SIAM J. COMPUT.
Vol. 48, No. 4, pp. 1425–1435

THE INDEPENDENCE NUMBER OF THE BIRKHOFF POLYTOPE GRAPH, AND APPLICATIONS TO MAXIMALLY RECOVERABLE CODES[*]

DANIEL KANE[†], SHACHAR LOVETT[†], AND SANKEERTH RAO[†]

Abstract. Maximally recoverable codes are codes designed for distributed storage which combine quick recovery from single node failure and optimal recovery from catastrophic failure. Gopalan et al. [*Maximally recoverable codes for grid-like topologies*, in Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 2017, pp. 2092–2108] studied the alphabet size needed for such codes in grid topologies and gave a combinatorial characterization for it. Consider a labeling of the edges of the complete bipartite graph $K_{n,n}$, with labels coming from $\mathbb{F}_2^d$, that satisfies the following condition: for any simple cycle, the sum of the labels over its edges is nonzero. The minimal d where this is possible controls the alphabet size needed for maximally recoverable codes in $n \times n$ grid topologies. Prior to the current work, it was known that d is between $(\log n)^2$ and $n \log n$. We improve both bounds and show that d is linear in n. The upper bound is a recursive construction which beats the random construction. The lower bound follows by first relating the problem to the independence number of the Birkhoff polytope graph, and then providing tight bounds for it using the representation theory of the symmetric group.

Key words. Birkhoff polytope, maximally recoverable codes, coding theory, graph theory, representation theory

AMS subject classifications. 05E10, 68P30, 05Cxx, 52Bxx

DOI. 10.1137/18M1205856

1. Introduction.

The Birkhoff polytope is the convex hull of $n \times n$ doubly stochastic matrices. The Birkhoff polytope graph is the graph associated with its 1-skeleton. This graph is well studied, as it plays an important role in combinatorics and optimization; see, for example, the book of Barvinok [2]. For us, this graph arose naturally in the study of certain maximally recoverable codes. Our main technical results are tight bounds on the independence number of the Birkhoff polytope graph, which translate to tight bounds on the alphabet size needed for maximally recoverable codes in grid topologies.

We start by describing the coding theory question that motivated the current work.

1.1. Maximally recoverable codes.

Maximally recoverable codes, first introduced by Gopalan et al. [6], are codes designed for distributed storage which combine quick recovery from single node failure and optimal recovery from catastrophic failure. More precisely, they are systematic linear codes which combine two types of redundancy symbols: local parity symbols, which allow for fast recovery from single symbol erasure, and global parity symbols, which allow for recovery from the maximal information theoretic number of erasures. This was further studied in [1, 7, 9, 11, 12].

[*] Received by the editors August 8, 2018; accepted for publication April 23, 2019; published electronically August 6, 2019. A preliminary version of this work appeared in the 58th Annual Symposium on Foundations of Computer Science (FOCS 2017).
https://doi.org/10.1137/18M1205856
Funding: The first author was supported by NSF CCF award 1553288 (CAREER). The first and second authors were supported by a Sloan research fellowship. The second and third authors were supported by NSF CCF award 1614023.

[†] Department of Computer Science, University of California, San Diego, La Jolla, CA 92093-0404 (dakane@ucsd.edu, slovett@ucsd.edu, skaringu@ucsd.edu).

The present paper is motivated by a recent work of Gopalan et al. [5], which studied the effect of the topology of the network on the code design. Concretely, they studied grid-like topologies. In the simplest setting, a codeword is viewed as an $n \times n$ array, with entries in a finite field $\mathbb{F}_{2^d}$, where there is a single parity constraint for each row and each column, and an additional global parity constraint. More generally, a $T_{n\times m}(a,b,h)$ maximally recoverable code has codewords viewed as an $n \times m$ matrix over $\mathbb{F}_2^d$, with a parity constraints per row, b parity constraints per column, and h additional global parity constraints. An important problem in this context is: How small can we choose the alphabet size 2^d and still achieve information theoretical optimal resiliency against erasures?

Gopalan et al. [5] gave a combinatorial characterization for this problem, in the simplest setting of $m = n$ and $a = b = h = 1$. Their characterization is in terms of labeling the edges of the complete bipartite graph $K_{n,n}$ by elements of $\mathbb{F}_2^d$, which satisfies the property that in every simple cycle, the sum is nonzero.

Let $[n] = \{1, \ldots, n\}$. Let $\gamma : [n] \times [n] \to \mathbb{F}_2^d$ be a labeling of the edges of the complete bipartite graph $K_{n,n}$ by bit vectors of length d.

DEFINITION 1.1. *A labeling $\gamma : [n] \times [n] \to \mathbb{F}_2^d$ is* simple cycle free *if for any simple cycle C in $K_{n,n}$ it holds that*

$$\sum_{e \in C} \gamma(e) \neq 0.$$

Gopalan et al. [5] showed that the question on the minimal alphabet size needed for maximally recoverable codes reduces to a question of how small can we take $d = d(n)$ so that a simple cycle free labeling exists. Concretely:

- The alphabet size needed for $T_{n\times n}(1,1,1)$ codes is $2^{d(n)}$.
- The alphabet size needed for $T_{n\times m}(a,b,h)$ codes is at least $2^{\min(d(n-a+1),d(m-b+1))/h}$.

