Tighter Bounds on the Independence Number
of the Birkhoff Graph

Leonardo Nagami Coregliano[*]   Fernando Granha Jeronimo[†]

June 20, 2020

Abstract

The Birkhoff graph $\mathcal{B}_n$ is the Cayley graph of the symmetric group S_n, where two permutations are adjacent if they differ by a single cycle. Our main result is a tighter upper bound on the independence number $\alpha(\mathcal{B}_n)$ of $\mathcal{B}_n$, namely, we show that $\alpha(\mathcal{B}_n) \leq O(n!/1.97^n)$ improving on the previous known bound of $\alpha(\mathcal{B}_n) \leq O(n!/\sqrt{2}^{\,n})$. With an explicit construction, we also improve the previous known lower bound on $\alpha(\mathcal{B}_n)$ by a factor of $n/2$. This construction is based on a proper coloring of $\mathcal{B}_n$, which gives an upper bound on the chromatic number $\chi(\mathcal{B}_n)$ of $\mathcal{B}_n$. Via known connections, the upper bound on $\alpha(\mathcal{B}_n)$ implies alphabet size lower bounds for a family of maximally recoverable codes on grid-like topologies.

---

[*]lenacore@uchicago.edu

[†]granha@uchicago.edu

Contents

1   Introduction

A celebrated theorem of Birkhoff [Bir46] characterizes the set of doubly stochastic matrices as forming a convex polytope whose extreme points are permutation matrices. More precisely, if M is an n-by-n matrix over the non-negative reals whose rows and column each sum to 1 (i.e., doubly stochastic), then M can be expressed as a convex combination of permutation matrices (i.e., matrices with exactly one entry 1 in each row and column and all other entries 0). It is well known (see e.g. [BA96, Section 2] and [Bar02, Section II.5]) that the skeleton of this polytope, called the Birkhoff graph $\mathcal{B}_n$, is the Cayley graph whose vertex set is the symmetric group S_n and two permutations σ and τ are adjacent if and only if they differ by a single cycle, that is, $\sigma\tau^{-1}$ is a cycle. For more properties of the Birkhoff graph and polytope we refer the reader to [Pak00, CM09].

Recently, connections between the Birkhoff graph and coding theory, more specifically, the theory of maximally recoverable codes [GHJY14], were pointed out by Kane, Lovett and Rao [KLR17], who showed that the alphabet size of a family of maximally recoverable codes on a grid-like topology (more precisely $T_{n\times n}(1,1,1)$ of [GHK$^+$17]) is at least the chromatic number $\chi(\mathcal{B}_n)$ of the Birkhoff graph, which in turn, by the trivial bound, is at least $n!/\alpha(\mathcal{B}_n)$, where $\alpha(\mathcal{B}_n)$ is the independence number of the Birkhoff graph[1]. Thus, upper bounds for the independence number $\alpha(\mathcal{B}_n)$ translate to lower bounds on the size of the alphabet needed for such codes.

A well-known spectral technique to bound the independence number of a graph is the Hoffman bound [Hof69], which uses only the largest and the smallest eigenvalues of the graph. However, Kane, Lovett and Rao [KLR17] observed that such spectral technique cannot yield a bound better than $\alpha(\mathcal{B}_n) \leq O((n-1)!)$. This was their motivation to use stronger techniques based on representation theory, which can be seen as a generalization of spectral theory. Using these techniques, they obtained the following upper bound on $\alpha(\mathcal{B}_n)$.

Theorem 1.1 ([KLR17, Theorem I.8]). *For every $n \in \mathbb{N}_+$, we have*

$$\alpha(\mathcal{B}_n) \leq \frac{n!}{\sqrt{2}^{\,n-4}}.$$

On the other side, the best lower bound known is given via an explicit construction of an independent set and yields the following result[2].

Theorem 1.2 ([KLR17, Theorem I.7]). *For every $n \in \mathbb{N}_+$ that is a power of 2, we have*

$$\alpha(\mathcal{B}_n) \geq \prod_{i=1}^{\log_2(n)-1} 2^i! \geq \frac{n!}{4^n}. \tag{1}$$

In this paper, we prove tighter asymptotic bounds on the independence number $\alpha(\mathcal{B}_n)$ of $\mathcal{B}_n$. Our main result is the following.

Theorem 1.3. *We have*

$$\alpha(\mathcal{B}_n) \leq O\left(\frac{n!}{1.97^n}\right).$$

---

[1]In [KLR17, Claim I.5], this bound is presented directly in terms of the independence number $\alpha(\mathcal{B}_n)$, but their proof in fact works for the chromatic number.

[2]Their construction yields an independent set whose size is exactly the product in (1), but only the rightmost bound is stated in [KLR17].

Our method consists of a generalization of KLR's representation theoretic approach combined with the solution of a particular linear programming problem with the aid of a computer solver. As discussed above, this also improves the lower bound for the size of the alphabet of a maximally recoverable code in the $T_{n \times n}(1, 1, 1)$ topology of [GHK+17] from $\Omega(\sqrt{2}^n)$ to $\Omega(1.97^n)$.

On the other side, we improve KLR's construction of an independent set by a factor of $n/2$ when n is a power of 2, and extend KLR's result for every n, which yields the following result.

Theorem 1.4. *For every $n \in \mathbb{N}_+$, we have*

$$\alpha(\mathcal{B}_n) \geq \prod_{i=1}^{\lfloor \log_2(n) \rfloor} \left\lfloor \frac{n}{2^i} \right\rfloor! \geq \frac{n!}{4^n} \cdot 2^{\Theta((\log(n))^2)}.$$

If n is a power of 2, then we can improve the bound above to

$$\alpha(\mathcal{B}_n) \geq \frac{n}{2} \prod_{i=1}^{\log_2(n)-1} 2^i!$$

In fact, we can construct a proper coloring of the Birkhoff graph such that each of the color classes is an independent set achieving the bounds above.

Theorem 1.5. *Let n be a positive integer. Then there is an explicit proper coloring establishing*

$$\chi(\mathcal{B}_n) \leq \prod_{i=0}^{\lceil \log_2(n) \rceil} \binom{\lceil n/2^i \rceil}{\lceil n/2^{i+1} \rceil} \leq \frac{4^n}{2^{\Theta((\log(n))^2)}}.$$

If n is a power of 2, then there is an explicit proper coloring strengthening the bound above to

$$\chi(\mathcal{B}_n) \leq \frac{2}{n} \prod_{i=1}^{\log_2(n)} \binom{2^i}{2^{i-1}}.$$

We believe that the techniques we introduce here for the upper bound should be strong enough to prove that $\alpha(\mathcal{B}_n) \leq O(n!/c^n)$ for any fixed constant $c \in (1, 2)$, but a full theoretic proof seems quite technical and elusive for now.

Conjecture 1.6. *There exists a constant $K > 0$ such that*

$$\alpha(\mathcal{B}_n) \leq K \cdot \frac{n!}{(2 - o(1))^n}.$$

The paper is organized as follows. In Section 2, we present an overview of the proof of our main result, Theorem 1.3. In Section 3, we establish some notation, recall some facts about representation theory and some results we need from [KLR17]. Our main result will be a combination of theoretical proofs done in Section 4 and computation which we describe in Section 5. Finally, we construct an explicit independent set and a proper coloring in Section 6 yielding Theorem 1.4 and Theorem 1.5.

2 Proof Strategy

In this section, we give an overview of our upper bound proof for the independence number $\alpha(\mathcal{B}_n)$. This proof builds on the representation theoretic techniques of Kane, Lovett and Rao [KLR17]. Roughly speaking, our approach can be seen as higher-order version of KLR. To establish an upper bound on $\alpha(\mathcal{B}_n)$ of $t(n)$ (e.g., $t(n) = n!/c_0^n$ for $c_0 \in (1,2)$), it is enough to show that any $A \subseteq S_n$ of size larger than $t(n)$ contains an edge in $\mathcal{B}_n$. KLR classify the set A in a pseudorandom versus structured dichotomy. If A meets the criteria of being pseudorandom, KLR use representation theory to count the number of edges within A corresponding to cycles of length precisely n and show that this number is positive. Otherwise, they show that A must have some structured subset $A' \subseteq A$ that can be embedded in a edge preserving way into a smaller symmetric group $S_{n'}$, with $n' < n$, and the proof concludes by an inductive argument. A crucial difference in our approach is that when A is pseudorandom, we are going to count edges corresponding to cycles of length[3] $n, n-1, \ldots, n-\ell_0$ for some constant ℓ_0 rather than only counting n-cycles. This is precisely the sense in which our approach is a higher-order version of KLR. By considering all these additional cycle lengths, the representation theoretic analysis becomes substantially more involved and more ingredients are used as we detail below.

To a set $A \subseteq S_n$ one can associate a class function φ_A whose precise definition is not important for this high-level discussion. We denote by $\Psi_\ell := \Psi_\ell(A)$ the number of edges within A corresponding to cycles of length exactly $n - \ell$. A simple representation theoretic argument (easily derivable from KLR) establishes that Ψ_ℓ is proportional to

$$\sum_{\lambda \vdash n} p_\lambda^\ell \cdot \chi^\lambda(\varphi_A),$$

where $\{p_\lambda^\ell\}_\lambda$ are (explicit) coefficients over the reals and χ^λ is the irreducible character associated to the partition $\lambda \vdash n$. In case A is pseudorandom, the characters must satisfy a set of linear constraints of the form

$$\left\{ \sum_{\lambda \vdash n} k_{\lambda,m} \cdot \chi^\lambda(\varphi_A) \leq c_m \right\}_{m \in \mathcal{M}}, \tag{2}$$

where $k_{\lambda,m}$ and c_m are (explicit) real coefficients. Set $M := \max_\ell \Psi_\ell$. In particular, if $M > 0$, then A is certainly not independent as some edge count $\Psi_\ell > 0$. Instead of working with the precise character evaluations $\chi^\lambda(\varphi_A)$ (abiding to representation theoretic rules), we can relax $\chi^\lambda(\varphi_A)$ to be arbitrary real variables[4] x_λ only bound to satisfy the linear constraints (2). By considering the objective function $\min_{x_\lambda} M$, we have a linear program on our hands. Similarly, if the optimum value of this linear program is positive, we are again certain that A is not independent since we are dealing with a relaxed minimization problem.

The linear program mentioned above is actually somewhat more delicate as the coefficients $k_{\lambda,m}$ and p_λ^ℓ depend on the degree n of S_n. Fortunately, for "low complexity" shapes $\lambda := (n - \sum_{i=2}^s \lambda_i, \ldots, \lambda_s)$, those in which $\sum_{i=2}^s \lambda_i$ is a constant, the coefficients $k_{\lambda,m}$ and p_λ^ℓ become fixed constants provided n is sufficiently large. To be able to work in this asymptotic regime where these coefficients of some low complexity shapes are fixed, we have to additionally require n large when

---

[3]In fact, we only need to consider cycles of length $n - 2i$ since we take A to have permutations of the same sign and we take n to be odd.

[4]Actually, the structure of φ_A forces $\chi^\lambda(\varphi_A) \geq 0$ and thus x_λ can be taken to be non-negative.

A is pseudorandom. For this reason, we will need to control the total loss in cardinality when A is structured. Recall that the structured case is handled by finding some $A' \subseteq A$ which is then embedded in $S_{n'}$ for some n' possibly much smaller than n. To ensure that n' is still arbitrarily large we exploit a density increment phenomenon in the structured case, namely, we observe that this A' satisfies $|A'|/n'! > c^{n-n'}|A|/n!$, where $c > 1$ is a constant related to the lack of pseudorandomness. To see that density increment gives a handle on the size of A, consider the following scenario. Initially, if A is not too small, say $|A| > n!/c_0^n$ for some constant $c_0 < c$, then any $A' \subseteq A$ must be ultimately embedded in $S_{n'}$ with degree $n' = \Omega_{c_0,c}(n)$, otherwise it is not difficult to show that the density of A' in $S_{n'}$ would be larger than 1 which is impossible.

We dealt with low complexity shapes above, but we need to explain how to analyze the remaining shapes. Following a similar argument of KLR, we show that provided the low complexity shapes are not too few, all the remaining shapes can be absorbed in a tail which crucially relies on $c_0 < 2$.

The steps above produce a family of linear programs with parameters c and ℓ_0 (but not depending on n). If for a given choice of these parameters the associated linear program has optimal value $M > 0$, then we conclude that $\alpha(\mathcal{B}_n) \leq K \cdot n!/(c - o(1))^n$ for some universal constant $K > 0$. We obtain our main result, Theorem 1.3, by computationally solving a carefully chosen set of parameters. Let us point out that solving these linear programs as ℓ_0 gets larger and c approaches 2 becomes quite challenging even computationally; this requires additional ideas, which we discuss in Section 5.

3   Preliminaries

We denote the set of natural numbers by $\mathbb{N} := \{0, 1, \ldots, \}$ and the set of positive integers by $\mathbb{N}_+ := \mathbb{N} \setminus \{0\}$. Given $n, k \in \mathbb{N}_+$ with $n \geq k$, we let $[n] := \{1, \ldots, n\}$ (and $[0] := \varnothing$) and let $[n]_k := \{(i_1, \ldots, i_k) \in [n]^k : |\{i_1, \ldots, i_k\}| = k\}$ be the set of k-tuples of elements of $[n]$ with no repeated coordinates. We also let $(n)_k := n(n-1)\cdots(n-k+1)$ denote the falling factorial so that $|[n]_k| = (n)_k$. Let S_n be the symmetric group on $[n]$. We denote the sign of permutation σ by $\mathrm{sgn}(\sigma)$. Let $\mathcal{C}_{n,\ell} \subseteq S_n$ be the set of all single cycles of length ℓ in S_n and let $\mathcal{C}_n := \bigcup_{\ell=2}^n \mathcal{C}_{n,\ell}$.

Definition 3.1 (Birkhoff Graph). The *Birkhoff Graph* $\mathcal{B}_n$ is the Cayley graph $\mathrm{Cay}(S_n, \mathcal{C}_n)$, i.e., the vertex set is S_n and $\sigma, \tau \in S_n$ are adjacent if and only if $\sigma\tau^{-1} \in \mathcal{C}_n$.

Remark 3.2. Recall that since $\mathcal{C}_n$ is closed under conjugation, for every $\sigma \in S_n$, both multiplication by σ maps $\tau \mapsto \tau\sigma$ and $\tau \mapsto \sigma\tau$ are automorphisms of $\mathcal{B}_n$.

