graphs for each $s \geqslant 2$ and large n.

4.4.1 A lower bound construction

We first recall Definition 1.20. Let $s \geqslant 2$ be the number of colours, and let $r \geqslant 2$ be an integer. Let z and $a \in \{0, \dots, r-2\}$ be the quotient and remainder of s when divided by $(r-1)$; that is, $z = \left\lfloor \frac{s}{r-1} \right\rfloor$ and $a = s-(r-1)z$. Define

$$g_s(r) := \left(\frac{r-1-a}{r}\right)\log(z) + \left(\frac{a}{r}\right)\log(z+1),$$

s	$R(s)$
$[2, 16]$	$\{2\}$
$[17, 76]$	$\{3\}$
$[77, 299]$	$\{4\}$
$[300, 1058]$	$\{5\}$
$[1059, 3544]$	$\{6\}$
$[3545, 11443]$	$\{7\}$
$[11444, 36023]$	$\{8\}$
$[36024, 90000]$	$\{9\}$

Table 4.1: $R(s)$ for $s \leqslant 90000$ generated by the script `optr.py`.

$$g(s) := \max_{r \in 2\mathbb{N}} g_s(r), \quad \text{and} \quad g^{\max}(s) := \max_{r \in \mathbb{N}} g_s(r)$$

as well as

$$R_2(s) := \{r \in 2\mathbb{N} : g_s(r) = g(s)\} \quad \text{and} \quad R(s) := \{r \in \mathbb{N} : g_s(r) = g^{\max}(s)\}.$$

Note that $g_s(r) = 0$ when $r - 1 \geqslant s$, and so $g(s)$ and $g^{\max}(s)$ exist (and hence $R(s)$ and $R_2(s)$ are non-empty), and this definition agrees with Definition 1.20 where we restricted to $r < s$. Recall that we listed $R_2(s)$ for $s \leqslant 10^7$ in Table 1.2. Table 4.1 shows values of $R(s)$ for $s \leqslant 90000$. Let also

$$\tilde{g}_s(x) := \frac{x-1}{x} \log\left(\frac{s}{x-1}\right), \qquad \text{for } x > 1 \text{ and } s \in \mathbb{R}^+, \text{ and let} \qquad \tilde{g}(s) := \sup_{x>1} \tilde{g}_s(x).$$

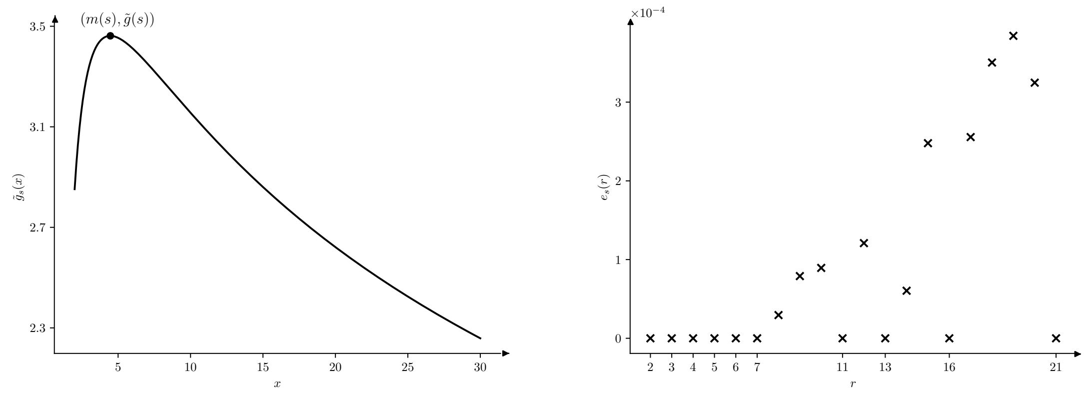
The function $\tilde{g}_s$ is a fractional approximation of g_s obtained by removing the rounding from $s/(r-1)$. So when s is divisible by $r - 1$, we have $\tilde{g}_s(r) = g_s(r)$. Finally, let

$$e_s(r) := \tilde{g}_s(r) - g_s(r).$$

We will show in Lemma 4.18 that $e_s(r)$ is very small and non-negative for near-optimal r, so $\tilde{g}_s$ is a good approximation to g_s.

Next we state two constructions of feasible triples with r parts for $r \in 2\mathbb{N}$, which we will show achieve a q-value of $g_s(r)$. The colourings generated by Construction 4.12 are equivalent to our earlier, less formal, description in Construction 1.19.

Construction 4.12. Let $r = 2t$ be even and let $M_1, \ldots, M_T$ be a list of all the perfect matchings of K_r, where $T = r!/(t!2^t)$. Let $\mathcal{A} = \{A_1, \ldots, A_T\}$ be a partition of $[s]$ (where some parts could be empty) with the property that $|b_e - b_{e'}| \leqslant 1$ for all $e, e' \in E(K_r)$, where $\phi_{\mathcal{A}}(e) := \bigcup_{\ell \in [T] : e \in M_\ell} A_\ell$ and $b_e := |\phi_{\mathcal{A}}(e)|$. Let $\boldsymbol{u} \in \Delta^r$ be uniform. This defines $(r, \phi_{\mathcal{A}}, \boldsymbol{u})$.



(a) (b)

Figure 4.4.1: Plot (A) shows $\tilde{g}_s(x)$ for $s = 300$ with maximum attained at $m(s) = W(s/e) + 1$. Plot (B) displays the difference $e_s(r) = \tilde{g}_s(r) - g_s(r)$ for $s = 300$. Note that the difference is 0 whenever s is divisible by $r - 1$.

As we will show shortly, this implies every $e \in E(K_r)$ satisfies

$$|\phi_{\mathcal{A}}(e)| \in \left\{ \left\lfloor \frac{s}{r^* - 1} \right\rfloor, \left\lceil \frac{s}{r^* - 1} \right\rceil \right\}. \tag{4.4.1}$$

The second construction is a special case of the first and shows that the set of partitions $\mathcal{A}$ satisfying the property above is non-empty.

Construction 4.13. Let $r = 2t$ be even and let $E(K_r) = M_1 \cup \ldots \cup M_{r-1}$ be a decomposition of the r-clique into perfect matchings and write $\mathcal{M} := \{M_1, \ldots, M_{r-1}\}$. Let $[s] = A_1 \cup \ldots \cup A_{r-1}$ be an equipartition.[1] For each $ij \in \binom{[r]}{2}$, let $\phi_{\text{dec}}(ij) := A_\ell$ where $ij \in M_\ell$. Let $\boldsymbol{u} \in \Delta^r$ be uniform. This defines $(r, \phi_{\text{dec}}, \boldsymbol{u})$.

In fact, these constructions are identical when $r \in \{2, 4\}$, since all pairs of perfect matchings of K_r are edge-disjoint. For all larger even r, Construction 4.12 yields strictly more ϕ.

We claim that the triples defined in both constructions are feasible. As the second is a special case of the first, it suffices to see why $(r, \phi_{\mathcal{A}}, \boldsymbol{u})$ in Construction 4.12 is feasible. For this, we just need to check that $\phi_{\mathcal{A}}$ is $K_3^{(2)}$-free. It suffices to show that there are no distinct $h, i, j \in [r]$ and $c \in [s]$ with $c \in \phi_{\mathcal{A}}(hi) \cap \phi_{\mathcal{A}}(hj)$. Let $\ell \in [T]$ be such that $c \in A_\ell$. Then $\phi_{\mathcal{A}}^{-1}(c) = M_\ell$, a matching, so we are done.

Let us calculate $q(\phi_{\mathcal{A}}, \boldsymbol{u})$. With a and z as defined earlier in this section, for each $i \in [r]$, since $|b_{ij} - b_{ij'}| \leqslant 1$ for any $j, j' \in [r] \setminus \{i\}$ and $\sum_{j \in [r] \setminus \{i\}} b_{ij} = s$, we have that

[1] An *equipartition* of a set has part sizes whose pairwise differences are at most 1.

a of the sets $\{\phi_{\mathcal{A}}(ij)\}_{j\in[r]\setminus\{i\}}$ have size $z+1$, while the remaining $r-1-a$ have size z (and, in particular, (4.4.1) holds). Thus

$$q(\phi_{\mathcal{A}}, \boldsymbol{u}) = 2\left(\frac{1}{r}\right)^2 \frac{r}{2}\left((r-1-a)\log(z) + a\log(z+1)\right) = g_s(r).$$

Thus also $q(\phi_{\mathrm{dec}}, \boldsymbol{u}) = g_s(r)$ for any ϕ_{dec} as in Construction 4.13. We have proved the following lemma.

Lemma 4.14. *For all integers $s \geqslant 2$, we have $Q(s) \geqslant g(s)$.* □

We end this section by introducing some terminology for s-colour templates ϕ on r parts that behave as in Construction 4.12. We say that ϕ is *uniform at i* for some $i \in [r]$ if the set $\{\phi(ij)\}_{j\in[r]\setminus\{i\}}$ is an equipartition of $[s]$. We further say that ϕ is *uniform* if it is uniform at i for all $i \in [r]$.

Remark 4.15. A quick check shows that $\phi \in \Phi_{P,s}(r)$ is uniform if and only if r is even and $\phi = \phi_{\mathcal{A}}$ for some $\mathcal{A}$ as in Construction 4.12.

4.4.2 Main optimisation result and proof sketch

The machinery we have developed for general colour patterns means that our main task here is to determine $\mathrm{OPT}^*(s)$ for all integers $s \geqslant 2$. We will show that, for all $s \geqslant 2$, every element of $\mathrm{OPT}^*(s)$ is of the form $(r, \phi_{\mathcal{A}}, (\frac{1}{r}, \ldots, \frac{1}{r}))$, where $r \in R_2(s)$ and $\phi_{\mathcal{A}}$ is as in Construction 4.12. This will show the following theorem, which we will then combine with Theorem 4.11 to prove Theorem 1.21.

Theorem 4.16. *Let $s \geqslant 2$ be an integer. Then $Q(s) = g(s)$, and $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \mathrm{OPT}^*(s)$ if and only if $r^* \in R_2(s)$, $\boldsymbol{\alpha}^* = (\frac{1}{r^*}, \ldots, \frac{1}{r^*})$ is uniform, and $\phi^* = \phi_{\mathcal{A}}$ for some $\mathcal{A}$ as in Construction 4.12.*

Remark 4.17. We are actually proving the following, where instead of maximising q over ϕ without dichromatic triangles, we maximise q over ϕ in which each colour class is a matching. Together with Lemma 4.5(ii), this implies Theorem 4.16.

Let $s \geqslant 2$ be an integer. Let $(r, \phi, \boldsymbol{\alpha})$ be such that r is an integer, $\phi : \binom{[r]}{2} \to 2^{[s]}$ and $\phi^{-1}(c)$ is a matching for all $c \in [s]$, and $\boldsymbol{\alpha} \in \Delta^r$. Then $q(\phi, \boldsymbol{\alpha}) \leqslant g(s)$ with equality if and only if $r \in R_2(s)$, $\boldsymbol{\alpha}$ is uniform and $\phi = \phi_{\mathcal{A}}$ for some $\mathcal{A}$ as in Construction 4.12.

We now sketch the proof of Theorem 4.16. Let $(r, \phi, \boldsymbol{\alpha})$ be a basic optimal solution. We recall from Lemma 4.3 that every part has optimal contribution, that is, $q_i(\phi, \boldsymbol{\alpha}) = \sum_{j\in[r]\setminus\{i\}} \alpha_j \log \phi_{ij} = Q(s)$. We will make use of this fact by analysing contributions in a number of ways. The first way is the following averaging trick for the contributions: for an optimal solution we have

$$q(\phi, \boldsymbol{\alpha}) = \frac{1}{r}\sum_{i\in[r]} q_i(\phi, \boldsymbol{\alpha}) = \sum_i \frac{1}{r}\sum_{j\neq i} \alpha_j \log \phi_{ij} = \frac{1}{r}\sum_j \alpha_j \sum_{i\neq j} \log \phi_{ij}$$
$$\leqslant \sum_j \alpha_j g_s(r) = g_s(r),$$

where the last inequality uses a simple fact about maximising the sum of logarithms (Lemma 4.23). This shows that solutions on an even number of parts have $q(\phi, \boldsymbol{\alpha}) \leqslant g(s)$ and some further analysis shows that this bound is only attained by the $(r, \phi, \boldsymbol{\alpha})$ described in Theorem 4.16.

It then only remains to rule out the case that r is odd for which we needed to work much harder. By Lemma 4.5(ii), every $\phi^{-1}(c)$ is a maximal matching, and hence we might intuitively expect r to be even so that the matchings can span the vertex set. When 1 is a largest part, again averaging over contributions, we see that

$$q(\phi, \boldsymbol{\alpha}) = \frac{1}{r}\sum_{i\in[r]} q_i(\phi, \boldsymbol{\alpha}) \leqslant \frac{2\alpha_1}{r} \sum_{ij\in\binom{[r]}{2}} \log \phi_{ij}.$$

But if r is odd, then every matching $\phi^{-1}(c)$ contains $\frac{r-1}{2}$ edges, and hence the sum of all ϕ_{ij} is at most $s \cdot \frac{r-1}{2}$ (if r is even, this sum is $s \cdot \frac{r}{2}$). If then α_1 is not significantly larger than $\frac{1}{r}$, this yields a contradiction to $(r, \phi, \boldsymbol{\alpha})$ being optimal (see Lemma 4.25 for details).

In Lemma 4.27, we maximise $q(\phi, \boldsymbol{\alpha})$ under the condition that 1 is a largest part of some fixed size x. A weight shifting process shows that the contribution $q_1(\phi, \boldsymbol{\alpha})$ is maximised when we have as many parts of size x as possible and possibly one smaller part. It turns out that when x is significantly larger than $\frac{1}{r}$, $q_1(\phi, \boldsymbol{\alpha})$ is not only smaller than $g_s(r)$, but actually smaller than either $g_s(r-1)$ or $g_s(r+1)$ (see Lemma 4.28). Thus the r-part solution is not optimal (and we do better by taking an even solution with one fewer part or one more part).

We give the full proof of Theorem 4.16 in Section 4.4.4. Before we can do this, in the following subsection we will investigate which r maximise $g_s(r)$, which will allow us to assume that $r \leqslant \log(s) + 2$ later.

4.4.3 Analytic estimates

In this section, we prove some estimates concerning the functions g_s, $\tilde{g}_s$, e_s, g, R and R_2.

Lemma 4.18. *For all integers s, r satisfying $2 \leqslant r \leqslant s-1$, we have*

$$0 \leqslant e_s(r) \leqslant \frac{1}{4}\left\lfloor \frac{s}{r-1} \right\rfloor^{-2}.$$

Proof. The fact that $e_s(r) \geqslant 0$ is a consequence of the concavity of $\log(x)$ and Jensen's inequality. Indeed, if we denote $s = (r-1)z + a$ for $0 \leqslant a \leqslant r-2$, we have

$$\begin{aligned} g_s(r) &= \frac{r-1}{r}\left(\frac{r-1-a}{r-1}\log(z) + \frac{a}{r-1}\log(z+1)\right) \\ &\leqslant \frac{r-1}{r}\log\left(\frac{r-1-a}{r-1}z + \frac{a}{r-1}(z+1)\right) = \tilde{g}_s(r). \end{aligned}$$

For the upper bound, it is also easy to see that $e_s(r) = 0$ when $r-1$ divides s, while otherwise we have

$$\begin{aligned} r \cdot e_s(r) &= a\log\left(\frac{s/(r-1)}{z+1}\right) + (r-1-a)\log\left(\frac{s/(r-1)}{z}\right) \\ &= a\log\left(1 + \frac{a-(r-1)}{(r-1)(z+1)}\right) + (r-1-a)\log\left(1 + \frac{a}{(r-1)z}\right) \\ &\leqslant \frac{a(r-1-a)}{r-1}\left(\frac{1}{z} - \frac{1}{z+1}\right), \end{aligned}$$

where for the inequality we used that $\log(1+x) \leqslant x$ for $x > -1$. Since $a(r-1-a) \leqslant \frac{1}{4}(r-1)^2$, we get

$$e_s(r) \leqslant \frac{1}{4}\left(\frac{1}{z} - \frac{1}{z+1}\right) = \frac{1}{4z(z+1)} \leqslant \frac{1}{4}\left\lfloor \frac{s}{r-1} \right\rfloor^{-2},$$

as required. □

Lemma 4.19. *For all integers r, s with $2 \leqslant r < (s/2)^{1/4}$ we have*

$$g_{s+1}(r+1) - g_{s+1}(r) > g_s(r+1) - g_s(r).$$

Proof. Lemma 4.18 implies that

$$\begin{aligned} -e_{s+1}(r+1) + e_s(r+1) - e_s(r) + e_{s+1}(r) &\geqslant -e_{s+1}(r+1) - e_s(r) \\ &\geqslant -\frac{2(r-1)^2}{4(s-r+1)^2} \geqslant -\frac{r^2}{s^2}, \end{aligned}$$

since $r \leqslant s/4$ which follows from $2 \leqslant r < (s/2)^{1/4}$. Using that $\log(1+x) \geqslant x/(1+x)$ for $x > 0$, we have

$$\begin{aligned} \tilde{g}_{s+1}(r+1) - \tilde{g}_s(r+1) + \tilde{g}_s(r) - \tilde{g}_{s+1}(r) = \frac{1}{r(r+1)}\log\left(1+\frac{1}{s}\right) &\geqslant \frac{1}{r(r+1)(s+1)} \\ &\geqslant \frac{1}{2r^2 s}. \end{aligned}$$

Combining the two previous bounds, we get

$$g_{s+1}(r+1) - g_s(r+1) + g_s(r) - g_{s+1}(r) \geqslant \frac{1}{2r^2 s} - \frac{r^2}{s^2} > 0,$$

since $2r^4 < s$, completing the proof. $\square$

The *Lambert W-function* is the inverse of $f(x) = xe^x$. That is, $y = xe^x$ for $y \geqslant 0$ if and only if $x = W(y)$.

By applying log to $y = W(y)e^{W(y)}$ and using standard bounds on log we have

$$\tfrac{1}{2}\log y < \log y - \log\log y < W(y) < \log y \quad \text{for } y > e. \tag{4.4.2}$$

For $s < e^2$, let $r(s) = r_2(s) := 2$, and for $s \geqslant e^2$, let

$$r(s) := \max\{r \in \mathbb{N} : (r-1)e^r \leqslant s\} \quad \text{and} \quad r_2(s) := \max\{r \in 2\mathbb{N} : (r-1)e^r \leqslant s\}.$$

(We already defined $r_2(s)$ in Definition 1.20.)

Then for $s \geqslant e^2$, we have $r_2(s) = 2\lfloor (W(s/e)+1)/2 \rfloor$ and (4.4.2) implies that $r(s) = \lfloor W(s/e) + 1 \rfloor \leqslant \log(s)$. Thus, for all $s \geqslant 2$,

$$r_2(s) \leqslant r(s) \leqslant \max\{2, \log(s)\} \tag{4.4.3}$$

and, for $s \geqslant 92$,

$$r(s) \geqslant \log s - 1 - \log(\log(s) - 1) \geqslant \log(s)/2. \tag{4.4.4}$$

Moreover, we note that $W'(y) = \frac{W(y)}{y(1+W(y))}$ for all $y > 0$ and hence

$$W'(y) = |W'(y)| \leqslant 1/y. \tag{4.4.5}$$

The next lemma concerns the analytic properties of the function $\tilde{g}_s$.

Lemma 4.20. *The following holds for any integer $s \geqslant 2$.*

(i) *$\tilde{g}_s$ has a unique maximum at $m(s) = W(s/e) + 1$ (so for $s \geqslant e^2$ we have $r(s) = \lfloor m(s) \rfloor$). Moreover, $\tilde{g}_s$ is strictly increasing on $(1, m(s))$ and strictly decreasing on $(m(s), \infty)$, and*

$$\tilde{g}_s'(x) = \frac{\log\left(\frac{s}{x-1}\right) - x}{x^2} \quad \textit{for } x > 1,$$

(ii) *$\tilde{g}(s) = W(s/e)$,*

(iii) $\tilde{g}_s(m(s)+a) - \tilde{g}_s(m(s)+b) \geqslant \frac{1}{16}(\log(s)+5/2)^{-2}$ *whenever* $s \geqslant e^2$; $m(s)+a, m(s)+b > 0$; $ab \geqslant 0$; $|a| < 2$ *and* $|a| + 1/2 \leqslant |b|$,

(iv) $\tilde{g}_s(m(s)) - \tilde{g}_s(m(s)+b) \leqslant 8|b|^2(\log(s)-4)^{-2}$ *whenever* $|b| \leqslant \min\{2, m(s)-2\}$ *and* $s \geqslant 55$.

Proof. To prove (i) and (ii), note that

$$\tilde{g}_s'(x) = \frac{\log\left(\frac{s}{x-1}\right) - x}{x^2},$$

and letting $m(s)$ be the unique solution to $\tilde{g}_s'(m(s)) = 0$, we have

$$s/e = (m(s)-1)e^{m(s)-1}$$

and hence

$$m(s) = W(s/e) + 1.$$

We also see that $\tilde{g}_s'(x)$ is positive for $x \in (1, m(s))$ and negative for $x \in (m(s), \infty)$, so $\tilde{g}_s$ is increasing on $(1, m(s))$ and decreasing on $(m(s), \infty)$ with $m(s)$ being the unique maximum. This proves (i).

Noting that $m(s) - 1 = s/e^{m(s)}$, we have

$$\tilde{g}(s) = \tilde{g}_s(m(s)) = \frac{m(s)-1}{m(s)} \log\left(\frac{s}{s/e^{m(s)}}\right) = m(s) - 1 = W(s/e),$$

proving (ii).

For part (iii), we only prove the case where $a, b \geqslant 0$, since the other case can be proved analogously (in fact, it is slightly easier). We will show that $\tilde{g}_s(m(s)+a') - \tilde{g}_s(m(s)+b') \geqslant \frac{1}{16}(\log(s)+5/2)^{-2}$ where $a' = a + 1/4$ and $b' = \min\{b, 5/2\}$, which implies (iii) since $a < a' < b' \leqslant b$ and $\tilde{g}_s$ is monotone decreasing on $[m(s), \infty)$. By the mean value theorem there is $x \in (m(s)+a', m(s)+b')$ such that

$$\begin{aligned}
\tilde{g}_s(m(s)+a') - \tilde{g}_s(m(s)+b') &= (b'-a')\frac{x - \log(\frac{s}{x-1})}{x^2} \\
&= (b'-a')\frac{-\log(\frac{s}{x-1}) + \log(\frac{s}{m(s)-1}) + x - m(s)}{x^2} \\
&= (b'-a')\frac{\log(\frac{x-1}{m(s)-1}) + x - m(s)}{x^2} \\
&\geqslant (b'-a')\frac{a'}{x^2} \geqslant \frac{1}{16(\log(s)+5/2)^2}
\end{aligned}$$

using $a' \geqslant 1/4$, $b' - a' \geqslant 1/4$, $b' \leqslant 5/2$ and the fact that $x \leqslant m(s) + b' \leqslant \log(s) + 5/2$ which follows from $\log(s) \geqslant 2$, (4.4.3) and part (i).

For part (iv), we only cover the case $b < 0$ which is trickier. Again by the mean value theorem there is some $x \in (m(s)+b, m(s))$ such that $\tilde{g}_s(m(s)) - \tilde{g}_s(m(s)+b) = |b|\tilde{g}'_s(x)$. Using that $-b = |b| \leqslant m(s) - 2 < m(s) - 1$ and $\log(1+y) \geqslant \frac{y}{1+y}$ for $y > -1$, we have

$$\left|\log\left(\frac{x-1}{m(s)-1}\right)\right| \leqslant \left|\log\left(\frac{m(s)+b-1}{m(s)-1}\right)\right| \leqslant \frac{|b|}{m(s)+b-1} \leqslant |b|.$$

Thus

$$\begin{aligned}\tilde{g}_s(m(s)) - \tilde{g}_s(m(s)+b) = |b|\frac{\log(\frac{m(s)-1}{x-1}) + m(s) - x}{x^2} &\leqslant |b|\frac{2|b|}{(m(s)-2)^2} \\ &\leqslant \frac{8|b|^2}{(\log(s)-4)^2},\end{aligned}$$

where we used that $m(s) - 2 \geqslant \log(s)/2 - 2 > 0$ by part (i), (4.4.2) and our lower bound $s \geqslant 55$. □

Proof of Lemma 4.22. Lemma 4.21(i) implies that we have $g(s) = g_s(r)$ for some $r \in \{r_2(s), r_2(s)+2\}$. Also $\tilde{g}_s(m(s)) = W(s/e)$ is the unique maximum of $\tilde{g}_s$ and $|r - m(s)| \leqslant 2$ by definition of $r_2(s)$. From Lemma 4.18 we have

$$W(s/e) - g(s) = \tilde{g}_s(m(s)) - g_s(r) \geqslant \tilde{g}_s(m(s)) - \tilde{g}_s(r) \geqslant 0.$$

From Lemma 4.18 and (4.4.3) we have

$$0 \leqslant e_s(r) \leqslant \frac{1}{4}\left(\frac{s}{\log(s)+1} - 1\right)^{-2} \leqslant \frac{(\log(s))^2}{s^2},$$

which yields

$$W(s/e) - g(s) = \tilde{g}_s(m(s)) - (\tilde{g}_s(r) - e_s(r)) \leqslant \frac{32}{(\log(s)-4)^2} + \frac{(\log(s))^2}{s^2} \leqslant \frac{600}{(\log(s))^2}.$$

where to bound the term $\tilde{g}_s(m(s)) - \tilde{g}_s(r)$ we used Lemma 4.20(iv) with $b = r - m(s)$. The upper and lower bounds on $W(s/e) - g(s)$ give the first part of the lemma. Since the upper bound tends to 0 as $s \to \infty$, and $e^{W(s/e)} = (s/e)/W(s/e)$, the second assertion follows. □

We are now ready to give a precise description of $R(s)$ and $R_2(s)$. While parts (i) and (ii) are used in later subsections to prove Theorem 1.21, part (iii) is used in the proof of Corollary 1.22 as well as providing additional structural information that tells us for fixed r what the sets $S(r) := \{s : r \in R(s)\}$ and $S_2(r) := \{s : r \in R_2(s)\}$ look like.