Before the current work, there were large gaps between upper and lower bounds on $d(n)$. For upper bounds, as the number of simple cycles in $K_{n,n}$ is $n^{O(n)}$, a random construction with $d = O(n \log n)$ succeeds with high probability. There are also simple explicit constructions matching the same bounds; see, e.g., [6]. In terms of lower bounds, it is simple to see that $d \geq \log n$ is necessary. The main technical lemma of Gopalan et al. [5] in this context is that, in fact, $d \geq \Omega(\log^2 n)$ is necessary. This implies a superpolynomial lower bound on the alphabet size 2^d in terms of n, which is one of their main results.

We improve on both upper and lower bounds and show that d is linear in n. We note that our construction improves upon the random construction, which for us was somewhat surprising. For convenience we describe it when n is a power of two, but note that it holds for any n with minimal modifications.

THEOREM 1.2 (explicit construction). *Let n be a power of two. There exists $\gamma : [n] \times [n] \to \mathbb{F}_2^d$ for $d = 3n$ which is simple cycle free.*

Our main technical result is a nearly matching lower bound.

THEOREM 1.3 (lower bound). *Let $\gamma : [n] \times [n] \to \mathbb{F}_2^d$ be simple cycle free. Then $d \geq n/2 - 2$.*

1.2. Labeling by general Abelian groups. The definition of simple cycles free labeling can be extended to labeling by general Abelian groups, not just $\mathbb{F}_2^d$. Let H be an Abelian group, and let $\gamma : [n] \times [n] \to H$. We say that γ is simple cycle free if for any simple cycle $C = (e_1, e_2, \dots, e_k)$,

$$\sum_{i \in [k]} (-1)^i \gamma(e_i) \neq 0.$$

We note that the analysis of Gopalan et al. [5] can be extended to nonbinary alphabets $\mathbb{F}_p$, in which case their combinatorial characterization extends to the one above with $H = \mathbb{F}_p$.

Theorem 1.4. *Let H be an Abelian group. Let $\gamma : [n] \times [n] \to H$ be simple cycle free. Then $|H| \geq 2^{n/2-2}$.*

As a side remark, we note that the study of graphs with nonzero circulations was instrumental in the recent construction of a deterministic quasi-polynomial algorithm for perfect matching in NC [4]. However, beyond some superficial similarities, the setup seems inherently different than ours. For starters, they study general bipartite graphs, while we study the complete graphs. Moreover, they need to handle certain families of cycles that are not not necessarily simple, while in this work we focus on simple cycles.

The proofs of Theorems 1.3 and 1.4 rely on the study of a certain Cayley graph of the permutation group, which encodes the property of simple cycle free labeling. Surprisingly, the corresponding graph is the Birkhoff polytope graph.

1.3. The Birkhoff polytope graph. Let S_n denote the symmetric group of permutations on $[n]$. A permutation $\tau \in S_n$ is said to be a *cycle* if, with the exception of its fixed points, it contains a single nontrivial cycle (in particular, the identity is not a cycle). We denote by $\mathcal{C}_n \subset S_n$ the set of cycles. The Cayley graph $\mathcal{B}_n = \mathrm{Cay}(S_n, \mathcal{C}_n)$ is a graph with vertex set S_n and edge set $\{(\pi, \tau\pi) : \pi \in S_n, \tau \in \mathcal{C}_n\}$. Note that this graph is undirected, because if $\tau \in \mathcal{C}_n$, then also $\tau^{-1} \in \mathcal{C}_n$.

The graph $\mathcal{B}_n$ turns out to be widely studied. It is also the graph of the Birkhoff polytope, which is the convex hull of all $n \times n$ permutation matrices; see, for example, [3] for a proof. Our analysis does not use this connection; we use the description of the graph as a Cayley graph.

The following claim shows that Theorem 1.4 reduces to bounding the chromatic number of the Birkhoff polytope graph.

Claim 1.5. *Let H be an Abelian group. Assume that $\gamma : [n] \times [n] \to H$ is simple cycle free. Then the chromatic number of $\mathcal{B}_n$ is at most $|H|$. In particular, $\mathcal{B}_n$ contains an independent set of size $\geq n!/|H|$.*

Proof. Define for $h \in H$ the set of vertices

$$A_h = \left\{ \pi \in S_n : \sum_{i=1}^{n} \gamma(i, \pi(i)) = h \right\}.$$

We will show that each A_h is an independent set in $\mathcal{B}_n$. In particular, choosing $A = A_h$ of maximal size shows that $\mathcal{B}_n$ has an independent set of size $|A| \geq n!/|H|$.

Assume that $A = A_h$ is not an independent set. Then there are two permutations $\pi, \pi' \in A$ such that $\tau = \pi(\pi')^{-1} \in \mathcal{C}_n$. Let $M_\pi = \{(i, \pi(i)) : i \in [n]\}$ denote the matching in $K_{n,n}$ associated with π, and define $M_{\pi'}$ analogously. Let $C = M_\pi \oplus M_{\pi'}$ denote their symmetric difference. The fact that $\tau \in \mathcal{C}_n$ has exactly one cycle is equivalent to C being a simple cycle. Let $C = (e_1, e_2, \dots, e_k)$; then

$$\sum_{i \in [k]} (-1)^i \gamma(e_i) = \sum_{e \in M_\pi} \gamma(e) - \sum_{e \in M_{\pi'}} \gamma(e) = h - h = 0.$$

This violates the assumption that γ is simple cycle free. $\square$

The construction of a simple cycle free labeling in Theorem 1.2, combined with Claim 1.5, implies that the Birkhoff polytope graph has a relatively small chromatic number.

