Representation-Theoretic Techniques for Independence Bounds of Cayley Graphs

Max Hopkins

Advisor: Madhu Sudan

A thesis submitted in partial fulfillment
of the requirements for the degree of
Bachelor of Arts in mathematics with honors

Harvard College, 2018

Abstract. The independence number of Cayley graphs has proved to be an important topic in discrete mathematics, information theory, and computer science [18, 2]. In early 2017, Kane, Lovett, and Rao provided a novel technique to compute such bounds via representation theory. While earlier works had used the representation theory of abelian groups to analyze independent sets [29, 3, 14], Kane et al. were the first to bring analysis of a non-abelian group, S_n, to the table, proving an independence bound for the Birkhoff graph. This work aims to provide background, intuition, and novel applications of Kane et al's new technique in order to show its efficacy and robustness across a variety of circumstances. In particular, we first show how the KLR technique builds off of the Hoffman bound which recovers special cases of closed-form bounds given by Delsarte's linear program on abelian Cayley graphs associated to $A_q(n, d)$. Second, we explain how the KLR technique employs structure vs randomness for non-abelian groups, and show that the technique generalizes to the hyperoctahedral group.

Contents

1. Introduction

The Maximum Independent Set problem (Max-Ind-Set), computing the maximal size of an independent set in a graph G, has been a long-standing and important problem in discrete mathematics and computer science.

Definition 1.1 (Independent Set)**.** *Given a graph $G(V, E)$, a subset $A \subseteq V$ is independent if $\forall v, v' \in A, (v, v') \notin E$. The size of the largest independent set in G is the graph's independence number $\alpha(G)$.*

Max-Ind-Set gave rise to a number of important topics in optimization, information theory, and complexity. In 1979, Lovasz showed that Max-Ind-Set may be estimated via a semi-definite program, providing the first example of an increasingly popular optimization scheme, and showed the importance of independence number for channel coding [27, 1]. In addition, Max-Ind-Set has a rich history in hardness, both as one of Karp's original NP-hard problems, and as one of the first problems in hardness of approximation [24, 19].

Cayley graphs, which provide a geometric interpretation of a group, have held a similar spot of importance.

Definition 1.2 (Cayley Graph)**.** *Given a group G and generator set $S \subseteq G$, the Cayley graph with base group G and generator set S, Cay(G,S), is the graph (V, E) where $V = G$, and $(g, g') \in E$ iff $g = sg'$ for some $s \in S$.*

Cayley graphs have far-reaching applications in algebraic graph and geometric group theory, such as in the study of Hamiltonicity, as well as in computer science in areas such as network architecture and graph theory [26, 11]. Cayley graphs are often used in the construction of expander graphs [2], and the spectra of any graph with a transitive automorphism group reduces to the spectra of a Cayley graph [6].

This paper focuses on the combination of these two concepts: the independence number of Cayley graphs. In the past, solving Max-Ind-Set even for specific Cayley graphs has proved important in discrete math, information theory, and computer science [2]. For instance, it is this problem that underlies the theory of error correcting codes. In 1950, Richard Hamming published a paper explaining how to build a code which allowed for the recovery of data despite a limited amount of corruption [18]. Hamming's "q-ary codes" are subsets of $\mathbb{Z}_q^n$ equipped with a metric called Hamming distance, such that no two elements in the subset are close to each other.

Definition 1.3 (Hamming Weight and Distance)**.** *The Hamming weight $|g|$ of an element $g \in \mathbb{Z}_q^n$ is the number of non-zero components of g. The Hamming distance $d(g, h)$ for $g, h \in \mathbb{Z}_q^n$ is the Hamming weight of their difference $|g - h|$.*

Formally, Hamming defines a q-ary code of distance d and size n (called an (n, d)-code) to be a subset of $\mathbb{Z}_q^n$ such that the minimum Hamming distance between any two elements (or codewords) is d. Intuitively, it is clear that such a code can correct up to $\lfloor \frac{d-1}{2} \rfloor$ errors, simply by finding the nearest codeword. Of course, a code with few elements is not so useful–we would like to encode as much data as possible! This raises the natural question of exactly how large a q-ary (n, d)-code can be, a number now referred to as $A_q(n, d)$ [31]. Hamming provided an initial upper bound to $A_q(n, d)$ via sphere-packing, and proved that in certain cases his constructed codes achieved the bound [18]. Computing improved bounds on $A_q(n, d)$ has remained one of the largest open problems in coding theory since the 1950's.

How does this relate to Cayley graphs or independent sets? Let $\mathbb{S}_q(n, d) = \{g \in \mathbb{Z}_q^n \mid |g| < d\}$. An (n, d)-code is exactly an independent set of $G = \text{Cay}(\mathbb{Z}_q^n, \mathbb{S}_q(n, d))$, and $A_q(n, d) = \alpha(G)$! In other words, this problem at the heart of coding theory is exactly the independence bound of an abelian Cayley graph.