3.1   Representation of the Symmetric Group

We recall some important definitions and results from the representation theory of S_n. For a thorough introduction, we point the reader to the book of Sagan [Sag13] (see [SS96] for an introduction to general representation theory). The irreducible representations of S_n are the so-called *Specht modules*, which are in one-to-one correspondence with *partitions* of n. Recall that a partition of n, $\lambda \vdash n$, is a tuple $\lambda := (\lambda_1, \ldots, \lambda_s)$ of positive integers with $\lambda_1 \geq \cdots \geq \lambda_s$ and $\sum_{i=1}^s \lambda_i = n$; the length s of λ as a sequence is called the *height* of λ and denoted by $\mathrm{ht}(\lambda)$ and the *size* of λ is denoted by $|\lambda| := n$. It will be convenient to visualize a partition via Young diagram (also known as Ferrers diagram), which is a left-adjusted box diagram in which the ith row has λ_i boxes (see Fig. 1a).

4

We denote by S^λ the Specht module corresponding to $\lambda \vdash n$ and by $\chi^\lambda \colon S_n \to \mathbb{R}$ its corresponding character. We let f_λ be the dimension of S^λ. Alternatively, f_λ can be computed as $\chi^\lambda(\mathrm{id}_n)$, which also corresponds to the number of *standard tableaux* on shape λ (see [Sag13, Section 2.5]). A standard tableau of shape $\lambda \vdash n$ is a filling of the Young diagram of λ, where each box is filled with a distinct number from $[n]$ so that each row and each column is increasing (see Fig. 1b). More generally, a semi-standard tableau of shape $\lambda \vdash n$ and content $\mu \vdash n$ is a filling of the Young diagram of λ with μ_i copies of i and which is (strictly) increasing in each column and non-decreasing in each row (see Fig. 1c). The number of tableaux of shape λ and content μ is called the *Kostka number* $K_{\lambda,\mu}$.

(a) Young/Ferrers diagram of the partition $(7,6,6,4,4,2,1) \vdash 30$.

(b) A standard tableau of the partition $(7,6,6,4,4,2,1) \vdash 30$.

(c) A semi-standard tableau of the partition $(7,6,6,4,4,2,1) \vdash 30$ and content $(5,4,4,3,3,3,3,3,2) \vdash 30$.
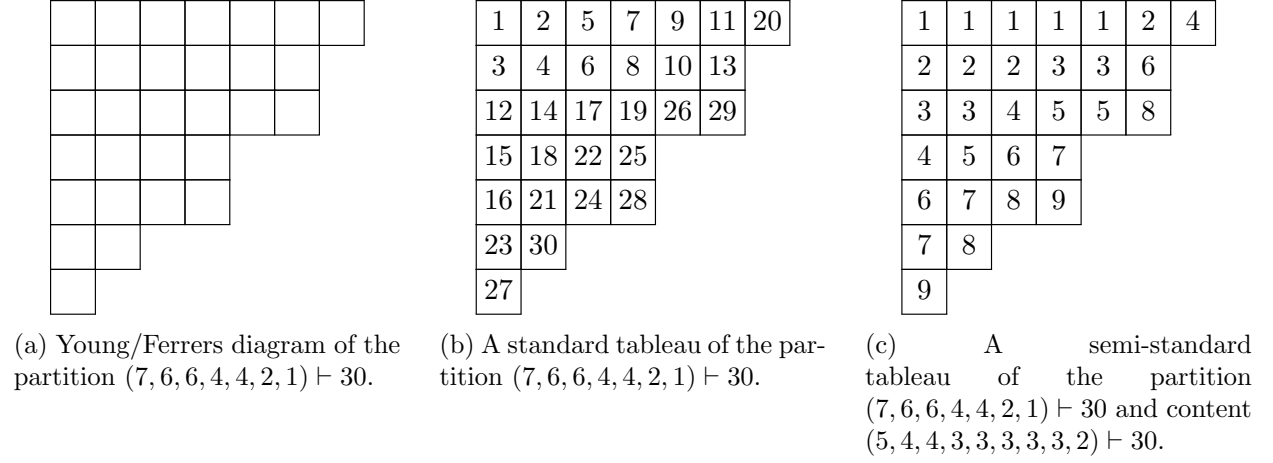
Figure 1: Young diagram and tableaux.

Another family of important modules, also indexed by partitions $\mu \vdash n$, is that of the *Young modules* M^μ. We only consider Young modules associated with shapes of the form $h_k^n := (n-k, 1^k)$ commonly referred to as *hooks*. The module $M^{h_k^n}$ corresponds to the natural action of S_n on $[n]_k$ given by $\sigma \cdot (i_1, \ldots, i_k) := (\sigma(i_1), \ldots, \sigma(i_k))$. The irreducible decomposition of M^μ is given by the Young's Rule, where the Kostka numbers give the multiplicities.

Theorem 3.3 (Young's Rule [Sag13, Theorem 2.11.2]). *Let $\mu \vdash n$. We have*

$$M^\mu \simeq \bigoplus_{\lambda \vdash n} K_{\lambda,\mu} \cdot S^\lambda.$$

The main tool we use to compute characters of Specht modules is the so-called Murnaghan–Nakayama rule. First, recall that characters are *class functions*, i.e., functions on S_n that are invariant under conjugation. Since conjugacy classes of S_n are in one-to-one correspondence with partitions of n (the conjugacy class of $\sigma \in S_n$ corresponds to the partition whose parts are the lengths of the cycles in the cycle decomposition of σ), we typically view characters as functions defined on partitions of n. To apply the Murnaghan–Nakayama rule, we will also need the notion of *rim hook*. Recall that a rim hook ξ of a shape $\lambda \vdash n$ is a contiguous region in the (right) border of λ that does not contain a two-by-two sub-shape and whose removal leaves a valid shape denoted $\lambda \setminus \xi$ (see Fig. 2).

5

(a) A valid rim hook.

(b) An invalid rim hook (it is not contiguous).

(c) An invalid rim hook (it contains a two-by-two subshape).

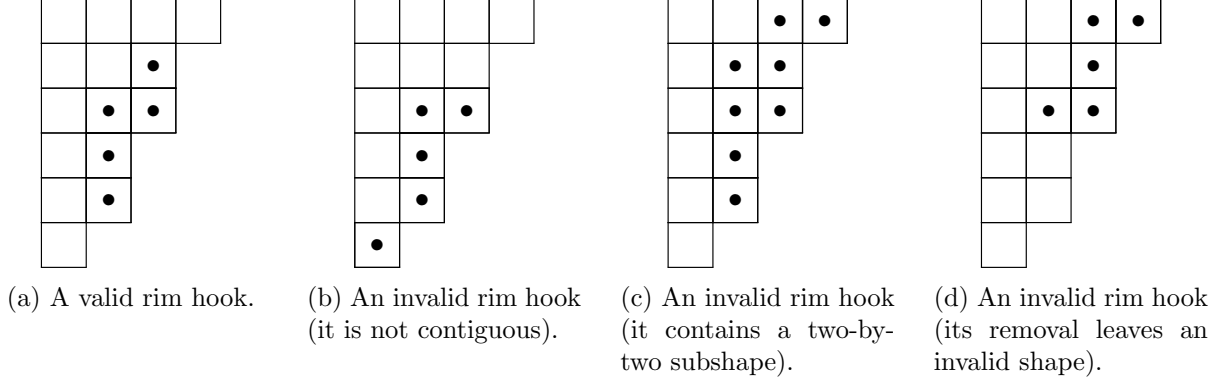(d) An invalid rim hook (its removal leaves an invalid shape).

Figure 2: Example of valid and invalid rim hooks in the partition $(4, 3, 3, 2, 2, 1) \vdash 15$.

Theorem 3.4 (Murnaghan–Nakayama Rule [Sag13]). *Let $\lambda \vdash n$ and $\mu := (\mu_1, \ldots, \mu_k)$ be a partition of n. Then for every $i \in [k]$, we have*

$$\chi^\lambda(\mu) = \sum_\xi (-1)^{\mathrm{ht}(\xi)-1} \cdot \chi^{\lambda \setminus \xi}((\mu_1, \ldots, \widehat{\mu_i}, \ldots, \mu_k)),$$

where ξ ranges over all the rim hooks of size μ_i in λ, $\mathrm{ht}(\xi)$ is the height of ξ (i.e., the number of rows of ξ) and $(\mu_1, \ldots, \widehat{\mu_i}, \ldots, \mu_k)$ is partition of $n - \mu_i$ obtained from μ by omitting part μ_i.

Remark. Note that the Murnaghan-Nakayama rule gives us the freedom to choose which part of μ to remove first. We will explore this flexibility in our proofs.

We will be working with elements in the group algebra $\mathbb{R}[S_n]$, which are formal $\mathbb{R}$-linear combinations of elements in S_n. Given $\phi := \sum_{\sigma \in S_n} \phi_\sigma \cdot \sigma \in \mathbb{R}[S_n]$ with $\phi_\sigma \in \mathbb{R}$, it is convenient to regard ϕ as a function $S_n \to \mathbb{R}$ defined by $\sigma \mapsto \phi_\sigma$. The space $\mathbb{R}[S_n]$ is equipped with the inner product

$$\langle \phi, \psi \rangle := \frac{1}{|S_n|} \sum_{\sigma \in S_n} \phi_\sigma \cdot \psi_\sigma.$$

Under this inner product, the characters of irreducible representations form an orthonormal basis of the sub-space of class functions [Sag13, Proposition 1.10.2], from which we can extract the following Parseval formula.

Fact 3.5 (Parseval). *Let $\phi, \psi \in \mathbb{R}[S_n]$ be class functions. Then*

$$\langle \phi, \psi \rangle = \frac{1}{|S_n|^2} \sum_{\lambda \vdash n} \chi^\lambda(\phi) \cdot \chi^\lambda(\psi),$$

where χ^λ is linearly extended to class functions.

By reflecting a shape $\lambda \vdash n$ along the diagonal, we obtain the transposed shape $\lambda^\top$ given by $\lambda_i^\top := |\{j : \lambda_j \geq i\}|$. We end this section with a simple, but useful fact of irreducible characters in S_n.

Fact 3.6. *Let $\lambda \vdash n$ and $\pi \in S_n$. We have*

$$\chi^{\lambda^\top}(\pi) = \mathrm{sgn}(\pi) \cdot \chi^\lambda(\pi).$$

3.2 Recalling KLR Results

In this section, we recall some key results of Kane, Lovett and Rao [KLR17], with minor generalizations when necessary. For the reader's convenience, the omitted proofs can be found in Appendix B.

For every non-empty set $A \subseteq S_n$, define the class function $\phi_A \in \mathbb{R}[S_n]$ as

$$\phi_A := \frac{1}{|S_n|} \frac{1}{|A|^2} \sum_{\substack{\sigma \in S_n \\ \pi, \pi' \in A}} \sigma \pi (\pi')^{-1} \sigma^{-1}.$$

For $\ell \in \{0, \ldots, n-2\}$, define another class function $\psi_\ell \in \mathbb{R}[S_n]$ as

$$\psi_\ell := \frac{1}{|C_{n,n-\ell}|} \sum_{\tau \in C_{n,n-\ell}} \tau,$$

where (we recall that) $C_{n,n-\ell}$ is the set of $(n-\ell)$-cycles of S_n.

The following claim says that the number of edges of $\mathcal{B}_n$ within A corresponding to cycles of length $n - \ell$ can be computed from the inner product of these class functions.

Claim 3.7. *We have*

$$\langle \phi_A, \psi_\ell \rangle = \frac{|E_\ell[A, A]|}{|A|^2 |S_n| |C_{n,n-\ell}|},$$

where

$$E_\ell[A, A] := \left\{ (\pi, \pi') \in A^2 : \pi(\pi')^{-1} \in C_{n,n-\ell} \right\}.$$

In particular, if $\sum_{\lambda \vdash n} \chi^\lambda(\phi_A) \chi^\lambda(\psi_\ell) > 0$, then A contains an edge of $\mathcal{B}_n$.

Proof. We have

$$|E_\ell[A, A]| = \sum_{\pi, \pi' \in A} \mathbb{1} \left[\pi(\pi')^{-1} \in C_{n,n-\ell} \right]$$

$$= \frac{1}{|S_n|} \sum_{\sigma \in S_n} \sum_{\pi, \pi' \in A} \mathbb{1} \left[\sigma \pi(\pi')^{-1} \sigma^{-1} \in C_{n,n-\ell} \right]$$

$$= |A|^2 |S_n| |C_{n,n-\ell}| \langle \phi_A, \psi_\ell \rangle,$$

where the second equality follows since $C_{n,n-\ell}$ is invariant under conjugation. Further, if we assume $\sum_{\lambda \vdash n} \chi^\lambda(\phi_A) \chi^\lambda(\psi_\ell) > 0$, then Fact 3.5 yields that A contains an edge of $\mathcal{B}_n$. $\square$

The following two facts give basic properties about $\chi^\lambda(\phi_A)$.

Fact 3.8 (From KLR). *Let $\lambda \vdash n$. Then $\chi^\lambda(\phi_A) \geq 0$.*

Fact 3.9 (Character folding). *Let $A \subseteq S_n$. If all permutations of A have the same sign, then*

$$\chi^{\lambda^\top}(\phi_A) = \chi^\lambda(\phi_A).$$

Proof. It follows from Fact 3.6 and $\mathrm{sgn}(\sigma\pi(\pi')^{-1}\sigma^{-1}) = \mathrm{sgn}(\pi)\,\mathrm{sgn}(\pi') = 1$ for every $\pi, \pi' \in A$ and every $\sigma \in S_n$. $\qquad\square$

To obtain an upper bound on $\chi^\lambda(\phi_A)$, we will use a notion of pseudorandomness for the set A.

Definition 3.10 (Pseudorandomness). Let $k \in [n]$ and $r > 0$. We say that a non-empty $A \subseteq S_n$ is (k, r)-*pseudorandom* if for every $I, J \in [n]_k$, we have

$$\Pr_{\pi \in A}[\pi(J) = I] < \frac{r}{(n)_k},$$

where $\pi \in A$ is chosen uniformly at random.

For $\theta \colon \mathbb{N}_+ \to (0, \infty)$ a non-decreasing function, we say that $A \subseteq S_n$ is θ-*pseudorandom* if A is $(k, \theta(k))$-pseudorandom for every *even* $k \in [n]$.