Lemma 4.21. *Let $s \geqslant 2$ be an integer. Then*

(i) $R(s) \subseteq \{r(s), r(s)+1\}$ *and* $R_2(s) \subseteq \{r_2(s), r_2(s)+2\}$. *In particular, we have* $\max R(s), \max R_2(s) \leqslant \log(s)+2$.

(ii) $g(s) > \min\{g_s(p), g_s(q)\}$ *for any two distinct odd integers* p, q.

(iii) *There are increasing sequences* $(s_2, s_3, s_4, \ldots)$ *and* $(\tilde{s}_2, \tilde{s}_4, \tilde{s}_6, \ldots)$ *with*

$$2 = s_2 < e^2 < s_3 < 2e^3 < \ldots < (r-2)e^{r-1} < s_r < (r-1)e^r < s_{r+1} < re^{r+1} < \ldots$$

and $s_{r-1} \leqslant \tilde{s}_r \leqslant s_r$ *such that the set* $S(r) := \{s : \ r \in R(s)\}$ *is either the interval* $[s_r, s_{r+1}-1]$ *or* $[s_r, s_{r+1}]$ *for all* $r \in \mathbb{N}$; *and the set* $S_2(r) := \{s : \ r \in R_2(s)\}$ *is either the interval* $[\tilde{s}_r, \tilde{s}_{r+2}-1]$ *or* $[\tilde{s}_r, \tilde{s}_{r+2}]$ *for all* $r \in 2\mathbb{N}$. *Moreover, for all* $r \in 2\mathbb{N}$ *there exists* $s \in \mathbb{N}$ *with* $R_2(s) = \{r\}$.

(iv) *The values in Table 1.2 and Table 4.1 hold.*

Before proving the lemma, we make some remarks on (ii). If $g_s(r)$ is maximised by an even r, then part (ii) follows from (i) since then $R(s)$ contains an even r and at most one of p, q. However, if $g_s(r)$ is only maximised by an odd r, then we could have $R(s) = \{p\}$ and $g_s(r) < g_s(p)$ where $r \in R_2(s)$, and must show $g_s(r) > g_s(q)$ for every odd $q \neq p$.

Proof. We will prove these assertions analytically for large s and computationally for small s. First we prove parts (i) and (ii). Suppose first that $s < 196$. We compute all values $g_s(r)$ for $r < s$ which determines $R(s), R_2(s)$ and verifies parts (i) and (ii). A script for this calculation is in the ancillary file `optr.py`.

Thus we may suppose that $s \geqslant 196$. We claim that for distinct integers $r_a = m(s)+a, r_b = m(s)+b \geqslant 2$ with $ab \geqslant 0$, $|a| < 2$ and $|a| < |b|$, we have $g_s(r_a) > g_s(r_b)$. To see this, note that Lemma 4.20(iii) gives $\tilde{g}_s(r_a) - \tilde{g}_s(r_b) \geqslant 1/(16(\log(s)+5/2)^2)$. Equation (4.4.3) shows that we have $r_a \leqslant \log(s)+2$. Thus Lemma 4.18 gives that

$$e_s(r_a) \leqslant \frac{1}{4\lfloor s/(r_a-1)\rfloor^2} \leqslant \frac{(r_a-1)^2}{4(s-r_a+1)^2} \leqslant \frac{(\log(s)+1)^2}{s^2}. \tag{4.4.6}$$

Putting the above observations together, we get

$$g_s(r_a) - g_s(r_b) \geqslant \tilde{g}_s(r_a) - \tilde{g}_s(r_b) - e_s(r_a) \geqslant \frac{1}{16(\log(s)+5/2)^2} - \frac{(\log(s)+1)^2}{s^2} > 0,$$

where the last inequality holds for $s \geqslant 196$. This proves the claim.

Parts (i) and (ii) follow easily from the claim. Indeed, to prove that $R(s) \subseteq \{r(s), r(s)+1\}$, it suffices to show that for all integers $r, r' \geqslant 2$ with $r < r(s) < r(s)+1 < r'$, we have $g_s(r(s)) > g_s(r)$ and $g_s(r(s)+1) > g_s(r')$. Given such r, let $r_a = r(s)$, $r_b = r$. Since $r(s) \leqslant m(s)$, a and b are both non-positive and the other conditions are also easy to verify. Given such r', the second inequality follows similarly since $r(s)+1 \geqslant m(s)$.

For the statement regarding $R_2(s)$, the claim implies that $g_s(r(s)+2) > g_s(r')$ when $r' > r(s)+2$.

For part (ii), suppose that p, q are both odd integers and note that there is an even integer r that satisfies $|r - m(s)| < 2$ and lies between $m(s)$ and, without loss of generality, p. By our claim applied with $r = r_a$ and $p = r_b$ we have that $g(s) \geqslant g_s(r) > g_s(p) \geqslant \min\{g_s(p), g_s(q)\}$ as desired.

Next, we prove part (iii). First we will consider $S^0(r) := S(r) \cap [2, 90000]$ and $S_2^0(r) := S_2(r) \cap [2, 90000]$. Suppose that $s < 90000$. By part (i) we can find $R(s)$ and $R_2(s)$ by computing $r(s)$ and $r_2(s)$ as approximations of $W(s/e)$ and comparing the two possible values of g_s. Then we can manually check that $S^0(2) = [2, 16], \ldots, S^0(8) = [11444, 36023]$, $S^0(9) = [36024, 90000]$ and $S_2^0(2) = [2, 27], \ldots, S_2^0(8) = [5857, 59470]$, $S_2^0(10) = [59471, 90000]$ are as shown in Table 1.2 and Table 4.1. The two tables contain the optimal r for $s \leqslant 10^7$ (for $R_2(s)$) and $s \leqslant 90000$ (for $R(s)$) for illustrative purposes. A script for this calculation is provided in the ancillary file `optr.py`.

Suppose now that $s \geqslant 90000$. We first show that $\min\{R(s+1)\} \geqslant \max\{R(s)\}$ and $\min\{R_2(s+1)\} \geqslant \max\{R_2(s)\}$. Let $r^* = \max\{R(s)\}$ and $r_2^* = \max\{R_2(s)\}$. By (4.4.3) and part (i) of this lemma, we have $r^*, r_2^* \leqslant r(s) + 2 \leqslant \log(s) + 2 < (s/2)^{1/4} + 1$, so Lemma 4.19 implies that for $1 \leqslant p < \max\{r^*, r_2^*\}$ we have $g_{s+1}(p+1) - g_{s+1}(p) > g_s(p+1) - g_s(p)$. For any even r with $r < r_2^*$, summing these equations over $r \leqslant p \leqslant r_2^* - 1$ yields

$$g_{s+1}(r_2^*) - g_{s+1}(r) > g_s(r_2^*) - g_s(r) \geqslant 0.$$

That is, for all even $r < r_2^*(s)$ we have $r \notin R_2(s+1)$. So $\min\{R_2(s+1)\} \geqslant \max\{R_2(s)\}$. Summing the equations over $r \leqslant p \leqslant r^*$, it can be proved analogously that $\min\{R(s+1)\} \geqslant \max\{R(s)\}$.

Combined with the $S^0(r)$ and $S_2^0(r)$ obtained above, this yields the claimed interval structure of $S(r)$ and $S_2(r)$ respectively and, defining s_r and $\tilde{s}_r$ as the minima of $S(r)$ and $S_2(r)$ respectively, it is easy to see that $s_{r-1} \leqslant \tilde{s}_r \leqslant s_r$. It now remains to show that for each integer $r \geqslant 2$, we have $s_r < s^* < s_{r+1}$ where $s^* := (r-1)e^r \geqslant 90000$ and for all $r \in \mathbb{N}$, there is $s \in \mathbb{N}$ with $R(s) = \{r\}$ (we can see that this holds for $r \in [2, 8]$, for which $s^* < 90000$). For this, it suffices to prove that $R(\lfloor s^* \rfloor) = \{r\}$. By definition, $r = W(s^*/e) + 1$. Let s satisfy $|s - s^*| \leqslant 1$. Due to (4.4.5) we have $|W'(x)| \leqslant 1/x$, so the mean value theorem gives $|m(s) - r| = |W(s/e) - W(s^*/e)| < 1/(s^* - 1)$. Then, by Lemma 4.20(iv) applied with $b = r - m(s)$,

$$\tilde{g}_s(m(s)) - \tilde{g}_s(r) \leqslant \frac{8}{(\log(s) - 4)^2 (s^* - 1)^2} \leqslant \frac{1}{s}.$$

On the other hand, for any integer $r' \neq r$ we have

$$|r' - m(s)| \geqslant |r' - r| - |r - m(s)| \geqslant 1 - 1/(s^* - 1) \geqslant 1/2$$

and thus, by Lemma 4.20(iii), $\tilde{g}_s(m(s)) - \tilde{g}_s(r') \geqslant \frac{1}{16}(\log(s) + 5/2)^{-2}$. Using (4.4.6) with r in place of r_a (which holds since $|m(s) - r| < 2$) and putting everything together, we get

$$\begin{aligned} g_s(r') \leqslant \tilde{g}_s(r') \leqslant \tilde{g}_s(m(s)) - \frac{1}{16(\log(s) + 5/2)^2} &\leqslant \tilde{g}_s(r) + \frac{1}{s} - \frac{1}{16(\log(s) + 5/2)^2} \\ \leqslant g_s(r) + \frac{(\log(s) + 1)^2}{s^2} + \frac{1}{s^* - 1} - \frac{1}{16(\log(s) + 5/2)^2} &< g_s(r), \end{aligned}$$

as desired. (This argument also shows that $R(\lfloor s^* \rfloor + 1) = \{r\}$.)

Finally, for (iv), we compute $g_s(r)$ for $r = r(s), r(s) + 1, r_2(s), r_2(s) + 2$ which, by (i), determines $R(s)$ and $R_2(s)$. See the ancillary file `optr.py`. □

The final result of this section provides asymptotics for $g(s)$. Its proof is similar to the previous one.

Lemma 4.22. *For all integers $s \geqslant 200$ we have*

$$0 \leqslant W(s/e) - g(s) \leqslant \frac{600}{(\log(s))^2}, \quad \textit{and} \quad \frac{e^{g(s)}}{(s/e)/W(s/e)} \to 1 \textit{ as } s \to \infty.$$

4.4.4 Proof of Theorem 4.16

In this section, we prove Theorem 4.16. We first need a simple lemma about maximising sums of logarithms.

Lemma 4.23. *Suppose that $a_1, \ldots, a_n$ are positive with $a_1 + \ldots + a_n = a$. Then, for any $x > 0$ and positive $x_1, \ldots, x_n$ with $x_1 + \ldots + x_n \leqslant x$, we have*

$$\sum_{i=1}^{n} a_i \log(x_i) \leqslant \sum_{i=1}^{n} a_i \log\left(\frac{x a_i}{a}\right).$$

Moreover, if $a_1 = a_2 = \ldots = a_n$, and the x_i are constrained to be integers, then the maximum is attained whenever $|x_i - x_j| \in \{0, 1\}$ for all $i, j \in [n]$.

Proof. We use the weighted AM-GM inequality. It implies that

$$x = \sum_{i \in [n]} \frac{a_i}{a} \cdot \frac{a x_i}{a_i} \geqslant \prod_{i \in [n]} \left(\frac{a x_i}{a_i} \right)^{a_i/a}$$

and therefore

$$\prod_{i \in [n]} x_i^{a_i} \leqslant x^a \prod_{i \in [n]} \left(\frac{a_i}{a} \right)^{a_i} = \prod_{i \in [n]} \left(\frac{x a_i}{a_i} \right)^{a_i}.$$

Taking logs (noting that all terms in both products are positive by our assumption) completes the proof of the first part.

Suppose now that we have positive integers $x_1, \ldots, x_n$ whose sum is x. It suffices to show that the maximum product $x_1 \ldots x_n$ is attained whenever $|x_i - x_j| \in \{0, 1\}$ for all $i, j \in [n]$. Suppose that $x_2 - x_1 \geqslant 2$. Let $y_1 = x_1 + 1$, $y_2 = x_2 - 1$, and $y_i = x_i$ for all $i \in [3, n]$. Then $y_1, \ldots, y_n$ are positive integers whose sum is x and

$$y_1 \cdots y_n = (x_1 x_2 + x_2 - x_1 - 1) x_3 \cdots x_n \geqslant x_1 \cdots x_n + x_3 \cdots x_n > x_1 \cdots x_n.$$

Thus if $x_1, \ldots, x_n$ does not satisfy the condition, then we can always increase the product. Finally, there is a unique multiset $\{x_1, \ldots, x_n\}$ for which $x_1, \ldots, x_n$ satisfies the condition, so any such $x_1, \ldots, x_n$ must attain the maximum. This completes the proof of the second part. □

The following lemma bounds $Q(s)$ by $g_s(r)$. This essentially proves Theorem 4.16 when all optimal solutions have an even number of parts (which we will eventually show is always the case), since for even r we have $g_s(r) \leqslant g(s)$.

Lemma 4.24. *Let $s \geqslant 2$. For any $(r, \phi, \boldsymbol{\alpha}) \in \text{OPT}^*(s)$ we have*

$$q(\phi, \boldsymbol{\alpha}) \leqslant g_s(r). \tag{4.4.7}$$

Moreover,

(i) *when r is odd, the inequality is strict;*

(ii) *when $r \in R_2(s)$ we have equality if and only if $\boldsymbol{\alpha}$ and ϕ are uniform.*

Proof. Suppose that $(r, \phi, \boldsymbol{\alpha}) \in \textsc{opt}^*(s)$. Then by Lemma 4.3, for all $i \in [r]$ we have that $q_i(\phi, \boldsymbol{\alpha}) = q(\phi, \boldsymbol{\alpha}) = Q(s)$. By Lemma 4.5(i), each $\phi^{-1}(c)$ is a matching and $\sum_{j \neq i} \phi_{ij} \leqslant s$. By Lemma 4.23 applied with $x_j = \phi_{ij}$, $a_j = 1$ and $x = s$ we have that $\sum_{j \neq i} \log \phi_{ij} \leqslant r g_s(r)$ with equality if and only if ϕ is uniform at i and i sees every colour. So, averaging over contributions, we have

$$\begin{aligned} q(\phi, \boldsymbol{\alpha}) = \frac{1}{r} \sum_{i \in [r]} q_i(\phi, \boldsymbol{\alpha}) &= \sum_i \frac{1}{r} \sum_{j \neq i} \alpha_j \log \phi_{ij} \\ &= \frac{1}{r} \sum_j \alpha_j \sum_{i \neq j} \log \phi_{ij} \leqslant \sum_j \alpha_j g_s(r) = g_s(r), \end{aligned} \tag{4.4.8}$$

proving the first part of the lemma. For part (ii), note that if r is odd, each $\phi^{-1}(c)$ is a matching missing at least one vertex, so there is some $i \in [r]$ with $\sum_{j \neq i} \phi_{ij} < s$. The average contribution is therefore strictly less than $g_s(r)$. On the other hand, if r is even and $\boldsymbol{\alpha}$ and ϕ are uniform, we have $q(\phi, \boldsymbol{\alpha}) = g_s(r)$, so it remains to prove the 'only if' direction of part (ii) for $r \in R_2(s)$.

Suppose that we have equality throughout (4.4.8) above. Then ϕ must be uniform by our application of Lemma 4.23. Moreover, if $r = 2$, we have $\phi_{12} = [s]$ and optimising over $\boldsymbol{\alpha}$ yields that $\boldsymbol{\alpha}$ must be uniform. So from now on, we may assume that $r \geqslant 4$ and that ϕ is uniform.

Suppose, without loss of generality, that $\alpha_1 = \max_i \alpha_i$ and let a be such that $0 \leqslant a < r - 1$ with $a \equiv s \bmod r - 1$. The fact that ϕ is uniform at 1 means that $\phi_{1i} = \lfloor s/(r-1) \rfloor =: z$ for $r - 1 - a$ values of i and $\phi_{1i} = z + 1$ for the remaining a values of i. Thus, without loss of generality,

$$\begin{aligned} q(\phi, \boldsymbol{\alpha}) = q_1(\phi, \boldsymbol{\alpha}) &= (\alpha_2 + \ldots + \alpha_{a+1}) \log(z+1) + (\alpha_{a+2} + \ldots + \alpha_r) \log(z) \\ &= (\alpha_2 + \ldots + \alpha_{a+1})(\log(z+1) - \log(z)) + (1 - \alpha_1) \log(z) \\ &\leqslant a\alpha_1(\log(z+1) - \log(z)) + (1 - \alpha_1) \log(z) =: p(\alpha_1). \end{aligned}$$

Note that $p(1/r) = g_s(r)$. We now show that p is strictly monotone decreasing. Using $a \leqslant r - 1$ and $z \geqslant (s - r + 1)/(r - 1)$, we obtain

$$\begin{aligned} p'(\alpha_1) = a \log\left(1 + \frac{1}{z}\right) - \log(z) &\leqslant a \log\left(1 + \frac{r-1}{s-r+1}\right) - \log\left(\frac{s-r+1}{r-1}\right) \\ &\leqslant \frac{(r-1)^2}{s-r+1} - \log\left(\frac{s-r+1}{r-1}\right). \end{aligned}$$

Since $r \geqslant 4$ we have $s \geqslant 27$ (recall Lemma 4.21(ii) and refer to Table 1.2) and, using Lemma 4.21(i), this yields $r - 1 \leqslant \log(s) + 1 \leqslant \sqrt{s} < s/2$. We can thus bound

$$p'(\alpha_1) \leqslant \frac{s}{s-r+1} - \log\left(\frac{s-r+1}{\sqrt{s}}\right) \leqslant 1 + \frac{2}{\sqrt{s}} - \log(\sqrt{s}-1) < 0$$

for $s \geqslant 27$, so we have shown that p is indeed strictly monotone decreasing.

To finish the proof, note that since p is strictly monotone decreasing and $\alpha_1 = \max_i \alpha_i \geqslant 1/r$, we have $q(\phi, \boldsymbol{\alpha}) \leqslant p(\alpha_1) \leqslant p(1/r) = g_s(r)$. We have equality throughout only if $\alpha_1 = 1/r$ which implies $\alpha_i = 1/r$ for all i, as required. □

The following lemmas help us to show that optimal solutions cannot have an odd number of parts. The first lemma shows that, if there were a basic optimal solution on an odd number of parts, the parts cannot be close to uniform, since the largest part must be significantly larger than $\frac{1}{r}$.

Lemma 4.25. *Let $s \geqslant 2$ be an integer. Let $(r, \phi, \boldsymbol{\alpha}) \in \mathrm{OPT}^*(s)$ and suppose that r is odd and $\alpha_1 = \max_i \alpha_i$. Then*

$$\alpha_1 \geqslant \frac{r}{r^2-1} - \frac{e_s(r+1)}{(r-1)\log(s/r)}.$$

Proof. Since $(r, \phi, \boldsymbol{\alpha})$ is optimal with no zero parts, we have that $Q(s) = q_i(\phi, \boldsymbol{\alpha})$ for all $i \in [r]$ by Lemma 4.3, so, averaging over contributions, we have

$$Q(s) = \frac{1}{r}\sum_{i\in[r]} q_i(\phi, \boldsymbol{\alpha}) = \frac{1}{r}\sum_{i\in[r]}\sum_{j\neq i} \alpha_j \log \phi_{ij} \leqslant \frac{2\alpha_1}{r} \sum_{ij\in\binom{[r]}{2}} \log \phi_{ij}.$$

Now, $\sum_{ij\in\binom{[r]}{2}} \phi_{ij} = \sum_{c\in[s]} |\phi^{-1}(c)| = s \cdot \frac{r-1}{2}$ as by Lemma 4.5(ii) each colour graph $\phi^{-1}(c)$ is a maximal matching. Thus, by concavity of the log function,

$$Q(s) \leqslant \frac{2\alpha_1}{r} \cdot \binom{r}{2} \log\left(\frac{\sum_{ij}\phi_{ij}}{\binom{r}{2}}\right) = \alpha_1(r-1)\log\left(\frac{s}{r}\right).$$

Since r is odd, and by Lemma 4.14, we have that $Q(s) \geqslant g(s) \geqslant g_s(r+1) = \tilde{g}_s(r+1) - e_s(r+1) = \frac{r}{r+1}\log\left(\frac{s}{r}\right) - e_s(r+1)$ and after rearrangement we obtain the required inequality. □

The next main lemma bounds from above the contribution of a largest part in any feasible solution in terms of the function f_s as defined below. We recall that, by Lemma 4.3, in a basic optimal solution, every part has contribution equal to $Q(s)$.

For integers $s, r \geqslant 2$ and $\frac{1}{r} \leqslant x < \frac{1}{r-1}$, let

$$f_{s,r}(x) = \max_{0 \leqslant t \leqslant s} f_{s,r,t}(x) \quad \text{where} \quad f_{s,r,t}(x) := (r-1)x g_{s-t}(r-1) + (1-(r-1)x) \log t$$

and

$$\tilde{f}_{s,r}(x) := (r-2)x \log\left(\frac{xs}{1-x}\right) + (1-(r-1)x) \log\left(\frac{(1-(r-1)x)s}{1-x}\right).$$

Let also $\tilde{f}_s$ be defined by setting $\tilde{f}_s(x) := \tilde{f}_{s,r}(x)$ for $\frac{1}{r} \leqslant x < \frac{1}{r-1}$ and similarly define f_s. See Figure 4.4.2 for a plot of $\tilde{f}_s$.

The function $f_{s,r,t}(x)$ is the contribution of a part of size x, say part 1, in a solution with $r-1$ parts of size x and one part, say part r, of size at most x and where $|\phi(1r)| = t$ while the other multiplicities $|\phi(1j)|$ are as equal as possible. The function $f_{s,r}(x)$ maximises this contribution in a solution of this form. The function $\tilde{f}_{s,r}(x)$ measures the same contribution, but, like $\tilde{g}_s(r)$, disregards the fact that the multiplicities $|\phi(1j)|$ have to be integral, which simplifies the optimisation of t (the optimal t is $\frac{(1-(r-1)x)s}{1-x}$). We start by giving some properties of f_s and $\tilde{f}_s$ that help us compare them with g_s. Ultimately, these comparisons will come together with the subsequent Lemma 4.27 to show that the objective function of a basic optimal solution on an odd number of parts can be bounded in terms of $g(s)$.

Lemma 4.26. *Let $s \geqslant 2$ be an integer.*

(i) *For all integers $r \geqslant 2$, on the domain $[\frac{1}{r}, \frac{1}{r-1}]$, $\tilde{f}_s$ and f_s are convex.*

(ii) *$f_{s-\ell}(x) < f_s(x) \leqslant \tilde{f}_s(x)$ for all $x \in (0,1)$ and integers $0 < \ell \leqslant s-2$.*

(iii) *$f_s(x) \leqslant \max\{g_s(r), g_s(r-1)\}$ for all $x \in (0,1)$, where r is the integer with $\frac{1}{r} \leqslant x < \frac{1}{r-1}$.*

Proof. For (i), we have $\tilde{f}''_{s,r}(x) = \frac{r-2}{(x-1)x((r-1)x-1)}$ which is positive for $0 < x < 1/(r-1)$, so $\tilde{f}_{s,r}$ is convex. Observe that $f_{s,r,t}(x)$ is a linear function of x, so $f_{s,r}$ is a maximum of convex functions and hence is convex.

For the first inequality in (ii), it suffices to show that $g_{s_1}(r-1) < g_{s_2}(r-1)$ for all $s_1 < s_2$ and $r \geqslant 2$. Indeed, if so, we have $f_{s-\ell,r,t}(x) < f_{s,r,t}(x)$ for all r, all $0 \leqslant t \leqslant s-\ell$ and all $\frac{1}{r} \leqslant x < \frac{1}{r-1}$, and hence $f_{s-\ell}(x) < f_s(x)$ for all $x \in (0,1)$. Now, $g_{s_1}(r-1)$ is the average of a multiset of $r-2$ numbers $\{a_1, \ldots, a_{r-2}\}$ while $g_{s_2}(r-1)$ is the average of a multiset of $r-2$ numbers $\{b_1, \ldots, b_{r-2}\}$ where $a_i \leqslant b_i$ for all $i \in [r]$ and at least one of these is strict. This follows since each multiset contains at most two different values and the sum s_1 of the former is strictly less than s_2, the sum of the latter. Thus $g_{s_1}(r-1) < g_{s_2}(r-1)$.