Corollary 1.6. *Let n be a power of two. Then the chromatic number of $\mathcal{B}_n$ is at most 9^n.*

We can improve the bound on the chromatic number from 9^n to 4^n via a construction which is not based on a simple cycle free labeling.

Theorem 1.7. *Let n be a power of two. Then the chromatic number of $\mathcal{B}_n$ is at most 4^n.*

The best previous bounds we are aware of are by Onn [8], who proved that $\mathcal{B}_n$ contains an independent set of size $\geq n^{\Omega(\sqrt{n})}$.

Our main technical result is an upper bound on the largest size of an independent set in the Birkhoff polytope graph.

Theorem 1.8. *The largest independent set in $\mathcal{B}_n$ has size $\leq n!/2^{(n-4)/2}$. In particular, the chromatic number of $\mathcal{B}_n$ is at least $2^{(n-4)/2}$.*

As an aside, we note that general bounds on the independence number of graphs, such as the Hoffman bound, give much weaker bounds. A standard application of the Hoffman bound gives a much weaker bound for the independence number of $\mathcal{B}_n$ of $O(n!)$, and if we restrict all permutations to have the same sign, the bound improves to $O((n-1)!)$. The reason is that the Hoffman bound (at least in its simplest form) directly relates to the minimal eigenvalues of the graph. However, in our case the eigenvalues are controlled by the irreducible representations of S_n, and the extreme eigenvalues are given by low dimensional representations. This prohibits obtaining strong bounds on the independence number directly.

In order to overcome this barrier, our analysis circumvents the effect of the low dimensional representations by appealing to a structure versus randomness dichotomy specialized for our setting. It allows us to either reduce the dimension of the ambient group, or restrict our attention to pseudo-random assumptions about the actions of the low dimensional representations.

Organization. We prove Theorem 1.2 in section 2 and Theorem 1.8 in section 3. Theorem 1.7 is proved in section 4.

2. A construction of a simple cycle free labeling.

We prove Theorem 1.2 in this section. We first introduce some notation. For $x \in [n]$ denote by $e_x^n \in \mathbb{F}_2^n$ the unit vector with 1 in coordinate x and 0 in all other coordinates. We let $0^n \in \mathbb{F}_2^n$ denote the all zero vector.

Let n be a power of two. We define recursively a labeling $\gamma_n : [n] \times [n] \to \mathbb{F}_2^{3n}$. For $n = 2$ set, for example,

$$\gamma_2(0,0) = e_1^6, \quad \gamma_2(0,1) = e_2^6, \quad \gamma_2(1,0) = e_3^6, \quad \gamma_2(1,1) = e_4^6.$$

Assume $n > 2$. Let $x' = x \mod (n/2)$ and $y' = y \mod (n/2)$, where $x', y' \in [n/2]$. Define $\gamma_n(x,y) \in \mathbb{F}_2^{3n}$ recursively as follows:

(i) The first n bits of $\gamma_n(x,y)$ are e_x^n if $y \leq n/2$, and otherwise they are 0^n.

(ii) The next $n/2$ bits of $\gamma_n(x,y)$ are $e_{y'}^{n/2}$ if $x \leq n/2$, and otherwise they are $0^{n/2}$.

(iii) The last $3n/2$ bits of $\gamma_n(x,y)$ are defined recursively to be $\gamma_{n/2}(x',y')$.

We claim that γ_n is indeed simple cycle free. For $n = 2$ it is simple to verify this directly, so assume $n > 2$.

Let C be a simple cycle in $K_{n,n}$, and assume towards a contradiction that $\sum_{e \in C} \gamma_n(e) = 0$. Assume C has $2k$ nodes for some $2 \leq k \leq n$, and let these be $C = (x_1, y_1, x_2, y_2, \ldots, x_k, y_k, x_1)$. We denote $X = \{x_1, \ldots, x_k\}$ and $Y = \{y_1, \ldots, y_k\}$. Define, furthermore, $L = \{1, \ldots, n/2\}$ and $U = \{n/2 + 1, \ldots, n\}$.

Claim 2.1. *Either $Y \subset L$ or $Y \subset U$.*

Proof. Assume that both $Y \cap L$ and $Y \cap U$ are nonempty. Then there must exist $i \in [k]$ with $y_i \in L$ and $y_{i+1} \in U$, where if $i = k$, then we take the subscript modulo k. Recall that x_{i+1} is the neighbor of y_i, y_{i+1} in C. Its contribution to the first n bits of the sum is $e^n_{x_{i+1}}$, since $y_i \leq n/2$ and $y_{i+1} > n/2$. Note that no other edge in C has a nonzero value in coordinate x_{i+1}. Thus the x_{i+1} coordinate in the sum over C is 1, which contradicts the assumption that the sum over C is zero. □

Thus we can assume from now on that either $Y \subset L$ or $Y \subset U$.