We will typically use functions of the form $k \mapsto c^k$ for some fixed $c > 1$ and in this case abuse notation by saying c^k-pseudorandom.

The pseudorandomness condition implies the following upper bound on Young module characters.

Claim 3.11 (Implicit in KLR). *If A is (k, r)-pseudorandom, then*

$$\mathrm{tr}(M^{h_k^n}(\phi_A)) \leq r.$$

An arbitrary non-trivial character can be bounded in terms of the pseudorandomness parameter and an appropriate Kostka number as follows.

Lemma 3.12 (Implicit in KLR). *Let $\lambda \vdash n$ be non-trivial (i.e., $\lambda \neq (1^n)$). If $A \subseteq S_n$ is (k, r)-pseudorandom and $K_{\lambda, h_k^n} \neq 0$, then*

$$\chi^\lambda(\phi_A) \leq \frac{r - 1}{K_{\lambda, h_k^n}}.$$

4 Theoretical Proofs

Recall from Section 2 that our proof strategy is to use a density increment argument to construct from a sufficiently large $A \subseteq S_n$ a pseudorandom set $C \subseteq S_{n'}$ containing at most as many edges as A in the Birkhoff graph $\mathcal{B}_n$. Translating edge counting representation theoretic arguments to linear programming, we will be able to deduce that pseudorandom sets are not independent, which in turn implies that large $A \subseteq S_n$ are also not independent. Therefore, quantifying how large A needs to be to make this approach viable provides upper bounds on $\alpha(\mathcal{B}_n)$. The density increment argument is formalized in Section 4.1, whereas linear programming arguments are formally treated in Section 4.2.

4.1 Dichotomy: Structure vs Randomness

We shed more light on the structure versus randomness dichotomy of [KLR17] by observing a density increment phenomenon. For $A \subseteq S_n$, let $d_n(A) := |A|/|S_n|$ be its *density*. The lemma below shows that if A fails to be (k, r)-pseudorandom, then we can find a set C in S_{n-k} with larger density than A and with at most as many edges as A.

Lemma 4.1 (density increment step). *If $A \subseteq S_n$ non-empty is not (k, r)-pseudorandom, then there exist $\sigma, \sigma' \in S_n$ and $B' \subseteq A$ such that $B := \sigma B' \sigma'$ satisfies*

(i) *For every $\tau \in B$ and every $i \in [n] \setminus [n-k]$, we have $\tau(i) = i$,*

(ii) $|B| \geq |A| \cdot r/(n)_k$.

In particular, by letting $C := \{\tau|_{[n-k]} \mid \tau \in B\} \subseteq S_{n-k}$, we have $d_{n-k}(C) \geq r \cdot d_n(A)$ and $|E_{\mathcal{B}_{n-k}}(C, C)| \leq |E_{\mathcal{B}_n}(A, A)|$. Furthermore, if all permutations of A have the same sign then all permutations of C also have the same sign.

Proof. Since A is not (k, r)-pseudorandom, there exist $I, J \in [n]_k$ such that

$$\Pr_{\pi \in A}[\pi(J) = I] \geq \frac{r}{(n)_k}.$$

Let $B' := \{\pi \in A \mid \pi(J) = I\}$. Take any $\sigma, \sigma' \in S_n$ such that $\sigma'(n-k+1, \ldots, n) = J$ and $\sigma(I) = (n-k+1, \ldots, n)$. Then $B := \sigma B' \sigma'$ satisfies (i) and we have

$$\frac{r}{(n)_k} \leq \Pr_{\pi \in A}[\pi(J) = I] = \frac{|B'|}{|A|} = \frac{|B|}{|A|};$$

thus item (ii) follows.

Note that item (i) implies that $C \subseteq S_{n-k}$ and

$$r \cdot d_n(A) = \frac{r \cdot |A|}{|S_n|} = \frac{r \cdot |A|}{(n)_k \cdot |S_{n-k}|} \leq \frac{|B|}{|S_{n-k}|} = \frac{|C|}{|S_{n-k}|} = d_{n-k}(C),$$

i.e., C is at least r times denser than A. Furthermore, by Remark 3.2 we have $|E_{\mathcal{B}_n}(B, B)| = |E_{\mathcal{B}_n}(B', B')| \leq |E_{\mathcal{B}_n}(A, A)|$. The assertions about C follow from the fact that the restriction map $f : \tau \mapsto \tau|_{[n-k]}$ is a bijection from B to C and if $\{\tau, \tau'\} \in E_{\mathcal{B}_{n-k}}(C, C)$ then $\{f^{-1}(\tau), f^{-1}(\tau')\} \in E_{\mathcal{B}_n}(B, B)$. Finally, if all permutations of A have the same sign, say s, then trivially so do all permutations in B'. By construction, the sign of all permutations in B and in C is $\mathrm{sgn}(\sigma) \cdot s \cdot \mathrm{sgn}(\sigma')$. $\square$

The following simple claim shows that we can pass from $A \subseteq S_n$ to $C \subseteq S_{n-k}$ without decreasing the density. In particular, this allow us to adjust the degree n' so that the n' cycles are even permutations.

Corollary 4.2 (density preserval). *If $A \subseteq S_n$ is non-empty and $k \in [n]$, then there exist $\sigma, \sigma' \in S_n$ and $B' \subseteq A$ such that $B := \sigma B' \sigma'$ satisfies*

(i) *For every $\tau \in B$ and every $i \in [n] \setminus [n-k]$, we have $\tau(i) = i$,*

(ii) $|B| \geq |A|/(n)_k$.

In particular, by letting $C := \{\tau|_{[n-k]} \mid \tau \in B\} \subseteq S_{n-k}$, we have $d_{n-k}(C) \geq d_n(A)$ and $|E_{\mathcal{B}_{n-k}}(C, C)| \leq |E_{\mathcal{B}_n}(A, A)|$. Furthermore, if every permutation of A has the same sign then every permutation of C has the same sign.

Proof. Follows from Lemma 4.1 by noting that every $A \subseteq S_n$ is not $(k, 1)$-pseudorandom. $\square$

9

Lemma 4.3 (density increment)**.** *For every $c_0 > c \geq 1$, every $n \in \mathbb{N}_+$ and $A \subseteq S_n$ with $d_n(A) \geq 1/c^n$, there exists a set $B \subseteq S_m$, where $m \geq (1 - \log_{c_0}(c))n$ and $m \equiv n \pmod 2$ such that*

    (i) B *is* c_0^k*-pseudorandom,*

    (ii) $|E_{\mathcal{B}_m}(B, B)| \leq |E_{\mathcal{B}_n}(A, A)|$,

    (iii) $d_m(B) \geq d_n(A)$, *and*

    (iv) *if all permutations in A have the same sign, then all permutations in B have the same sign.*

Proof. We construct inductively $A_0 \subseteq S_{n_0}, A_1 \subseteq S_{n_1}, \ldots$ through the following algorithm.

1. Set $n_0 := n$ and $A_0 := A$,

2. Given $A_t \subseteq S_{n_t}$, if A_t is c_0^k-pseudorandom, stop and set $T := t$ and $B := A_T$; otherwise, let k_t be such that A_t is not $(k_t, c_0^{k_t})$-pseudorandom with k_t even, let $n_{t+1} := n_t - k_t$ and by Lemma 4.1, let $A_{t+1} \subseteq S_{n_{t+1}}$ be such that

       (i) $d_{n_{t+1}}(A_{t+1}) \geq c_0^{k_t} \cdot d_{n_t}(A_t)$,

       (ii) $|E_{\mathcal{B}_{n_{t+1}}}(A_{t+1}, A_{t+1})| \leq |E_{\mathcal{B}_{n_t}}(A_t, A_t)|$,

       (iii) if all permutations in A_t in all have the same sign, then all permutations in A_{t+1} have the same sign.

We claim that the above procedure stops. Indeed, using induction, for every t such that A_t is constructed we have

$$1 \geq d_{n_t}(A_t) \geq c_0^{\sum_{i=0}^{t-1} k_i} \cdot d_n(A) \geq \frac{c_0^{\sum_{i=0}^{t-1} k_i}}{c^n},$$

which implies that $t \leq \sum_{i=0}^{t-1} k_i \leq n \log_{c_0}(c)$. Hence, the claim follows.

Let $k := \sum_{i=0}^{T-1} k_i$ and $m := n_T = n - k$ so that $m \geq (1 - \log_{c_0}(c))n$. By construction, the k_i's are necessarily even, from which $m \equiv n \pmod 2$. Finally, observe that $|E_{\mathcal{B}_m}(B, B)| = |E_{\mathcal{B}_m}(A_T, A_T)| \leq \cdots \leq |E_{\mathcal{B}_n}(A_0, A_0)| = |E_{\mathcal{B}_n}(A, A)|$ and analogously we also have $d_m(B) \geq d_n(A)$ since $c_0 \geq 1$. By induction, we also have that if all permutations of $A = A_0$ have the same sign, then all permutations of $B = A_T$ have the same sign. $\qquad \square$

4.2 Linear Programs

We define three (families of) linear programs in order to establish an upper bound on $\alpha(\mathcal{B}_n)$. The first is a convex relaxation closely capturing the representation theoretic argument for counting edges of pseudorandom sets of S_n while the last one does not depend on n and captures the asymptotic properties of the Birkhoff graph family. Each such linear program can be roughly described as follows.

Linear Program I. It "contains" all pseudorandom sets $A \subseteq S_n$ as feasible points and its objective value being positive implies that pseudorandom sets are not independent. This allows us to deduce an upper bound on $\alpha(\mathcal{B}_n)$ using the density increment results of Section 4.1. Its number of variables, its coefficients and its number of constraints depend on n.

Linear Program II. Its objective value being positive implies that the first linear program has positive objective value. Its number of variables is independent of n, but its coefficients and its number of constraints depend on n.

Linear Program III. Similarly, its objective value being positive implies that the linear program II has positive objective value for every n sufficiently large. This third linear program is completely independent of n.

4.2.1 Linear Program I

In the definition below, we present a family of linear programs such that each c^k-pseudorandom $A \subseteq S_n$ yields a feasible solution. If further the objective value of such solution is positive, then we will show this implies that A is not independent. These properties of this family of linear programs are established in the next two lemmas.

Definition 4.4 (Linear Program I). Given an odd positive integer n, a real $c > 1$ and a non-negative even integer $\ell_0 \leq n$, let $P_n^{\ell_0}(c)$ be the following linear program.

$$
\begin{array}{rll}
\text{minimize} & M & \\
\text{s.t.} & M \geq \Psi_\ell & \forall \ell \leq \ell_0 \text{ even,} \quad (3) \\
\text{(Parseval)} & \Psi_\ell = \sum_{\lambda \vdash n} \chi^\lambda((n - \ell)) \cdot x_\lambda & \forall \ell \leq \ell_0 \text{ even,} \quad (4) \\
\text{(Young)} & \sum_{\lambda \vdash n} K_{\lambda, h_m^n} \cdot x_\lambda \leq c^m & \forall m \leq n \text{ even,} \quad (5) \\
\text{(Transposition)} & x_\lambda = x_{\lambda^\top} & \forall \lambda \vdash n, \quad (6) \\
\text{(Unit)} & x_{(n)} = x_{(1^n)} = 1, & (7) \\
& x_\lambda \geq 0 & \forall \lambda \vdash n, \quad (8)
\end{array}
$$

where $\chi^\lambda((n - \ell))$ stands for χ^λ evaluated at a cycle of length $n - \ell$ and the variables are M, $(\Psi_\ell)_\ell$ and $(x_\lambda)_{\lambda \vdash n}$.

Lemma 4.5. Let n be an odd positive integer and $c > 1$. If $A \subseteq S_n$ is c^k-pseudorandom and all permutations of A have the same sign, then taking $x_\lambda := \chi^\lambda(\phi_A)$, defining Ψ_ℓ by (4) and letting $M := \max_\ell \Psi_\ell$ gives a feasible solution $P_n^{\ell_0}(c)$.

Proof. By definition, constraints (3) and (4) are trivially satisfied. Note that $x_\lambda \geq 0$ by Fact 3.8, so the constraints (8) are also satisfied.

Now, we proceed to show that constraints (5) are also satisfied. Since A is c^m-pseudorandom, Claim 3.11 gives that $\text{tr}(M^{h_m^n}(\phi_A)) \leq c^m$ for every even $m \in [n]$. Combining with Young's rule, Theorem 3.3, we have

$$
c^m \geq \text{tr}(M^{h_m^n}(\phi_A)) = \sum_{\lambda \vdash n} K_{\lambda, \mu} \cdot \chi^\lambda(\phi_A),
$$

showing that constraints (5) are satisfied.

Since all the permutations of A have the same sign and $\chi^\lambda(\phi_A) = \sum_{\pi, \pi' \in A} \chi^\lambda(\pi(\pi')^{-1})/|A|^2$, the transposition constraint (6) follows from Fact 3.9. Finally, note that $\chi^{(n)}(\phi_A) = 1$ and $\chi^{(1^n)}(\phi_A) = 1$, where the latter follows from the transposition constraint (6). $\qquad \square$

11

Lemma 4.6. *Let n be an odd positive integer, $c > 1$. Suppose $\mathrm{OPT}(P_n^{\ell_0}(c)) > 0$. If $A \subseteq S_n$ is c^k-pseudorandom and all permutations of A have the same sign, then A is not independent in $\mathcal{B}_n$.*

Proof. By Lemma 4.5, setting $x_\lambda \coloneqq \chi^\lambda(\phi_A)$ and Ψ_ℓ and M as in the lemma gives a feasible solution of $P_n^{\ell_0}(c)$, which must have a positive objective value since $\mathrm{OPT}(P_n^{\ell_0}(c)) > 0$. In particular, there exists an even integer $\ell \leq \ell_0$ such that $\Psi_\ell > 0$, i.e.,

$$\sum_{\lambda \vdash n} \chi^\lambda(\phi_A) \cdot \chi^\lambda(\psi_\ell) > 0.$$

By Claim 3.7, this implies that A is not independent in $\mathcal{B}_n$. □

Putting together the above two lemmas with the density increment results of Section 4.1, we get the following asymptotic upper bound on the independence number $\alpha(\mathcal{B}_n)$.