The second part of (ii) follows from the concavity of log. Indeed, let $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{r-1}) \in \mathbb{R}^{r-1}_{\geqslant 0}$ with $\rho_1 + \ldots + \rho_{r-1} \leqslant s$, and let

$$q_{s,r,x}(\boldsymbol{\rho}) := x(\log \rho_1 + \ldots + \log \rho_{r-2}) + (1-(r-1)x) \log \rho_{r-1}.$$

Then, subject to every $\rho_i \in \mathbb{N}$ and $\rho_{r-1} = t \leqslant s$, $q_{s,r,x}(\boldsymbol{\rho})$ is maximised by taking $\rho_1, \ldots, \rho_{r-2}$ as equal as possible, by the second part of Lemma 4.23. This maximum is precisely $f_{s,r,t}(x)$. Thus the maximum of $q_{s,r,x}(\boldsymbol{\rho})$ subject to all $\rho_i \in \mathbb{N}$ is $f_{s,r}(x)$. By the first part of Lemma 4.23, the maximum of $q_{s,r,x}(\boldsymbol{\rho})$ with no additional constraints is $\tilde{f}_{s,r}(x)$. Thus $f_s \leqslant \tilde{f}_s$.

Finally, to prove (iii), since $f_{s,r}$ is convex by part (i), any maxima lie at the extreme points of its domain, so $f_{s,r}(x) \leqslant \max\{f_{s,r}(\frac{1}{r}), f_{s,r}(\frac{1}{r-1})\}$, and $f_{s,r}(\frac{1}{r}) = g_s(r)$ and $f_{s,r}(\frac{1}{r-1}) = g_s(r-1)$. □

Lemma 4.27. *Let $s \geqslant 2$ be an integer and let $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_2(s)$ with $\alpha_1 = \max_i \alpha_i$. Then $q_1(\phi, \boldsymbol{\alpha}) \leqslant f_s(\alpha_1)$ with equality only if $\frac{1}{r} \leqslant \alpha_1 < \frac{1}{r-1}$.*

Proof. Let $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}_2(s)$ with $x = \alpha_1 = \max_i \alpha_i$ be given. Without loss of generality we may further assume that $\phi_{12} \geqslant \ldots \geqslant \phi_{1r}$. Let $r' \geqslant 2$ be the unique integer with $\frac{1}{r'} \leqslant x < \frac{1}{r'-1}$, so $r' \geqslant r$, and let $y := 1-(r'-1)x$. Setting

$$\boldsymbol{\beta} := (\underbrace{x, \ldots, x}_{r'-1}, y, 0, \ldots, 0) \in \Delta^r,$$

we get $q_1(\phi, \boldsymbol{\alpha}) \leqslant q_1(\phi, \boldsymbol{\beta})$. To see this, note that due to the assumed monotonicity of the ϕ_{1i}, moving weight from α_j to α_i for $2 \leqslant i < j \leqslant r$ does not decrease $q_1(\phi, \boldsymbol{\alpha})$. Since all parts have size at most $\alpha_1 = x$ at the start, we can move weight in this way until we reach $\boldsymbol{\beta}$.

Lemma 4.5(i) implies that $\sum_{i \in [2,r]} \phi_{1i} \leqslant s$. Fixing $\boldsymbol{\beta}$, $t := \phi_{1r'}$ and $\ell := \sum_{i \in [r'+1,r]} \phi_{1i}$, Lemma 4.23 implies that $\sum_{i \in [2,r'-1]} \beta_i \log(\phi_{1i})$ is maximised when the $\phi_{12}, \ldots, \phi_{1,r'-1}$ are as equal as possible, with sum $s-t-\ell$. This yields that

$$\begin{aligned} q_1(\phi, \boldsymbol{\beta}) &= y \log t + x \sum_{i \in [2,r'-1]} \log(\phi_{1i}) \leqslant y \log t + (r'-1)x g_{s-t-\ell}(r'-1) \\ &= f_{s-\ell,r',t}(x) \leqslant f_{s-\ell,r'}(x) \leqslant f_{s,r'}(x) = f_s(x). \end{aligned}$$

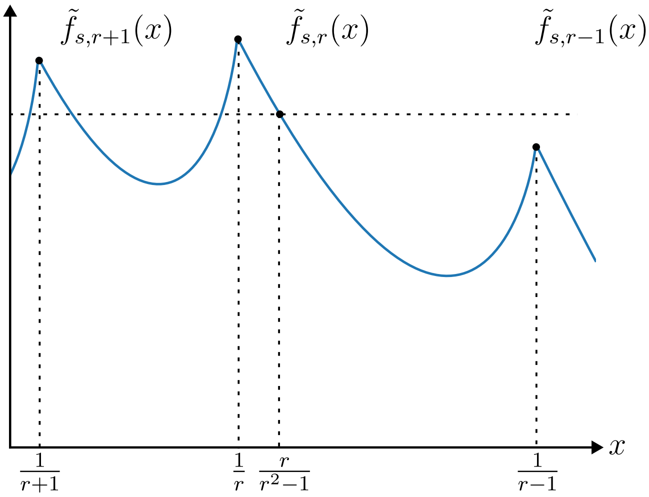


Figure 4.4.2: Plot of $\tilde{f}_s(x)$ illustrating the hypothesis of Lemma 4.28 for values $s = 800$ and $r = 5$ (noting that $\tilde{f}_s(x)$ and $f_s(x)$ are very close to each other).

where we used Lemma 4.26(ii) for the last inequality.

When $\alpha_1 \geqslant 1/(r-1)$, we have $r' < r$, so $\ell \geqslant \phi_{1r} \geqslant 2$ (and, e.g. $s - \ell \geqslant \phi_{12} \geqslant 2$), so Lemma 4.26(ii) gives the strict inequality $f_{s-\ell,r'}(x) < f_{s,r'}(x)$ in the above calculation, as required. □

The final main lemma in this section shows, roughly speaking, in a feasible solution on r parts, that if there is a part at least almost as large as $\frac{r}{r^2-1}$, this solution cannot be optimal since it is beaten by one on $r-1$ or $r+1$ parts. As an illustration, for a fixed set of parameters, we refer the reader to Figure 4.4.2. Combined with Lemma 4.25 which says that there is always such an inflated part in an optimal solution on r parts when r is odd, this will imply that there cannot be an optimal solution with an odd number of parts.

Lemma 4.28. *Let $s \geqslant 2$ be an integer. Then, for all $r \in R(s)$ with $r \geqslant 3$, we have*

$$f_{s,r}(x) < \max\{g_s(r-1), g_s(r+1)\} \quad \textit{for all} \quad \frac{r}{r^2-1} - \frac{e_s(r+1)}{(r-1)\log(s/r)} \leqslant x < \frac{1}{r-1}.$$

Proof. For $s \leqslant 1100$ we verify the lemma directly. A script for this calculation is provided in the ancillary file `small_s.py`. For the rest of the proof, we assume that $s > 1100$. Let $r \in R(s)$. By Lemma 4.21(iv) we have $r \geqslant 6$. Moreover, by

Lemma 4.21(i), (4.4.3) and (4.4.4), we have $\log(s)/2 \leqslant r \leqslant \log(s)+1$. Let

$$x_0 := \frac{r}{r^2-1} \quad \text{and} \quad \varepsilon := \frac{e_s(r+1)}{(r-1)\log(s/r)}.$$

By Lemma 4.18 and our bounds on r and s we have

$$1/\varepsilon \geqslant 4\left\lfloor \frac{s}{r-1} \right\rfloor^2 (r-1)\log(s/r) \geqslant 4 \cdot \frac{1}{2}\left(\frac{s}{r}\right)^2 \frac{5r}{6} \cdot 4 = \frac{20s^2}{3r} \tag{4.4.9}$$

and so

$$x_0 - \varepsilon \geqslant \frac{r}{r^2-1} - \frac{3r}{20s^2} > \frac{1}{r}. \tag{4.4.10}$$

To prove the lemma, we need to show that

$$f_{s,r}(x) < \max\{g_s(r-1), g_s(r+1)\} \quad \text{for all} \quad x \in [x_0 - \varepsilon, \tfrac{1}{r-1}). \tag{4.4.11}$$

Claim 4.29. *It suffices to show that*

$$\tilde{f}_{s,r}(x_0) \leqslant \max\{\tilde{g}_s(r-1), \tilde{g}_s(r+1)\} - 1/(5r^2). \tag{4.4.12}$$

Proof of claim. First we will show that (4.4.11) is implied by

$$\tilde{f}_{s,r}(x_0 - \varepsilon) < \max\{g_s(r-1), g_s(r+1)\}. \tag{4.4.13}$$

To see this, let $x \in [x_0 - \varepsilon, \frac{1}{r-1})$. Then (4.4.10) implies that there is $\lambda \in [0,1)$ such that $x = (1-\lambda)(x_0 - \varepsilon) + \lambda/(r-1)$. By Lemma 4.26(i), $f_{s,r}$ is convex on $[\frac{1}{r}, \frac{1}{r-1}] \supseteq [x_0 - \varepsilon, \frac{1}{r-1})$, so $f_{s,r}(x) \leqslant (1-\lambda)f_{s,r}(x_0 - \varepsilon) + \lambda f_{s,r}(\frac{1}{r-1})$. But $f_s(\frac{1}{r-1}) = g_s(r-1)$ and by Lemma 4.26(ii), we have $f_{s,r}(x_0 - \varepsilon) \leqslant \tilde{f}_{s,r}(x_0 - \varepsilon)$, so $f_{s,r}(x) \leqslant (1-\lambda)\tilde{f}_{s,r}(x_0 - \varepsilon) + \lambda g_s(r-1)$. Now (4.4.13) implies (4.4.11).

Thus it remains to show that (4.4.12) implies (4.4.13). This will follow from the fact that $\tilde{f}_{s,r}$ is close to $f_{s,r}$ and g_s is close to $\tilde{g}_s$; more specifically, it suffices to show the following three inequalities:

$$\tilde{f}_{s,r}(x_0 - \varepsilon) - \tilde{f}_{s,r}(x_0) \leqslant \varepsilon(r^2-1) + \varepsilon(r-1)\log\left(\frac{s}{r-1}\right) + \varepsilon^2 r^3, \tag{4.4.14}$$

$$\tilde{g}_s(r-1) \leqslant g_s(r-1) + \frac{1}{4}\left\lfloor \frac{s}{r-2} \right\rfloor^{-2} \leqslant g_s(r-1) + \frac{1}{4}\left\lfloor \frac{s}{r} \right\rfloor^{-2}, \tag{4.4.15}$$

$$\frac{1}{5r^2} \geqslant \frac{1}{4\lfloor s/r \rfloor^2} + \varepsilon(r-1) + \varepsilon(r-1)\log\left(\frac{s}{r-1}\right) + \varepsilon^2 r^3. \tag{4.4.16}$$

To show (4.4.14), we estimate $\tilde{f}_{s,r}(x_0 - \varepsilon)$ in terms of $\tilde{f}_{s,r}(x_0)$ and ε. We have

$$\tilde{f}_{s,r}(x_0) = (r-2)x_0 \log\left(\frac{sx_0}{1-x_0}\right) + (1-(r-1)x_0)h(x_0)$$

and

$$\tilde{f}_{s,r}(x_0-\varepsilon) \leqslant (r-2)x_0 \log\left(\frac{sx_0}{1-x_0}\right) + (1-(r-1)(x_0-\varepsilon))h(x_0-\varepsilon)$$

where $h(y) := \log\left(\frac{(1-(r-1)y)s}{1-y}\right)$. Using the mean value theorem, for $x_0 = \frac{r}{r^2-1}$ and $\varepsilon < x_0 - 1/r$, we have

$$\begin{aligned} h(x_0-\varepsilon) &\leqslant h(x_0) + \varepsilon \cdot \max_{y\in(x_0-\varepsilon,x_0)} \frac{r-2}{(1-y)(1-(r-1)y)} \\ &= h(x_0) + \varepsilon(r-2)\frac{1}{(1-r/(r^2-1))(1-r/(r+1))} \\ &\leqslant h(x_0) + \varepsilon(r^2-1). \end{aligned}$$

Thus, combining inequalities, we get

$$\begin{aligned} \tilde{f}_{s,r}(x_0-\varepsilon) - \tilde{f}_{s,r}(x_0) &\leqslant (1-(r-1)x_0)\varepsilon(r^2-1) + \varepsilon(r-1)h(x_0) + \varepsilon^2(r-1)^2(r+1) \\ &\leqslant \varepsilon(r^2-1) + \varepsilon(r-1)\log\left(\frac{s}{r-1}\right) + \varepsilon^2 r^3, \end{aligned}$$

where for the last inequality we used that $h(y)$ is decreasing on $[\frac{1}{r}, \frac{1}{r-1}]$, so $h(x_0) \leqslant h(\frac{1}{r}) = \log(\frac{s}{r-1})$. Thus (4.4.14) holds.

Next, (4.4.15) follows from Lemma 4.18.

To see (4.4.16), we first recall from (4.4.9) that $\varepsilon \leqslant 3r/20s^2$. We can bound $20r^2$ times each of the four terms on the right hand side by 1 using $r \leqslant \log(s)+1$ and the fact that s is large, as follows:

- for the first term we have $\frac{20r^2}{4\lfloor s/r\rfloor^2} \leqslant 5\left(\frac{r^2}{s-r}\right)^2 \leqslant 1$ which holds for all $s \geqslant 65$;
- the second term is at most the third;
- for the third term we have $20r^2\varepsilon(r-1)\log\left(\frac{s}{r-1}\right) \leqslant \frac{3r^4}{s^2}\log(s) < 1$ for $s \geqslant 133$; and
- for the fourth term we have $20r^2\varepsilon^2 r^3 < \frac{9r^7}{20s^2} < 1$, using $s > 881$.

This completes the proof that (4.4.14)–(4.4.16) hold, and hence completes the proof of the claim. □

It remains to show that (4.4.12) holds. Let

$$a := \frac{s}{r-2}, \quad b := \frac{sr}{r^2-r-1}, \quad c := \frac{s(r-1)}{r^2-r-1}, \quad d := \frac{s}{r}.$$

Then $a > b > c > d$. We have

$$\tilde{g}_s(r-1) = \frac{r-2}{r-1}\log(a) \quad \text{and} \quad \tilde{g}_s(r+1) = \frac{r}{r+1}\log(d).$$

Now,

$$\tilde{f}_{s,r}(x_0) = \frac{r(r-2)}{(r-1)(r+1)}\log(b) + \frac{1}{r+1}\log(c) = \begin{cases} \tilde{g}_s(r-1) + \frac{1}{r^2-1}(\log(a) - u_{r-1}) \\ \tilde{g}_s(r+1) - \frac{1}{r^2-1}(\log(d) - u_{r+1}) \end{cases}$$

where

$$u_{r-1} = r(r-2)\log\left(\frac{r^2-r-1}{r^2-2r}\right) + (r-1)\log\left(\frac{r^2-r-1}{r^2-3r+2}\right) \quad \text{and}$$
$$u_{r+1} = r(r-2)\log\left(\frac{r^2}{r^2-r-1}\right) + (r-1)\log\left(\frac{r^2-r}{r^2-r-1}\right).$$

Then using inequalities approximating $\log(1+x)$ by x we get that $u_{r-1} - u_{r+1}$ is

$$r(r-2)\log\left(\frac{(r^2-r-1)^2}{(r^2-2r)r^2}\right) + (r-1)\log\left(\frac{(r^2-r-1)^2}{(r^2-3r+2)(r^2-r)}\right) \geqslant 0.9$$

for all $r \geqslant 6$ (and indeed a Taylor expansion shows that $u_{r-1} - u_{r+1} \approx 1$ for large r). Suppose that (4.4.12) does not hold. Then

$$\log(a) \geqslant u_{r-1} - (r^2-1)/(5r^2) \quad \text{and} \quad \log(d) \leqslant u_{r+1} + (r^2-1)/(5r^2).$$

Subtracting the first inequality from the second, we have

$$\frac{2(r^2-1)}{5r^2} \geqslant \log\left(\frac{r-2}{r}\right) + u_{r-1} - u_{r+1} \geqslant \log\left(\frac{r-2}{r}\right) + 0.9$$

which yields a contradiction for every $r \geqslant 6$. Thus (4.4.12) holds, completing the proof of the lemma. □

We can now put everything together to prove Theorem 4.16 on solutions to the optimisation problem for dichromatic triangles.

Proof of Theorem 4.16. Let $s \geqslant 2$ be an integer and let $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(s)$. Now, Lemma 4.14 and Lemma 4.24 imply that

$$g(s) \overset{4.14}{\leqslant} Q(s) = q(\phi^*, \boldsymbol{\alpha}^*) \overset{4.24}{\leqslant} g_s(r^*). \tag{4.4.17}$$

If r^* is even, we have $g_s(r^*) \leqslant g(s)$, so we must have equality throughout. Having $g_s(r^*) = g(s)$ implies by definition that $r^* \in R_2(s)$ and Lemma 4.24(ii) implies that

ϕ^* and $\boldsymbol{\alpha}^*$ are uniform. Recall that by Remark 4.15, ϕ^* being uniform implies that $\phi^* = \phi_{\mathcal{A}}$ for some $\mathcal{A}$ as in Construction 4.12. This completes the proof of the even case.

Now suppose that r^* is odd. We claim that this never happens, so we will obtain a contradiction. From Lemma 4.24(i) we have the strict inequality

$$q(\phi^*, \boldsymbol{\alpha}^*) < g_s(r^*). \tag{4.4.18}$$

Suppose, without loss of generality, that $\alpha_1^* \geqslant \alpha_i^*$ for all $2 \leqslant i \leqslant r^*$, and let r' be the unique integer with $\frac{1}{r'} \leqslant \alpha_1^* < \frac{1}{r'-1}$. Since $\alpha_1^* \geqslant \frac{1}{r^*}$, we have $r' \leqslant r^*$. By Lemmas 4.3, 4.27 and 4.26(iii), we have

$$g(s) \leqslant Q(s) \overset{4.3}{=} q_1(\phi^*, \boldsymbol{\alpha}^*) \overset{4.27}{\leqslant} f_s(\alpha_1^*) \overset{4.26}{\leqslant} \max\{g_s(r'), g_s(r'-1)\}, \tag{4.4.19}$$

where, by definition, $f_s(\alpha_1^*) = f_{s,r'}(\alpha_1^*)$.

We now claim that r' must be equal to r^*. Indeed, suppose that $r' < r^*$, and let $r_{\text{even}}, r_{\text{odd}}$ be the even and odd integer in $\{r', r'-1\}$, respectively. In this case $\{r_{\text{odd}}, r^*\}$ are two distinct odd integers so by Lemma 4.21(iii) we have $\min\{g_s(r^*), g_s(r_{\text{odd}})\} < g(s)$. From (4.4.17) and (4.4.18) we have that $g(s) < g_s(r^*)$ so we must have $g_s(r_{\text{odd}}) < g(s)$. Thus the right-hand side of (4.4.19) is at most $g(s)$ and again we have equality throughout. Lemma 4.27 implies that $\alpha_1^* < 1/(r^*-1)$, which contradicts $r' < r^*$. Thus from now on we may assume that $r' = r^*$.

We now proceed to compare $g_s(r')$ and $g_s(r'-1)$ which we now know are equal to $g_s(r^*)$ and $g_s(r^*-1)$, respectively. From (4.4.17) and (4.4.18) it follows that $g_s(r^*) > g(s)$ while $g_s(r^*-1) \leqslant g(s)$ since r^*-1 is even, so $g_s(r^*-1) < g_s(r^*)$ and the maximum in (4.4.19) is strictly attained by r^*. Note that $g_s(r^*) > g(s)$ implies that $R(s)$ can consist solely of odd numbers.

Further to the above, we must also have $r^* \in R(s)$, for otherwise there is some (odd) $r \in R(s)$ distinct from r^* with both $g_s(r^*)$ and $g_s(r)$ exceeding $g(s)$, contradicting Lemma 4.21(iii).

Finally, by Lemma 4.25 we have

$$\alpha_1^* \geqslant \frac{r^*}{(r^*)^2 - 1} - \frac{e_s(r^*+1)}{(r^*-1)\log(s/r^*)}.$$

Then Lemma 4.28 implies that $f_{s,r^*}(\alpha_1^*) < \max\{g_s(r^*-1), g_s(r^*+1)\} \leqslant g(s)$, where the last inequality follows from the definition of $g(s)$. Together with (4.4.19), this contradicts our initial assumption that r^* is odd and finishes the proof. □

4.4.5 Proof of Theorem 1.21

Now that we have determined the set of basic optimal solutions for the dichromatic triangle problem (Theorem 4.16), we need to show that it has various properties in order to apply our general exact result, Theorem 4.11. Most of these are easy to see; that the extension property holds is slightly more involved.

Lemma 4.30. *Let $s \geqslant 2$ be an integer. The dichromatic triangle family $\mathcal{X} = (K_3^{(2)}, s)$*

(i) *is bounded,*

(ii) *is hermetic (and hence stable inside),*

(iii) *has the extension property,*

(iv) *has the property that for every basic optimal solution $(r^*, \phi^*, \boldsymbol{\alpha}^*)$, there is some $z \geqslant r^*$ such that each multiplicity ϕ^*_{ij} equals z or $z+1$.*

Proof. Theorem 4.16 implies that for any $(r^*, \phi^*, \boldsymbol{u}_{r^*}) \in \text{OPT}^*(\mathcal{X})$, we have that r^* is even, $\boldsymbol{u}_{r^*}$ is uniformly $1/r^*$, all colour classes of ϕ^* are perfect matchings, the values ϕ^*_{ij} differ by at most one, and $r^* \in R_2(s)$, which implies $r^* \leqslant \log(s)+2$ due to Lemma 4.21(i). We calculated $R_2(s)$ for small values of s in Table 1.2 (we showed these values hold in Lemma 4.21(iv); in particular, $r^*=2$ for $2 \leqslant s \leqslant 26$ and $r^* \leqslant 4$ for $s \leqslant 496$.

Part (i) was proved in Lemma 4.5(iii).

Next we prove (ii). Let $(r^*, \phi^*, \boldsymbol{u}_{r^*}) \in \text{OPT}^*(\mathcal{X})$ and suppose that $\phi \in \Phi_{\mathcal{X},1}(r^*+1)$ extends ϕ^*, i.e., $\phi|_{\binom{[r^*]}{2}} = \phi^*$. Let $c \in \phi(1, r^*+1)$. Then there is some $i \in [r^*]$ such that $c \in \phi(1i)$ since the edges of colour c form a perfect matching in ϕ^*. To avoid $\mathcal{X}$ we must have $\phi(i, r^*+1) = \{c\}$. However, since $\phi^*_{1i} \geqslant 2$, there is some $c' \in \phi_{1i}$, giving a dichromatic triangle on $1, i, r^*+1$. Thus $\mathcal{X}$ is hermetic.

For (iii), let $(r^*, \phi^*, \boldsymbol{u}_{r^*}) \in \text{OPT}^*(\mathcal{X})$ and $\phi \in \Phi_{\mathcal{X}}(r^*+1)$ such that $\phi|_{\binom{[r^*]}{2}} = \phi^*$ and

$$\text{ext}(\phi, \boldsymbol{u}_{r^*}) = \frac{1}{r^*} \sum_{i \in [r^*]} \log \phi_{i,r^*+1} = Q(s). \tag{4.4.20}$$

We first note that r^*+1 can see each colour at most once: if colour c was contained in $\phi(i, r^*+1)$ and $\phi(j, r^*+1)$, there would be a dichromatic triangle in ϕ on parts i, j, r^*+1 since $\phi^*_{ij} \geqslant 2$. In fact r^*+1 sees each colour exactly once, for otherwise $\text{ext}(\phi, \boldsymbol{u}_{r^*})$ cannot equal $Q(s)$. Let $\ell := |\{i \in [r^*] : \phi_{i,r^*+1} \neq \emptyset\}|$. Note that $\mathcal{X}$ is hermetic so $\ell \leqslant r^*-1$.

We claim that $\ell = r^* - 1$. This is immediate when $r^* = 2$, so we may assume $s \geqslant 27$. Equation (4.4.20) and Lemma 4.23 imply that

$$Q(s) \leqslant (\ell/r^*) \log(s/\ell). \tag{4.4.21}$$

Suppose for a contradiction that $\ell \leqslant r^* - 2$. We have that

$$\frac{s}{r^* - 1} \geqslant 6 \quad \text{and} \quad \frac{(r^* - 1)^2}{s - r^* + 1} \leqslant 1/2; \tag{4.4.22}$$

these assertions are true for $27 \leqslant s \leqslant 496$ by Table 1.2 where $r^* \leqslant 4$ and for larger s they follow from $r^* \leqslant \log(s) + 2$ (recall (4.4.3)). Our assumption implies that $(s - r^* + 1)/(r^* - 1) < s/\ell$. The mean value theorem implies that there is $\frac{s-r^*+1}{r^*-1} < u < \frac{s}{\ell}$ such that

$$\Lambda := \ell \left(\log \left(\frac{s - r^* + 1}{r^* - 1} \right) - \log \left(\frac{s}{\ell} \right) \right) = \frac{\ell}{u} \left(\frac{s - r^* + 1}{r^* - 1} - \frac{s}{\ell} \right) > \ell - \frac{s(r^* - 1)}{s - r^* + 1}.$$

By Lemma 4.3, we have $Q(s) = q_1(\phi^*, \boldsymbol{u}_{r^*}) = \frac{r^*-1}{r^*} \sum_{2 \leqslant i \leqslant r^*} \log \phi^*_{1i}$. By (4.4.1), for all $ij \in \binom{[r^*]}{2}$, we have $\phi^*_{ij} \in \{z, z+1\}$ where $z := \lfloor s/(r^* - 1) \rfloor$. Thus

$$\begin{aligned} r^* \cdot Q(s) - \ell \log \left(\frac{s}{\ell} \right) &\geqslant (r^* - 1) \log \left(\frac{s}{r^* - 1} - 1 \right) - \ell \log \left(\frac{s}{\ell} \right) \\ &= (r^* - 1 - \ell) \log \left(\frac{s - r^* + 1}{r^* - 1} \right) + \Lambda \\ &> (r^* - 1 - \ell) \log \left(\frac{s - r^* + 1}{r^* - 1} \right) - \frac{s(r^* - 1)}{s - r^* + 1} + \ell \\ &= (r^* - 1 - \ell) \left(\log \left(\frac{s}{r^* - 1} - 1 \right) - 1 \right) - \frac{(r^* - 1)^2}{s - r^* + 1} \\ &\overset{(4.4.22)}{\geqslant} \log(5) - 1 - 1/2 > 0, \end{aligned}$$

contradicting (4.4.21) and hence proving the claim.