Claim 2.2. *Either $X \subset L$ or $X \subset U$.*

Proof. Assume that $Y \subset L$, and the case of $Y \subset U$ is identical. Assume that both $X \cap L$ and $X \cap U$ are nonempty. Then there must exist $i \in [k]$ with $x_i \in L$ and $x_{i+1} \in U$. Recall that y_i is the neighbor of x_i, x_{i+1} in C. Its contribution to the second batch (of $n/2$ bits) of the sum is $e^{n/2}_{y'_i}$, since $x_i \leq n/2$ and $x_{i+1} > n/2$. Note that no other edge in C has a nonzero value in coordinate $n + y'_i$, whereas here we need the assumption that $Y \subset L$ or $Y \subset U$. Thus the $n + y'_i$ coordinate in the sum over C is 1, which contradicts the assumption that the sum over C is zero. □

Thus we have that $X \subset U$ or $X \subset L$, and similarly $Y \subset U$ or $Y \subset L$. Thus, C is a simple cycle in $K_{n/2,n/2}$ embedded in $K_{n,n}$ in one of four disjoint ways: $L \times L$, $L \times U$, $U \times L$, or $U \times U$. Observe that in each of these copies, the last $3n/2$ coordinates of the sum are precisely $\gamma_{n/2}$, so by induction C cannot have zero sum.

3. The independence number of the Birkhoff polytope graph.

We prove Theorem 1.8 in this section. Let A be an independent set in $\mathcal{B}_n$. We prove an upper bound on the size of A. Concretely, we will show that $|A| \leq \frac{a}{c^n} n!$ for some absolute constants $a, c > 1$. As we will see at the end, the choice of $a = 4, c = \sqrt{2}$ works.

The proof relies on representation theory, in particular representation theory of the symmetric group. We refer readers to the excellent book of Sagan [10], which provides a thorough introduction to the topic. We will try to adhere to the notation in that book whenever possible.

Overall strategy. Our basic plan will be to break our analysis into two cases based on whether or not the action of A on m-tuples is nearly uniform for all m. This will be in analogy with standard structure versus randomness arguments. If the action on m-tuples is highly nonuniform, this will allow us to take advantage of this nonuniformity to reduce to a lower dimensional case. On the other hand, if A acts nearly uniformly on m-tuples, this suggests that it behaves somewhat randomly. This intuition can be cashed out usefully by considering the Fourier-analytic considerations of this condition, which will allow us to prove that some pair of elements of A differ by a simple cycle using Fourier analysis on S_n.

Nonuniform action on tuples. Let $[n]_m = \{(i_1, \ldots, i_m) : i_1, \ldots, i_m \in [n] \text{ distinct}\}$ denote the family of ordered m-tuples of distinct elements of $[n]$. Its size is $(n)_m = n(n-1) \cdots (n-m+1)$. A permutation $\pi \in S_n$ acts on $[n]_m$ by sending $I = (i_1, \ldots, i_m)$ to $\pi(I) = (\pi(i_1), \ldots, \pi(i_m))$. Below, when we write $\Pr_{\pi \in A}[\cdot]$ we always mean the probability of an event under a uniform choice of $\pi \in A$.

Notice that if $\Pr_{\pi \in A}[\pi(I) = J] \geq c^m/(n)_m$ for some pair $I, J \in [n]_m$, this will allow us to reduce to a lower dimensional version of the problem. In particular, if we let $A' = \{\pi \in A : \pi(I) = J\}$, we note that $|A| \leq |A'|(n)_m/c^m$. On the other hand, after multiplying on the left and right by appropriate permutations (an operation which doesn't impact our final problem), we can assume that $I = J = \{n-m+1, \ldots, n\}$. Then, if A were an independent set for $\mathcal{B}_n$, A' would correspond to an independent set for $\mathrm{Cay}(S_{n-m}, \mathcal{C}_{n-m})$. Then, if we could prove the bound that $|A'| \leq \frac{a}{c^{n-m}}(n-m)!$, we could inductively prove that $|A| \leq \frac{a}{c^n} n!$.

Uniform action on tuples. When the action of A on m-tuples is near uniform for all m, we will attempt to show that two elements of A differ by a simple cycle using techniques from the Fourier analysis of S_n. In fact, we will show the stronger statement that some pair of elements of A differ by a single cycle of length n.

Some slight complications arise here when parity of the permutations is considered. In particular, all n-cycles have the same parity. This is actually a problem for n even, as all such cycles will be odd, and thus our statement will fail if A consists only of permutations with the same parity. Thus, we will have to consider our statement only in the case of n odd. Even in this case, though, parity will still be relevant. In particular, note that the difference between two permutations in A can be a cycle of length n only if the initial permutations had the same parity. Thus, we lose very little by restricting our attention to only elements of A with the more common parity. Thus we lose a factor of 2 in the size of A, but will make our analysis somewhat easier. We are now prepared to state our main technical proposition.

PROPOSITION 3.1. *Let n be an odd integer, and let $c > 1$ be a sufficiently small constant. Let $A \subset S_n$ be a set of permutations satisfying the following:*

(i) *All elements of A are of the same sign.*

(ii) *For any even $m < n$ and any $I, J \in [n]_m$, $\Pr_{\pi \in A}[\pi(I) = J] < \frac{c^m}{(n)_m}$.*

Then there exist two elements of A that differ by a cycle of length n. In particular, we can take $c = \sqrt{2}$.