Proposition 4.7. *Let $c_0 > 1$ and ℓ_0 be a non-negative even integer. Suppose $n_0 \in \mathbb{N}$ is such that for every $n \geq n_0$ odd, we have $\mathrm{OPT}(P_n^{\ell_0}(c_0)) > 0$. Then for every $c \in (1, c_0)$ and every integer $n \geq 1 + n_0/(1 - \log_{c_0}(c))$,*

$$\alpha(\mathcal{B}_n) \leq 2 \cdot \frac{n!}{c^{n-1}}.$$

Proof. Suppose that $A_0 \subseteq S_n$ is an independent set in $\mathcal{B}_n$ with $|A_0| \geq 2 \cdot n!/c^{n-1}$, i.e., $d_n(A_0) \geq 2/c^{n-1}$. Let s be the most frequent permutation sign in A_0 and $A_1 \coloneqq \{\sigma \in A_0 \mid \mathrm{sgn}(\sigma) = s\}$. If n is odd, let $A_2 \coloneqq A_1$ and $n_2 \coloneqq n$. Otherwise, let $n_2 \coloneqq n - 1$ and apply Corollary 4.2 to obtain a $A_2 \subseteq S_{n_2}$. By construction, we have $n_2 \geq n - 1$, A_2 is an independent set of $\mathcal{B}_{n_2}$, all permutations of A_2 have the same sign and $d_{n_2}(A_2) \geq d_n(A_0)/2 \geq 1/c^{n_2}$.

By Lemma 4.3, we can find $B \subseteq S_m$ with $m \geq (1 - \log_{c_0}(c))n_2 \geq n_0$ and such that

(i) B is c_0^k-pseudorandom,

(ii) B is independent in $\mathcal{B}_m$,

(iii) $d_m(B) \geq d_{n_2}(A_2) \geq 1/c^{n_2}$ (in particular B is non-empty), and

(iv) all permutations in B have the same sign.

Since $m \geq n_0$, we have $\mathrm{OPT}(P_m^{\ell_0}(c_0)) > 0$. Therefore, by Lemma 4.6 the set B must have an edge in $\mathcal{B}_m$ contradicting the assumption that A_0 is independent. □

4.2.2 Linear Program II

The next step is to define a second family of simpler linear programs with the number of variables being independent of n. To do so we classify partitions based on their leg length and belly as defined below and we use some basic properties of irreducible characters and Kostka constants to remove variables with large leg length or large belly. We will show in Proposition 4.17 that when n is sufficiently large, a positive optimum value in this family implies a positive optimum value for the preceding family $P_n^{\ell_0}(c)$.

Notation 4.8. We denote the hook of size n and leg length k by $h_k^n := (n-k, 1^k)$. More generally, if $\beta := (\beta_1, \beta_2, \ldots, \beta_t)$ is a partition and $n \geq |\beta| + k + \beta_1 + 1$, then let

$$b_{k,\beta}^n := (n - |\beta| - k, \beta_1 + 1, \beta_2 + 1, \ldots, \beta_t + 1, 1^{k-t}).$$

In this case, β is called the *belly* of $b_{k,\beta}^n$ and k is called the *leg length*. See Fig. 3 for a pictorial image of $b_{k,\beta}^n$.

Remark. Note that $(h_k^n)^\top = h_{n-k-1}^n$ and $(b_{k,\beta}^n)^\top = b_{n-|\beta|-k-1,\beta^\top}^n$.
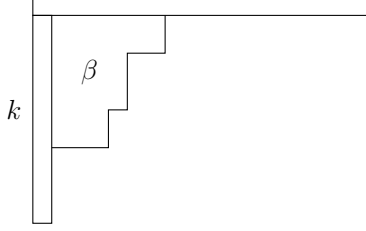

Figure 3: Pictorial image of the shape $b_{k,\beta}^n$.

Definition 4.9 (Linear Program II). Given a positive odd integer n, a real $c > 1$, a non-negative even integer $\ell_0 \leq n$ and a positive odd integer k_0, we define $P_n^{\ell_0, k_0}(c)$ as follows.

$$
\begin{aligned}
\text{minimize} \quad & M \\
\text{s.t.} \quad & M \geq \Psi_\ell && \forall \ell \leq \ell_0 \text{ even}, && (9) \\
\text{(Parseval)} \quad & \Psi_\ell = 2 + 2 \cdot \sum_{i=0}^{\ell_0} \sum_{\beta \vdash i} \sum_{\substack{k=\mathrm{ht}(\beta) \\ k \geq 1}}^{k_0} \chi^{b_{k,\beta}^n}((n-\ell)) \cdot x_{b_{k,\beta}^n} - 2 \cdot T_n^{\ell, k_0}(c) && \forall \ell \leq \ell_0 \text{ even}, && (10) \\
\text{(Young)} \quad & \sum_{i=0}^{\ell_0} \sum_{\beta \vdash i} \sum_{k=\mathrm{ht}(\beta)}^{k_0} K_{b_{k,\beta}^n, h_m^n} \cdot x_{b_{k,\beta}^n} \leq c^m && \forall m \leq n \text{ even}, && (11) \\
\text{(Unit)} \quad & x_{(n)} = 1, && && (12) \\
& x_{b_{k,\beta}^n} \geq 0 && \forall i \in \{0, \ldots, \ell_0\}, && (13) \\
& && \forall \beta \vdash i, && \\
& && \forall k \in \{\mathrm{ht}(\beta), \ldots, k_0\}, &&
\end{aligned}
$$

where the variables are M, $(\Psi_\ell)_\ell$ and $(x_{b_{k,\beta}^n})_{k,\beta}$ and we have

$$T_n^{\ell,k_0}(c) := \sum_{\substack{k=k_0+2 \\ k \text{ odd}}}^{\frac{n-3\ell-1}{2}} \frac{c^{2k+2\ell}}{\binom{2k+2\ell}{k+\ell}} + \sum_{\substack{k=\frac{n-3\ell-1}{2}+1 \\ k \text{ odd}}}^{\frac{n-\ell-1}{2}} \frac{c^{n-1}}{\binom{n-\ell-1}{k+\ell}}.$$

First, we show that if the belly size of a partition is larger than ℓ, then its coefficient in the Parseval (4) is zero, so the corresponding variable in the linear program can be ignored.

Claim 4.10. If $|\beta| > \ell$, then $\chi^{b_{k,\beta}^n}((n-\ell)) = 0$.

Proof. We apply the Murnaghan–Nakayama rule, Theorem 3.4, removing the ℓ fixed points first. Let λ be the resulting shape after their removal in a particular derivation path. Note that $|\lambda| = n-\ell$ and we still need to remove a rim hook of size $n - \ell$. The contribution of the corresponding path is non-zero only if λ is a hook. But for λ to be a hook, the belly β needs to be completely removed by some of the ℓ removed fixed points and thus $|\beta| \leq \ell$. $\qquad\square$

Next, we show that the coefficients $\chi^{b_{k,\beta}^n}((n-\ell))$ are still zero as long as the belly size is smaller than ℓ but the leg length is not too small nor too large. It will be convenient to use the following notation for two distinguished cells of a partition.

Notation 4.11. Let $\lambda \vdash n$. We call *hand* of λ the rightmost cell in the first row of λ. We call *foot* of λ the lowest cell in the first column of λ.

Claim 4.12. *If $|\beta| < \ell \leq n/2 - 1$ and $\ell \leq k \leq n - |\beta| - \ell - 1$, then $\chi^{b_{k,\beta}^n}((n-\ell)) = 0$.*

Proof. We apply the Murnaghan–Nakayama rule, Theorem 3.4, removing $(n - \ell)$ first. To leave a valid shape after this removal, the following must hold: (i) if a cell is removed in the first row of $b_{k,\beta}^n$, then all cells to its right must also be removed. (ii) if a cell is removed in the first column of $b_{k,\beta}^n$, then all cells below it must also be removed.

Our goal is to show that no removal leaves a valid shape. Now we have four cases to consider when we remove a valid rim hook of size $n - \ell$ from $b_{k,\beta}^n$.

Case 1. The removal occurred completely inside the belly β. This implies $n - \ell \leq |\beta|$. Since $\ell \leq n/2 - 1$, this contradicts $|\beta| < \ell$.

Case 2. The hand and the foot were removed. In this case, the remaining shape must be β so $|\beta| = \ell$. This contradicts the assumption $|\beta| < \ell$.

Case 3. The hand was removed but not the foot, which implies that $k + 1 \leq \ell$ since no cell in the first column was removed. This contradicts the assumption $\ell \leq k$.

Case 4. The foot was removed but not the hand, which implies that $n - |\beta| - k \leq \ell$ since no cell in the first row was removed. This contradicts the assumption $k \leq n - |\beta| - \ell - 1$.

Since no removal leaves a valid shape, we have $\chi^{b_{k,\beta}^n}((n - \ell)) = 0$. $\qquad\square$

For the size of the belly being exactly ℓ, the coefficient $\chi^{b_{k,\beta}^n}((n - \ell))$ is not zero but can be computed with the following claim.

Claim 4.13. *If $\beta \vdash \ell$, then*

$$\chi^{b_{k,\beta}^n}((n - \ell)) = (-1)^k f_\beta.$$

Proof. We apply the Murnaghan–Nakayama rule, Theorem 3.4, removing $(n - \ell)$ first. Note the sign of the rim hook is $(-1)^k$ and the remaining shape is β from which the claim readily follows. $\qquad\square$

Observe that Claim 4.13 says that a number of partitions depending on n have non-zero coefficient in the Parseval (4). To be able to remove most of the corresponding variables, we will make use of the Young restrictions (5), which will require bounding the Kostka constants. In one regime, we also compute the Kostka constant exactly since this will be used later. In the process, we will need the following derived shape.

Notation 4.14. Let $\beta := (\beta_1, \ldots, \beta_t)$ be a shape and $k \geq t$ be an integer. We define $\mu_{k,\beta} := (\beta_1 + 1, \beta_2 + 1, \ldots, \beta_t + 1, 1^{k-t})$ to be shape obtained from $b^n_{k,\beta}$ by removing the first row. Note that this does not depend on n.

Claim 4.15. Let $\beta := (\beta_1, \ldots, \beta_t)$ be a partition. If $m \leq n - \beta_1 - 1$, then

$$K_{b^n_{k,\beta}, h^n_m} = \binom{m}{k + |\beta|} f_{\mu_{k,\beta}} \geq \binom{m}{k + |\beta|} f_\beta.$$

Proof. First, we prove the equality. Under the assumption $m \leq n - \beta_1 - 1$, the palette h^n_m has enough ones to fill the first $\beta_1 + 1$ positions of the first row of $b^n_{k,\beta}$ (and necessarily all the ones go in this first row). Then we can choose $k + |\beta|$ colors out of $\{2, \ldots, m + 1\}$ to fill the positions *not* in the first row of $b^n_{k,\beta}$ giving a total of $\binom{m}{k+|\beta|}$ possibilities. For each such possibility, there are exactly $f_{\mu_{k,\beta}}$ ways of filling these positions with these colors, since $\mu_{k,\beta}$ is their shape.

The inequality follows by observing that the shape β is contained in the shape $\mu_{k,\beta}$ resulting in $f_{\mu_{k,\beta}} \geq f_\beta$. $\qquad\square$

The following claim provides a defective bound when $m = n - 1$ and the condition of Claim 4.15 is not met.

Claim 4.16. Let $\beta := (\beta_1, \ldots, \beta_t) \vdash \ell$. Then

$$K_{b^n_{k,\beta}, h^n_{n-1}} \geq \binom{n - \ell - 1}{k + \ell} f_{\mu_{k,\beta}} \geq \binom{n - \ell - 1}{k + \ell} f_\beta.$$

Proof. We start by the first inequality. Since we are only interested in a lower bound, we only consider the fillings which place the numbers in $[\beta_1 + 1]$ of the palette h^n_{n-1} in the first $\beta_1 + 1$ positions of the first row of $b^n_{k,\beta}$. Since $|\beta| = \ell$, this leaves at least $n - \ell - 1$ colors out of which we choose to fill the remaining rows of $b^n_{k,\beta}$ accounting for $k + \ell$ cells. As before, for each such possibility, there are exactly $f_{\mu_{k,\beta}}$ ways of filling these positions with these colors, since $\mu_{k,\beta}$ is their shape and we have filled the first $\beta_1 + 1$ cells of the first row of $b^n_{k,\beta}$ with the numbers in $[\beta_1 + 1]$.

The second inequality follows again by $f_{\mu_{k,\beta}} \geq f_\beta$ since the shape β is contained in the shape $\mu_{k,\beta}$. $\qquad\square$

Now, we relate the optimum values of the second linear program from Definition 4.9 and the first one from Definition 4.4 for n sufficiently large.

Proposition 4.17. Let n be an odd positive integer, $c > 1$ and ℓ_0 be non-negative even integer. Suppose k_0 is a positive integer such that $\ell_0 \leq k_0 \leq (n - \ell_0 - 3)/2$. Then

$$\mathrm{OPT}(P^{\ell_0, k_0}_n(c)) \leq \mathrm{OPT}(P^{\ell_0}_n(c)).$$

Proof. Fix an optimum solution (M, Ψ, x) to $P^{\ell_0}_n(c)$. Let x' be the tuple obtained from x by ignoring variables of the form $x_{b^n_{k,\beta}}$ where $k \geq k_0 + 1$. We compute Ψ' from x' using (10) and we let $M' := \max_\ell \Psi'_\ell$.

We claim that (M', Ψ', x') is a feasible solution of $P^{\ell_0, k_0}_n(c)$ (not necessarily with the same value as (M, Ψ, x)). Trivially, restrictions (9), (10), (12) and (13) are satisfied. Also observe that restrictions (11) follow from the fact that x is non-negative and it satisfied the restrictions (5) from $P^{\ell_0}_n(c)$.