Let $i \in [r^*]$ be the unique index such that $\phi_{i,r^*+1} = \emptyset$. We claim that $r^* + 1$ must be a clone of i. To see this, note that for any $j \in [r^*] \setminus \{i\}$ and any $c \in \phi(j, r^* + 1)$, we must have $c \in \phi^*(ij)$, for otherwise the perfect matching on $[r^*]$ in colour c will contain the edge jj' for some $j' \neq i$ and $r^* + 1, j, j'$ will form a dichromatic triangle. Conversely, if $c \in \phi^*(ij)$, $r^* + 1$ necessarily sees c but we cannot have $c \in \phi(j', r^* + 1)$ for any $j' \neq j$. This is because according to ϕ^* colour c forms a perfect matching on $[r^*]$, one of whose edges must be $j'j''$ for some $j'' \neq i, j$, and this would yield a dichromatic triangle on $j', j'', r^* + 1$. Therefore, $r^* + 1$ is a clone of i, as claimed.

Finally, for (iv), recall that each multiplicity ϕ^*_{ij} equals z or $z+1$ by (4.4.1). Since $r^* \leqslant \log(s)+2$, each multiplicity is therefore at least $\lfloor s/(\log(s)+1)\rfloor$. This is at least $\log(s)+2 \geqslant r^*$ when $s \geqslant 27$. For $2 \leqslant s \leqslant 26$, we have $r^*=2$ by Table 1.2. In this case, $z=s \geqslant 2 = r^*$, as required. □

The properties of the dichromatic triangle family guaranteed by the previous lemma now allow us to apply our general exact result, Theorem 4.11. This result immediately tells us that every extremal graph is a complete partite graph with the right number of parts with roughly the correct sizes, so it remains to prove that the sizes are in fact as equal as possible.

Proof of Theorem 1.21. The statement about $R_2(s)$ is Lemma 4.21(i). Recall that by Theorem 4.16, $\textsc{opt}^*(s)$ consists of triples $(r^*, \phi^*, \boldsymbol{u}_{r^*})$ with $r^* \in R_2(s) \subseteq 2\mathbb{N}$, uniform $\boldsymbol{u}_{r^*}$ and ϕ^* where every ${\phi^*}^{-1}(c)$ is a perfect matching, and the multiplicities ϕ^*_{ij} over pairs ij are as equal as possible. By Lemma 4.21(i) and (4.4.3), we have $r^* \leqslant \log(s)+2$. Given $r^* \in R_2(s)$, let Y_{r^*} consist of all ϕ^* such that $(r^*, \phi^*, \boldsymbol{u}_{r^*}) \in \textsc{opt}^*(s)$. In other words, Y_{r^*} is the set of r^*-partite colour templates obeying Construction 4.12.

It remains to prove the statements about n-vertex extremal graphs for sufficiently large n. Suppose that $1/n \ll \varepsilon \ll 1/k, 1/s$ and let G be an n-vertex extremal graph. By Theorem 4.11 (which is applicable by Lemma 4.30(i)–(iii)) there is $(r^*, \phi^*, \boldsymbol{u}_{r^*}) \in \textsc{opt}^*(s)$ so that G is a complete r^*-partite graph whose parts $W_1, \ldots, W_{r^*}$ all have size between $n/(2r^*)$ and $2n/r^*$, and moreover, for at least a $(1-e^{-\varepsilon n})$ proportion of the valid colourings $\chi \in F(G;(K_3^{(2)}, s))$ there exists $\phi^* \in Y_{r^*}$ such that χ follows ϕ^* perfectly. For a complete r^*-partite graph H and $\phi^* \in Y_{r^*}$, denote the set of such perfect colourings of H by $X_H(\phi^*)$.

Before determining the exact part sizes of G, we quickly show that $F(G;(K_3^{(2)}, s))$ is approximately equal to $\sum_{\phi^* \in Y_{r^*}} |X_G(\phi^*)|$. Intuitively, this is because very few perfect colourings can be assigned to more than one $\phi^* \in Y_{r^*}$. Indeed, if $\chi \in X_G(\phi^*) \cap X_G(\phi)$ for some $\phi^*, \phi \in Y_{r^*}$, then there is $ij \in \binom{[r^*]}{2}$ where $\phi^*(ij) \neq \phi(ij)$ and $\chi(xy) \in \phi^*(ij) \cap \phi(ij)$ for all $x \in W_i, y \in W_j$, so the number of such colourings χ for a fixed ϕ^* is at most

$$\phi^*_{ij}\left(\frac{\phi^*_{ij}-1}{\phi^*_{ij}}\right)^{(n/(2r^*))^2} |X_G(\phi^*)| \leqslant s\left(1-\frac{1}{s}\right)^{(n/(2r^*))^2} |X_G(\phi^*)| \leqslant e^{-n}|X_G(\phi^*)|,$$

say. Thus the number of colourings of G which follow ϕ^* perfectly and do not follow another ϕ is at least $(1-e^{-n})|X_G(\phi^*)|$, so

$$F(G;(K_3^{(2)},s)) = (1 \pm e^{-\varepsilon n}) \sum_{\phi^* \in Y_{r^*}} |X_G(\phi^*)|. \tag{4.4.23}$$

We now show that $|X_H(\phi^*)|$ is maximised when H is the Turán graph.

Claim 4.4.5.1. *Let H be a sufficiently large complete r^*-partite graph with parts $V_1, \ldots, V_{r^*}$. Then for every $\phi^* \in Y_{r^*}$, we have $|X_T(\phi^*)|/|X_H(\phi^*)| \geqslant 2$ whenever H is not isomorphic to the Turán graph $T := T_{r^*}(n)$.*

Proof. Suppose not, and, without loss of generality, we may suppose that $|V_2|-|V_1| \geqslant 2$ is the largest among all class size differences. Consider the complete multipartite graph $\tilde{H}$ obtained by removing a vertex from V_2 and adding a vertex to V_1. Fix $\phi^* \in Y_{r^*}$ and for each integer m, let $J_m := \{ij \in \binom{[r^*]}{2} : \phi^*_{ij} = m\}$. Since ϕ^* is optimal, (4.4.1) implies that $E(K_{r^*}) = J_z \cup J_{z+1}$ where $z := \lfloor s/(r^*-1) \rfloor$. Note that each J_m is regular since the multigraph which is the union of the ${\phi^*}^{-1}(c)$ is regular.

Let $\ell := \phi^*_{12}$. Passing from H to $\tilde{H}$, $|X(\phi^*)|$ increases by a factor of

$$D := \frac{|X_{\tilde{H}}(\phi^*)|}{|X_H(\phi^*)|} = \ell^{|V_2|-|V_1|-1} \prod_{y \in [r^*] \setminus \{1,2\}} \left(\frac{\phi^*_{1y}}{\phi^*_{2y}}\right)^{|V_y|}.$$

It suffices to show that $D \geqslant 2$, then the claim follows by a telescoping product.

To prove this, suppose first that $r^* = 2$. Then $D = s^{|V_2|-|V_1|-1} \geqslant s \geqslant 2$, as required. Thus we may assume $r^* \geqslant 4$, and so $s \geqslant 27$, and $z \geqslant 9$. Note that the factor ϕ^*_{1y}/ϕ^*_{2y} equals $\frac{z}{z+1}$ if $y \in A := N_{J_z}(1) \cap N_{J_{z+1}}(2)$, equals $\frac{z+1}{z}$ if $y \in B := N_{J_{z+1}}(1) \cap N_{J_z}(2)$, and equals 1 otherwise. Since J_z and J_{z+1} are regular graphs partitioning the edge set of K_{r^*}, we have $|A| = |B|$. Thus

$$D \geqslant \left(\frac{z}{z+1}\right)^{\sum_{y\in B}|V_y| - \sum_{y \in A}|V_y|} \ell^{|V_2|-|V_1|-1} \geqslant \left(\frac{z}{z+1}\right)^{|A|(|V_2|-|V_1|)} \ell^{|V_2|-|V_1|-1},$$

as $|V_2|-|V_1|$ is the maximum difference between the size of two parts. For the next step, we note that since $\frac{r^*-2}{2z} \leqslant \frac{(r^*)^2}{s} \leqslant 1$ we have, using $\log(1+x) \geqslant \frac{x}{x+1}$ for $x > -1$,

$$\ell \left(\frac{z}{z+1}\right)^{(r^*-2)/2} \geqslant \exp\left(\log \ell - \frac{r^*-2}{2z}\right) \geqslant \exp\left(\log z - \frac{r^*-2}{2z}\right)$$
$$\geqslant \exp(\log z - 1) > 1.$$

Using $|V_2|-|V_1|\geqslant 2$ and $|A|\leqslant (r^*-2)/2$, this allows us to bound

$$D \geqslant \ell^{-1}\left(\ell\left(\frac{z}{z+1}\right)^{(r^*-2)/2}\right)^{|V_2|-|V_1|} \geqslant \ell^{-1}\left(\ell^{\frac{2}{r^*-2}}\frac{z}{z+1}\right)^{r^*-2}$$
$$= \begin{cases} z^{r^*-1}/(z+1)^{r^*-2} & \text{if } \ell = z \\ z^{r^*-2}/(z+1)^{r^*-3} & \text{if } \ell = z+1. \end{cases}$$

Suppose that $r^* = 4$. Then $s \geqslant 27$ so $z \geqslant 9$ and hence $z^3/(z+1)^2 > 7$, and $z^2/(z+1) > 8$, as required. Thus we may assume that $r^* \geqslant 6$ and hence $s \geqslant 497$. Since $\frac{r^*-2}{z} \leqslant \frac{2(r^*)^2}{s} \leqslant 1$, we have for $i \in \{2,3\}$ that

$$D \geqslant z\left(\frac{z}{z+1}\right)^{r^*-i} \geqslant \exp\left(\log(z) - \frac{r^*-2}{z}\right) \geqslant \exp(\log(z)-1) > 2,$$

completing the claim. □

The maximality of $F(G;(K_3^{(2)},s))$ and the claim now imply that $G \cong T_{r^*}(n)$, that is, the parts W_i form an equipartition of $V(G)$.

It remains to accurately estimate $|X_{T_{r^*}(n)}(\phi^*)|$ for every $\phi^* \in Y_{r^*}$ to obtain the formula for $F(G;(K_3^{(2)},s))$. For this, let m, f be the unique positive integers with $0 \leqslant f < r^*$ such that $n = r^*m + f$. We claim that

$$e(T_{r^*}(n)) =: t_{r^*}(n) = \binom{r^*}{2}m^2 + (r^*-1)mf + \binom{f}{2}. \tag{4.4.24}$$

To see this, note that f parts $W_1,\ldots,W_f$ have size $m+1$ and r^*-f parts $W_{f+1},\ldots,W_{r^*}$ have size m. Choose an arbitrary labelling $W_i = U_i \cup \{v_i\}$ for each large part W_i with $i \in [f]$ and $W_i = U_i$ for each small part with $i > f$. The terms in the required expression are, in order, the number of edges between the U_i, between each v_i and U_j for $i \neq j$, and between the v_i.

Fix an optimal $\phi^* \in Y_{r^*}$ and again let $z = \lfloor \frac{s}{r^*-1} \rfloor$ and $a = s - z(r^*-1)$. The number of colourings χ of $K_{r^*}(U_1,\ldots,U_{r^*})$ which follow ϕ^* is exactly

$$z^{m^2\left(\binom{r^*}{2}-ar^*/2\right)}(z+1)^{m^2ar^*/2}.$$

For each v_i with $i \in [f]$, the number of colourings following ϕ^* of edges v_ix with $x \in \bigcup_{j\neq i} U_j$ is exactly $z^{m(r^*-1-a)}(z+1)^{ma}$. The number of colourings of $\{v_iv_j : ij \in \binom{[f]}{2}\}$ is a function C_ϕ which is at least $z^{\binom{f}{2}}$ and at most $(z+1)^{\binom{f}{2}}$, and where, due to the symmetric form of Y_{r^*}, $\sum_{\phi \in Y_{r^*}} C_\phi$ depends only on s, r^* and the remainder f of n

modulo r^*. Using (4.4.24), we see that the number of colourings of $T_{r^*}(n)$ following ϕ^* is

$$\begin{aligned}|X_{T_{r^*}(n)}(\phi^*)| &= (z+1)^{m^2ar^*/2}z^{m^2\left(\binom{r^*}{2}-ar^*/2\right)}\cdot\left((z+1)^{ma}z^{m(r^*-1-a)}\right)^f\cdot C_\phi\\ &= (z+1)^{\left(t_{r^*}(n)-\binom{f}{2}\right)a/(r^*-1)}z^{\left(t_{r^*}(n)-\binom{f}{2}\right)(r^*-1-a)/(r^*-1)}\cdot C_\phi\\ &= e^{\frac{r^*}{r^*-1}g_s(r^*)\left(t_{r^*}(n)-\binom{f}{2}\right)}\cdot C_\phi = C'_\phi e^{\frac{r^*}{r^*-1}g(s)t_{r^*}(n)}\end{aligned}$$

where $C'_\phi = C_\phi e^{-\frac{r^*}{r^*-1}g(s)\binom{f}{2}}$, so that, again, $C=\sum_{\phi\in Y_{r^*}}C'_\phi$ is a function of only s, r^*, f. By (4.4.23) we have

$$F(T_{r^*}(n);(K_3^{(2)},s)) = (1+o(1))\sum_{\phi\in Y_{r^*}}C'_\phi e^{\frac{r^*}{r^*-1}g(s)t_{r^*}(n)} = (C+o(1))e^{\frac{r^*}{r^*-1}g(s)t_{r^*}(n)}, \tag{4.4.25}$$

as desired. The number $|Y_{r^*}|$ of optimal ϕ^* is a function of r^*, s. In fact,

$$|Y_{r^*}| \geqslant \binom{s}{\underbrace{z,\ldots,z}_{r^*-1-a}\,\underbrace{z+1,\ldots,z+1}_{a}}$$

which is a lower bound on the number of ϕ obeying Construction 4.12. We note for future reference that the constant C in the statement of the theorem therefore satisfies

$$C \geqslant \frac{s!}{(z+1)!^{r^*-1}}z^{\binom{f}{2}}e^{-\frac{r^*}{r^*-1}g(s)\binom{f}{2}}. \tag{4.4.26}$$

It remains to prove the last sentence of the statement. The set $R_2(s)$ contains a single value for all $s\in[2,10^7]\setminus\{27\}$ from Table 1.2 (proved in Lemma 4.21(iv)), while $R_2(27)=\{2,4\}$. Thus, to complete the proof, it remains to prove the assertion about $s=27$. Clearly $F(T_2(n);(K_3^{(2)},27)) = 27^{t_2(n)}$ since every colouring is valid.

Suppose $s=27$ and $r^*=4$. Then $z=9$ and since 3 divides 27 we have $C_\phi = 9^{\binom{f}{2}}$ and $e^{\frac{3}{4}g(27)}=9$ so $C'_\phi=1$. Appealing to the remark after Construction 4.12, $|Y_4|$ is the number of triples (A_1,A_2,A_3) where $A_1\cup A_2\cup A_3=[27]$ is an equipartition. Thus $|Y_4|=\binom{27}{9,9,9}$ and therefore $F(T_4(n);(K_3^{(2)},27)) = (1+o(1))\binom{27}{9,9,9}9^{t_4(n)}$, which is many times greater than $F(T_2(n);(K_3^{(2)},27))=27^{t_2(n)}$. So $T_4(n)$ is the unique optimal graph for $s=27$ and $F(n;(K_3^{(2)},27)) = (1+o(1))\binom{27}{9,9,9}9^{t_4(n)}$. □

Proof of Corollary 1.22. Theorem 1.21 implies that for each s, the set of $(K_3^{(2)},s)$-extremal graphs is a subset of $\{T_r(n): r\in R_2(s)\}$, so it suffices to show that each set $S_2(r)=\{s: r\in R_2(s)\}$ for $r\in 2\mathbb{N}$ contains an interval $[s^-(r),s^+(r)]$ that is disjoint from all other $S_2(r')$, $r'\neq r$. Lemma 4.21(iii) implies that for all $r\in 2\mathbb{N}$, there is s such that $R_2(s) = \{r\}$ is a singleton. We also have that the interval $S_2(r) = [\tilde{s}_r, \tilde{s}_{r+2} - 1]$ or $[\tilde{s}_r, \tilde{s}_{r+2}]$ overlaps with each of $S_2(r-2)$ and $S_2(r+2)$ in at most one element, so we can take $s^-(r) = \tilde{s}_r$ or $\tilde{s}_r + 1$ and $s^+(r) = \tilde{s}_{r+2}$ or $\tilde{s}_{r+2} - 1$. The bounds on s_r in Lemma 4.21(iii) yield

$$(r-3)e^{r-2} < s_{r-1} \leqslant \tilde{s}_r \leqslant s_r < (r-1)e^r,$$

as required. □

4.5 Forbidding improperly coloured cliques

In this section, we prove Theorem 1.24 on the improper pattern, which follows with some extra work from the proof of Theorem 1.21. We first determine the set of basic optimal solutions and show that the pattern satisfies the hypotheses of Theorem 4.11.

Lemma 4.31. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X} = \mathcal{X}^{\wedge}_{k,s}$ be the family of all improper s-edge-colourings of K_k. Then $\mathcal{X}$*

(i) *has $Q(\mathcal{X}) = \max\{\frac{k-2}{k-1}\log(s), g(s)\}$ and basic optimal solutions*

$$\mathrm{OPT}^*(\mathcal{X}) = \begin{cases} \{(k-1, \phi_{[s]}, \boldsymbol{u})\} & \text{if } \frac{k-2}{k-1}\log(s) > g(s) \\ \{(k-1, \phi_{[s]}, \boldsymbol{u})\} \cup \mathrm{OPT}^*(K_3^{(2)}, s) & \text{if } \frac{k-2}{k-1}\log(s) = g(s) \\ \mathrm{OPT}^*(K_3^{(2)}, s) & \text{if } \frac{k-2}{k-1}\log(s) < g(s) \end{cases}$$

where $\phi_{[s]} \equiv [s]$ and $\boldsymbol{u}$ is uniform; and in particular $\mathcal{X}$ is bounded,

(ii) *is hermetic (and hence stable inside),*

(iii) *has the extension property.*

Proof. Let $(r, \phi, \boldsymbol{\alpha}) \in \mathrm{FEAS}_{\mathcal{X},2}(s)$. Suppose first that $r \leqslant k-1$. Then $q(\phi, \boldsymbol{\alpha}) = 2\sum_{ij} \alpha_i\alpha_j \log\phi_{ij} \leqslant 2\log(s)\sum_{ij}\alpha_i\alpha_j \leqslant (1-\frac{1}{r})\log(s) \leqslant (\frac{k-2}{k-1})\log(s)$. This is uniquely attained by the solution $(k-1, \phi_{[s]}, \boldsymbol{u})$ where $\phi_{[s]} \equiv [s]$ and $\boldsymbol{u}$ is uniform. Suppose instead that $r \geqslant k$. Then $\phi^{-1}(c)$ is a matching for all $c \in [s]$, otherwise there would be an improperly coloured clique (with two adjacent edges of colour c). By Remark 4.17, we have $q(\phi, \boldsymbol{\alpha}) \leqslant g(s)$ with equality if and only if $r \in R_2(s)$, $\boldsymbol{\alpha}$ is uniform and $\phi = \phi_{\mathcal{A}}$ for some $\mathcal{A}$ as in Construction 4.12. Thus $Q(\mathcal{X})$ and the set of basic optimal solutions are as stated. Clearly $\mathcal{X}$ is bounded since the dichromatic triangle pattern is bounded.

For parts (ii) and (iii), since the dichromatic triangle family $(K_3^{(2)}, s)$ is hermetic and has the extension property, it suffices to check the conditions for the solution

$(k-1, \phi_{[s]}, \boldsymbol{u})$ assuming that $Q(\mathcal{X}) = \frac{k-2}{k-1}\log(s)$. Let $\phi \in \Phi_{\mathcal{X},1}(k)$ be an extension of $\phi_{[s]}$. Let $c \in \phi(1k)$. Then $c \in \phi(12) = [s]$ so if $|\phi(ik)| \geqslant 1$ for all $i \in [k-1]$, there is a k-clique with adjacent edges of colour c, a contradiction. Thus $\mathcal{X}$ is hermetic. We have also shown that any extension ϕ must have $\phi(ik) = \emptyset$ for some $1 \leqslant i \leqslant k-1$. Without loss of generality, suppose $\phi(1k) = \emptyset$. Then the contribution of k is $\sum_{2 \leqslant i \leqslant k-1} \frac{1}{k-1} \log|\phi(ik)| \leqslant \frac{k-2}{k-1}\log(s) = Q(\mathcal{X})$ with equality if and only if $\phi(ik) = [s]$ for all $2 \leqslant i \leqslant k$. In other words, if and only if k is a clone of 1 under $\phi_{[s]}$. Thus $\mathcal{X}$ has the extension property. □

Given our proof of Theorem 1.21, it is now an easy task to prove our second main result.

Proof of Theorem 1.24. Let $\mathcal{X} = \mathcal{X}^{\wedge}_{k,s}$ be the family of all improper s-edge-colourings of K_k. In Lemma 4.31 we determined $Q(\mathcal{X})$ and $\text{OPT}^*(\mathcal{X})$ and showed that the hypotheses of Theorem 4.11 hold.

Let $0 < 1/n \ll \varepsilon \ll \delta \ll 1/s, 1/k$ and let G be an $\mathcal{X}$-extremal graph on n vertices. Theorem 4.11 implies that there is $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$ such that G is a complete r^*-partite graph whose i-th part W_i has size $(\alpha^*_i \pm \delta)n$ for all $i \in [r^*]$, and (SE2) holds.

If $(r^*, \phi^*, \boldsymbol{\alpha}^*) = (k-1, \phi_{[s]}, \boldsymbol{u})$, then G is K_k-free so every colouring of G is valid, so $F(n; \mathcal{X}) = s^{\sum_{ij \in \binom{[k-1]}{2}} |W_i||W_j|}$ which is uniquely maximised when $||W_i| - |W_j|| \leqslant 1$ for all i, j; that is, we must have $G \cong T_{k-1}(n)$ which has exactly $s^{t_{k-1}(n)}$ valid colourings. On the other hand, if $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(K_3^{(2)}, s) \setminus \{(k-1, \phi_{[s]}, \boldsymbol{u})\}$, Theorem 1.21 implies that $G \cong T_{r^*}(n)$ for some $r^* \in R_2(s)$ and $F(G; (K_3^{(2)}, s)) = (C + o(1))e^{\frac{r^*}{r^*-1}g(s)t_{r^*}(n)}$ where C satisfies (4.4.26) and $g(s) = g_s(r^*)$. Recall that by Lemma 4.21(i) and (4.4.3) we have

$$r^* \leqslant \log(s) + 2. \tag{4.5.1}$$

We now first show that

$$\text{if } r^* \geqslant k \text{ and } g_s(r^*) \geqslant \frac{k-2}{k-1}\log(s), \text{ then } g_{s+1}(r^*) > \frac{k-2}{k-1}\log(s+1). \tag{4.5.2}$$

We verify (4.5.2) manually for all $k \geqslant 3$ and $k-1 \leqslant s \leqslant 1231$ and $r^* \in R_2(s) \setminus \{k-1\}$ by a computer search given in the ancillary file `improper_patterns.py`. Suppose now that $s \geqslant 1232$. Then

$$g_{s+1}(r^*) - g_s(r^*) - \frac{k-2}{k-1}\log\left(\frac{s+1}{s}\right) \geqslant \left(\frac{r^*-1}{r^*} - \frac{k-2}{k-1}\right)\log\left(\frac{s+1}{s}\right) - e_{s+1}(r^*)$$
$$> \left(\frac{1}{k-1} - \frac{1}{r^*}\right)\frac{1}{s+1} - \frac{1}{4}\left\lfloor\frac{s+1}{r^*-1}\right\rfloor^{-2} > 0$$

using $k \leqslant r^* \leqslant \log s + 2$ and $s \geqslant 1232$ in the final inequality. So (4.5.2) holds for all s. Define

$$\overline{s}(k) := \min\left\{s \in \{2,3,\ldots\} : g(s) \geqslant \frac{k-2}{k-1}\log(s) \text{ and } \min R_2(s) \geqslant k\right\}.$$

We will show that $s(k) = \overline{s}(k) - 1$ satisfies the theorem. Note that if there is $r^* \in R_2(s)$ with $r^* < k$, then $g(s) \leqslant \frac{k-2}{k-1}\log(s)$. Thus, by comparing the leading terms in the exponent of the formulas we gave above for $F(T_{k-1}(n); \mathcal{X})$ and $F(T_{r^*}(n); \mathcal{X})$, we obtain from (4.5.2) that

- if $s < \overline{s}(k)$, we have $g(s) \leqslant \frac{k-2}{k-1}\log(s)$ and $G \cong T_{k-1}(n)$;
- if $s > \overline{s}(k)$, we have $g(s) > \frac{k-2}{k-1}\log(s)$ and thus $G \cong T_{r^*}(n)$.