Remark. In the second condition above, we consider only even m. This is because if this condition fails, we are going to use our other analysis to recursively consider permutations of $[n-m]$, and would like $n-m$ to also be odd.

We prove Proposition 3.1 and then show that it implies Theorem 1.8.

Proof. First, note that by replacing all $\pi \in A$ by $\pi\sigma$ for some odd permutation σ if necessary, it suffices to assume that all $\pi \in A$ are even. We will assume this henceforth.

Rephrasing the problem using class functions. Let $\mathcal{C}'_n$ denote the set of n-cycles in S_n. Define two class functions $\varphi, \psi \in \mathbb{R}[S_n]$ as

$$\varphi = \frac{1}{|S_n||A|^2} \sum_{\sigma \in S_n, \pi, \pi' \in A} \sigma\pi(\pi')^{-1}\sigma^{-1}, \qquad \psi = \frac{1}{|\mathcal{C}'_n|} \sum_{\tau \in \mathcal{C}'_n} \tau.$$

It is easy to see that our conclusion is equivalent to showing that $\langle \varphi, \psi \rangle > 0$.

Let $\lambda \vdash n$ denote a partition of n, namely $\lambda = (\lambda_1, \ldots, \lambda_k)$, where $\lambda_1 \geq \cdots \geq \lambda_k \geq 1$ and $\sum \lambda_i = n$. The irreducible representations of S_n are the Specht modules, which are indexed by partitions $\{S^\lambda : \lambda \vdash n\}$. Let $\chi^\lambda : S_n \to \mathbb{R}$ denote their corresponding characters. Their action extends linearly to $\mathbb{R}[S_n]$. Namely, if $\zeta \in \mathbb{R}[S_n]$ is given by $\zeta = \sum_{\pi \in S_n} \zeta_\pi \pi \in \mathbb{R}[S_n]$ where $\zeta_\pi \in \mathbb{R}$, then $\chi^\lambda(\zeta) = \sum_{\pi \in S_n} \zeta_\pi \chi^\lambda(\pi)$.

As $\varphi, \psi \in \mathbb{R}[S_n]$ are class functions, their inner product equals

$$\langle \varphi, \psi \rangle = \sum_{\lambda \vdash n} \chi^\lambda(\varphi) \chi^\lambda(\psi). \tag{1}$$

Let $(n) \in \mathcal{C}'_n$ be a fixed cycle of length n. As all elements in ψ are conjugate to (n), we have $\chi^\lambda(\psi) = \chi^\lambda((n))$ and we can simplify (1) to

$$\langle \varphi, \psi \rangle = \sum_{\lambda \vdash n} \chi^\lambda(\varphi) \chi^\lambda((n)). \tag{2}$$

Thus, we are led to explore the action of the irreducible characters on the full cycle (n).

Characters' action on the full cycle. The Murnaghan–Nakayama rule is a combinatorial method to compute the value of a character χ^λ on a conjugacy class, which in our case is (n). In this special case it is very simple. It equals zero unless λ is a hook, e.g., its corresponding tableaux has only one row and one column, and otherwise its either -1 or 1. Concretely, let $h_m = (n-m, 1, 1, \ldots, 1)$ for $0 \le m \le n-1$ denote the partition corresponding to a hook. Then

$$\chi^\lambda((n)) = \begin{cases} (-1)^m & \text{if } \lambda = h_m, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Thus we can simplify (2) to

$$\langle \varphi, \psi \rangle = \sum_{m=0}^{n-1} (-1)^m \chi^{h_m}(\varphi). \tag{4}$$

Bounding the characters on $\boldsymbol{\varphi}$. The character h_0 corresponds to the trivial representation, and by our definition of φ it equals $\chi^{h_0}(\varphi) = 1$. Observe that we can simplify $\chi^\lambda(\varphi)$ as

$$\chi^\lambda(\varphi) = \frac{1}{|A|^2 |S_n|} \sum_{\pi, \pi' \in A, \sigma \in S_n} \chi^\lambda(\sigma \pi (\pi')^{-1} \sigma^{-1}) = \frac{1}{|A|^2} \sum_{\pi, \pi' \in A} \chi^\lambda(\pi (\pi')^{-1}). \tag{5}$$

First, we argue that the evaluation of characters on φ is always nonnegative.

Claim 3.2. *$\chi^\lambda(\varphi) \ge 0$ for all $\lambda \vdash n$.*

Proof. Let $\zeta \in \mathbb{R}[S_n]$ be given by $\zeta = \frac{1}{|A|} \sum_{\pi \in A} \pi$. Then

$$\chi^\lambda(\varphi) = \frac{1}{|A|^2} \sum_{\pi, \pi' \in A} \operatorname{Tr}\left(S^\lambda(\pi) S^\lambda((\pi')^{-1})\right) = \operatorname{Tr}\left(S^\lambda(\zeta) S^\lambda(\zeta)^T\right) = \|S^\lambda(\zeta)\|_F^2,$$

where for a matrix M its Frobenius norm is given by $\|M\|_F^2 = \sum |M_{i,j}|^2$. In particular it is always nonnegative. ∎

The following lemma bounds $\chi^{h_m}(\varphi)$. Observe that, in particular, for $c = 1$ it gives $\chi^{h_m}(\varphi) = 0$. However, we would use it to obtain effective bounds when $c > 1$.