15

We will now prove that $M' \leq M$. For this, it is enough to show that $\Psi'_\ell \leq \Psi_\ell$ for every non-negative even $\ell \leq \ell_0$. Fix some such ℓ. Since $k_0 \leq (n - \ell_0 - 3)/2$, if $k \leq k_0$ and β is a shape of size at most ℓ, then $(b^n_{k,\beta})^\top = b^n_{n-|\beta|-k-1,\beta^\top}$ and $n - |\beta| - k - 1 \geq k_0 + 1$ implying that the variable $x_{(b^n_{k,\beta})^\top}$ is not in x'. Now note that by Fact 3.9 we have $\chi^{b^n_{k,\beta}}((n-\ell)) = \chi^{(b^n_{k,\beta})^\top}((n-\ell))$ since $\mathrm{sgn}((n-\ell))$ is positive. By Claim 4.10, we have $\chi^{b^n_{k,\beta}}((n-\ell)) = 0$ whenever $|\beta| > \ell$. On the other hand, when $|\beta| < \ell \leq \ell_0$, since $k_0 \geq \ell_0$, by Claim 4.12 we have $\chi^{b^n_{k,\beta}}((n-\ell)) = 0$ for every integer k such that $k_0 + 1 \leq k \leq n - |\beta| - k_0 - 2$. From these and restriction (6) ($x_\lambda = x_{\lambda^\top}$) from $P^{\ell_0}_n(c)$ and $x_{(n)} = x_{(1^n)} = 1$, we conclude that

$$\Psi_\ell - \Psi'_\ell = \sum_{\beta \vdash \ell} \sum_{k=k_0+1}^{n-i-k_0-2} \chi^{b^n_{k,\beta}}((n-\ell)) \cdot x_{b^n_{k,\beta}} + 2 \cdot T^{\ell,k_0}_n(c).$$

Using Claim 4.13, we have

$$\Psi_\ell - \Psi'_\ell = \sum_{\beta \vdash \ell} \sum_{k=k_0+1}^{n-\ell-k_0-2} (-1)^k f_\beta \cdot x_{b^n_{k,\beta}} + 2 \cdot T^{\ell,k_0}_n(c)$$

$$\geq - \sum_{\beta \vdash \ell} \sum_{\substack{k=k_0+1 \\ k \text{ odd}}}^{n-\ell-k_0-2} f_\beta \cdot x_{b^n_{k,\beta}} + 2 \cdot T^{\ell,k_0}_n(c) \tag{14}$$

$$\geq -2 \sum_{\beta \vdash \ell} \sum_{\substack{k=k_0+1 \\ k \text{ odd}}}^{(n-\ell-1)/2} f_\beta \cdot x_{b^n_{k,\beta}} + 2 \cdot T^{\ell,k_0}_n(c),$$

where the last inequality follows from (6) ($x_\lambda = x_{\lambda^\top}$) and note that we are double counting the cases where $\lambda = \lambda^\top$, namely, when $k = (n - \ell - 1)/2$.

Fix an odd integer k in $[k_0 + 1, n - \ell - k_0 - 2]$. We consider two cases.

The first case is when $k \leq (n-3\ell-1)/2$, where we use the Young restriction (11) for $m = 2(k+\ell)$ (and the fact that $x \geq 0$) to get that

$$\sum_{\beta \vdash \ell} K_{b^n_{k,\beta}, h^n_{2(k+\ell)}} \cdot x_{b^n_{k,\beta}} \leq c^{2(k+\ell)}.$$

By our choice of k and by Claim 4.15, we can simplify the above equation as

$$\sum_{\beta \vdash \ell} f_\beta \cdot x_{b^n_{k,\beta}} \leq \frac{c^{2(k+\ell)}}{\binom{2(k+\ell)}{k+\ell}}. \tag{15}$$

The second case is when $k > (n-3\ell-1)/2$, where we use the Young restriction (11) for $m = n-1$ (and the fact that $x \geq 0$) to get that

$$\sum_{\beta \vdash \ell} K_{b^n_{k,\beta}, h^n_{n-1}} \cdot x_{b^n_{k,\beta}} \leq c^{n-1}.$$

By Claim 4.16, we can simplify the above equation as

$$\sum_{\beta \vdash \ell} f_\beta \cdot x_{b^n_{k,\beta}} \leq \frac{c^{n-1}}{\binom{n-\ell-1}{k+\ell}}. \tag{16}$$

16

Combining (14), (15) and (16), we conclude that $\Psi_\ell - \Psi'_\ell \geq 0$ for every non-negative even $\ell \leq \ell_0$ implying

$$\text{OPT}(P_n^{\ell_0,k_0}(c)) \leq M' \leq M = \text{OPT}(P_n^{\ell_0}(c)),$$

which concludes the proof. $\qquad\square$

4.2.3 Linear Program III

To define the third family of linear programs, we first show that the coefficients of the second family of linear programs $P_n^{\ell_0,k_0}(c)$ for ℓ_0, k_0 fixed become independent of n as long as n is sufficiently large. The main ingredient for the stabilization of the Kostka coefficients is Claim 4.15, whereas for Parseval coefficients it will be convenient to work with the following generalization of shape.

Definition 4.18. Given a shape β, an integer $k \geq \text{ht}(\beta)$ and an integer $\ell \geq |\beta|$, we let $\xi_{k,\beta,\ell}$ be the (not necessarily valid) shape obtained from $b_{k,\beta}^n$ by removing the rim hook of size $n-\ell$ that contains the hand of $b_{k,\beta}^n$, where $n \geq |\beta| + k + \beta_1 + 1$, and we let $t_{k,\beta,\ell}$ be the height of this removed rim (note that $\xi_{k,\beta,\ell}$ and $t_{k,\beta,\ell}$ do not depend on the choice of n). Note that if $|\beta| = \ell$, then $\xi_{k,\beta,\ell} = \beta$ and $t_{k,\beta,\ell} = k + 1$. See Fig. 4 for some examples.
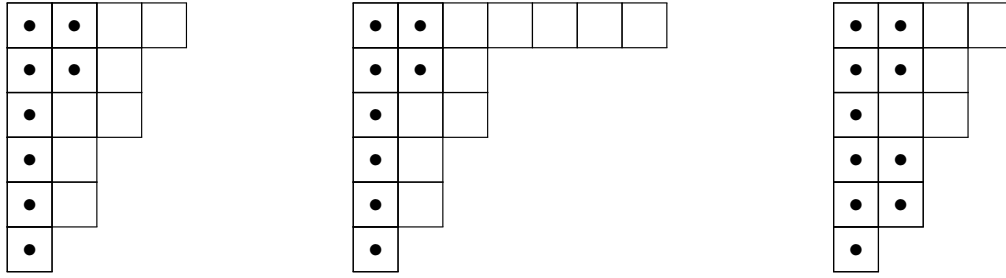
(a) A valid shape $\xi_{5,(2,2,1,1),8}$ obtained from $b_{5,(2,2,1,1)}^{15}$.

(b) A valid shape $\xi_{5,(2,2,1,1),8}$ obtained from $b_{5,(2,2,1,1)}^{18}$.

(c) An invalid shape $\xi_{5,(2,2,1,1),10}$ obtained from $b_{5,(2,2,1,1)}^{15}$.

Figure 4: Examples of derived shapes $\xi_{k,\beta,\ell}$. The independence from n is illustrated by (a) and (b).

Lemma 4.19. Suppose $n \geq |\beta|+k+\beta_1+1$ (so that $b_{k,\beta}^n$ is well-defined) and $\ell \geq |\beta|$. If $n-\ell > |\beta|+k$, then

$$\chi^{b_{k,\beta}^n}((n-\ell)) = \begin{cases} 0, & \text{if } \xi_{k,\beta,\ell} \text{ is not a valid shape} \\ (-1)^{t_{k,\beta,\ell}-1} \cdot f_{\xi_{k,\beta,\ell}}, & \text{otherwise.} \end{cases}$$

In particular, the value above does not depend on n.

Proof. We apply the Murnaghan–Nakayama rule, Theorem 3.4, removing $(n-\ell)$ first. To leave a valid shape after this removal, the following must hold: (i) if a cell is removed in the first row of $b_{k,\beta}^n$, then all cells to its right (including the hand) must also be removed. (ii) if a cell is removed in the first column of $b_{k,\beta}^n$, then all cells below it (including the foot) must also be removed.

If the hand is not removed, then nothing in the first row was removed. Since there are $k + |\beta| < n - \ell$ cells not in the first row, this case is impossible. Hence, the hand must be removed, the shape

17

remaining after this removal is $\xi_{k,\beta,\ell}$ and the removed rim hook has height $t_{k,\beta,\ell}$. If $\xi_{k,\beta,\ell}$ is a valid shape, then the removal of $(n-\ell)$ gives a sign of $(-1)^{t_{k,\beta,\ell}-1}$ and the removal of the remaining ℓ fixed points gives a factor of $f_{\xi_{k,\beta,\ell}}$. If $\xi_{k,\beta,\ell}$ is not a valid shape, then $\chi^{b^n_{k,\beta}}((n-\ell)) = 0$. $\qquad\square$

Since $\lim_{n\to\infty} \chi^{b^n_{k,\beta}}((n-\ell))$ is well-defined, we can give a name to this limit.

Definition 4.20. Lemma 4.19 above states that $\chi^{b^n_{k,\beta}}((n-\ell))$ does not depend on n as long as $n \geq \max\{|\beta| + k + \beta_1 + 1, |\beta| + k + \ell\}$, so we let $\chi_\ell^{b_{k,\beta}} := \lim_{n\to\infty} \chi^{b^n_{k,\beta}}((n-\ell))$.

The third family of linear programs, which is completely independent of n, is defined as follows.

Definition 4.21 (Linear Program III)**.** Given a positive odd integer k_0, a real $c \in (1,2)$, a non-negative even integer ℓ_0 and a positive even integer m_0, we let $P^{\ell_0,k_0,m_0}(c)$ be the following linear program.

$$
\begin{aligned}
\text{minimize} \quad & M \\
\text{s.t.} \quad & M \geq \Psi_\ell && \forall \ell \leq \ell_0 \text{ even,} && (17) \\
\text{(Parseval)} \quad & \Psi_\ell = 2 + 2 \cdot \sum_{i=0}^{\ell_0} \sum_{\beta \vdash i} \sum_{\substack{k=\mathrm{ht}(\beta) \\ k \geq 1}}^{k_0} \chi_\ell^{b_{k,\beta}} \cdot x_{k,\beta} - 2 \cdot T^{\ell,k_0}(c) && \forall \ell \leq \ell_0 \text{ even,} && (18) \\
\text{(Young)} \quad & \sum_{i=0}^{\ell_0} \sum_{\beta \vdash i} \sum_{\substack{k=\mathrm{ht}(\beta) \\ k \geq 1}}^{k_0} \binom{m}{k+|\beta|} f_{\mu_{k,\beta}} \cdot x_{k,\beta} \leq c^m - 1 && \forall m \leq m_0 \text{ even,} && (19) \\
& x_{k,\beta} \geq 0 && \forall i \in \{0,\ldots,\ell_0\}, \\
& && \forall \beta \vdash i, \\
& && \forall k \in \{\mathrm{ht}(\beta),\ldots,k_0\},
\end{aligned}
$$

where the variables are M, $(\Psi_\ell)_\ell$ and $(x_{k,\beta})_{k,\beta}$ and we have

$$
T^{\ell,k_0}(c) := 1.5 \cdot \frac{(2(k_0+\ell)+4)\left(\frac{c}{2}\right)^{2(k_0+\ell)+4} - 2(k_0+\ell)\left(\frac{c}{2}\right)^{2(k_0+\ell)+8}}{\left(1 - \left(\frac{c}{2}\right)^4\right)^2}.
$$

We now connect the optimum objective values from the third family to the second family of linear programs.

Proposition 4.22. *Let $c \in (1,2)$ and ℓ_0 be a non-negative even integer. Let k_0 be a positive odd integer such that $\ell_0 \leq k_0$.*

If $\mathrm{OPT}(P^{\ell_0,k_0,m_0}(c)) > 0$, then there exists an integer n_0 large enough such that $\mathrm{OPT}(P_n^{\ell_0,k_0}(c)) > 0$ for every odd integer $n \geq n_0$.

Proof. Fix a positive odd integer n. Suppose $s_n := (M^n, (\Psi_\ell^n)_\ell, (x_{b^n_{k,\beta}}^n)_{k,\beta})$ is an optimum solution of $P_n^{\ell_0,k_0}(c)$.

We will construct a solution $\widehat{s}_n := (\widehat{M}^n, (\widehat{\Psi}_\ell^n)_\ell, (\widehat{x}_{k,\beta}^n)_{k,\beta})$ of $P^{\ell_0,k_0,m_0}(c)$. Set $\widehat{x}_{k,\beta}^n := x_{b^n_{k,\beta}}^n$. We compute $\widehat{\Psi}^n$ from $\widehat{x}^n$ using (18) and we let $\widehat{M}^n := \max_\ell \widehat{\Psi}_\ell^n$.

Using Claim 4.15, we have that the restrictions (19) are exactly the same as the restrictions (11) (since $f_{\mu_{0,()}} = x_{(1^n)}^n = 1$). Hence, $\widehat{s}_n$ is a feasible solution of $P^{\ell_0,k_0,m_0}(c)$.

Now we compare the objective values $\widehat{M}^n$ and M^n. For this, it is enough to compare $\widehat{\Psi}_\ell^n$ and Ψ_ℓ^n for every non-negative even integer $\ell \leq \ell_0$. By Lemma 4.19 and Definition 4.20, for $n \geq 2\ell_0 + k_0 + 1$, we have

$$\widehat{\Psi}_\ell^n - \Psi_\ell^n = -2 \cdot T^{\ell,k_0}(c) + 2 \cdot T_n^{\ell,k_0}(c). \tag{20}$$

A straightforward computation done in Claim A.1 (in Appendix A) establishes that

$$\lim_{\substack{n \to \infty \\ \text{odd}}} T_n^{\ell,k_0}(c) \leq T^{\ell,k_0}(c). \tag{21}$$

For every positive odd integer n, let $\widehat{\ell}_n, \ell_n$ be such that

$$\widehat{\Psi}_{\widehat{\ell}_n}^n = \max_{\substack{\ell \leq \ell_0 \\ \ell \text{ even}}} \widehat{\Psi}_\ell^n, \qquad\qquad \Psi_{\ell_n}^n = \max_{\substack{\ell \leq \ell_0 \\ \ell \text{ even}}} \Psi_\ell^n,$$

then we have

$$\widehat{M}^n - M^n = \widehat{\Psi}_{\widehat{\ell}_n}^n - \Psi_{\ell_n}^n = \widehat{\Psi}_{\widehat{\ell}_n}^n - \Psi_{\widehat{\ell}_n}^n + \Psi_{\widehat{\ell}_n}^n - \Psi_{\ell_n}^n \leq \widehat{\Psi}_{\widehat{\ell}_n}^n - \Psi_{\widehat{\ell}_n}^n,$$

hence

$$\limsup_{\substack{n \to \infty \\ \text{odd}}} (\widehat{M}^n - M^n) \leq \limsup_{\substack{n \to \infty \\ \text{odd}}} (\widehat{\Psi}_{\widehat{\ell}_n}^n - \Psi_{\widehat{\ell}_n}^n) \leq 0,$$

where the last inequality follows from (20) and (21). Therefore, we obtain

$$\mathrm{OPT}(P^{\ell_0,k_0,m_0}(c)) \leq \liminf_{\substack{n \to \infty \\ \text{odd}}} \widehat{M}^n \leq \liminf_{\substack{n \to \infty \\ \text{odd}}} M^n = \liminf_{\substack{n \to \infty \\ \text{odd}}} \mathrm{OPT}(P_n^{\ell_0,k_0}(c)),$$

which implies the result. $\qquad\square$

The results of this section culminate in the next theorem, which reduces the problem of asymptotically upper bounding $\alpha(\mathcal{B}_n)$ to showing that a *single* linear program from the third family Definition 4.21 has a positive objective value for a given $c \in (1, 2)$.