It remains to show that for $s = \overline{s}(k)$, we have $G \cong T_{r^*}(n)$. Suppose $g(s) > \frac{k-2}{k-1}\log(s)$. Then we can deduce as before that $G \cong T_{r^*}(n)$ for some $r^* \in R_2(s)$.

So suppose $g(s) = \frac{k-2}{k-1}\log(s)$ and $\min R_2(s) \geqslant k$. First note that by our definition of $\overline{s}(k)$ we have $\overline{s}(k) \geqslant 27$ for all k since $R_2(s) = \{2\}$ for $s \leqslant 26$. Thus for the calculations that follow we may assume that $s \geqslant 27$. By making the appropriate substitutions we can deduce from (4.4.24) that for an integer r, the number of edges in the Turán graph $T_r(n)$ is

$$t_r(n) = \frac{r-1}{r}\frac{n^2}{2} - \frac{i}{2}\left(1 - \frac{i}{r}\right),$$

where $0 \leqslant i \leqslant r-1$ and $i \equiv n \mod r$. Thus, letting f, b be such that $0 \leqslant f \leqslant r^* - 1$ and $f \equiv n \mod r^*$ and $0 \leqslant b \leqslant k-2$ and $b \equiv n \mod (k-1)$, and setting $z := \left\lfloor\frac{s}{r^*-1}\right\rfloor$, we have, using $e^{g(s)} = s^{\frac{k-2}{k-1}}$,

$$\frac{F(T_{r^*}(n); \mathcal{X})}{F(T_{k-1}(n); \mathcal{X})} \overset{(4.4.26)}{\geqslant} \frac{\frac{s!}{((z+1)!)^{r^*-1}} z^{\binom{f}{2}} \exp\left(-\frac{r^*}{r^*-1}g(s)\binom{f}{2} - \frac{r^*}{r^*-1}g(s)\frac{f}{2}\left(1 - \frac{f}{r^*}\right)\right)}{(1+o(1))\exp\left(-\frac{b}{2}\left(1 - \frac{b}{k-1}\right)\log(s)\right)}$$
$$\geqslant \frac{1}{2}\frac{s!}{((z+1)!)^{r^*-1}} z^{\binom{f}{2}} e^{-\log(s)f^2/2}.$$

The second inequality follows by expanding the numerator, substituting $g(s) \leqslant \tilde{g}(s) \leqslant \log(s)$ (which holds due to Lemma 4.18, Lemma 4.20(ii) and (4.4.2)), and noting that the exponent in the denominator is non-positive. Also, using $s \geqslant 27$ and $z \geqslant \frac{s}{\log(s)+1} - 1 \geqslant \frac{s}{2\log(s)}$ and $f \leqslant \log(s) + 1$ (which follow from (4.5.1)),

$$\binom{f}{2} \log(z) - \frac{f^2}{2} \log(s) \geqslant \binom{f}{2} \log\left(\frac{s}{2\log(s)}\right) - \frac{f^2}{2} \log(s) \geqslant -2\log^2(s) \log\log(s).$$

Putting this together with the Stirling bounds $e(\ell/e)^\ell \leqslant \ell! \leqslant e\ell(\ell/e)^\ell$ applied to $\ell = s, z$, and (4.5.1), we obtain

$$\frac{F(T_{r^*}(n); \mathcal{X})}{F(T_{k-1}(n); \mathcal{X})} > 1,$$

as required. □

4.6 Proofs of general results

This section contains the proofs of Theorem 4.8–Theorem 4.11. These proofs are mainly adaptations of proofs in [74, 75], which are versions of our results for monochromatic colour patterns. The arguments transfer almost directly from the monochromatic setting – the key point being that in both settings, the family of forbidden colourings are on *cliques*. This means that, informally, colours at two non-adjacent vertices are 'independent' since they cannot be in a forbidden clique together, and thus it can be shown that the following holds:

(S) for any pair u, v of non-adjacent vertices, either replacing u by a twin of v or v by a twin of u (the operation of *symmetrisation*) gives a graph with at least as many valid colourings.

This is stated in e.g. [19, Lemma 2.7]. Symmetrisation was originally introduced by Zykov in the 1950s [86, 87] to give a new proof of Turán's theorem.

The original proofs in [74, 75] are rather long. We therefore give only a sketch for those parts which are not new, and refer the reader to the relevant reference.

4.6.1 Tools

We start by collecting some tools concerning the optimisation problem. Versions of most of these for monochromatic patterns appear in [74–76], but usually the same proof goes through verbatim. We limit ourselves to brief remarks where this is not quite true.

The first lemma, a generalisation of [74, Lemma 2.8], states that sizes of parts of basic optimal solutions are bounded below. It is proved via a compactness argument, which requires that there is a finite number of possible ϕ^*. This holds when $\mathcal{X}$ is bounded (we recall from Section 4.2.3 that the monochromatic pattern is bounded).

Lemma 4.32. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a bounded family of s-edge-colourings of K_k that has the extension property. Then there exists $\mu > 0$ such that for all $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$, we have $\alpha_i^* > \mu$ for all $i \in [r^*]$.*

The next lemma states that q is continuous in its second argument and follows from simple calculus.

Lemma 4.33 (Proposition 11,[76])**.** *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a family of s-edge-colourings of K_k. Then for any $r \in \mathbb{N}$, $\phi \in \Phi_{\mathcal{X},0}(r)$ and $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \Delta^r$ we have*

$$|q(\phi, \boldsymbol{\alpha}) - q(\phi, \boldsymbol{\beta})| < 2\log(s)\|\boldsymbol{\alpha} - \boldsymbol{\beta}\|_1.$$

Next, we have that almost optimal solutions have the property that the vertex weighting can be perturbed slightly to produce an optimal solution. This is another compactness argument and generalises [76, Claim 15].

Lemma 4.34. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a bounded family of s-edge-colourings of K_k. For all $\delta > 0$ there exists $\varepsilon > 0$ such that the following holds. Let $(r, \phi, \boldsymbol{\alpha}) \in \textsc{feas}_1(\mathcal{X})$ be such that $q(\phi, \boldsymbol{\alpha}) \geqslant Q(\mathcal{X}) - \varepsilon$. Then there exists $\boldsymbol{\alpha}^* \in \Delta^r$ such that $\|\boldsymbol{\alpha} - \boldsymbol{\alpha}^*\|_1 < \delta$ and $(r, \phi, \boldsymbol{\alpha}^*) \in \textsc{opt}_1(\mathcal{X})$.*

Our next tool is a version of [74, Lemma 2.10] which states that when a bounded family has the extension property, any vertex attached to a basic optimal solution with almost optimal contribution must be a clone of an existing vertex. (Recall that the extension property (Definition 4.4) guarantees this to be true when the contribution is optimal rather than almost optimal.)

Lemma 4.35. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a bounded family of s-edge-colourings of K_k that has the (strong) extension property. Then there exists $\eta > 0$ such that the following holds. Let $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$ and $\phi \in \Phi_{\mathcal{X},0}(r^* + 1)$ be such that $\phi|_{\binom{[r^*]}{2}} = \phi^*$ and $\mathrm{ext}(\phi, \boldsymbol{\alpha}^*) > Q(\mathcal{X}) - \eta$. Then $r^* + 1$ is a (strong) clone of some $i \in [r^*]$ under ϕ.*

The final result is a new lemma which allows us to find a common optimal vertex weighting given several optimal solutions with similar vertex weightings. It will be used in the proof of Theorem 4.11.

Lemma 4.36. *Let $s \geqslant 2$ and $k \geqslant 3$ be integers and let $\mathcal{X}$ be a family of s-edge-colourings of K_k. Let $r \geqslant 2$ be an integer and suppose that $\Phi \subseteq \Phi_{\mathcal{X},0}(r)$, and that for each $\phi \in \Phi$, we have $\boldsymbol{\alpha}_\phi \in \Delta^r$ such that $(r, \phi, \boldsymbol{\alpha}_\phi) \in \text{OPT}(\mathcal{X})$. Then for any $\varepsilon > 0$ there is $\delta > 0$ such that the following holds. If for all $\phi, \psi \in \Phi$ we have $\|\boldsymbol{\alpha}_\phi - \boldsymbol{\alpha}_\psi\|_1 < \delta$, then there is $\boldsymbol{\alpha}^* \in \Delta^r$ such that for all $\phi \in \Phi$ we have $(r, \phi, \boldsymbol{\alpha}^*) \in \text{OPT}(\mathcal{X})$ and $\|\boldsymbol{\alpha}_\phi - \boldsymbol{\alpha}^*\|_1 < \varepsilon$.*

Proof. Suppose that the lemma does not hold for some $\varepsilon > 0$. Then for each integer $m \geqslant 1$, there is $\Phi_m \subseteq \Phi_{\mathcal{X},0}(r)$ and a family $(r, \phi, \boldsymbol{\alpha}_\phi^{(m)})_{\phi \in \Phi_m}$ with $\|\boldsymbol{\alpha}_\phi - \boldsymbol{\alpha}_\psi\|_1 < \frac{1}{m}$ for all $\phi, \psi \in \Phi_m$ and such that, if there is $\boldsymbol{\alpha}^* \in \Delta^r$ so that $(r, \phi, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$ for all $\phi \in \Phi_m$, then $\|\boldsymbol{\alpha}_\phi^{(m)} - \boldsymbol{\alpha}^*\|_1 > \varepsilon$ for some $\phi \in \Phi_m$.

We can restrict to a subsequence where some ϕ either never appears or appears infinitely often. Since $\Phi_{\mathcal{X},0}(r)$ is finite, we may iteratively do this for all ϕ to obtain a subsequence $(m_\ell)_{\ell \in \mathbb{N}}$ such that for all $\ell \geqslant 1$ the set Φ_{m_ℓ} is equal to some fixed set Φ. Owing to the compactness of Δ^r, by restricting the sequence $(m_\ell)_{\ell \in \mathbb{N}}$ further, we may assume that for all $\phi \in \Phi$, $(\boldsymbol{\alpha}_\phi^{(m_\ell)})_\ell$ converges. Since for any $\phi, \phi' \in \Phi$ we have $\|\boldsymbol{\alpha}_\phi^{(m_\ell)} - \boldsymbol{\alpha}_{\phi'}^{(m_\ell)}\|_1 \to 0$ as $\ell \to \infty$, we get the same limit $\boldsymbol{\alpha}^*$ of the sequences $(\boldsymbol{\alpha}_\phi^{(m_\ell)})_\ell$ for $\phi \in \Phi$. Because $(r^*, \phi, \boldsymbol{\alpha}_\phi^{(m_\ell)})$ is optimal for all $\phi \in \Phi$ and $\ell \geqslant 1$, the continuity of q implies that $(r^*, \phi, \boldsymbol{\alpha}^*) \in \text{OPT}(\mathcal{X})$ for all $\phi \in \Phi$. This yields a contradiction, since we can choose L so that $\|\boldsymbol{\alpha}_\phi - \boldsymbol{\alpha}^*\|_1 < \varepsilon$ for all $\phi \in \Phi = \Phi_{m_L}$. □

4.6.2 Proof of Theorem 4.9

The proof of Theorem 4.9 is largely the same as that of [74, Lemma 3.1] to which the reader is referred for details. The idea is to repeatedly 'symmetrise' an almost optimal solution, and then show that actually not much changed during this process.

Sketch proof of Theorem 4.9. Let $\mathcal{X}$ be a bounded family of s-edge-colourings of K_k which has the extension property. Write $Q := Q(\mathcal{X})$. Let $\nu > 0$ be given. Let μ be the constant obtained from Lemma 4.32, let $0 < \gamma \ll \delta \ll \mu$ and obtain $\varepsilon^{1/4}$ by applying Lemma 4.34 with γ^2. Without loss of generality, we have $0 < \varepsilon \ll \gamma \ll \delta \ll \mu \ll \nu, 1$. Suppose that $(r, \phi, \boldsymbol{\alpha}) \in \text{FEAS}(\mathcal{X})$ has $q(\phi, \boldsymbol{\alpha}) > Q - \varepsilon$. By splitting up parts we can assume that every part α_i has equal size $1/r \ll \varepsilon$. The first main step in the proof is the *forwards symmetrisation* procedure which is the same as in [74]. The only property required of solutions $(r, \phi, \boldsymbol{\alpha})$ is a version (S′) of (S) stating that symmetrising produces a feasible solution and q does not decrease (we already used this to prove Fact 4.1).

(S′) In a feasible solution $(r, \phi, \boldsymbol{\alpha})$, for any $hj \in \binom{[r]}{2}$ with $\phi_{hj} \leqslant 1$, such that, after possibly swapping the labels h, j, then defining $\boldsymbol{\alpha}' \in \Delta^r$ with $\alpha'_h := \alpha_h + \alpha_j$ and $\alpha'_j := 0$ and $\alpha'_\ell := \alpha_\ell$ for all $\ell \in [r] \setminus \{h, j\}$ (the operation of *symmetrisation*), we have that $(r, \phi, \boldsymbol{\alpha}')$ is a feasible solution and $q(\phi, \boldsymbol{\alpha}') \geqslant q(\phi, \boldsymbol{\alpha})$.

This procedure yields Claim 4.6.2.1, which we will state after introducing notation. Let $\mathcal{V}_0 = \{\{1\}, \ldots, \{r\}\}$ and $\phi_0 = \phi$. Given $r' \leqslant r$ and colour templates $\psi' \in \Phi_{\mathcal{X},0}(r'), \psi \in \Phi_{\mathcal{X},0}(r)$ and a partition $\mathcal{V} = \{V_1, \ldots, V_{r'}\}$ of $[r]$ into r' parts, we say that $\psi =_{\mathcal{V}} \psi'$ if $\psi(i'j') = \psi'(ij)$ for all $i' \in V_i$, $j' \in V_j$ and $ij \in \binom{[r']}{2}$, and $\psi(i'i'') = \emptyset$ whenever $i', i'' \in V_i$ for all $i \in [r']$.

Claim 4.6.2.1 (See Claim 3.1.1 in [74])**.** *There is $f \in \mathbb{N}$ and $2 \leqslant r_f \leqslant \ldots \leqslant r_0 = r$ such that, after relabelling $[r]$, for all $i = 0, \ldots, f$, there is a partition $\mathcal{V}_i$ of $[r]$ with $|\mathcal{V}_i| = r_i$ and a colour template $\phi_i \in \Phi_{\mathcal{X},0}(r)$ such that the following hold.*

- *There is a single $x_i \in [r]$ such that $\mathcal{V}_i$ consists of the same elements as $\mathcal{V}_{i-1}$, except that x_i has been moved from one part to another and any empty parts are deleted;*
- *$\phi_i =_{\mathcal{V}_i} \psi_i$ where $\psi_i := \phi|_{\binom{[r_i]}{2}}$ and $\psi_f \in \Phi_{\mathcal{X},2}(r_f)$;*
- *$q(\phi_i, \boldsymbol{\alpha}) - q(\phi_{i-1}, \boldsymbol{\alpha}) \geqslant 0$.*

The claim is obtained by repeatedly applying (S′). We start with $\phi_0, \psi_0 = \phi_0$ and $\boldsymbol{\alpha}_0 = \boldsymbol{\alpha}$ and $\mathcal{V}_0$. While $\psi_i \notin \Phi_{\mathcal{X},2}(r_i)$, let $hj \in \binom{[r_i]}{2}$ be such that $|\psi_i(hj)| \leqslant 1$ as in (S′) and take x_{i+1} in the j-th class of $\mathcal{V}_i$. If x_{i+1} is a singleton, set $r_{i+1} := r_i - 1$, otherwise $r_{i+1} := r_i$. Let $\boldsymbol{\alpha}_{i+1}$ be obtained by moving $\alpha_{x_{i+1}}$ from $\alpha_{i,j}$ to $\alpha_{i,h}$ as in (S′). Take ϕ_{i+1} to be ϕ_i with the edges incident to x_{i+1} replaced with the edges incident to any vertex in the h-th class of $\mathcal{V}_i$ (which are all clones of each other). It can be shown that $q(\phi_{i+1}, \boldsymbol{\alpha}) = q(\psi_{i+1}, \boldsymbol{\alpha}_{i+1})$, which together with (S′) yields $q(\phi_{i+1}, \boldsymbol{\alpha}) \geqslant q(\phi_i, \boldsymbol{\alpha})$.

We next describe *backwards symmetrisation*, which goes backwards through the steps of forwards symmetrisation, removing parts with small contribution and assigning parts with large contribution into new groups depending on where they ended up in the forwards process. For $(r, \phi, \boldsymbol{\alpha}) \in \Phi_{\mathcal{X},0}(r)$ and $P \subseteq [r]$, write $q_x(P, \phi) :=$

$\frac{1}{r}\sum_{y\in P\setminus\{x\}}\log\phi_{xy}$ (recalling each $\alpha_y = \frac{1}{r}$). Let $\mathcal{V}_i := \{V_{i,j} : j \in [r_i]\}$. We do the following (backwards symmetrisation). Let us define $\mathcal{U}_f := \{U_f^0, \ldots, U_f^{r_f}\}$ by setting $U_f^j := V_{f,j}$ for all $j \in [r_f]$ and $U_f^0 := \emptyset$, and let $U_f := [r]$. For each $i = f-1, \ldots, 0$ define U_i and $\mathcal{U}_i = \{U_i^0, \ldots, U_i^{r_f}\}$ inductively as follows. Initialise with $U_i = [r]$ and $U_i^0 = \emptyset$.

1. *Add vertices with small contribution to the exceptional set*:

   (a) If $\frac{r}{|U_i|}q_{x_i}(U_i, \phi_i) < Q - \sqrt{\varepsilon}$, move x_i from U_i into U_i^0.

   (b) Repeat the following until no longer possible, updating U_i each time: If there is $y \in U_i$ such that $\frac{r}{|U_i|}q_y(U_i, \phi_i) < Q - \sqrt{\varepsilon}$, move y into U_i^0. (Note that y could be x_i if its contribution becomes too small after moving other vertices into U_i^0).

2. *Group x_i and newly non-exceptional vertices via comparison to ψ_f*: For each $j \in [r_f]$, let U_i^j be the restriction of U_{i+1}^j to $U_i \setminus \{x_i\}$. For each z in

$$B_i := (U_{i+1}^0 \cup \{x_i\}) \cap U_i,$$

   add z to the part U_i^j such that z looks most like a ϕ_f-clone of j under $\phi_i|_{U_i}$. That is, choose the j such that

$$|\{y \in U_i^j : |\phi_i(zy)| \geqslant 2\}| + \sum_{j'\in[r_f]\setminus\{j\}} |\{y \in U_i^{j'} : \phi_i(zy) \neq \phi_f(j'j)\}| \qquad (4.6.1)$$

   is minimal.

We note that after moving vertices we still have $U_i = [r] \setminus U_i^0$. The next claim is also unchanged. Its proof follows from the fact that any vertex moved into U_i^0 in (1) has much smaller than average contribution since the average contribution of a vertex is at least $Q - \varepsilon$. (It is closely related to the familiar procedure in graph theory of repeatedly removing a vertex of small degree, which cannot remove too many vertices in a graph with many edges.)

Claim 4.6.2.2 (See Claim 3.1.2, [74]). *For all $i = f, \ldots, 0$, we have $|U_i^0| \leqslant 2\sqrt{\varepsilon}r$.*

It suffices to show that the parts $U_0^1, \ldots, U_0^{r_f}$ are essentially the ones required by the theorem (so in particular, almost all parts are subparts of one of these). For this, we argue inductively for $i = f, \ldots, 0$. Let G_i be the complete graph with vertex set U_i. For each $x \in U_i$, let $j_x \in [r_f]$ be such that $x \in U_i^{j_x}$. Then we colour each $xy \in E(G_i)$ as follows:

- xy is *red* if $j_x \neq j_y$ and $\phi_i(xy) \subsetneq \psi_f(j_x j_y)$, so there are missing colours,
- xy is *blue* if either $j_x \neq j_y$ and $\phi_i(xy) \setminus \psi_f(j_x j_y) \neq \emptyset$, or $j_x = j_y$ and $|\phi_i(xy)| > 1$, so there are extra colours,
- xy is *green* otherwise.

The next claim is the heart of the proof.

Claim 4.6.2.3 (See Claim 3.1.3, [74])**.** *G_i has no blue edges.*

Sketch proof of claim. As promised, the proof is by backwards induction on i. After sketching the first part of the proof, which is almost identical to [74], we give more details at the end in showing that any blue edge leads to a contradiction, where the argument diverges slightly.

First, the claim is true for $i = f$ as every edge of G_f is green. Fix $i < f$ and let $n_i := |U_i| = r - |U_i^0| \geqslant (1 - 2\sqrt{\varepsilon})r$. Since G_{i+1} and G_i only differ at B_i, the induction hypothesis implies that every blue edge of G_i is incident with B_i. Let $\boldsymbol{\beta}_i := (|U_i^1|/n_i, \ldots, |U_i^{r_f}|/n_i) \in \Delta^{r_f}$. By construction, every $x \in U_i$ satisfies $\frac{r}{|U_i|} q_x(U_i, \phi_i) \geqslant Q - \sqrt{\varepsilon}$; that is, every non-exceptional vertex has large contribution. Since red edges at a vertex reduce its contribution, summing these contributions implies that

$$e_{\mathrm{red}}(G_i) \leqslant \varepsilon^{1/3} n_i^2 \quad \text{and} \quad q(\psi_f, \boldsymbol{\beta}_i) \geqslant Q - \varepsilon^{1/3},$$

so $\boldsymbol{\beta}_i$ is close to an optimal vertex weighting. By our choice of ε there is a vertex weighting $\boldsymbol{\alpha}'_i$ with $\|\boldsymbol{\alpha}'_i - \boldsymbol{\beta}_i\|_1 \leqslant \gamma^2$ such that $(r_f, \psi_f, \boldsymbol{\alpha}'_i)$ is in fact optimal. Relabelling parts and removing zero parts, we have a basic optimal solution $(\tilde{r}_i, \psi_f, \tilde{\boldsymbol{\alpha}}_i) \in \mathrm{OPT}^*(\mathcal{X})$, which is a slight abuse of notation since we really mean the restriction of ψ_f to $[\tilde{r}_i] \subseteq [r_f]$.

Next, we show that, in ϕ_i, every $z \in B_i$ is $\frac{\delta}{2}$-*close* to being a ψ_f-clone of some $j \in [\tilde{r}_i]$, meaning that one can change at most $\frac{\delta}{2}$ proportion of pairs at z so that its attachment is the same as the attachment of j in ψ_f. Here is where we need the extension property (Definition 4.4), which states that any vertex added to a basic optimal solution with *maximal* contribution Q is in fact a clone of one of the existing vertices. A version of this remains true for vertices which have *almost* maximal contribution (Lemma 4.35), and for a small perturbation of a basic optimal solution, as ϕ_i is of ψ_f (see [74, Lemmas 2.10–2.12] which are for monochromatic patterns but generalise straightforwardly). This version can be used to show that indeed any $z \in B_i$ is $\frac{\delta}{2}$-close to being a clone of some vertex as claimed above. Owing to the way we allocated z to a part in the backwards symmetrisation, this yields that z is added to a part $U_i^{j_z}$ such that z is $\frac{\delta}{2}$-close to a ψ_f-clone of j_z under ϕ_i. Since the closeness measure (4.6.1) counts non-green edges, a short calculation reveals that almost every edge in G_i incident to z is green:

$$d_{\text{green}}(z) \geqslant (1-\delta)n_i \quad \text{for all } z \in B_i.$$

The next step is to show that every $x \in U_i \setminus B_i$ has small red degree in G_i (and hence again almost every incident edge is green). Indeed, x can only have blue edges incident with the small set B_i, and thus too many red edges mean the contribution of x in ϕ_i is too small. We have $d_{\text{red}}(x) \leqslant \gamma n_i$ and hence $d_{\text{green}}(x) \geqslant (1-2\gamma)n_i$ for all $x \in U_i \setminus B_i$. Altogether, we have shown that

$$d_{\text{green}}(x) \geqslant (1-\delta)n_i \quad \text{for all } x \in U_i.$$

A short calculation shows that this implies that for every $y \in U_i$ we have $q_{j_y}(\psi_f, \boldsymbol{\alpha}'_i) \geqslant Q - \sqrt{\delta}$. The extension property now implies that $\boldsymbol{\beta}_i$ has the same support as $\boldsymbol{\alpha}'_i$ (which is the support of $\tilde{\boldsymbol{\alpha}}_i$ by definition); i.e. $[\tilde{r}_i]$. In particular, $U_i = \bigcup_{j \in [\tilde{r}_i]} U_i^j$. To see this implication, since every vertex in U_i^j with $\tilde{r}_i < j \leqslant r_f$ has large contribution and the parts outside of $[\tilde{r}_i]$ are small, j must have large contribution to ψ_f induced on $[\tilde{r}_i]$. So, by Lemma 4.35, j is a clone under ψ_f of some vertex j^* in $[\tilde{r}_i]$. In particular, $|\psi_f(jj^*)| \leqslant 1$ which contradicts $\psi_f \in \Phi_{\mathcal{X},2}(r_f)$.