Lemma 3.3. *Let $m \in \{1, \ldots, n-1\}$. For any even $k \in \{m, \ldots, n\}$ it holds that $\chi^{h_m}(\varphi) \le \frac{c^k - 1}{\binom{k}{m}}$.*

Proof. Let M^μ denote the (not irreducible) Young module associated with a partition $\mu \vdash n$. In the case of $\mu = h_k$ it corresponds to the action of S_n on $[n]_k$, that is, for any $\pi \in S_n$ we have that $M^{h_k}(\pi)$ is a matrix whose rows and columns are indexed by $I, J \in [n]_k$, respectively, where $M^{h_k}(\pi)_{I,J} = 1_{\pi(I)=J}$. Observe that $M^{h_k}(\pi^{-1}) = \left(M^{h_k}(\pi)\right)^T$. We extend this action to $\mathbb{R}[S_n]$ linearly.

Recall that $\zeta = \frac{1}{|A|}\sum_{\pi\in A}\pi \in \mathbb{R}[S_n]$. By assumption (ii) in Proposition 3.1 we have

$$\left(M^{h_k}(\zeta)\right)_{I,J} = \Pr_{\pi\in A}[\pi(I) = J] \le \frac{c^k}{(n)_k}.$$

Thus, we can bound the Frobenius norm of $M^{h_k}(\zeta)$ by

$$\|M^{h_k}(\zeta)\|_F^2 = \sum_{I,J} |\left(M^{h_k}(\zeta)\right)_{I,J}|^2 \le \left(\frac{c^k}{(n)_k}\right)\sum_{I,J} |\left(M^{h_k}(\zeta)\right)_{I,J}| = c^k.$$

This is useful as

$$\mathrm{Tr}(M^{h_k}(\varphi)) = \mathrm{Tr}\left(M^{h_k}(\zeta)\left(M^{h_k}(\zeta)\right)^T\right) = \|M^{h_k}(\zeta)\|_F^2 \le c^k.$$

The Kostka numbers $K_{\lambda,\mu}$ denote the multiplicity of the Specht module S^λ in the Young module M^μ. We can thus decompose

$$\mathrm{Tr}(M^\mu(\varphi)) = \sum_\lambda K_{\lambda,\mu}\chi^\lambda(\varphi).$$

We saw that $\chi^\lambda(\varphi) \ge 0$ for all λ. By Young's rule, $K_{\lambda,\mu}$ equals the number of semistandard tableaux of shape λ and content μ. In particular, it is always a nonnegative integer. In the special case of $\lambda = h_m$ and $\mu = h_k$ for $k \ge m$, Young's rule is simple to compute and gives

$$K_{h_m,h_k} = \binom{k}{m}.$$

Recall that χ^{h_0} is the trivial representation, for which $K_{h_0,h_k} = 1$ and $\chi^{h_0}(\varphi) = 1$. Thus

$$1 + \binom{k}{m}\chi^{h_m}(\varphi) \le \sum_\lambda K_{\lambda,h_k}\chi^\lambda(\varphi) = \mathrm{Tr}(M^{h_k}(\varphi)) \le c^k.$$

This completes the proof. ∎

We next apply Lemma 3.3 to bound $\chi^{h_m}(\varphi)$ for all $1 \le m \le n-1$. If $m \le n/2$, then we can apply Lemma 3.3 for $k = 2m$ and obtain the bound

$$\chi^{h_m}(\varphi) \le \frac{c^{2m}-1}{\binom{2m}{m}}.$$

For $m > n/2$ we need the following claim relating χ^{h_m} to $\chi^{h_{n-1-m}}$.

Claim 3.4. *For any $1 \le m \le n-1$ it holds that $\chi^{h_m}(\varphi) = \chi^{h_{n-1-m}}(\varphi)$.*

Proof. For any partition λ, let λ' denote the transpose (also known as conjugate) partition. It satisfies $\chi^{\lambda'}(\pi) = \chi^\lambda(\pi)\mathrm{sign}(\pi)$ for all $\pi \in S_n$, where sign : $S_n \to \{-1, 1\}$ is the sign representation. As all elements in A are even permutations, it holds by the definition of φ that

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

In particular, if $\lambda = h_m$, then $\lambda' = h_{n-1-m}$. □

Next, we lower bound $\langle \varphi, \psi' \rangle$ as follows. The dominant terms are $\chi^{h_0}(\varphi) = \chi^{h_{n-1}}(\varphi) = 1$. For any $1 \le m \le (n-1)/2 - 1$, the corresponding term in (4) appears twice, once as $(-1)^m \chi^{h_m}(\varphi)$ and once as $(-1)^{n-1-m}\chi^{h_{n-1-m}}(\varphi) = (-1)^m \chi^{h_m}(\varphi)$. The term for $m = (n-1)/2$ appears once.