Theorem 4.23. *Let $c_0 > 1$, ℓ_0 be a non-negative even integer, k_0 be a positive odd integer and m_0 be a positive even integer. Suppose we have $\mathrm{OPT}(P^{\ell_0,k_0,m_0}(c_0)) > 0$. Then for every $c \in (1, c_0)$, there exists $n_0 := n_0(c_0, c, \ell_0, k_0, m_0) \in \mathbb{N}$ such that for every integer $n \geq n_0$*

$$\alpha(\mathcal{B}_n) \leq 2 \cdot \frac{n!}{c^{n-1}}.$$

Proof. Given a linear program $P^{\ell_0,k_0,m_0}(c_0)$ in the third family such that $\mathrm{OPT}(P^{\ell_0,k_0,m_0}(c_0)) > 0$, by Proposition 4.22 the linear program $P_n^{\ell_0,k_0}(c_0)$ in the second family satisfies $\mathrm{OPT}(P_n^{\ell_0,k_0}(c_0)) > 0$ for n sufficiently large. By Proposition 4.17, this in turn implies that the linear program $P_n^{\ell_0}(c_0)$ in the first family satisfies $\mathrm{OPT}(P_n^{\ell_0}(c_0)) > 0$ for n sufficiently large. Finally, the bound on $\alpha(\mathcal{B}_n)$ follows from Proposition 4.7. $\qquad\square$

5   Computational Part

From Theorem 4.23, to claim an asymptotic upper bound of $\alpha(\mathcal{B}_n) \leq K \cdot n!/(c - o_n(1))^n$ with $K = 2c$, it is enough to find a single choice of parameters ℓ_0, k_0 and m_0 that makes the linear program $P^{\ell_0,k_0,m_0}(c)$ (from Definition 4.21) have positive optimum value. However, a full theoretical analysis of these linear programs remains elusive. Nevertheless, we can still computationally solve these programs for various specific choices of parameters. The closer the target c is to 2, the larger the parameter ℓ_0 (and k_0) needs to be to yield a positive optimum value. This computational approach poses its own challenges and getting $c = 1.971$ requires careful consideration on how to solve these linear programs.

The first challenge is the poor dependence of the size of these programs on the parameter ℓ_0, both in terms of number of variables and bit complexity of coefficients. More specifically, the number of variables grows at least as fast as the number of partitions of ℓ_0, which is asymptotically

$$\frac{1}{4\ell_0\sqrt{3}} \cdot \exp\left(\pi\sqrt{\frac{2\ell_0}{3}}\right),$$

by a celebrated theorem of Hardy–Ramanujan [HR18].

The second challenge is that the linear programs $P^{\ell_0,k_0,m_0}(c)$ seem to be very sensitive to numerical rounding errors preventing the use of conventional LP solvers. We believe that such sensitivity comes from the large difference in magnitude of the linear program coefficients (e.g., Kostka constants, see Claim 4.15). To avoid approximation errors and obtain an exact optimum solution, we implemented the Simplex method with support to exact rational computations (note that this is enough since $P^{\ell_0,k_0,m_0}(c)$ has rational coefficients as long as $c \in \mathbb{Q}$). Our implementation is available at https://github.com/lenacore/birkhoff_code.

In this section, we address these challenges.

5.1   Dual Linear Program

We actually solve the dual of the third linear program Definition 4.21, which is presented in Definition 5.1. By strong linear programming duality, the dual program has positive optimum value if and only if the primal has positive optimum value, so there is no loss in working with the dual program.

Definition 5.1 (Dual Linear Program). Given a positive odd integer k_0, $c \in (1, 2)$, a non-negative

even integer ℓ_0 and a positive even integer m_0, we let $D^{\ell_0,k_0,m_0}(c)$ be the following linear program.

$$\text{maximize} \quad 2 - 2 \cdot \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{\ell_0} T^{\ell,k_0}(c) \cdot w_\ell - \sum_{\substack{m \leq m_0 \\ m \text{ even}}} (c^m - 1) \cdot y_m$$

$$\text{s.t.} \quad \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{\ell_0} w_\ell = 1$$

$$(\text{Restriction } b_{k,\beta}) \quad 2 \cdot \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{\ell_0} \chi_\ell^{b_{k,\beta}} \cdot w_\ell + \sum_{\substack{m \leq m_0 \\ m \text{ even}}} \binom{m}{k+|\beta|} f_{\mu_{k,\beta}} \cdot y_m \geq 0 \qquad \forall i \in \{0,\dots,\ell_0\}, \quad (22)$$
$$\forall \beta \vdash i,$$
$$\forall k \in [k_0], k \geq \text{ht}(\beta),$$

$$w_\ell \geq 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall \ell \leq \ell_0 \text{ even},$$

$$y_m \geq 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall m \leq m_0 \text{ even},$$

where the variables are $(w_\ell)_\ell$ and $(y_m)_m$ and $T^{\ell,k_0}(c)$ is as in Definition 4.21.

There are two reasons for working with the dual linear program. First, we will be able to replace several inequalities corresponding to shapes $b_{k,\beta}$ with large leg k by a few provably more stringent inequalities, an approach that we dub "*joint large leg*" and is carried out in Section 5.2. The second reason is due to a heuristic to speed the computation, called "*fragmented heuristic*", which is explained in Section 5.3.

5.2 Joint Large Leg

To reduce the number of restrictions of the dual linear program of Definition 5.1, for a given $k \geq \ell_0$ and s we replace all restrictions associated to partitions $b_{k,\beta}$, where $|\beta| = s$, with a single more stringent restriction. Note that this modification can only decrease the objective value, which still allows us to deduce asymptotic upper bounds on $\alpha(\mathcal{B}_n)$ using Theorem 4.23.

Lemma 5.2. *Let $k \geq \ell_0 \geq s$ be non-negative integers with $k \geq 1$ and $(y_m)_m$ be non-negative. The inequality*

$$2 \cdot (-1)^k \cdot w_s + \sum_{\substack{m \leq m_0 \\ m \text{ even}}} \binom{m}{k+s} \cdot y_m \geq 0$$

implies that the inequalities (22) associated with $b_{k,\beta}$ in Definition 5.1 for every $\beta \vdash s$ are satisfied, i.e.,

$$2 \cdot \sum_{\substack{\ell=0 \\ \ell \text{ even}}}^{\ell_0} \chi_\ell^{b_{k,\beta}} \cdot w_\ell + \sum_{\substack{m \leq m_0 \\ m \text{ even}}} \binom{m}{k+|\beta|} f_{\mu_{k,\beta}} \cdot y_m \geq 0. \qquad (23)$$

Proof. Since $k \geq \ell_0 \geq \ell$, to obtain $\xi_{k,\beta,\ell}$ from some $b_{k,\beta}^n$ we must have removed some cell in the first column; thus $\xi_{k,\beta,\ell}$ is a valid shape if and only if the foot of $b_{k,\beta}^n$ was removed, which in turn is equivalent to $|\beta| = \ell$. By Lemma 4.19, inequality (23) becomes

$$2 \cdot (-1)^k \cdot f_\beta \cdot w_s + \sum_{\substack{m \leq m_0 \\ m \text{ even}}} \binom{m}{k+|\beta|} f_{\mu_{k,\beta}} \cdot y_m \geq 0.$$

Then the result follows by noticing that the y_m are non-negative and $f_{\mu_{k,\beta}} \geq f_\beta$ since β is contained in $\mu_{k,\beta}$. □

5.3 Speeding the Computation

We briefly explain three heuristics used to speed the computations. We stress that with any combination of these heuristics if the objective value of the resulting linear program is positive, then the objective value of the original dual linear program (Definition 5.1) is also positive.

- The first heuristic consists in setting some y_m to 0 (this corresponds to dropping restrictions in the primal) to decrease the size of the problem. Note that this can only decrease the optimum value.

- To reduce the bit complexity of the program, we round up $(c^m - 1)$ in the objective and we round down the Kostka constant $\binom{m}{k+|\beta|} f_{\mu_{k,\beta}}$. Similarly, this modification can only decrease the optimum value.

- The final heuristic consists in solving a small "fragment" of the dual linear program containing much fewer restrictions. Of course, this can increase the optimum value. However, we can then check if an optimum solution to this fragment problem is feasible (therefore optimum) for original dual linear program. For reference, our best result used a fragment containing only restrictions associated with partitions $b^n_{k,\beta}$ for β having height at most 1 or being the partition $(1,1)$. This suggests that some optimum solutions of the dual program might have enough structure to be analyzable completely symbolically.

5.4 Computational Results

We finish this section by presenting some computational results in Table 1, which contains some parameters for which the dual linear program has positive optimum value.

We remark that solving $P^{\ell_0,k_0,m_0}(c)$ (or $D^{\ell_0,k_0,m_0}(c)$) with $\ell_0 = 0$ can be viewed as (essentially) the Kane–Lovett–Rao approach [KLR17]. In this case, we obtain an improved $c = 1.49$ over $c = \sqrt{2}$ from [KLR17] since we work with slightly stronger inequalities. It is interesting to see that for $\ell_0 = 0$ making $k_0 > 19$ does not allow us to obtain a larger c for which the dual has positive optimum value. This means that increasing ℓ_0 is crucial to obtain better values of c.

Combining the theoretical results from Section 4.2 and the computational results of this section, we obtain our main result.

Theorem 1.3. *We have*

$$\alpha(\mathcal{B}_n) \leq O\left(\frac{n!}{1.97^n}\right).$$

Proof. Follows from Theorem 4.23 and $\mathrm{OPT}(D^{\ell_0,k_0,m_0}(1.971)) > 0$ for $\ell_0 = 74$, $k_0 = 469$ and $m_0 = 1086$. □

6 Explicit Constructions

In this section we provide explicit constructions of independent sets and proper colorings of the Birkhoff graph. Although [KLR17] only constructs independent sets and only when n is a power

ℓ_0	c	k_0	
0	1.49	19	
2	1.69	29	
4	1.72	29	
6	1.78	39	
8	1.80	39	
10	1.82	49	
12	1.85	59	
14	1.87	79	
20	1.90	89	$*$
30	1.93	139	$*$
50	1.95	199	$**$
70	1.97	539	$**$
74	1.971	469	$**$

Table 1: List of parameters that yield positive optimum values for $D^{\ell_0,k_0,m_0}(c)$ (in all cases, we take $m_0 = 2(\ell_0 + k_0)$). Entries marked with $*$ were computed using the fragment heuristic. Entries marked with $**$ were computed using all three heuristics and joint large leg. For reference, the instance with $\ell_0 = 74$ has approximately 3×10^9 restrictions even after the joint large leg heuristic (before, the number was approximately 2.4×10^{10}), whereas the number of restrictions for $\ell_0 = 14$ and $\ell_0 = 30$ are approximately 3.8×10^4 and 3×10^6, respectively.

of 2, our constructions build on similar ideas. However, since we adopt a simpler group theoretical language, this enables us to achieve a modest improvement of an $n/2$ factor whenever n is a power of 2. Even though an explicit independent set achieving the same bound can be deduced from our coloring, we first directly present an independent set construction as it is simpler and serves as a warm up for the coloring construction.

6.1 Independent Set

We start by presenting in Lemma 6.1 the recursive step of a construction of an independent set that works in any size n. Such construction step will later be improved in Lemma 6.2 by a factor of 2 conditioned on n being divisible by 4.

Lemma 6.1. *Let $n \geq 2$ be an integer and suppose A is an independent set of $\mathcal{B}_{\lceil n/2 \rceil}$. Then the following is an independent set of $\mathcal{B}_n$ (under the natural inclusion of $S_{\lfloor n/2 \rfloor} \times S_{\lceil n/2 \rceil}$ in S_n)*

$$A' := \{(\sigma, (\sigma')^{-1}\tau) \in S_{\lfloor n/2 \rfloor} \times S_{\lceil n/2 \rceil} \mid \sigma \in S_{\lfloor n/2 \rfloor}, \tau \in A\},$$

where σ' is the natural extension of σ to $[\lceil n/2 \rceil]$ (by possibly fixing $\lceil n/2 \rceil$). In particular, we have

$$|A'| = \left\lfloor \frac{n}{2} \right\rfloor! \cdot |A|.$$

Proof. Note that in $S_{\lfloor n/2 \rfloor} \times S_{\lceil n/2 \rceil}$ a single cycle must either act only on the first part or only on the second part. This means that if $(\sigma_1, (\sigma_1')^{-1}\tau_1), (\sigma_2, (\sigma_2')^{-1}\tau_2) \in A'$ are adjacent in $\mathcal{B}_n$, then either $\sigma_1 = \sigma_2$ and $(\sigma_1')^{-1}\tau_1 \cdot \tau_2^{-1}\sigma_2'$ is a single cycle; or $\sigma_1 \cdot \sigma_2^{-1}$ is a single cycle and $(\sigma_1')^{-1}\tau_1 = (\sigma_2')^{-1}\tau_2$.