Since $\|\boldsymbol{\beta}_i - \boldsymbol{\alpha}'_i\|_1 < \gamma^2$, and since our choice of μ implies that $\tilde{\alpha}_i \geqslant \mu$ for all $i \in [\tilde{r}_i]$, it follows that $|U_i^j| = \beta_{i,j}|U_i| \geqslant \mu n_i/2$ for all $j \in [\tilde{r}_i]$. Until here the proof is essentially the same as in [74].

We now complete the claim by comparing ϕ_i and the partition $\bigcup_{j \in [\tilde{r}_i]} U_i^j$ of U_i to $(\tilde{r}_i, \psi_f, \tilde{\boldsymbol{\alpha}}_i) \in \text{OPT}^*(\mathcal{X})$. Suppose for a contradiction that there is a blue edge y_1y_2 between parts $U_i^{j_1}, U_i^{j_2}$ where $j_1, j_2 \in [\tilde{r}_i]$ are distinct. Then there is some $c \in \phi_i(y_1y_2) \setminus \psi_f(j_1j_2)$. Since adding c to $\psi_f(j_1j_2)$ to form ψ'_f increases q, there is a forbidden pattern in ψ'_f (this is Fact 4.2), say on the vertices $j_1, j_2, j_3, \ldots, j_k$. Then as $d_{\text{green}}(x) \geqslant (1-\delta)n_i$ for all $x \in U_i$ and $|U_i^j| \geqslant \mu n_i/2 \geqslant k\delta n_i$ for all $j \in [\tilde{r}_i]$, there are $y_3, \ldots, y_k$ with $y_\ell \in U_i^{j_\ell}$ such that all the edges $y_\ell y_{\ell'}$ for $\ell, \ell' \in [k]$ are green except for y_1y_2. Thus, in ϕ_i, we can find the same forbidden pattern on $y_1, \ldots, y_k$ as found in ψ'_f above, which contradicts $\phi_i \in \Phi_{\mathcal{X},0}(r)$.

Suppose for a contradiction that there is a blue edge y_1y_2 inside a part U_i^j for $j \in [\tilde{r}_i]$. Then $|\phi_i(y_1y_2)| \geqslant 2$ and the argument is very similar: To obtain $(\tilde{r}_i + 1, \psi'_f, \tilde{\boldsymbol{\alpha}}'_i)$ from

$(\tilde{r}_i, \psi_f, \tilde{\boldsymbol{\alpha}}_i)$, add a clone j' of j in ψ_f and the colours in $\phi_i(y_1y_2)$ between j and j' and evenly split the previous weight of j between j and j', i.e. let $\tilde{\alpha}'_{i,j} = \tilde{\alpha}'_{i,j'} = \tilde{\alpha}_{i,j}/2$. We then have $q(\tilde{r}_i + 1, \psi'_f, \tilde{\boldsymbol{\alpha}}'_i) - q(\tilde{r}_i, \psi_f, \tilde{\boldsymbol{\alpha}}_i) \geqslant (\tilde{\alpha}_{i,j})^2 \log(2)/4 > 0$, hence ψ'_f contains a forbidden pattern, say on vertices $j, j', j_3, \ldots, j_k$. Then as above we find $y_3, \ldots, y_k$ with $y_\ell \in U_i^{j_\ell}$ with all the edges $y_\ell y_{\ell'}$ for $\ell, \ell' \in [k]$ green except for y_1y_2. In ϕ_i we can then find the same forbidden configuration on $y_1, \ldots, y_k$ as found in ψ'_f, which is again a contradiction. This completes the proof of the claim. $\square$

We showed in the claim that for each $i = 0, \ldots, f$ there is a partition $\bigcup_{j\in[\tilde{r}_i]} U_i^j$ of $[r]$ where $\tilde{r}_i \leqslant r_f$ such that $\sum_{j\in[\tilde{r}_i]} ||U_i^j| - \tilde{\alpha}_{i,j}n_i| \leqslant \gamma^2 n_i$ for some $(\tilde{r}_i, \psi_f, \tilde{\boldsymbol{\alpha}}_i) \in \text{OPT}^*(\mathcal{X})$. We now claim that setting $(r^*, \phi^*, \boldsymbol{\alpha}^*) := (\tilde{r}_0, \psi_f, \tilde{\boldsymbol{\alpha}}_0)$, and defining the sets $Y_j := U_0^j$ for all $j \in [r^*]$ and $Y_0 := [r] \setminus \bigcup_{j\in[r^*]} Y_j$ yields a partition satisfying (SO1)–(SO3). For part (SO1), setting $\beta_j := \sum_{j'\in Y_j} \alpha_{j'}$, we have that $r\beta_j = |U_0^j| = \beta_{0,j}n_0$ (recall that without loss of generality, we assumed $\alpha_i = 1/r$ for all i) so the inequality $\sum_{j\in[\tilde{r}_i]} ||U_i^j| - \tilde{\alpha}_{i,j}n_i| \leqslant \gamma^2 n_i$ for $i = 0$ implies $\sum_{j\in[r^*]} |r\beta_j/n_0 - \alpha_j^*| \leqslant \gamma^2$. Substituting $(1 - 2\sqrt{\varepsilon})r \leqslant n_0 \leqslant r$ and $\gamma, \varepsilon \ll \nu$ yields (SO1). For (SO2), by Claim 4.6.2.3 for $i = 0$, the auxiliary graph G_0 has no blue edges, that is, if $i' \in U_0^i = Y_i$ and $j' \in U_0^j = Y_j$ for $ij \in \binom{[r_f]}{2} \supseteq \binom{[r^*]}{2}$, then $\phi(i'j') \setminus \phi^*(ij) = \emptyset$ (recall that $\phi_0 = \phi$ and $\phi^* = \psi_f$).

Property (SO3) requires a bit of extra work. We know from Claim 4.6.2.3 that pairs within parts see at most one colour in ϕ_i but not that they all see the same colour. The following paragraph is true for all intermediate partitions $\mathcal{U}_i$, in particular for $i = 0$, which will yield the final property (SO3).

We start by showing that in each part U_i^j with $j \in [\tilde{r}_i]$ we see at most one colour. Suppose that there are $j \in [\tilde{r}_i]$, vertices $y_1, y_2, z_1, z_2 \in U_i^j$ and distinct colours c, c' such that $c \in \phi_i(y_1y_2), c' \in \phi_i(z_1z_2)$. Splitting part j into parts j_1, j_2 in $\tilde{\boldsymbol{\alpha}}_i$, each of size $\tilde{\alpha}_{i,j}/2$, and adding colours c, c' to the edge between j_1, j_2 increases $q(\psi_f, \tilde{\boldsymbol{\alpha}}_i)$ above Q. Therefore this colour template must contain a forbidden pattern involving j_1 and j_2. Assume that it uses colour c between j_1, j_2 and let $j_3, \ldots, j_k$ be the other vertices of the pattern. Since $d_{\text{green}}(x) \geqslant (1-\delta)n$ for all $x \in U_i$, we then find $y_3, \ldots, y_k$ with $y_\ell \in U_i^\ell$ such that the same forbidden pattern appears on $y_1, \ldots, y_k$ in ϕ_i, a contradiction. Thus for each $j \in [\tilde{r}_i]$ there is a unique colour c_j such that all edges inside U_i^j see no other colour than c_j.

Finally, for the second part of (SO3), suppose that ϕ^* is stable inside, in which case we must show that there are no colours inside any part. Indeed, suppose there are $x, y \in U_i^j$ with $c \in \phi_i(xy)$ for some colour c. Since the pattern is stable inside, splitting

j into j_1 and j_2 as above and letting $\phi_i(j_1j_2) = \{c\}$ creates a forbidden pattern on parts $j_1, j_2, \ldots, j_k$. Analogously to previous arguments we can show that this results in a forbidden pattern in ϕ_i, a contradiction. $\square$

4.6.3 Proof of Theorem 4.10

The idea of this proof is to apply the multicolour regularity lemma, and the coloured regularity partition thus obtained can be approximated by an almost optimal solution. We can then apply Theorem 4.9 to this solution.

We use standard terminology and tools related to Szemerédi regularity. Given a graph G, disjoint $A, B \subseteq V(G)$ and $0 \leqslant \delta \leqslant d \leqslant 1$, we say that $G[A, B]$ is *δ-regular* if $|d_G(X, Y) - d_G(A, B)| \leqslant \delta$ for all $X \subseteq A$, $Y \subseteq B$ with $|X|/|A|, |Y|/|B| \geqslant \delta$. Recall the definition from the beginning of Section 4.3.3 that $G[A, B]$ is *(δ, d)-regular* if it is δ-regular and $d_G(A, B) = d \pm \delta$. We further say that $G[A, B]$ is *$(\delta, \geqslant d)$-regular* if it is δ-regular and $d_G(A, B) \geqslant d - \delta$. Subscripts are omitted where they are clear from the context.

We use the multicolour regularity lemma, and the Embedding lemma, which says that if $0 < \varepsilon \ll d, 1/k$ and G is a sufficiently large k-partite graph with vertex classes $V_1, \ldots, V_k$ such that each $G[V_i, V_j]$ is $(\varepsilon, \geqslant d)$-regular, then $K_k \subseteq G$. See [61, Theorems 1.18 and 2.1] for full statements and relevant definitions.

Proof of Theorem 4.10. Let $\mathcal{X}$ be a bounded family of s-edge-colourings of K_k which has the extension property and let $\delta > 0$ be given. Let μ be the constant guaranteed by Lemma 4.32. We may assume without loss of generality that $0 < \delta \ll \mu \ll 1$. Let $0 < 1/n_0 \ll \gamma_1 \ll \gamma_2 \ll \varepsilon \ll \nu \ll \gamma \ll \delta \ll \mu \ll 1/s, 1/k$, and suppose that G is a graph on $n \geqslant n_0$ vertices satisfying the theorem hypothesis. As in [74], the multicolour regularity lemma yields a map RL from the set of all $\mathcal{X}$-free s-colourings $\chi : E(G) \to [s]$ to the set of pairs $(\mathcal{U}, \phi)$ consisting of an equipartition $\mathcal{U} = U_1 \cup \ldots \cup U_r$ of $V(G)$ and $\phi : \binom{[r]}{2} \to 2^{[s]}$ given by $\phi(ij) = \{c : \chi^{-1}(c)[U_i, U_j] \text{ is a } (\gamma_1, \geqslant \gamma_2)\text{-regular pair}\}$ such that

(R1) at most $s\gamma_1\binom{r}{2}$ pairs of sets in $\mathcal{U}$ are γ_1-irregular with respect to at least one colour,

(R2) between pairs U_i, U_j which are γ_1-regular with respect to all colours, at most $s\gamma_2|U_i||U_j|$ edges have colours not in $\phi(ij)$

and r does not depend on n. Thus we may suppose that $1/n \ll 1/r \ll \gamma_1$. Again as in [74, Definition 4.9, Proposition 4.10] we call $(\mathcal{U}, \phi)$ *popular* if its pre-image under RL satisfies $|\mathrm{RL}^{-1}(\mathcal{U}, \phi)| \geqslant e^{-3\varepsilon n^2} F(G; \mathcal{X})$, and define $\mathrm{Col}(G)$ to be the set of colourings χ mapping to a popular pair. By an adaptation of [74, Proposition 4.10] we have that

$$|\mathrm{Col}(G)| \geqslant (1 - e^{-2\varepsilon n^2}) F(G; \mathcal{X}). \tag{4.6.2}$$

In particular, the set of popular pairs is non-empty.

We now repeat [74, Claim 4.1] to show that if $\mathrm{RL}(\chi) = (\mathcal{U}, \phi)$ is popular, then setting $\boldsymbol{\alpha} := (|U_1|/n, \ldots, |U_r|/n)$ yields a near-optimal feasible solution $(r, \phi, \boldsymbol{\alpha}) \in \textsc{feas}_{\mathcal{X}}(s)$.

Claim 4.6.3.1. *Suppose that* $\mathrm{RL}(\chi) = (\mathcal{U}, \phi)$ *is popular. Then* $(r, \phi, \boldsymbol{\alpha}) \in \textsc{feas}_{\mathcal{X}}(s)$*. Moreover,* $q(\phi, \boldsymbol{\alpha}) \geqslant Q(\mathcal{X}) - 8\varepsilon$ *and there are at most* $s\gamma_2 n^2$ *edges* $xy \in E(G)$*,* $x \in U_i, y \in U_j$ *such that* $i = j$ *or* $\chi(xy) \notin \phi(ij)$*.*

Proof of claim. Feasibility is immediate: $\mathcal{U}$ is a partition of $V(G)$, so $\sum_i \alpha_i = 1$. Also, if ϕ contained an element of $\mathcal{X}$ then by the Embedding lemma, G coloured according to χ would contain the same element of $\mathcal{X}$, a contradiction.

The number of 'atypical' edges, i.e. edges in pairs specified by (R1) and (R2), is at most $s\gamma_1 \binom{r}{2} \times \left(\frac{n}{r}\right)^2 + \binom{r}{2} \times s\gamma_2 \left(\frac{n}{r}\right)^2 \leqslant 2s\gamma_2 n^2/3$. Together with edges internal to the U_i, this adds up to at most $2s\gamma_2 n^2/3 + r \times \binom{n/r}{2} \leqslant s\gamma_2 n^2$. All other edges $xy \in E(G)$ are between sets with $\chi(xy) \in \phi(ij)$, proving the last part of the claim.

Every colouring χ in $\mathrm{RL}^{-1}(\mathcal{U}, \phi)$ satisfies (R1) and (R2) above, so

$$\begin{aligned}|\mathrm{RL}^{-1}(\mathcal{U}, \phi)| \leqslant \underbrace{s^{r \times \left(\frac{n}{r}\right)^2}}_{\text{edges in each } U_i} &\times \underbrace{\binom{\binom{r}{2}}{\leqslant s\gamma_1 \binom{r}{2}} s^{s\gamma_1 \binom{r}{2} \times \left(\frac{n}{r}\right)^2}}_{\text{edges in irregular pairs}} \\ &\times \underbrace{\binom{\left(\frac{n}{r}\right)^2}{\leqslant s\gamma_2 \left(\frac{n}{r}\right)^2} s^{s\gamma_2 \left(\binom{r}{2} \times \frac{n}{r}\right)^2}}_{\text{low-density regular pairs}} \times \prod_{ij:\ \phi_{ij} \geqslant 1} \phi_{ij}^{e(G[U_i, U_j])}.\end{aligned}$$

Now, using standard properties of binomial coefficients, the first three terms multiply to at most $e^{\varepsilon n^2/2}$, whereas the last term is at most $e^{n^2 \sum_{ij} \alpha_i \alpha_j \log \phi_{ij}} = e^{q(\phi, \boldsymbol{\alpha}) \frac{n^2}{2}}$. Putting the last observations together, we get $|\mathrm{RL}^{-1}(\mathcal{U}, \phi)| \leqslant e^{(q(\phi, \boldsymbol{\alpha}) + \varepsilon)\binom{n}{2}}$. Since $(\mathcal{U}, \phi)$ is popular, we therefore have

$$F(G; \mathcal{X}) \leqslant e^{3\varepsilon n^2} |\mathrm{RL}^{-1}(\mathcal{U}, \phi)| \leqslant e^{3\varepsilon n^2} \times e^{(q(\phi, \boldsymbol{\alpha}) + \varepsilon)\binom{n}{2}} \leqslant e^{Q(\mathcal{X})\binom{n}{2} + 8\varepsilon\binom{n}{2}}, \tag{4.6.3}$$

giving a matching upper bound to (4.2.2) and hence a proof of Theorem 4.8 (note that we have not used the extension property, which is an assumption of Theorem 4.10 but not Theorem 4.8). Together with our assumption on $F(G;\mathcal{X})$ we have $q(\phi,\boldsymbol{\alpha}) \geqslant Q(\mathcal{X}) - 8\varepsilon$, as required. □

Let us fix a popular $(\mathcal{U},\phi)$ and its corresponding $\boldsymbol{\alpha}$. By Theorem 4.9 there exists a partition $[r] = Y_0 \cup \ldots \cup Y_{r^*}$ and a triple $(r^*,\phi^*,\boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$ which have the following properties:

(i) $\alpha_i^* > \mu$ for all $i \in [r^*]$ (from Lemma 4.32).

(ii) Setting $\beta_i := \sum_{j\in Y_i} \alpha_j$ for all $i \in [0,r^*]$ and $\boldsymbol{\beta} := (\beta_1,\ldots,\beta_{r^*})$, we have $\|\boldsymbol{\beta} - \boldsymbol{\alpha}^*\|_1 \leqslant \nu$.

(iii) For all $ij \in \binom{[r^*]}{2}, i' \in Y_i, j' \in Y_j$ we have $\phi(i'j') \subseteq \phi^*(ij)$.

(iv) For all $i \in [r^*]$ there is a colour $c_i \in [s]$ such that $\phi(i'j') \subseteq \{c_i\}$ for every $i'j' \in \binom{Y_i}{2}$. If $\mathcal{X}$ is stable inside, then for all $i \in [r^*]$ and every $i'j' \in \binom{Y_i}{2}$, we have $\phi(i'j') = \emptyset$.

We now merge each group Y_i of clusters of $\mathcal{U}$ to form a partition $\mathcal{W} = \{W_0,\ldots,W_{r^*}\}$ of $V(G)$ with $W_i = \bigcup_{j\in Y_i} U_j$. We have $|W_i| = \beta_i n$. We now want to show that most colourings χ in $\mathrm{RL}^{-1}(\mathcal{U},\phi)$ are regular with respect to $(\mathcal{W},\phi^*)$.

Note that $|W_0| = \beta_0 n = \sum_{i\in[r^*]}(\alpha_i^* - \beta_i^*)n \leqslant \|\boldsymbol{\alpha}^* - \boldsymbol{\beta}\|_1 n \leqslant \nu n$. Similarly, the other sets have sizes $|W_i| = \beta_i n \geqslant (\alpha_i^* - \nu)n \geqslant \mu n/2$. We also have that the number of edges in cluster W_i not coloured in c_i is small. More precisely, from the proof of Claim 4.6.3.1 at most $s\gamma_2 n^2$ edges $xy \in E(G)$ are within the clusters of $\mathcal{U}$ or between them with $\chi(xy) \notin \phi(ij)$. Together with property (iv) above, this gives

$$\sum_{i\in[r^*]} (e(W_i) - |\chi^{-1}(c_i)[W_i]|) \leqslant s\gamma_2 n^2, \tag{4.6.4}$$

or, if the pattern is stable inside, $\sum_{i\in[r^*]} e(W_i) \leqslant s\gamma_2 n^2$.

The following claim is based on a standard calculation that shows that the probability that a random bipartite graph is not γ-regular is exponentially small in n^2.

Claim 4.6.3.2. *At least* $(1-e^{-\Omega(\gamma^3 n^2)})|\mathrm{RL}^{-1}(\mathcal{U},\phi)|$ *colourings* χ *in* $\mathrm{RL}^{-1}(\mathcal{U},\phi)$ *satisfy:*

(†) *for all* $ij \in \binom{[r^*]}{2}$ *and* $c \in \phi^*(ij)$ *the bipartite graph* $\chi^{-1}(c)[W_i,W_j]$ *is* $(\gamma, \geqslant {\phi_{ij}^*}^{-1})$*-regular.*

Proof of claim. If $\chi \in \mathrm{RL}^{-1}(\mathcal{U},\phi)$ does not satisfy (†), there exist $i^*j^* \in \binom{[r^*]}{2}$ and $c^* \in \phi^*(i^*j^*)$ such that $\chi^{-1}(c^*)[W_{i^*},W_{j^*}]$ is not $(\gamma, \geqslant {\phi^*_{ij}}^{-1})$-regular. By [74, Proposition 4.4] this implies the existence of $X \subseteq W_{i^*}$ and $Y \subseteq W_{j^*}$ of size $\gamma|W_{i^*}|$ and $\gamma|W_{j^*}|$ respectively such that $d(\chi^{-1}(c^*)[X,Y]) < (1-\gamma/2){\phi^*_{i^*j^*}}^{-1}$.

Therefore, the number of such colourings is at most

$$\underbrace{s^{\nu n^2}}_{\substack{\text{edges incident}\\ \text{to } W_0}} \times \underbrace{s^{s\gamma_2 n^2}}_{\substack{\text{edges in } W_i\\ \chi(e)\neq c_i}} \times \underbrace{\binom{r^*}{2} s \binom{|W_{i^*}|}{\gamma|W_{i^*}|}\binom{|W_{j^*}|}{\gamma|W_{j^*}|}}_{\text{choice for } i^*j^*, c^*, X, Y} \times C \times D \times \prod_{ij\neq i^*j^*} {\phi^*_{ij}}^{|W_i||W_j|}$$

where

$$\begin{aligned} C &= \sum_{\ell \leqslant ({\phi^*_{i^*j^*}}^{-1} - \gamma/2s)|X||Y|} \binom{|X||Y|}{\ell} (\phi^*_{i^*j^*}-1)^{|X||Y|-\ell} \\ &\leqslant e^{-\gamma^2 \phi^*_{i^*j^*}|X||Y|/12s^2} {\phi^*_{i^*j^*}}^{|X||Y|} \qquad \text{(using [74, Corollary 4.8])} \\ &\leqslant e^{-\gamma^2\mu^2 n^2/48s} {\phi^*_{i^*j^*}}^{|X||Y|} \end{aligned}$$

serves as an upper bound for the number of colour choices for $G[X,Y]$ from above while $D = {\phi^*_{i^*j^*}}^{|W_{i^*}||W_{j^*}|-|X||Y|}$ bounds the number of colour choices for the rest of $G[W_{i^*},W_{j^*}]$ from above. Similarly to the previous claim, the product of the terms except $e^{-\gamma^2\mu^2n^2/48s}\prod {\phi^*_{ij}}^{|W_i||W_j|}$ is at most $e^{O(\nu n^2)}$ and we have

$$\begin{aligned} \prod {\phi^*_{ij}}^{|W_i||W_j|} &= e^{q(\phi^*,\boldsymbol{\beta})\binom{n}{2}} \leqslant e^{(q(\phi^*,\boldsymbol{\alpha}^*)+2\nu\log(s))\binom{n}{2}} \qquad \text{(using Lemma 4.33)} \\ &\overset{(4.2.1)}{\leqslant} F(n;\mathcal{X}) \cdot e^{2\nu\log(s)\binom{n}{2}+O(n)} \\ &\leqslant |\mathrm{RL}^{-1}(\mathcal{U},\phi)| \times e^{7\varepsilon\binom{n}{2}} \times e^{3\nu\log(s)\binom{n}{2}} \quad \text{(since } (\mathcal{U},\phi) \text{ is popular).} \end{aligned}$$

Thus the number of colourings not satisfying (†) is at most

$$e^{O(\nu n^2)-\gamma^2\mu^2n^2/48s}|\mathrm{RL}^{-1}(\mathcal{U},\phi)| \leqslant e^{-\Omega(\gamma^3 n^2)}|\mathrm{RL}^{-1}(\mathcal{U},\phi)|,$$

as required. □

We are now almost done. It remains to adjust $\mathcal{W}$ so that $|W_i| = \alpha^*_i n$ (recall that we suppress rounding) and W_0 is empty. For each $i \geqslant 1$, let V_i be obtained by taking any $\min\{\alpha^*_i n, |W_i|\}$ vertices from W_i, and then distributing the remaining vertices of G arbitrarily so that $|V_i| = \alpha^*_i n$. This gives property (SG1) from the theorem statement. Recall that $|W_i| = \beta_i n$ with $\|\boldsymbol{\beta}-\boldsymbol{\alpha}^*\|_1 \leqslant \nu$, so $\sum_i |V_i \triangle W_i| = \|\boldsymbol{\beta}-\boldsymbol{\alpha}^*\|_1 n \leqslant \nu n$. Since the parts of $\boldsymbol{\alpha}^*$ have size at least μ and $\nu \ll \gamma \ll \mu$, each colouring $\chi \in \mathrm{R}L^{-1}(\mathcal{U},\phi)$ satisfying (†) with respect to $\mathcal{W}$ now satisfies it with respect to $\mathcal{V}$ with a larger regularity parameter 2γ (see [74, Proposition 4.5]), that is, each $\chi^{-1}(c)[V_i, V_j]$ is $(2\gamma, \geqslant {\phi^*_{ij}}^{-1})$-regular. This gives property (SG2) from the theorem statement. Finally, recall that by (4.6.4) at most $s\gamma_2 n^2$ edges of G are within the clusters of $\mathcal{W}$ and not in colour c_i, so at most $s\gamma_2 n^2 + \nu n^2 \leqslant \delta n^2$ are within the clusters of $\mathcal{V}$ and not in colour c_i, giving property (SG3). For patterns that are stable inside, there are simply at most $s\gamma_2 n^2$ edges within clusters of $\mathcal{W}$, and hence at most δn^2 within the clusters of $\mathcal{V}$, giving the second part of property (SG3).