Furthermore, as $\chi^{h_m}(\varphi) \ge 0$ for all m by Claim 3.2, the only negative terms correspond to odd $1 \le m \le (n-1)/2$. Thus we can lower bound

$$\frac{1}{2}\langle \varphi, \psi' \rangle \ge 1 - \sum_{m \ge 1,\ m \text{ odd}} \frac{c^{2m} - 1}{\binom{2m}{m}}. \tag{6}$$

It is not hard to show that this is positive if $c > 1$ is small enough. If we take $c = \sqrt{2}$, the right-hand side of (6) is slightly negative for large enough m (the limit as $m \to \infty$ is $-0.02451\ldots$). However, when $n \ge 8$, the second term can be replaced by $\frac{c^8-1}{\binom{8}{3}}$ rather than $\frac{c^6-1}{\binom{6}{3}}$, making our lower bound on $\frac{1}{2}\langle \varphi, \psi' \rangle$ at least 0.057. This completes our proof. □

We are now prepared to prove Theorem 1.8.

Proof. We first prove that if n is odd and if all permutations in A have the same sign, then $|A| \le \frac{n!}{2^{(n-1)/2}}$.

We proceed by induction on n. First, we note that if $n = 1$, the bound follows trivially.

For odd $n > 1$, we note that unless there is some even $m < n$ and some $I, J \in [n]_m$ with $\Pr_{\pi \in A}[\pi(I) = J] \ge 2^{m/2}/(n)_m$, then our result follows immediately from Proposition 3.1. Otherwise, we may assume without loss of generality that $I = J = (n-m+1, \ldots, n)$. It then follows that letting $A' = \{\pi \in A : \pi(I) = J\}$, we can think of A' as a set of permutations on $[n-m]$. Also, note that A being an independent set for $\mathcal{B}_n$ implies that A' is an independent set for $\mathrm{Cay}(S_{n-m}, \mathcal{C}_{n-m})$. Therefore, by the inductive hypothesis,

$$|A| \le (n)_m 2^{-m/2}|A'| \le (n)_m 2^{-m/2}(n-m)!/2^{(n-m-1)/2} = n!/2^{(n-1)/2}.$$

We now need to reduce to the case of n odd and A consisting only of permutations of the same sign. First, restricting A to only permutations of the most common sign, we can assume that all permutations in A have the same sign, losing only a factor of 2 in $|A|$. Now, if n is odd, we are done. Otherwise, let j be the most likely value of $\pi(n)$ for π taken from A. We have that $\Pr_{\pi \in A}[\pi(n) = j] \ge 1/n$. Without loss of generality $j = n$, and we can let $A' = \{\pi \in A : \pi(n) = n\}$. Since A' is an independent set in $\mathrm{Cay}(S_{n-1}, \mathcal{C}_{n-1})$, and since $n-1$ is odd, we have

$$|A| \le n|A'| \le n(n-1)!/2^{(n-2)/2} = n!/2^{n/2-1}.$$ □

4. Construction of a larger independent set.

We prove Theorem 1.7 in this section. Assume that $n = 2^m$. We construct a coloring of S_n with at most 4^n colors such that each color class is an independent set in $\mathcal{B}_n$.

Let $T_{i,j} = \{2^{m-i}(j-1)+1, \ldots, 2^{m-i}j\}$ for $0 \le i \le m, 1 \le j \le 2^i$. Note that $\{T_{i,j} : j \in [2^i]\}$ is a partition of $[n]$ for every i, that $|T_{i,j}| = 2^{m-i}$, and that $T_{i,2j-1} \cup T_{i,2j}$ is a partition of $T_{i-1,j}$. For $1 \le i \le m$, let $M_i = \binom{2^{m-i+1}}{2^{m-i}}$. For any set

R of size $|R| = 2^{m-i+1}$, let $\operatorname{ind}_i(R, \cdot)$ be a bijection between subsets of R of size 2^{m-i} and $\mathbb{Z}_{M_i}$.

Fix $a_i \in \mathbb{Z}_{M_i}$ for $i = 1, \ldots, m$. Each tuple $(a_1, \ldots, a_m)$ will define a color class. Given such a tuple, we define a sequence of subsets of S_n as follows. Define $A_0 = S_n$ and

$$A_i = \left\{ \pi \in A_{i-1} : \sum_{j=1}^{2^{i-1}} \operatorname{ind}_i(\pi(T_{i-1,j}), \pi(T_{i,2j-1})) \equiv a_i \mod M_i \right\}.$$

Since each value mod M_i occurs equally often as an $\operatorname{ind}_i(\pi(T_{i-1,j}), \pi(T_{i,2j-1}))$ for each j, and since these values are independent of one another, $|A_i| = |A_{i-1}|/M_i$. Finally, set $A = A_m$. The following claim (applied for $i = m$) shows that A is an independent set in $\mathcal{B}_n$.

CLAIM 4.1. *Let $1 \le i \le m$. Let $\pi, \pi' \in A_i$ be such that $\tau = \pi(\pi')^{-1} \in \mathcal{C}_n$. Then there exists $j_i \in [2^i]$ such that:*

1. *$\tau(T_{i,j_i}) = T_{i,j_i}$.*
2. *$\tau(x) = x$ for all $x \in T_{i,j}, j \ne j_i$.*

Proof. We prove the claim by induction on i. The case of $i = 1$ follows from the definition of A_1. By assumption π, π' fix both $T_{1,1}$ and $T_{1,2}$. However, as $\tau = \pi(\pi')^{-1}$ is a cycle, it must be contained in either $T_{1,1}$ or $T_{1,2}$. This implies that $\tau(x) = x$ for all $x \in T_{1,1}$ or all $x \in T_{1,2}$.