23

In the first case, since we also have $\sigma_1' = \sigma_2'$, it follows that $\tau_1 \tau_2^{-1}$ must also be a single cycle, contradicting the assumption that A is independent in $\mathcal{B}_{\lceil n/2 \rceil}$. In the second case, we have $\tau_1 \cdot \tau_2^{-1} = \sigma_1' \cdot (\sigma_2')^{-1}$, which must be a single cycle (as $\sigma_1 \cdot \sigma_2^{-1}$ is so), generating the same contradiction. $\square$

Lemma 6.2. *Let $n \geq 4$ be an integer divisible by 4, let γ be the product of transpositions $\gamma :=$ $\prod_{i=1}^{n/2}(i, n/2 + i)$ and suppose A is an independent set of $\mathcal{B}_{n/2}$ containing only permutations of positive sign. Let*

$$A' := \{(\sigma, \sigma^{-1}\tau) \in S_{n/2} \times S_{n/2} \mid \sigma \in S_{n/2}, \tau \in A\}.$$

Then $A' \cup \gamma A'$ is an independent set of $\mathcal{B}_n$ (under the natural inclusion of $S_{n/2} \times S_{n/2}$ in S_n) containing only permutations of positive sign. In particular, we have

$$|A' \cup \gamma A'| = 2 \left(\frac{n}{2}\right)! \cdot |A|.$$

Proof. Since n is divisible by 4, we have $\mathrm{sgn}(\gamma) = 1$, so all permutations of $A' \cup \gamma A'$ have positive sign. By Lemma 6.1, we know that A' is an independent set of $\mathcal{B}_n$ and since $\pi \mapsto \gamma \pi$ is an automorphism of $\mathcal{B}_n$, it follows that $\gamma A'$ is also an independent set of $\mathcal{B}_n$.

This means that if $A' \cup \gamma A'$ is not independent in $\mathcal{B}_n$, it must contain an edge between some $\gamma \cdot (\sigma_1, \sigma_1^{-1}\tau_1) \in \gamma A'$ and some $(\sigma_2, \sigma_2^{-1}\tau_2) \in A'$, that is, the permutation

$$\pi := \gamma \cdot (\sigma_1 \sigma_2^{-1}, \sigma_1^{-1}\tau_1 \tau_2^{-1}\sigma_2)$$

must be a single cycle. But note that from the definition of γ, the permutation π cannot have any fixed points, so π must be a full cycle, which in particular implies that $\mathrm{sgn}(\pi) = -1$ (as n is even). But this contradicts the fact that $\gamma \cdot (\sigma_1, \sigma_1 \tau_1)$ and $(\sigma_2, \sigma_2 \tau_2)$ both have positive sign.

Note that $\gamma A'$ is contained in the left coset $\gamma(S_{n/2} \times S_{n/2})$, so it must be disjoint from $A' \subseteq S_{n/2} \times S_{n/2}$. $\square$

Equipped with these two lemmas, we can now prove Theorem 1.4 (restated below). When n is a power of 2, the factor of 2 advantage of Lemma 6.2 will compound to a total advantage of $n/2$ in the final construction.

Theorem 1.4. *For every $n \in \mathbb{N}_+$, we have*

$$\alpha(\mathcal{B}_n) \geq \prod_{i=1}^{\lfloor \log_2(n) \rfloor} \left\lfloor \frac{n}{2^i} \right\rfloor! \geq \frac{n!}{4^n} \cdot 2^{\Theta((\log(n))^2)}.$$

If n is a power of 2, then we can improve the bound above to

$$\alpha(\mathcal{B}_n) \geq \frac{n}{2} \prod_{i=1}^{\log_2(n)-1} 2^i!$$

Proof. The first part of the theorem follows by a simple induction in n using Lemma 6.1 (the base case of $n = 1$ consists of an independent set of size 1 in $\mathcal{B}_1$). The second part follows by induction in $\log_2(n)$ using Lemma 6.2 instead and base cases of $n = 1$ and $n = 2$, in which the independent sets have size 1. $\square$

6.2 Coloring

Just as in the case of the independent set, we start by presenting in Lemma 6.3 the recursive step of a construction that works in any size n and later improve this construction in Lemma 6.4 by a factor of 2 when n is divisible by 4.

Lemma 6.3. *Let $n \geq 2$ be an integer and suppose $f \colon S_{\lceil n/2 \rceil} \to \mathcal{X}$ is a proper coloring of $\mathcal{B}_{\lceil n/2 \rceil}$. Then there exists an explicit proper coloring of $\mathcal{B}_n$ with*

$$\binom{n}{\lceil n/2 \rceil} \cdot |\mathcal{X}|$$

colors.

Proof. Set $A \coloneqq [\lceil n/2 \rceil]$ and $B \coloneqq [n] \setminus A$. Let $H \coloneqq S_A \times S_B \subseteq S_n$ and let $\mathcal{T} \coloneqq \{t_1, \ldots, t_k\}$ be a set of representatives of (left) cosets of H in S_n. Note that $k = \binom{n}{\lceil n/2 \rceil}$. Let ι_A, ι_B be natural injections of S_A, S_B in $S_{[\lceil n/2 \rceil]}$ (preserving the cycle type). For convenience, we use $\widehat{h}_A \coloneqq \iota_A(h_A)$ for $h_A \in S_A$ and similarly for S_B. Define the coloring $f' \colon S_n \to \mathcal{T} \times \mathcal{X}$ as

$$f'(t_i h_A h_B) \coloneqq (t_i, f(\widehat{h}_A \widehat{h}_B)),$$

for every $i \in [k]$, every $h_A \in S_A$ and every $h_B \in S_B$.

Now we prove that f' is a proper coloring of $\mathcal{B}_n$. By construction, permutations of different cosets of H receive different colors. Let σ and τ be permutations in the same coset $t_i H$ such that $\pi \coloneqq \sigma \tau^{-1}$ is a non-trivial cycle, i.e., σ and τ are adjacent in $\mathcal{B}_n$. Write $\sigma = t_i \cdot g_A g_B$ and $\tau = t_i \cdot h_A h_B$ for some $g_A, h_A \in S_A$ and $g_B, h_B \in S_B$ so that

$$\pi = t_i g_A g_B \cdot h_B^{-1} h_A^{-1} t_i^{-1} = t_i (g_A h_A^{-1})(g_B h_B^{-1}) t_i^{-1},$$

where the second equality follows because elements of S_A commute with elements of S_B. Since $t_i^{-1} \pi t_i$ is also a single cycle, exactly one of $(g_A h_A^{-1})$ or $(g_B h_B^{-1})$ must be a single cycle and the other the identity. We show that $f'(\sigma) \neq f'(\tau)$ by showing that $f(\widehat{g}_A \widehat{g}_B) \neq f(\widehat{h}_A \widehat{h}_B)$. Suppose first that $(g_A h_A^{-1})$ is a cycle and $g_B = h_B$. Then

$$\widehat{g}_A \widehat{g}_B \cdot (\widehat{h}_A \widehat{h}_B)^{-1} = \widehat{g}_A (\widehat{h}_A)^{-1}$$

is a cycle and thus $f(\widehat{g}_A \widehat{g}_B) \neq f(\widehat{h}_A \widehat{h}_B)$. In the second case we have $g_A = h_A$ and $(g_B h_B^{-1})$ is a cycle, so

$$\widehat{g}_A \widehat{g}_B \cdot (\widehat{h}_A \widehat{h}_B)^{-1} = \widehat{g}_A (\widehat{g}_B \widehat{h}_B^{-1}) \widehat{g}_A^{-1}$$

is also a cycle and again $f(\widehat{g}_A \widehat{g}_B) \neq f(\widehat{h}_A \widehat{h}_B)$. Therefore, f' is a proper coloring of $\mathcal{B}_n$ with

$$\binom{n}{\lceil n/2 \rceil} \cdot |\mathcal{X}|$$

colors. $\square$

The idea to improve the coloring construction above by a factor of 2 is a small generalization of the idea of Lemma 6.2 for the independent set.

Lemma 6.4. *Let $n \geq 4$ be an integer divisible by 4 and suppose $f \colon S_{n/2} \to \mathcal{X}$ of $\mathcal{B}_{n/2}$ is a proper coloring that respects signs in the sense that permutations in the same color class have the same sign. Then there exists an explicit proper coloring f' of $\mathcal{B}_n$ that respects signs and with*

$$\frac{1}{2} \cdot \binom{n}{n/2} \cdot |\mathcal{X}|$$

colors.

Proof. We proceed as in the proof of Lemma 6.3, but instead of assigning a color to each (left) coset of $H := S_A \times S_B$ (where $A := [n/2]$ and $B := [n] \setminus [n/2]$) we will be able to assign the same color to every two cosets, thereby using only half as many colors. Again, we let ι_A, ι_B be natural injections of S_A, S_B in $S_{[\lceil n/2 \rceil]}$ and we use the notation $\widehat{h}_A := \iota_A(h_A)$ for $h_A \in S_A$ and similarly for S_B.

Let γ be the product of transpositions $\gamma := \prod_{i=1}^{n/2}(i, n/2 + i)$ and note that since n is divisible by 4, we have $\mathrm{sgn}(\gamma) = 1$.

Note that since $\gamma \notin H$, it follows that for every $u \in S_n$, we have $uH \neq u\gamma H$. In particular, for $k := \binom{n}{n/2}$, we can find $u_1, \ldots, u_{k/2} \in S_n$ so that the cosets of H are precisely

$$u_1 H, u_2 H, \ldots, u_{k/2} H, u_1 \gamma H, u_2 \gamma H, \ldots, u_{k/2} \gamma H.$$

Let $\mathcal{U} := \{u_1, \ldots, u_{k/2}\}$ and define the coloring $f' \colon S_n \to \mathcal{U} \times \mathcal{X}$ as

$$f'(u_i \cdot h_A h_B) := (u_i, f(\widehat{h}_A \widehat{h}_B));$$
$$f'(u_i \gamma \cdot h_A h_B) := (u_i, f(\widehat{h}_A \widehat{h}_B));$$

for every $i \in [k/2]$, every $h_A \in S_A$ and every $h_B \in S_B$.

Let us prove that f' is a proper coloring. We classify the edges of $\mathcal{B}_n$ into the following six types.

(i) $\{u_i \cdot g_A g_B, u_j \cdot h_A h_B\}$ for some $i, j \in [k/2]$ with $i \neq j$, some $g_A, h_A \in S_A$ and some $g_B, h_B \in S_B$.

(ii) $\{u_i \gamma \cdot g_A g_B, u_j \gamma \cdot h_A h_B\}$ for some $i, j \in [k/2]$ with $i \neq j$, some $g_A, h_A \in S_A$ and some $g_B, h_B \in S_B$.

(iii) $\{u_i \gamma \cdot g_A g_B, u_j \cdot h_A h_B\}$ for some $i, j \in [k/2]$ with $i \neq j$, some $g_A, h_A \in S_A$ and some $g_B, h_B \in S_B$.

(iv) $\{u_i \cdot g_A g_B, u_i \cdot h_A h_B\}$ for some $i \in [k/2]$, some $g_A, h_A \in S_A$ and some $g_B, h_B \in S_B$.

(v) $\{u_i \gamma \cdot g_A g_B, u_i \gamma \cdot h_A h_B\}$ for some $i \in [k/2]$, some $g_A, h_A \in S_A$ and some $g_B, h_B \in S_B$.

(vi) $\{u_i \gamma \cdot g_A g_B, u_i \cdot h_A h_B\}$ for some $i \in [k/2]$, some $g_A, h_A \in S_A$ and some $g_B, h_B \in S_B$.

Edges of the types (i), (ii) and (iii) are not monochromatic by observing the first coordinate of f'. Edges of the types (iv) and (v) are not monochromatic by an argument completely analogous to that of Lemma 6.3.

Let us then consider an edge of type (vi). Since

$$u_i \gamma \cdot g_A g_B \cdot h_B^{-1} h_A^{-1} \cdot u_i^{-1}$$

is a cycle, it follows that

$$\pi := \gamma \cdot g_A g_B \cdot h_B^{-1} h_A^{-1}$$

is a cycle. But the definition of γ implies that π does not have any fixed points, so it must be a full cycle, hence $\mathrm{sgn}(\pi) = -1$. Therefore, exactly one of $g_A g_B$ and $h_A h_B$ must have negative sign, hence $f(\widehat{g_A}\widehat{g_B}) \neq f(\widehat{h_A}\widehat{h_B})$ as f respects signs, which implies that $f'(u_i \gamma \cdot g_A g_B) \neq f'(u_i \cdot h_A h_B)$.

It remains to show that f' respects signs. It is enough to show that for every $i \in [k/2]$, every $g_A, h_A \in S_A$ and every $g_B, h_B \in S_B$ such that $f(\widehat{g_A}\widehat{g_B}) = f(\widehat{h_A}\widehat{h_B})$, the following permutations have the same sign

$$u_i \cdot g_A g_B, \qquad u_i \gamma \cdot g_A g_B, \qquad u_i \cdot h_A h_B, \qquad u_i \gamma \cdot h_A h_B.$$

Since $\mathrm{sgn}(\gamma) = 1$, the first two permutations have the same sign. The same argument shows the last two have the same sign. Hence, it is enough to show that $\mathrm{sgn}(u_i \cdot g_A g_B) = \mathrm{sgn}(u_i \cdot h_A h_B)$, which is equivalent to $\mathrm{sgn}(g_A g_B) = \mathrm{sgn}(h_A h_B)$ and this follows from the fact that $f(\widehat{g_A}\widehat{g_B}) = f(\widehat{h_A}\widehat{h_B})$ and that f respects signs. □

Similarly to the independence number, when n is a power of 2, the factor of 2 advantage of Lemma 6.4 compounds to a total advantage of $n/2$ in the final construction of a proper coloring.