For each popular $(\mathcal{U}, \phi)$, the above holds for at least $(1 - e^{-\Omega(\gamma^3 n^2)})|\mathrm{RL}^{-1}(\mathcal{U}, \phi)|$ colourings in $\mathrm{RL}^{-1}(\mathcal{U}, \phi)$. Thus, using (4.6.2), we get a total of

$$\begin{aligned}\sum_{\mathcal{U},\phi}(1 - e^{-\Omega(\gamma^3 n^2)})|\mathrm{RL}^{-1}(\mathcal{U}, \phi)| &= (1 - e^{-\Omega(\gamma^3 n^2)})|\mathrm{Col}(G)| \\ &\geqslant (1 - e^{-\Omega(\gamma^3 n^2)})(1 - e^{-2\varepsilon n^2})F(G; \mathcal{X}) \\ &\geqslant (1 - e^{-\varepsilon n^2})F(G; \mathcal{X})\end{aligned}$$

colourings for which properties (SG1)–(SG3) hold. □

4.6.4 Proof of Theorem 4.11

Proof of Theorem 4.11. We are given constants k, s, δ and may choose additional constants $\eta, \xi, \varepsilon, \delta_1, \delta_2, n_0$ so that, without loss of generality, we have $0 < 1/n_0 \ll \eta \ll \xi \ll \varepsilon \ll \delta_1 \ll \delta_2 \ll \delta \ll \mu, 1/k, 1/s, 1/R$, where for every $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \mathrm{OPT}^*(\mathcal{X})$, we have $r^* \leqslant R$ (by boundedness), and $\alpha^*_i \geqslant \mu$ for all $i \in [r^*]$ (by Lemma 4.32), and also $Q(\mathcal{X}) - o(1) \leqslant (\log F(n; \mathcal{X}))/\binom{n}{2} < Q(\mathcal{X}) + \varepsilon$ for all $n \geqslant n_0$ (by Theorem 4.8).

Claim 4.6.4.1 (See Lemma 2.14, [75]). *If there is an $\mathcal{X}$-extremal graph which is not complete partite, then there is an $\mathcal{X}$-extremal graph which is complete partite with a part of size one.*

Sketch proof of claim. The proof of [75, Lemma 2.14] goes through verbatim, by property (S) at the beginning of this section. The proof proceeds by *symmetrising*, that is, successively replacing a vertex u of an $\mathcal{X}$-extremal graph G by a twin of another vertex v which is not adjacent to u. The graph G_{uv} obtained in this way is such that *both* G_{uv} and G_{vu} are extremal. Indeed, if $C(J)$ is the set of valid colourings of a graph J, we have $F(G; \mathcal{X}) = \sum_{\chi \in C(G-u-v)} \chi_u \chi_v$ where e.g. χ_u is the number of extensions of χ to $G - v$, since u and v are not adjacent and every forbidden colouring in $\mathcal{X}$ is on a clique. Thus $F(G_{uv};\mathcal{X}) = \sum_{\chi\in C(G-u-v)} \chi_v^2$ and so

$$0 \leqslant \sum_{\chi\in C(G-u-v)} (\chi_u - \chi_v)^2 = F(G_{uv};\mathcal{X}) - 2F(G;\mathcal{X}) + F(G_{vu};\mathcal{X}) \leqslant 0.$$

The choice of which pair u, v to symmetrise at each step can be made so that the final graph has a part of size one. □

Let G be an $\mathcal{X}$-extremal graph on $n \geqslant n_0$ vertices. By Claim 4.6.4.1 either $G' := G$ is complete partite or there is a complete partite extremal graph G' with a part of size one with the same number of vertices. Let $\mathcal{W} = W_1 \cup \ldots \cup W_r$ be the partition of $V(G')$. Apply Theorem 4.10 with regularity parameter δ_1 to G'. Let $\chi : E(G') \to [s]$ be one of the $(1 - e^{-\varepsilon n^2})F(G';\mathcal{X})$ 'typical' $\mathcal{X}$-free colourings of G' described in Theorem 4.10. Then there exist $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \mathrm{OPT}^*(\mathcal{X})$ and a partition $\mathcal{V} = V_1 \cup \ldots \cup V_{r^*}$ of $V(G')$ satisfying properties (SG1)–(SG3).

Since $\mathcal{X}$ is hermetic and hence stable inside, (SG3) implies that one can change at most $\delta_1 {r^*}^2 n^2$ adjacencies to obtain the complete r^*-partite graph $K[V_1, \ldots, V_{r^*}]$. Thus our partition $W_1, \ldots, W_r$ must be close to $V_1, \ldots, V_{r^*}$: after relabelling we have $|W_i \triangle V_i| \leqslant 2r^*\sqrt{s\delta_1}n$ and so $|W_i| = (\alpha_i^* \pm \delta_2)n > \mu n/2$ for all $i \in [r^*]$ (in particular, $r \geqslant r^*$) and the union of the $r - r^*$ other parts W_j has size at most $r^*\sqrt{s\delta_1}n \leqslant \delta_2 n$.

We now proceed to show that G' is complete r^*-partite, which will also imply that $G = G'$ and together with the previous observations give (SE1). Suppose for a contradiction that $r > r^*$, i.e. that G' has at least one small part W_{r^*+1}. Let $x \in W_{r^*+1}$ be an arbitrary vertex in the small part. We define an extension ϕ of ϕ^* to $r^* + 1$ parts which records the majority colour in χ between x and each of $W_1, \ldots, W_{r^*}$. More precisely, let $\phi : \binom{[r^*+1]}{2} \to 2^{[s]}$ be such that $\phi(ij) := \phi^*(ij)$ for all $ij \in \binom{[r^*]}{2}$ and $\phi(i, r^* + 1) := \{c \in [s] :\ d_{\chi^{-1}(c)}(x, W_i) \geqslant |W_i|/s\}$.

Claim 4.6.4.2. $\phi \in \Phi_{\mathcal{X},1}(r^* + 1)$.

Proof of claim. We have that $\phi^* \in \Phi_{\mathcal{X},2}(r^*)$ so every pair ij in $\binom{[r^*]}{2}$ satisfies $\phi_{ij} = \phi^*_{ij} \geqslant 2$. There are $|W_i|$ edges between x and W_i coloured in s colours, so by the pigeonhole principle at least one colour will appear $|W_i|/s$ times, thus $|\phi(i, r^*+1)| \geqslant 1$.

We now need to show that ϕ is $\mathcal{X}$-free. Suppose not. Then since ϕ^* is $\mathcal{X}$-free, there is a copy of some $X \in \mathcal{X}$ containing the vertex $r^* + 1$, say with vertex set $\{r^* + 1, i_1, \ldots, i_{k-1}\}$ and colouring σ. Suppose that $a \in V(X)$ is the vertex mapped to $r^* + 1$. For each $j \in [k - 1]$, consider the neighbourhood N_j of x in colour $\sigma(r^* + 1, i_j)$

in V_{i_j}. Each N_j has size $|N_j| \geqslant |W_{i_j}|/s - |W_{i_j} \triangle V_{i_j}| \geqslant |V_{i_j}|/2s$. Here we are using $(|V_{i_j}| \pm 1)/n = \alpha_i^* \geqslant \mu \gg \delta_1, 1/s, 1/r^*$ and the fact that $|W_{i_j} \triangle V_{i_j}| \leqslant \delta n$. For each pair of indices i_h, i_j with colour $c := \sigma(i_h i_j)$ in X we have that $c \in \phi^*(i_h i_j)$, so the bipartite graph $\chi^{-1}(c)[V_{i_h}, V_{i_j}]$ is $(\delta_1, {\phi^*_{i_h i_j}}^{-1})$-regular by property (SG2). Then $\chi^{-1}(c)[N_h, N_j]$ is $(2s\delta_1, {\phi^*_{i_h i_j}}^{-1})$-regular. The Embedding lemma then gives a copy of $X - a$ in the neighbourhood of x, which together with x gives a copy of X in G'. This contradicts the fact that χ is $\mathcal{X}$-free, thus proving the claim. □

However, the family $\mathcal{X}$ is hermetic, so no such ϕ exists. Thus $r = r^*$ and since $|W_i| \geqslant \mu n/2$ for all $i \in [r^*]$, G' has no part of size 1 and thus $G = G'$. Thus we have shown that for each typical χ there is $(r^*, \phi^*, \boldsymbol{\alpha}^*)$ such that

1. (SE1) holds with a smaller error term δ_2, i.e. G is a complete r^*-partite graph whose i-th part W_i has size $(\alpha_i^* \pm \delta_2)n$ for all $i \in [r^*]$, and
2. (SE2.1) holds, i.e. $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$, and (SE2.2) holds with a smaller error δ_1, i.e. $\chi^{-1}(c)[W_i, W_j]$ is $(\delta_1, {\phi^*_{ij}}^{-1})$-regular for all $ij \in \binom{[r^*]}{2}$ and $c \in \phi^*(ij)$.

Note that thus far, $(r^*, \phi^*, \boldsymbol{\alpha}^*)$ may depend on χ. Observe that G being r^*-partite, where r^* came from an arbitrary χ, means that r^* in fact is independent of the choice of χ. We now argue that we may also choose $\boldsymbol{\alpha}^*$ independently of χ (which is at least a priori possible as (1) means each of the $\boldsymbol{\alpha}^*$ are all very close). For this, note that for a given ϕ^*, we may use the same choice of $\boldsymbol{\alpha}^*_{\phi^*}$ for any χ that uses a triple involving ϕ^*, maintaining the above properties. This gives rise to a family $\Phi \subseteq \Phi_{\mathcal{X},2}(r^*)$ such that for each typical χ, there is $(r^*, \phi^*, \boldsymbol{\alpha}^*_{\phi^*})$ with $\phi^* \in \Phi$ for which (1) and (2) hold for χ. Lemma 4.36 applies to show that there is $\boldsymbol{\alpha}^*$ such that for all $\phi^* \in \Phi$, $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}(\mathcal{X})$ and $\|\boldsymbol{\alpha}_{\phi^*} - \boldsymbol{\alpha}^*\|_1 < \delta/2$. In fact, we get $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \textsc{opt}^*(\mathcal{X})$ since for each $i \in [r^*]$ we have $\alpha^*_{\phi^*,i} > \mu$ and $|\alpha_i^* - \alpha^*_{\phi^*,i}| < \delta/2$, so $\alpha_i^* > \mu/2$. Since for each typical χ and $i \in [r^*]$ we have $||W_i| - \alpha_i^*| \leqslant ||W_i| - \alpha^*_{\phi^*,i}| + |\alpha^*_{\phi^*,i} - \alpha_i^*| \leqslant \delta_2 + \delta/2 \leqslant \delta$, we have found choices of r^* and $\boldsymbol{\alpha}^*$ independent of χ such that the properties (SE1), (SE2.1) and (SE2.2) hold. It remains to show that for most of the typical colourings χ, our choice of $(r^*, \phi^*, \boldsymbol{\alpha}^*)$ satisfies (SE2.3), which is somewhat harder.

We say that a vertex $x \in G$ has *large contribution* if

$$\log F(G; \mathcal{X}) \geqslant \log F(G - x; \mathcal{X}) + (Q(\mathcal{X}) - \xi)n.$$

Claim 4.6.4.3 (See Lemma 2.15, [75]). *Every vertex in G has large contribution.*

Sketch proof of claim. Again, this proof goes through verbatim. If x did not have large contribution, then since it has many twins all necessarily with the same contribution, we would have many vertices with small contribution. A calculation using the Cauchy-Schwarz inequality shows that $\log F(G;\mathcal{X}) \leqslant \log F(G-T;\mathcal{X}) + (Q(\mathcal{X})-\xi)|T|n$, where T is the set of twins of x and satisfies $\mu n/2 \leqslant |T| \leqslant (1-\mu/2)n$ since every $\alpha_i^* \geqslant \mu$. Apply Theorem 4.8 to $G-T$ with parameter η (noting that $|G-T| \geqslant \mu n_0/2 \gg 1/\eta$) to obtain $\log F(G-T;\mathcal{X}) \leqslant (Q(\mathcal{X})+\eta)\binom{n-|T|}{2}$. Together with $\log F(G;\mathcal{X}) \geqslant Q(\mathcal{X})\binom{n}{2} - O(n)$, we obtain a contradiction since $\eta \ll \xi$. □

Applying the claim twice yields that for every $x, y \in V(G)$ we have

$$\log F(G;\mathcal{X}) \geqslant \log F(G-x-y;\mathcal{X}) + (Q(\mathcal{X})-\xi)(n+n-1). \tag{4.6.5}$$

We say that an $\mathcal{X}$-free colouring χ of G is *locally good* if there is ϕ^* such that $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \text{OPT}^*(\mathcal{X})$ and

(LG1) for all $ij \in \binom{[r^*]}{2}$ and $c \in \phi^*(ij)$, we have that $\chi^{-1}(c)[W_i, W_j]$ is $(\delta_1, {\phi_{ij}^*}^{-1})$-regular;

(LG2) for all $hij \in \binom{[r^*]}{3}$ and vertices $x \in W_i$ and $y \in W_j$, and colours $c \in \phi^*(ih)$ and $c' \in \phi^*(jh)$, we have $|N_{\chi^{-1}(c)}(x, W_h) \cap N_{\chi^{-1}(c')}(y, W_h)| \geqslant \delta_2 |W_h|$.

Claim 4.6.4.4 (See Lemma 2.11, [75]). *Every locally good valid colouring χ of G is* perfect, *meaning that $\chi(xy) \in \phi^*(ij)$ for all $x \in W_i$ and $y \in W_j$ and $ij \in \binom{[r^*]}{2}$.*

Proof of claim. Let χ be such a colouring and suppose, for contradiction, that there are $x \in W_i$ and $y \in W_j$ with $\chi(xy) = c \notin \phi^*(ij)$. By Fact 4.2, adding c to $\phi^*(ij)$ creates $K \cong K_k$ coloured by $\sigma : E(K) \to [s]$ which is an element of $\mathcal{X}$. Note that $i, j \in V(K)$, $\sigma(ij) = c$ and $\sigma(hf) \in \phi^*(hf)$ for all other pairs $hf \in \binom{[r^*]}{2} \setminus \{ij\}$. We now show that xy lies in a copy of K in G coloured by χ.

For every $h \in V(K) \setminus \{i, j\}$ let $W'_h := N_{\chi^{-1}(\sigma(ih))}(x, W_h) \cap N_{\chi^{-1}(\sigma(jh))}(y, W_h)$. By (LG2), we have $|W'_h| \geqslant \delta_2 |W_h|$. By (LG1), for any $h, h' \in V(K) \setminus \{i, j\}$ we have that $\chi^{-1}(\sigma(hh'))[W_i, W_j]$ is $(\delta_1/\delta_2, {\phi_{hh'}^*}^{-1})$-regular. Thus by the Embedding lemma we can find a copy of $K - i - j$ coloured according to σ with each vertex $h \in V(K) \setminus \{i, j\}$ mapped to W'_h. By definition of the sets W'_h, any such copy, together with x and y, forms a copy of K coloured according to σ. □

Thus, to complete the proof of (SE2.2), it suffices to show that most colourings are locally good.

Claim 4.6.4.5 (See [75], Claims 2.12.2–2.12.4)**.** *At least* $(1-e^{-\varepsilon n})F(G;\mathcal{X})$ *valid colourings* χ *of* G *are locally good.*

Sketch proof of claim. We note that the locally good condition in [75] is stronger than the one here. Recall that for $(1-e^{-\varepsilon n^2})F(G;\mathcal{X})$ valid colourings there is ϕ^* (which depends on the colouring) for which (LG1) holds, so it suffices to show that $(1-e^{-\delta_1 n})F(G;\mathcal{X})$ valid colourings satisfy (LG2) for that same ϕ^*. So suppose this does not hold. Then, by averaging, there are two vertices x, y such that at least $\binom{n}{2}^{-1}e^{-\delta_1 n}F(G;\mathcal{X})$ valid colourings of G do satisfy (LG1) but do not satisfy (LG2) at x, y. By (4.6.5) the number of such colourings is at least

$$\binom{n}{2}^{-1} e^{-\delta_1 n}\left(F(G-x-y;\mathcal{X})\cdot e^{(Q(\mathcal{X})-\xi)(n+n-1)}\right) \geqslant F(G-x-y;\mathcal{X})\cdot e^{(Q(\mathcal{X})-2\delta_1)(n+n-1)}.$$

Again by averaging, there is a valid colouring χ of $G-x-y$ and associated colour template $\phi^* : \binom{[r^*]}{2} \to 2^{[s]}$ from (LG1), with at least $e^{(Q(\mathcal{X})-3\delta_1)(n+n-1)}$ valid extensions $\overline{\chi}$ to G which do not satisfy the (LG2) condition at x, y. Averaging once more, there are $c_x, c_y \in [s]$ and $h^* \in [r^*]$ such that at least $e^{(Q(\mathcal{X})-4\delta_1)(n+n-1)}$ valid extensions of χ violate (LG2) at x, y with part W_{h^*} and colours c_x, c_y. For each such extension $\overline{\chi}$, we define colour templates $\phi^*_{\overline{\chi},x}, \phi^*_{\overline{\chi},y}$ on r^*+1 parts which are extensions of ϕ^* by taking popular colours at x, y, as follows: for every $h \in [r^*]$, let $\phi^*_{\overline{\chi},x}(hf) := \phi^*(hf)$ for all $hf \in \binom{[r^*]}{2}$, and let $\phi^*_{\overline{\chi},x}(\{h, r^*+1\}) := \{c \in [s] : |N_{\overline{\chi}^{-1}(c)}(x) \cap W_h| \geqslant \delta_1|W_h|\}$, and similarly for $\phi_{\overline{\chi},y}$. These colour templates are valid, via an almost identical argument to Claim 4.6.4.2. Choose ϕ^*_x, ϕ^*_y that appear most frequently among the $\phi^*_{\overline{\chi},x}, \phi^*_{\overline{\chi},y}$. The number of extensions $\overline{\chi}$ with these ϕ^*_x, ϕ^*_y is at least $e^{(Q(\mathcal{X})-5\delta_1)(n+n-1)}$.

We next argue that this lower bound on the number of valid extensions implies that $\mathrm{ext}(\phi^*_x, \boldsymbol{\alpha}^*) \geqslant Q(\mathcal{X}) - \sqrt{\delta}$, say (and similarly for ϕ^*_y). For this implication, note that since ϕ^*_x records all colours appearing in at least a δ_1 fraction and we have part sizes $|W_i| = (\alpha^*_i \pm \delta)n$, the logarithm of the number of valid extensions of χ to G (or, more precisely, to $G - xy$) corresponds closely to $\mathrm{ext}(\phi^*_x, \boldsymbol{\alpha}^*) + \mathrm{ext}(\phi^*_y, \boldsymbol{\alpha}^*)$. This gives the desired conclusion since $\mathrm{ext}(\phi^*_y, \boldsymbol{\alpha}^*) \leqslant Q(\mathcal{X})$ (see Proposition 2.6 in [74]).

Recall that $\mathcal{X}$ has the strong extension property since it has the extension property and is hermetic. Thus we can apply Lemma 4.35 to see that in r^*+1 is a strong clone of i' under ϕ^*_x for some $i' \in [r^*]$, and r^*+1 is a strong clone of j' under ϕ^*_y for some $j' \in [r^*]$. If i, j are such that $x \in W_i$ and $y \in W_j$, then $i = i'$ and $j = j'$ since, for example, $i \neq i'$ would imply $\emptyset = \phi^*_x(\{i, r^*+1\}) = \phi^*_x(ii') = \phi^*(ii')$ which is a contradiction to $\phi^* \in \textsc{opt}^*(\mathcal{X})$.

Finally we bound from above the number of valid extensions of χ to $\overline{\chi}$ that follow ϕ^*_x, ϕ^*_y to get a contradiction: we choose at most $s\delta_1 n$ vertices z in $V(G) \setminus \{x, y\}$ where at least one of xz, yz is not coloured according to the colour template ϕ^*, and at most δ_2 vertices z in W_{h^*} for which $(\chi(xz), \chi(yz)) = (c, c')$. The remaining at least $(1 - \delta_2)|W_{h^*}| - s\delta_1 n > \mu n/2$ vertices z in W_{h^*} are coloured according to the colour template except that $(\chi(xz), \chi(yz)) = (c, c')$ is forbidden. Since, for most vertices, there are at most $|\phi^*(ih^*)||\phi^*(jh^*)| - 1$ choices of $(\chi(xz), \chi(yz))$ as (c, c') is forbidden, rather than the $|\phi^*(ih^*)||\phi^*(jh^*)|$ we expect, we obtain fewer than $e^{(Q(\mathcal{X})+O(\delta))(n+n-1)} \cdot \left(\frac{s-1}{s}\right)^{\mu n/2} < e^{(Q(\mathcal{X})-5\delta_1)(n+n-1)}$ extensions, a contradiction. □

This completes the proof of the theorem. □

4.7 Concluding remarks

In this chapter we studied the *generalised Erdős-Rothschild problem* on maximising the number of s-edge-colourings of graphs on n vertices where each colouring we count must not contain any forbidden colouring in some given family $\mathcal{X}$ of s-edge-coloured cliques K_k. Our main results were for two specific forbidden families: the dichromatic triangle family (Theorem 1.21) and the family of improperly coloured cliques (Theorem 1.24). Along the way, we adapt results of [74, 75] on monochromatic colour patterns to obtain some general results (Section 4.3) which are likely to be useful for further cases.

4.7.1 A more general framework

A key property of families $\mathcal{X}$ in many of our arguments was boundedness: that there is at least one basic optimal solution and every basic optimal solution has a bounded number of parts. For example, Ramsey's theorem guarantees that the monochromatic pattern $K_k^{(1)}$ is bounded (recall Section 4.2.3).

As we mentioned in Section 4.2, our framework does not capture all of the colour patterns we would like to study. Indeed, in Lemma 4.7, we showed that the family $(K_3^{(3)}, 3)$ for the rainbow triangle with three colours is not bounded.

We can introduce a generalisation Problem $Q_t^\bullet(\mathcal{X})$ of Problem $Q_t(\mathcal{X})$ that extends the set of feasible solutions by allowing vertex weights (the bullet represents the loops at vertices present in the new problem).

Suppose we have $r \in \mathbb{N}$ and a function $\phi : \binom{[r]}{\leqslant 2} \to 2^{[s]}$, which maps singletons and pairs of vertices to sets of colours. Let $\boldsymbol{a} = (a_1, \ldots, a_r) \in \mathbb{N}^r$ be an integer vector. We define the *$\boldsymbol{a}$-blow-up of ϕ* to be the edge-multicoloured graph $G_\phi(\boldsymbol{a})$ with parts $V_1, \ldots, V_r$ such that $|V_i| = a_i$, and each $G[V_i]$ has no edges if $\phi(i) = \emptyset$, and otherwise it is a clique and every edge receives all of the colours $\phi(i)$; and for every distinct $i, j \in [r]$, $G[V_i, V_j]$ has no edges if $\phi(ij) = \emptyset$, and otherwise it is a complete bipartite graph and every edge receives all of the colours $\phi(ij)$. Given a family $\mathcal{X}$ of forbidden colourings of K_k, we say that ϕ is *$\mathcal{X}$-free* if the $(k, \ldots, k)$-blow up $G_\phi((k, \ldots, k))$ of ϕ is $\mathcal{X}$-free.

We define $\Phi^\bullet_{\mathcal{X},t}(r)$ to be the set of $\mathcal{X}$-free colourings $\phi : \binom{[r]}{\leqslant 2} \to 2^{[s]}$ such that $|\phi(ij)| \geqslant t$ for all $i \neq j$ and each $\phi(i)$ satisfies $|\phi(i)| = 0$ or $|\phi(i)| \geqslant t$.

Problem $Q^\bullet_t(\mathcal{X})$.
Maximise

$$q(\phi, \boldsymbol{\alpha}) := \sum_{i \in [r]} \alpha_i^2 \log |\phi(i)| + 2 \sum_{ij \in \binom{[r]}{2} : \phi(ij) \neq \emptyset} \alpha_i \alpha_j \log |\phi(ij)|$$

subject to $r \in \mathbb{N}$ and $\phi \in \Phi^\bullet_{\mathcal{X},t}(r)$ and $\boldsymbol{\alpha} \in \Delta^r$.

Any feasible solution of Problem $Q_t(\mathcal{X})$ is a feasible solution of Problem $Q^\bullet_t(\mathcal{X})$, by mapping all singletons to the empty set, so $Q_t(\mathcal{X}) \leqslant Q^\bullet_t(\mathcal{X})$. Again we can show that

$$F_{\mathcal{X}}(n, s) \geqslant e^{Q^\bullet_0(\mathcal{X})\binom{n}{2} + O(n)}. \tag{4.7.1}$$

To remove degenerate solutions, we define the set of *basic optimal solutions* $\mathrm{OPT}^*_\bullet(\mathcal{X})$ to be the set of $(r^*, \phi^*, \boldsymbol{\alpha}^*)$ such that

- $\phi^* \in \Phi^\bullet_2(\mathcal{X})$,
- $\alpha^*_i > 0$ for all $i \in [r^*]$,
- whenever there are distinct $i, j \in [r^*]$ with $\phi^*(i) = \phi^*(ij) = \phi^*(j)$, there is $h \in [r^*] \setminus \{i, j\}$ such that $\phi^*(hi) \neq \phi^*(hj)$.

We say that $\mathcal{X}$ is *$\bullet$-bounded* if $\mathrm{OPT}^*_\bullet(\mathcal{X})$ is non-empty and there is some $R > 0$ such that $r^* \leqslant R$ for all $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \mathrm{OPT}^*_\bullet(\mathcal{X})$.