Consider next the case of $i > 1$. By induction $\pi(T_{i-1,j}) = \pi'(T_{i-1,j})$ for all $j \in [2^{i-1}]$. Moreover, there exists $j' = j_{i-1}$ such that $\pi(x) = \pi'(x)$ for all $x \in T_{i-1,j}, j \ne j'$. This implies that $\pi(T_{i,j}) = \pi'(T_{i,j})$ for all $j \notin \{2j' - 1, 2j'\}$.

Next, the assumption that $\pi, \pi' \in A_i$ guarantees that

$$\sum_{j=1}^{2^{i-1}} \operatorname{ind}_i(\pi(T_{i-1,j}), \pi(T_{i,2j-1})) \equiv \sum_{j=1}^{2^{i-1}} \operatorname{ind}_i(\pi'(T_{i-1,j}), \pi'(T_{i,2j-1})) \equiv a_i, \quad \mod M_i.$$

For any $j \ne j'$ we know that $\pi(T_{i-1,j}) = \pi'(T_{i-1,j})$ and $\pi(T_{i,2j-1}) = \pi'(T_{i,2j-1})$, so $\operatorname{ind}_i(\pi(T_{i-1,j}), \pi(T_{i,2j-1})) = \operatorname{ind}_i(\pi'(T_{i-1,j}), \pi'(T_{i,2j-1}))$. Thus we obtain that also $\operatorname{ind}_i(\pi(T_{i-1,j'}), \pi(T_{i,2j'-1})) = \operatorname{ind}_i(\pi'(T_{i-1,j'}), \pi'(T_{i,2j'-1}))$. Moreover, as we also know that $\pi(T_{i-1,j'}) = \pi'(T_{i-1,j'})$ and that $\operatorname{ind}_i(\pi(T_{i-1,j'}), \cdot)$ is a bijection to $\mathbb{Z}_{M_i}$, it must be the case that $\pi(T_{i,2j'-1}) = \pi'(T_{i,2j'-1})$ and hence also $\pi(T_{i,2j'}) = \pi'(T_{i,2j'})$. Thus we conclude that $\pi(T_{i,j}) = \pi'(T_{i,j})$ for all $j \in [2^i]$.

To conclude, as $\tau = \pi(\pi')^{-1}$ is a cycle, it must be contained in either $T_{i,2j'-1}$ or $T_{i,2j'}$. Thus, τ must fix all points in $T_{i,2j'-1}$ or all points in $T_{i,2j'}$. We set $j_i \in \{2j' - 1, 2j'\}$ accordingly. □

Finally, we compute the size of A. As $|A_i| = |A_{i-1}|/M_i$ and $M_i = \binom{2^{m-i+1}}{2^{m-i}} \le 2^{2^{m-i+1}}$, we obtain that

$$|A| \ge \frac{n!}{\prod_{i=1}^m 2^{2^i}} \ge \frac{n!}{2^{2^{m+1}}} = \frac{n!}{4^n}.$$

Acknowledgments. We thank Ran Gelles and Sergey Yekhanin for useful discussions on the problem and for comments on a preliminary version of this paper. We thank Igor Pak for bringing to our attention that the Cayley graph which we study is the Birkhoff polytope graph.

REFERENCES

[1] S. Balaji and P. V. Kumar, *On partial maximally-recoverable and maximally-recoverable codes*, in Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), IEEE, 2015, pp. 1881–1885.

[2] A. Barvinok, *A Course in Convexity*, Grad. Stud. Math. 54, Amer. Math. Soc., Providence, RI, 2002.

[3] L. J. Billera and A. Sarangarajan, *The combinatorics of permutation polytopes*, in Formal Power Series and Algebraic Combinatorics, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 24, Amer. Math. Soc., Providence, RI, 1996, pp. 1–23.

[4] S. Fenner, R. Gurjar, and T. Thierauf, *Bipartite perfect matching is in quasi-nc*, in Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2016, pp. 754–763.

[5] P. Gopalan, G. Hu, S. Kopparty, S. Saraf, C. Wang, and S. Yekhanin, *Maximally recoverable codes for grid-like topologies*, in Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 2017, pp. 2092–2108.

[6] P. Gopalan, C. Huang, B. Jenkins, and S. Yekhanin, *Explicit maximally recoverable codes with locality*, IEEE Trans. Inform. Theory, 60 (2014), pp. 5245–5256.

[7] V. Lalitha and S. V. Lokam, *Weight enumerators and higher support weights of maximally recoverable codes*, in, Proceedings of the 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, 2015, pp. 835–842.

[8] S. Onn, *Geometry, complexity, and combinatorics of permutation polytopes*, J. Combin. Theory Ser. A, 64 (1993), pp. 31–49.

[9] J. S. Plank, M. Blaum, and J. L. Hafner, *Sd codes: erasure codes designed for how storage systems really fail*, in FAST, San Jose, CA, 2013, pp. 95–104.

[10] B. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*, Grad. Texts Math. 203, Springer-Verlag, New York, 2013.

[11] I. Tamo and A. Barg, *Bounds on locally recoverable codes with multiple recovering sets*, in Proceedings of the 2014 IEEE International Symposium on Information Theory (ISIT), IEEE, 2014, pp. 691–695.

[12] I. Tamo and A. Barg, *A family of optimal locally recoverable codes*, IEEE Trans. Inform. Theory, 60 (2014), pp. 4661–4676.