Theorem 1.5. *Let n be a positive integer. Then there is an explicit proper coloring establishing*

$$\chi(\mathcal{B}_n) \leq \prod_{i=0}^{\lceil \log_2(n) \rceil} \binom{\lceil n/2^i \rceil}{\lceil n/2^{i+1} \rceil} \leq \frac{4^n}{2^{\Theta((\log(n))^2)}}.$$

If n is a power of 2, then there is an explicit proper coloring strengthening the bound above to

$$\chi(\mathcal{B}_n) \leq \frac{2}{n} \prod_{i=1}^{\log_2(n)} \binom{2^i}{2^{i-1}}.$$

Proof. The first part of the theorem follows by a induction in n using Lemma 6.3 (the base case of $n = 1$ consists of a trivial coloring). The second part follows by induction in $\log_2(n)$ using Lemma 6.4 instead and base cases of $n = 1$ and $n = 2$, in which the coloring is rainbow. □

Acknowledgements

References

[BA96]   L. Billera and Sarang Aravamuthan. The combinatorics of permutation polytopes. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 24, 01 1996. 1

[Bar02]   A. Barvinok. *A Course in Convexity*. Graduate studies in mathematics. American Mathematical Society, 2002. 1

[Bir46]   G. Birkhoff. Tres observaciones sobre el algebra lineal. *Universidad Nacional de Tucuman, Revista. Serie A*, 5, 1946. 1

[CM09]   E. Rodney Canfield and Brendan D. McKay. The asymptotic volume of the Birkhoff polytope. *Online J. Anal. Comb.*, (4):4, 2009. 1

[GHJY14]   P. Gopalan, C. Huang, B. Jenkins, and S. Yekhanin. Explicit maximally recoverable codes with locality. *IEEE Transactions on Information Theory*, 60(9):5245–5256, Sep. 2014. 1

[GHK+17]   Parikshit Gopalan, Guangda Hu, Swastik Kopparty, Shubhangi Saraf, Carol Wang, and Sergey Yekhanin. Maximally recoverable codes for grid-like topologies. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA 17, 2017. 1, 2

[Hof69]   Alan J. Hoffman. On the eigenvalues and coloring of graphs. In *Graph Theory and its Applications*, 1969. 1

[HR18]   G. H. Hardy and S. Ramanujan. Asymptotic Formula in Combinatory Analysis. *Proceedings of the London Mathematical Society*, s2-17(1):75–115, 01 1918. 20

[KLR17]   Daniel Kane, Shachar Lovett, and Sankeerth Rao. The independence number of the birkhoff polytope graph, and applications to maximally recoverable codes. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 252–259, 2017. 1, 2, 3, 7, 8, 22, 27, 31

[Pak00]   Igor Pak. Four questions on birkhoff polytope. *Annals of Combinatorics*, 4, 2000. 1

[Rob55]   Herbert Robbins. A remark on Stirling's formula. *The American Mathematical Monthly*, 62(1):26–29, 1955. 29

[Sag13]   B.E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. Graduate Texts in Mathematics. Springer New York, 2013. 4, 5, 6

[SS96]   L.L. Scott and J.P. Serre. *Linear Representations of Finite Groups*. Graduate Texts in Mathematics. Springer New York, 1996. 4

A   Tail Bounds Asymptotics

The tail values $T_n^{\ell,k_0}(c)$ from the linear program II, in Definition 4.9, can be bounded by $T^{\ell,k_0}(c)$ from the linear program III in Definition 4.21 as long as n is sufficiently large. More precisely, our goal in this section is to prove the following.

Claim A.1. *Let k_0 be an odd positive integer and ℓ be a non-negative even integer. Then*

$$\lim_{\substack{n \to \infty \\ odd}} T_n^{\ell,k_0}(c) \leq T^{\ell,k_0}(c).$$

First, we recall reasonably sharp bounds on the Stirling's approximation in Appendix A.1, then we derive a simple result about a series closely related to the geometric series in Appendix A.2 and we finally relate $T_n^{\ell,k_0}(c)$ and $T^{\ell,k_0}(c)$ in Appendix A.3.

A.1 Stirling Approximation

We rely on Robbins' version of Stirling's approximation.

Theorem A.2 (Robbins' version of Stirling's approximation [Rob55])**.** *For $n \in \mathbb{N}_+$, we have*

$$n! = \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n \cdot e^{F(n)},$$

where

$$\frac{1}{12n+1} < F(n) < \frac{1}{12n}.$$

The approximation above gives us the following approximation of the *"middle binomial"*.

Corollary A.3. *For $n \in \mathbb{N}_+$, we have*

$$\binom{2n}{n} = \frac{2^{2n}}{\sqrt{n\pi}} e^{F(2n)-2F(n)},$$

where $F(n)$ is as in Theorem A.2. In particular, we have

$$\binom{2n}{n} \geq \frac{2^{2n}}{\sqrt{n\pi}} e^{-\frac{2}{15n}},$$

A.2 Series Related to the Geometric Series

We will need a closed form expression for a series related to the geometric series.

Claim A.4. *For $|q| < 1$, $a \in \mathbb{N}_+$ and $n_0, b \in \mathbb{N}$, we have*

$$\sum_{n \geq n_0} (an+b)q^{an+b} = \frac{q^b}{(1-q^a)^2}((an_0+b)q^{an_0} + (a-an_0-b)q^{an_0+a}).$$

Proof. Indeed

$$\sum_{n \geq n_0}(an+b)q^{an+b} = q\sum_{n \geq n_0}(an+b)q^{an+b-1}$$

$$= q\left(\frac{d}{dt}\sum_{n \geq n_0}t^{an+b}\right)\Bigg|_{t=q}$$

$$= q\left(\frac{d}{dt}\frac{t^{an_0+b}}{1-t^a}\right)\Bigg|_{t=q}$$

$$= q\frac{(an_0+b)q^{an_0+b-1}(1-q^a) + aq^{an_0+b+a-1}}{(1-q^a)^2},$$

which simplifies to the claimed bound. $\square$

A.3 Comparing Tail Bounds

We proceed to relate $T_n^{\ell,k_0}(c)$ and $T^{\ell,k_0}(c)$, but first we will need two technical claims.

Claim A.5. *Let k_0 be a positive odd integer. We have*

$$\sum_{\substack{k=k_0+2 \\ k\,odd}}^{\infty} \frac{c^{2k+2\ell}}{\binom{2k+2\ell}{k+\ell}} \leq T^{\ell,k_0}(c).$$

Proof. Indeed, we have

$$
\begin{aligned}
\sum_{\substack{k=k_0+2 \\ k \text{ odd}}}^{\infty} \frac{c^{2k+2\ell}}{\binom{2k+2\ell}{k+\ell}} &\leq \sum_{\substack{k=k_0+2 \\ k \text{ odd}}}^{\infty} \frac{c^{2(k+\ell)}}{\frac{2^{2(k+\ell)}}{\sqrt{\pi(k+\ell)}} e^{-\frac{2}{15(k+\ell)}}} && \text{(By Corollary A.3.)} \\
&\leq \sqrt{\pi} \cdot e^{\frac{2}{15(k_0+\ell+2)}} \sum_{\substack{k=k_0+2 \\ k \text{ odd}}}^{\infty} \sqrt{k+\ell}\left(\frac{c}{2}\right)^{2(k+\ell)} \\
&\leq \sqrt{\pi} \cdot e^{\frac{2}{15(k_0+\ell+2)}} \sum_{\substack{k=k_0+2 \\ k \text{ odd}}}^{\infty} (k+\ell)\left(\frac{c}{2}\right)^{2(k+\ell)} \\
&= \frac{\sqrt{\pi}}{2} \cdot e^{\frac{2}{15(k_0+\ell+2)}} \sum_{i=(k_0+1)/2}^{\infty} (4i+2\ell+2)\left(\frac{c}{2}\right)^{4i+2\ell+2} && \text{(Change } k \text{ to } 2i+1.) \\
&= \frac{\sqrt{\pi}}{2} \cdot e^{\frac{2}{15(k_0+\ell+2)}} \cdot \frac{(2(k_0+\ell)+4)\left(\frac{c}{2}\right)^{2(k_0+\ell)+4} - 2(k_0+\ell)\left(\frac{c}{2}\right)^{2(k_0+\ell)+8}}{\left(1-\left(\frac{c}{2}\right)^4\right)^2} && \text{(By Claim A.4.)} \\
&\leq T^{\ell,k_0}(c).
\end{aligned}
$$

$\square$

We will need a bound on the ratio between a middle binomial and a defective middle binomial as follows.

Claim A.6. *If d and $s \geq 2d$ are non-negative integers, then*

$$\frac{\binom{2s}{s}}{\binom{2s-d}{s}} \leq 3^d.$$

Proof. Since $s \geq 2d$, we get

$$\frac{\binom{2s}{s}}{\binom{2s-d}{s}} = \frac{(2s)!(s-d)!}{s!(2s-d)!} = \frac{(2s)_d}{(s)_d} = \prod_{i=0}^{d-1} \frac{2s-i}{s-i} = \prod_{i=0}^{d-1}\left(2+\frac{i}{s-i}\right) \leq 3^d. \qquad \square$$

Now we are ready to prove the main result of this section, which we restate below for convenience.

Claim A.1. *Let k_0 be an odd positive integer and ℓ be a non-negative even integer. Then*

$$\lim_{\substack{n\to\infty \\ odd}} T_n^{\ell,k_0}(c) \leq T^{\ell,k_0}(c).$$

Proof. For a positive odd integer n, let

$$T := \sum_{\substack{k=k_0+2 \\ k \text{ odd}}}^{\infty} \frac{c^{2k+2\ell}}{\binom{2k+2\ell}{k+\ell}} \qquad\qquad T'_n := \sum_{\substack{k=\frac{n-3\ell-1}{2}+1 \\ k \text{ odd}}}^{\frac{n-\ell-1}{2}} \frac{c^{n-1}}{\binom{n-\ell-1}{k+\ell}}.$$

In the sum of T'_n, note that $(n-3\ell-1)/2 \le k \le (n-\ell-1)/2$ implies $|(n-\ell-1)-2(k+\ell)| \le 2\ell$ and $|(n-1)-2(k+\ell)| \le \ell$. By Claim A.6 for $n \ge 9\ell+1$, we have

$$T'_n \le c^{\ell} \cdot 3^{2\ell} \sum_{\substack{k=\frac{n-3\ell-1}{2}+1 \\ k \text{ odd}}}^{\frac{n-\ell-1}{2}} \frac{c^{2(k+\ell)}}{\binom{2(k+\ell)}{k+\ell}}.$$

Observe that if $n \ge 2k_0 + 3\ell + 3$, the sum above is contained in the tail of T, which is a convergent series by Claim A.5. Hence, $T'_n \xrightarrow[\text{odd}]{n\to\infty} 0$. Therefore, we obtain

$$\lim_{\substack{n\to\infty \\ \text{odd}}} T_n^{\ell,k_0}(c) = \lim_{\substack{n\to\infty \\ \text{odd}}} \sum_{\substack{k=k_0+2 \\ k \text{ odd}}}^{\frac{n-3\ell-1}{2}} \frac{c^{2k+2\ell}}{\binom{2k+2\ell}{k+\ell}} + \lim_{\substack{n\to\infty \\ \text{odd}}} \sum_{\substack{k=\frac{n-3\ell-1}{2}+1 \\ k \text{ odd}}}^{\frac{n-\ell-1}{2}} \frac{c^{n-1}}{\binom{n-\ell-1}{k+\ell}}$$

$$= T + \lim_{\substack{n\to\infty \\ \text{odd}}} T'_n = T \le T^{\ell,k_0}(c),$$

where the last inequality follows from Claim A.5 again. $\qquad\square$

B KLR Proofs

For the reader's convenience we recall some proofs either from [KLR17] or implicit in it.

Fact 3.8 (From KLR). *Let $\lambda \vdash n$. Then $\chi^\lambda(\phi_A) \ge 0$.*

Proof. We have

$$\chi^\lambda(\phi_A) = \frac{1}{|A|^2} \sum_{\pi,\pi' \in A} \chi^\lambda(\pi(\pi')^{-1})$$

$$= \frac{1}{|A|^2} \sum_{\pi,\pi' \in A} \mathrm{tr}(S^\lambda(\pi) S^\lambda(\pi')^\top) = \mathrm{tr}(S^\lambda(\xi) S^\lambda(\xi)^\top) \ge 0,$$

where $\xi := \sum_{\pi \in A} \pi / |A|$. $\qquad\square$

Using the pseudorandomness condition, we can bound the character of M^μ on ϕ_A. In order to do so, observe that the action of S_n on $[n]_k$ corresponds precisely to the action of S_n on the Young module $M^{h_k^n}$ corresponding to the hook h_k^n of leg k (since the leg of a tabloid $[T]$ of shape h_k^n corresponds to a tuple in $[n]_k$ and the order of elements in the first row of $[T]$ is arbitrary). In this case, we refer to tuples and tabloids interchangeably.

Claim 3.11 (Implicit in KLR)**.** *If A is (k, r)-pseudorandom, then*

$$\operatorname{tr}(M^{h_k^n}(\phi_A)) \leq r.$$

Proof. Let $\mu := h_k^n$. We explore the uniformity of the action of S_n on $[n]_k$ given by the (k, r)-pseudorandomness assumption. Consider the matrix representation of M^μ indexed by k-tuples. More precisely, for $\pi \in S_n$ and $I, J \in [n]_k$, we have

$$M_{I,J}^\mu(\pi) := 1_{[\pi(J)=I]}.$$

Set $\xi := \sum_{\pi \in A} \pi / |A|$. Then (k, r)-pseudorandomness yields

$$M_{I,J}^\mu(\xi) = \Pr_{\pi \in A}[\pi(J) = I] < \frac{r}{(n)_k}.$$

Hence

$$\begin{aligned}
\operatorname{tr}(M^\mu(\phi_A)) &= \operatorname{tr}(M^\mu(\xi) M^\mu(\xi)^\top) \\
&= \sum_{I,J \in [n]_k} M^\mu(\xi)_{I,J} \cdot M^\mu(\xi)_{I,J} \\
&\leq \sum_{I,J \in [n]_k} \frac{r}{(n)_k} \cdot M^\mu(\xi)_{I,J} = r,
\end{aligned}$$

where the last equality follows from M^μ having exactly one entry of value 1 and all others zero in each column. $\square$

We can bound an arbitrary non-trivial character in terms of the pseudorandomness parameter and an appropriate Kostka number.

Lemma 3.12 (Implicit in KLR)**.** *Let $\lambda \vdash n$ be non-trivial (i.e., $\lambda \neq (1^n)$). If $A \subseteq S_n$ is (k, r)-pseudorandom and $K_{\lambda, h_k^n} \neq 0$, then*

$$\chi^\lambda(\phi_A) \leq \frac{r-1}{K_{\lambda, h_k^n}}.$$

Proof. By the Young's rule, Theorem 3.3, we obtain

$$\operatorname{tr}(M^{h_k^n}(\phi_A)) = \sum_{\lambda' \vdash n} K_{\lambda', h_k^n} \cdot \chi^{\lambda'}(\phi_A).$$

Using the bound on $\operatorname{tr}(M^{h_k^n}(\phi_A))$ from Claim 3.11 gives

$$\sum_{\lambda' \vdash n} K_{\lambda', h_k^n} \cdot \chi^{\lambda'}(\phi_A) \leq r.$$

Since $\chi^{(1^n)}(\phi_A) = K_{(1^n), h_k^n} = 1$ and $\chi^{\lambda'}(\phi_A) \geq 0$ for every $\lambda' \vdash n$ from Fact 3.8, we have

$$1 + K_{\lambda, h_k^n} \cdot \chi^\lambda(\phi_A) \leq r,$$

and the bound follows. $\square$