We will show that $(K_3^{(3)}, 3)$ is $\bullet$-bounded. Suppose that there is $(r^*, \phi^*, \boldsymbol{\alpha}^*) \in \mathrm{OPT}^*_\bullet(K_3^{(3)}, 3)$ in which all three colours appear on pairs. Then, without loss of generality, there are distinct $h, i, j \in [r^*]$ such that $\phi^*(hi) = \{1, 2\}$ and $\phi^*(hj) = \{1, 3\}$. But then $\phi^*(ij) = \emptyset$, a contradiction. Thus, again without loss of generality, $\phi^*(ij) = \{1,2\}$ for all pairs ij in $[r^*]$. The colour 3 cannot appear in any $\phi^*(i)$, so by optimality we have $\phi^*(i) = \{1,2\}$ for all $i \in [r^*]$. But then we must have $r^* = 1$ and $\phi^*(1) = \{1,2\}$ by the third property of $\text{OPT}^*_\bullet$ solutions. So $Q^\bullet_0(K_3^{(3)}, 3) = \log(2)$. It was shown in [10] that K_n is the graph with the most 3-edge-colourings free of rainbow triangles, and that the number of such colourings is $(\binom{3}{2} + o(1))2^{\binom{n}{2}}$, which matches $Q^\bullet_0(K_3^{(3)}, 3)$.

The question of whether $\mathcal{X}$ is $\bullet$-bounded is a Ramsey-type question. If it has a positive answer, then this increases the motivation for studying Problem $Q^\bullet(\mathcal{X})$ and it could lead us to general results that apply to an even wider class of colour patterns $\mathcal{X}$.

Problem 4.37. *For all integers $s \geqslant 2$ and $k \geqslant 3$ and every (symmetric) family $\mathcal{X}$ of s-edge-colourings of K_k, is $\mathcal{X}$ $\bullet$-bounded?*

To show that the monochromatic family $(K_k^{(1)}, s)$ and the dichromatic triangle family $(K_3^{(2)}, s)$ are bounded, we showed the stronger statement that there is some R for which $r < R$ for every *feasible* solution $(r, \phi, \boldsymbol{\alpha})$ where all multiplicities are at least 2. As we know, for general patterns, this is not true (e.g. for the rainbow triangle and 3 colours), so this approach will not suffice to prove $\bullet$-boundedness.

4.7.2 A more general exact result

The 'hermetic' assumption in Theorem 4.11 allows us to simplify the proof in [75] significantly, and is sufficient to prove our main results, Theorem 1.21 and Theorem 1.24. However, the assumption is fairly strong. We see no obstacle to proving an analogue of the exact result in [75] which would replace the assumption that $\mathcal{X}$ has the extension property and is hermetic with the assumption that $\mathcal{X}$ has the strong extension property (recall Definition 4.4). Such a result should allow the recovery of the results in the literature on non-monochromatic patterns. We did not attempt this in the paper this chapter is based on since it was already rather long and would only yield new results where we could solve the corresponding optimisation problem.

4.7.3 The dichromatic triangle problem

We have essentially solved the dichromatic triangle problem completely, for large n. The outstanding part concerns the set $R_2(s)$ of optimal r in the optimisation problem, which we show contains at most two values $r_2(s) = 2\lfloor (W(s/e)+1)/2 \rfloor$ and $r_2(s)+2$. Given any s, one can simply calculate $g_s(r)$ for both these values r to see which is optimal. We strongly suspect that, apart from for $s = 27$, the set contains a single element. This is a number theoretic statement: we conjecture that the equation $g_s(r) = g_s(r+2)$, or equivalently,

$$z^{(r-1-a)(r+2)}(z+1)^{a(r+2)} = y^{(r+1-b)r}(y+1)^{br}$$

where $z = \lfloor \frac{s}{r-1} \rfloor$, $y = \lfloor \frac{s}{r+1} \rfloor$, $a = s - (r-1)z$, $b = s - (r+1)y$ has no solutions for any integer $s \neq 27$, where $r = r_2(s)$. It is not too hard to show that, for example, there is no other solution when $(r-1)(r+1) \mid s$, in other words, no other solution to $\tilde{g}_s(r) = \tilde{g}_s(r+2)$.

In Lemma 4.21(iii) we showed that the set $S_2(r)$ of s for which $r \in R_2(s)$ is an interval, and furthermore that $S_2(r)$ and $S_2(r+2)$ overlap in at most one value of s. We provided some fairly weak bounds for the startpoint $\tilde{s}_r$ (and endpoint) of $S_2(r)$. From Table 1.2 which shows $R_2(s)$ for small values of s, it seems that, as $r \to \infty$, we have $W(\tilde{s}_{r+2}/e) - r \to 0$. That is, the rough value of $\tilde{s}_{r+2}$ when $r+2$ becomes optimal instead of r is about when $W(\tilde{s}_{r+2}/e) = r$, which yields $\tilde{s}_{r+2} \approx re^{r+1}$.

Even if we cannot prove that $|R_2(s)| = 1$ for all $s \neq 27$, one could try to show that there is nevertheless a unique extremal graph by comparing the number of colourings $c(n,r) := (C + o(1))e^{\frac{r}{r-1}g(s)t_r(n)}$ for $r = r_2(s), r_2(s)+2$ arising from the two potential extremal graphs $T_{r_2(s)}(n)$ and $T_{r_2(s)+2}(n)$ in the case that $R_2(s) = \{r_2(s), r_2(s)+2\}$, for which it can be shown that $e^{\frac{r-1}{r}g(s)t_r(n)}$ and $e^{\frac{r+1}{r+2}g(s)t_{r+2}(n)}$ differ by a multiplicative factor independent of n. Determining the constant C which depends on r, s and $n \pmod r$, is straightforward for small values of the parameters but is in general difficult. Even with the favourable divisibility conditions $r \mid n$ and $(r-1) \mid s$, in which case all parts of $T_r(n)$ have the same size and all multiplicities are equal, C equals the number of decompositions of the multigraph $\frac{s}{r-1}K_r$ into perfect matchings (see Construction 4.12).

4.7.4 Other colour patterns

We hope that the ideas and methods presented in this chapter will be useful for other colour patterns, including the central case of monochromatic patterns. For the family $(K_k^{(1)}, s)$, it is expected that, typically, every basic optimal solution $(r^*, \phi^*, \boldsymbol{\alpha}^*)$ satisfies that $(k-1) \mid r^*$, $\phi^{*-1}(c) \cong T_{k-1}(r^*)$ for all $c \in [s]$ and $\boldsymbol{\alpha}^*$ is uniform. (Note that this is not always true; there is a single known example $(K_3^{(1)}, 5)$ where there are basic optimal solutions without these properties.) Recall that our key idea for the dichromatic triangle problem was to consider the contribution of a largest part, and to show that if r^* is odd then this contribution is non-optimal. A similar idea may rule out $(k-1) \nmid r^*$ for instances of the monochromatic problem.

We conjecture that for every pattern, every extremal graph for the generalised Erdős-Rothschild problem is complete partite.

Conjecture 4.38. *For all integers $k \geqslant 3$ and $s \geqslant 2$ and every colour pattern P of K_k, whenever n is sufficiently large, every n-vertex (P, s)-extremal graph is complete partite.*

The results we presented in this chapter together with the results of [2, 9, 19] imply that we may assume P is monochromatic, in which case this conjecture is implied by [76, Conjecture 16] (which is slightly stronger).

As we have noted, results are very sporadic and there are many patterns which have not been studied. Monochromatic patterns are still of central interest. We finish with some questions about specific patterns of particular interest.

Problem 4.39. *What is the extremal graph for the following?*

1. *The monochromatic triangle pattern $K_3^{(1)}$ for $s \geqslant 8$ colours. It is believed $s = 11$ could be the next most tractable case (see [75]).*
2. *Rainbow patterns $K_t^{(\binom{t}{2})}$ for $t \geqslant 4$. As discussed in the introduction, Hàn, Hoppen, Müller, and Schmidt [47] showed that $T_3(n)$ is the extremal graph for $s \geqslant 12$. To fully confirm the conjecture posed in [45], it remains to show that K_n is the extremal graph for $s \leqslant 11$.*
3. *The dichromatic family $K_4^{(2)}$ as well as $K_4^{(\geqslant \ell)}$ and $K_4^{(\leqslant \ell)}$ for $2 \leqslant \ell \leqslant 6$.*

Bibliography

[1] N. Alon. Ranking tournaments. *SIAM Journal on Discrete Mathematics*, **20**(1):137–142, 2006. ↑67

[2] N. Alon, J. Balogh, P. Keevash and B. Sudakov. The number of edge colorings with no monochromatic cliques. *J. London Math. Soc. (2)*, **70**(2):273–288, 2004. ↑12, 13, 14, 16, 17, 82, 136

[3] N. Alon, E. Powierski, M. Savery, A. Scott and E. Wilmer. Invertibility of digraphs and tournaments. *SIAM Journal on Discrete Mathematics*, **38**(1):327–347, 2024. ↑iv, 11

[4] N. Alon and R. Yuster. The number of orientations having no fixed tournament. *Combinatorica*, **26**(1):1–16, 2006. ↑16

[5] P. Araújo, F. Botler and G. O. Mota. Counting graph orientations with no directed triangles. *arXiv preprint arXiv:2005.13091*, 2020. ↑16

[6] G. Aubian, F. Havet, F. Hörsch, F. Klingelhoefer, N. Nisse, C. Rambaud and Q. Vermande. Problems, proofs, and disproofs on the inversion number. *arXiv:2212.09188* preprint, 2022. ↑6, 11, 69

[7] V. Bafna and P. A. Pevzner. Genome rearrangements and sorting by reversals. *SIAM Journal on Computing*, **25**(2):272–289, 1996. ↑66

[8] P. Balister, E. Powierski, A. Scott and J. Tan. Counting partitions of $G_{n,1/2}$ with degree congruence conditions. *Random Structures & Algorithms*, **62**(3):564–584, 2023. ↑iv, 5

[9] J. Balogh. A remark on the number of edge colorings of graphs. *European J. Combin.*, **27**(4):565–573, 2006. ↑12, 13, 14, 17, 136

[10] J. Balogh and L. Li. The typical structure of Gallai colorings and their extremal

graphs. *SIAM J. Discrete Math.*, **33**(4):2416–2443, 2019. ↑13, 15, 17, 18, 73, 81, 134

[11] J. Bang-Jensen, J. C. F. da Silva and F. Havet. On the inversion number of oriented graphs. *Discrete Mathematics & Theoretical Computer Science*, **23**(2), 2022. ↑5, 6, 7, 8, 9, 10, 11, 48, 49, 51, 56, 57, 65, 66, 67, 68

[12] A. J. d. O. Bastos, H. Lefmann, A. Oertel, C. Hoppen and D. R. Schmidt. Maximum number of r-edge-colorings such that all copies of K_k are rainbow. *Procedia Computer Science. Amsterdam. Vol. 195 (2021), p. 419-426*, 2021. ↑17

[13] J. d. O. Bastos, F. S. Benevides, G. O. Mota and I. Sau. Counting Gallai 3-colorings of complete graphs. *Discrete Math.*, **342**(9):2618–2631, 2019. ↑13

[14] J. d. O. Bastos, C. Hoppen, H. Lefmann, A. Oertel and D. R. Schmidt. Graphs with many edge-colorings such that complete graphs are rainbow. *Discrete Appl. Math.*, **333**:151–164, 2023. ↑13, 15, 17

[15] N. Behague. Research statement. https://nb453.user.srcf.net/wp-content/uploads/2024/10/Research_plan.pdf. ↑69

[16] H. Belkhechine. *Indécomposabilité des graphes et des tournois.* Theses, Université Claude Bernard - Lyon I ; Université de Sfax. Faculté des sciences, July 2009. ↑5

[17] H. Belkhechine, M. Bouaziz, I. Boudabbous and M. Pouzet. Inversions in tournaments. Unpublished. ↑5, 65, 68

[18] H. Belkhechine, M. Bouaziz, I. Boudabbous and M. Pouzet. Inversion dans les tournois. *Comptes Rendus Mathematique*, **348**(13):703–707, 2010. ↑5, 8, 10, 63, 64, 65, 67, 68

[19] F. S. Benevides, C. Hoppen and R. M. Sampaio. Edge-colorings of graphs avoiding complete graphs with a prescribed coloring. *Discrete Math.*, **340**(9):2143–2160, 2017. ↑13, 14, 17, 115, 136

[20] E. Berlekamp. On subsets with intersections of even cardinality. *Canadian Mathematical Bulletin*, **12**(4):471–474, 1969. ↑61

[21] B. Bollobás. *Random graphs.* Cambridge Studies in Advanced Mathematics, Vol. **73** (second edition), Cambridge University Press, Cambridge, 2001. ↑25

[22] F. Botler, J. Corsten, A. Dankovics, N. Frankl, H. Hàn, A. Jiménez and J. Skokan. Maximum number of triangle-free edge colourings with five and six colours. *Acta Math. Univ. Comenian. (N.S.)*, **88**(3):495–499, 2019. ↑12, 13, 14, 17, 18, 73

[23] M. Bucić, O. Janzer and B. Sudakov. Counting H-free orientations of graphs. *Math. Proc. Cambridge Philos. Soc.*, **174**(1):79–95, 2023. ↑16

[24] A. Caprara. Sorting by reversals is difficult. In *Proceedings of the First Annual International Conference on Computational Molecular Biology*, RECOMB '97, page 75–83, New York, NY, USA, 1997. Association for Computing Machinery. ↑66

[25] Y. Caro. On induced subgraphs with odd degrees. *Discrete Math.*, **132**:23–28, 1994. ↑2

[26] Y. Caro, I. Krasikov and Y. Roditty. Zero-sum partition theorems for graphs. *Internat. J. Math. & Math. Sci.*, **17**(4):697–702, 1994. ↑2

[27] P. Charbit, S. Thomassé and A. Yeo. The minimum feedback arc set problem is NP-hard for tournaments. *Combinatorics, Probability and Computing*, **16**(1):1–4, 2007. ↑67

[28] J. Chen, Y. Liu, S. Lu, B. O'sullivan and I. Razgon. A fixed-parameter algorithm for the directed feedback vertex set problem. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 177–186, 2008. ↑67

[29] Y. Cheng, Y. Jing, L. Li, G. Wang and W. Zhou. Integer colorings with forbidden rainbow sums. *J. Combin. Theory Ser. A*, **199**:Paper No. 105769, 28, 2023. ↑16

[30] D. Clemens, S. Das and T. Tran. Colourings without monochromatic disjoint pairs. *European J. Combin.*, **70**:99–124, 2018. ↑16

[31] L. Colucci, E. Győri and A. Methuku. Edge colorings of graphs without monochromatic stars. *Discrete Math.*, **343**(12):112140, 8, 2020. ↑16

[32] L. d. O. Contiero, C. Hoppen, H. Lefmann and K. Odermann. Rainbow Erdős-Rothschild problem for the Fano plane. *SIAM J. Discrete Math.*, **35**(3):1927–1951, 2021. ↑16

[33] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein. *Introduction to Algorithms*. The MIT Press, 3rd edition, 2009. ↑7

[34] S. Das, R. Glebov, B. Sudakov and T. Tran. Colouring set families without monochromatic k-chains. *J. Combin. Theory Ser. A*, **168**:84–119, 2019. ↑16

[35] W. F. de la Vega. On the maximum cardinality of a consistent set of arcs in a random tournament. *Journal of Combinatorial Theory, Series B*, **35**(3):328–332, 1983. ↑67

[36] R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity.* Springer London, 2013. ↑9, 57

[37] P. Erdős. Some new applications of probability methods to combinatorial analysis and graph theory. In *Proceedings of the Fifth Southeastern Conference on Combinatorics, Graph Theory and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1974)*, volume No. X of *Congress. Numer.*, pages 39–51. Utilitas Math., Winnipeg, MB, 1974. ↑12, 16

[38] A. Ferber, L. Hardiman and M. Krivelevich. *Induced subgraphs with prescribed degrees modulo q.* Talk given by A. Ferber in the EPC Webinar 15/6/2020, available at https://www.youtube.com/watch?v=LEHccogFjhk&ab_channel=EPC-Webinar, 2020. ↑3

[39] A. Ferber, L. Hardiman and M. Krivelevich. On subgraphs with degrees of prescribed residues in the random graph. *Random Structures & Algorithms*, **63**(1):192–214, 2023. ↑3

[40] A. Ferber and M. Krivelevich. Every graph contains a linearly sized induced subgraph with all degrees odd. *arXiv:2009.05495* preprint, 2020. ↑2

[41] G. Fertin, A. Labarre, I. Rusu, E. Tannier and S. Vialette. *Combinatorics of Genome Rearrangements.* The MIT Press, 2009. ↑66

[42] T. Gallai. Transitiv orientierbare Graphen. *Acta Math. Acad. Sci. Hungar.*, **18**:25–66, 1967. ↑15

[43] T. Gallai. Transitiv orientierbare Graphen. *Acta Mathematica Academiae Scientiarum Hungarica*, **18**:25–66, 1967. ↑67

[44] J. E. Graver. Boolean designs and self-dual matroids. *Linear Algebra and its Applications*, **10**(2):111–128, 1975. ↑61

[45] P. Gupta, Y. Pehova, E. Powierski and K. Staden. A framework for the generalised Erdős-Rothschild problem and a resolution of the dichromatic triangle case. *arXiv preprint arXiv:2502.12291*, 2025. ↑iv, 15, 22, 136

[46] A. Gyárfás and G. Simonyi. Edge colorings of complete graphs without tricolored triangles. *J. Graph Theory*, **46**(3):211–216, 2004. ↑15

[47] H. Hàn, C. Hoppen, N. M. Müller and D. R. Schmidt. Maximum number of edge colorings avoiding rainbow copies of k_4. *arXiv preprint arXiv:2503.19244*, 2025. ↑15, 17, 136

[48] H. Hàn and A. Jiménez. Improved bound on the maximum number of clique-free colorings with two and three colors. *SIAM J. Discrete Math.*, **32**(2):1364–1368, 2018. ↑16, 73

[49] S. Hannenhalli and P. Pevzner. To cut… or not to cut (applications of comparative physical maps in molecular evolution). In *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '96, page 304–313, USA, 1996. Society for Industrial and Applied Mathematics. ↑66

[50] S. Hannenhalli and P. A. Pevzner. Transforming cabbage into turnip: polynomial algorithm for sorting signed permutations by reversals. *Journal of the ACM*, **46**(1):1–27, 1999. ↑66

[51] C. Hoppen, Y. Kohayakawa and H. Lefmann. Hypergraphs with many Kneser colorings. *European J. Combin.*, **33**(5):816–843, 2012. ↑16

[52] C. Hoppen, Y. Kohayakawa and H. Lefmann. Edge-colorings of graphs avoiding fixed monochromatic subgraphs with linear Turán number. *European J. Combin.*, **35**:354–373, 2014. ↑16

[53] C. Hoppen and H. Lefmann. Edge-colorings avoiding a fixed matching with a prescribed color pattern. *European J. Combin.*, **47**:75–94, 2015. ↑13, 16

[54] C. Hoppen and H. Lefmann. Remarks on an edge-coloring problem. In *The proceedings of Lagos 2019, the tenth Latin and American Algorithms, Graphs and Optimization Symposium (LAGOS 2019)*, volume 346 of *Electron. Notes Theor. Comput. Sci.*, pages 511–521. Elsevier Sci. B. V., Amsterdam, 2019. ↑13, 15, 17, 18, 21

[55] C. Hoppen, H. Lefmann and D. Nolibos. An extension of the rainbow Erdős-Rothschild problem. *Discrete Math.*, **344**(8):Paper No. 112443, 18, 2021. ↑13, 15, 17

[56] C. Hoppen, H. Lefmann and K. Odermann. A coloring problem for intersecting vector spaces. *Discrete Math.*, **339**(12):2941–2954, 2016. ↑16

[57] C. Hoppen, H. Lefmann and K. Odermann. A rainbow Erdős-Rothschild problem. *SIAM J. Discrete Math.*, **31**(4):2647–2674, 2017. ↑13, 15, 16, 17, 18

[58] C. Hoppen, H. Lefmann, K. Odermann and J. Sanches. Edge-colorings avoiding rainbow stars. *J. Graph Theory*, **87**(4):399–429, 2018. ↑13, 16

[59] C. Hoppen, H. Lefmann and D. R. Schmidt. Edge-colorings avoiding patterns in a triangle. *Discrete Math.*, **347**(4):Paper No. 113896, 16, 2024. ↑13, 15, 17, 18, 21

[60] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller, J. W. Thatcher and J. D. Bohlinger, editors, *Complexity of Computer Computations: Proceedings of a symposium on the Complexity of Computer Computations, held March 20–22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, and sponsored by the Office of Naval Research, Mathematics Program, IBM World Trade Corporation, and the IBM Research Mathematical Sciences Department*, pages 85–103, Boston, MA, 1972. Springer US. ↑67

[61] J. Komlós, A. Shokoufandeh, M. Simonovits and E. Szemerédi. The regularity lemma and its applications in graph theory. In *Theoretical aspects of computer science (Tehran, 2000)*, volume 2292 of *Lecture Notes in Comput. Sci.*, pages 84–112. Springer, Berlin, 2002. ↑123

[62] R. Kudelić. *Feedback Arc Set.* SpringerBriefs in Computer Science. Springer Cham, 1st edition, 2022. ↑67

[63] B. J. Latka. Structure theorem for tournaments omitting N_5. *Journal of Graph Theory*, **42**(3):165–192, 2003. ↑67

[64] F. Lazebnik. On the greatest number of 2 and 3 colorings of a (v, e)-graph. *J. Graph Theory*, **13**(2):203–214, 1989. ↑22

[65] H. Lefmann and Y. Person. Exact results on the number of restricted edge colorings for some families of linear hypergraphs. *J. Graph Theory*, **73**(1):1–31, 2013. ↑16

[66] H. Lefmann, Y. Person, V. Rödl and M. Schacht. On colourings of hypergraphs without monochromatic Fano planes. *Combin. Probab. Comput.*, **18**(5):803–818, 2009. ↑16

[67] H. Lefmann, Y. Person and M. Schacht. A structural result for hypergraphs with many restricted edge colorings. *J. Comb.*, **1**(3-4):441–475, 2010. ↑16

[68] H. Liu, M. Sharifzadeh and K. Staden. On the maximum number of integer colourings with forbidden monochromatic sums. *Electron. J. Combin.*, **28**(1):Paper No. 1.59, 35, 2021. ↑16

[69] P.-S. Loh, O. Pikhurko and B. Sudakov. Maximizing the number of q-colorings. *Proc. Lond. Math. Soc. (3)*, **101**(3):655–696, 2010. ↑22

[70] L. Lovász. *Combinatorial Problems and Exercises*. North-Holland, Amsterdam, 1993. ↑1

[71] J. Ma and H. Naves. Maximizing proper colorings on graphs. *J. Combin. Theory Ser. B*, **115**:236–275, 2015. ↑22

[72] J. MacWilliams. Orthogonal matrices over finite fields. *The American Mathematical Monthly*, **76**(2):152–164, 1969. ↑42, 63

[73] F. Maffray and M. Preissmann. A translation of Gallai's paper: "Transitiv orientierbare Graphen". In J. L. Ramírez-Alfonsín and B. A. Reed, editors, *Perfect Graphs*, chapter 3, pages 25–66. Wiley, 2001. ↑67

[74] O. Pikhurko and K. Staden. Stability for the Erdős-Rothschild problem. *Forum Math. Sigma*, **11**:Paper No. e23, 43, 2023. ↑13, 16, 73, 74, 77, 82, 115, 116, 117, 118, 119, 120, 121, 123, 124, 126, 127, 131, 132

[75] O. Pikhurko and K. Staden. Exact solutions to the Erdős-Rothschild problem. *Forum Math. Sigma*, **12**:Paper No. e8, 45, 2024. ↑12, 13, 14, 16, 17, 18, 73, 74, 83, 115, 127, 129, 130, 131, 132, 134, 136

[76] O. Pikhurko, K. Staden and Z. B. Yilma. The Erdős-Rothschild problem on edge-colourings with forbidden monochromatic cliques. *Math. Proc. Cambridge Philos. Soc.*, **163**(2):341–356, 2017. ↑16, 73, 74, 82, 115, 116, 136

[77] O. Pikhurko and Z. B. Yilma. The maximum number of K_3-free and K_4-free edge 4-colorings. *J. Lond. Math. Soc. (2)*, **85**(3):593–615, 2012. ↑12, 13, 17, 18

[78] S. Poljak, V. Rödl and J. Spencer. Tournament ranking with expected profit in polynomial time. *SIAM Journal on Discrete Mathematics*, **1**(3):372–376, 1988. ↑67

[79] M. Pouzet, H. S. Kaddour and B. Thatte. On the Boolean dimension of a graph and other related parameters. *Discrete Mathematics & Theoretical Computer Science*, **23**(2), 2022. ↑5, 6, 7, 8, 10, 53

[80] A. D. Scott. Large induced subgraphs with all degrees odd. *Combinatorics, Probability and Computing*, **1**:335–349, 1992. ↑2

[81] A. D. Scott. On induced subgraphs with all degrees odd. *Graphs and Combinatorics*, **17**:539–553, 2001. ↑2, 3, 4

[82] J. Spencer. Optimal ranking of tournaments. *Networks*, **1**(2):135–138, 1971. ↑67

[83] J. Spencer. Optimally ranking unrankable tournaments. *Periodica Mathematica Hungarica*, **11**(2):131–144, 1980. ↑67

[84] H. Wang, Y. Yang and M. Lu. The inversion number of dijoins and blow-up digraphs. *arXiv preprint arXiv:2404.14937*, 2024. ↑69

[85] R. Yuster. The number of edge colorings with no monochromatic triangle. *J. Graph Theory*, **21**(4):441–452, 1996. ↑12

[86] A. A. Zykov. On some properties of linear complexes. *Mat. Sbornik N.S.*, **24/66**:163–188, 1949. ↑115

[87] A. A. Zykov. On some properties of linear complexes. *Amer. Math. Soc. Translation*, **1952**(79):33, 1952. ↑115