This paper focuses on a novel technique introduced last year by Kane, Lovett, and Rao (henceforth referred to as *the KLR technique*) to compute an independence bound on Cayley graphs via representation theory. In particular, Kane et al. study a Cayley graph of S_n called the Birkhoff graph $B_n = Cay(S_n, C)$, where $C \subset S_n$ is the set of pure cycles, and prove the following upper bound:

Theorem 1.4 (Theorem 1.8 in [23]). *For any independent set $A \subseteq B_n$, $|A| \leq \frac{n!}{2^{\frac{n-4}{2}}} = \frac{|B_n|}{2^{\frac{n-4}{2}}}$*

The best previously known bound was $\alpha(B_n) \leq \frac{n!}{n^{log(n)}}$, super-polynomial in the denominator, but not exponential [16]. Further, KLR prove this bound is near-tight.

The use of representation theory itself in this area is hardly new. Earlier works, for instance, used the representation theory of $\mathbb{F}_2^n$ to improve bounds on $A_2(n,d)$ [29], or $\mathbb{F}_q^n$ to examine structural properties of independent sets [3, 14]. What makes the KLR technique special is its ability to deal with more complex, and in particular non-abelian groups. Our work provides intuition, explanation, and examples of the KLR technique, and in doing so recovers or improves previous independence bounds on Cayley graphs with varying base groups and generator sets. With this in mind, we argue that the KLR technique may in some sense be the "correct" way to view the independence number of Cayley graphs.

1.1. **Our Results.** We begin by examining the Hoffman bound, an independence bound for d-regular graphs, and its applications to $\mathrm{Cay}(\mathbb{Z}_q^n, \mathbb{S}_q(n,d))$. Further, we show that the KLR technique may be viewed as an extension of the Hoffman bound which allows for the exploitation of structure vs randomness via representation theory. Recall a graph is d-regular if every vertex has exactly d edges, and a graph's adjacency matrix is the indicator function for its edge set.

Theorem 1.5 (Hoffman). *Given a d-regular graph $G(V,E)$, $|V| = n$, whose adjacency matrix A has eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$, an independent set S must satisfy:*

$$|S| \leq n\frac{-\lambda_n}{d-\lambda_n}$$

The eigenvalues of $\mathrm{Cay}(\mathbb{Z}_q^n, \mathbb{S}_q(n,d))$ have a well known form as Krawtchouk polynomials:

Lemma 1.6 (Prop 25 in [22]). *Let $K_{n,k}^q(x)$ be the Krawtchouk polynomials, defined by:*

$$K_{n,k}^q(x) = \sum_{i=0}^{k}(-1)^i(q-1)^{k-i}\binom{x}{i}\binom{n-x}{k-i}$$

Given $g \in \mathbb{Z}_q^n$ the corresponding eigenvalue λ_g of eigenvector χ^g is the Krawtchouk polynomial:

$$\lambda_g = K_{n-1,d-1}^q(|g|-1) - 1$$

How good is the Hoffman bound in this case? Here we require a bit more background into the history of $A_q(n,d)$. Hamming's bound stood largely untouched until the 1970s, when Delsarte found a linear program which provided uniform improvement [12]. Delsarte's bound has stood relatively untouched up to today, with only slight improvements given by semi-definite over linear programs [28]. We look at the specific case of distance 3, $q \neq 2$, and show that Theorem 1.5 recovers Delsarte's bound for $n \equiv 1, 2 \pmod q$, and Hamming's bound for $n \equiv 3 \pmod q$.

Corollary 1.7. *For $n \equiv 1, 2 \pmod q$, Theorem 1.5 gives*

$$A_q(aq+r,3) = q^n\frac{(q-1)n - r(q-r)}{((q-1)n+r)((q-1)n+r-q)}$$

which is exactly the closed-form of Delsarte's Linear Program for non-binary codes proved in [31].

We do not prove an explicit relation between the Hoffman bound and Delsarte's Linear Program, but we briefly discuss the connection between the two methods. Unfortunately, the minimum eigenvalue grows too fast past $d = 3$ for the Hoffman bound to be tight. This is at least in part because the method is naive: it assumes the worst distribution of weights across characters in our computation. In particular, it assumes that the entirety of the weight outside of the trivial character occurs at the same position as the minimum of the Krawtchouk polynomials, which forces our bound to be large. In some cases, this is the true distribution of the weight–in fact it follows from Corollary 1.7 that this is the case for some perfect codes (codes which attain the Hamming bound). However, in many cases it may be possible to find a psuedorandomness condition on A which assures a better distribution of weight and improves our bound, as will be the method for non-abelian groups. We leave this as an area of further research.

Our second theorem is an analog to that of Kane et al. and proves that the KLR technique can be extended to further non-abelian groups. In particular, we examine a Cayley graph of the hyperoctahedral group $S(2,n) = \mathbb{Z}_2 \wr S_n$ (see Section 7 for an in-depth definition), a group with representations structurally similar to S_n.

Theorem 1.8. *Let C be the set of pure cycles in the hyperoctahedral group $S(2,n)$. For any independent set $A \subseteq Cay(S(2,n), C) = G$, $|A| \leq \frac{n!2^n}{2^{\frac{n-6}{2}}} = \frac{|G|}{2^{\frac{n-6}{2}}}$*

While this result is an immediate corollary of [23], recovering the analogous near-tight bound using the representation theory of a more complex group shows the robustness of the KLR technique.

2. Representation Theory

This section provides an overview of general concepts in representation theory essential to understanding the KLR technique and to continue work in the area. For further explanation of this material, we suggest referring to Artin's Algebra [5]. If the reader is familiar with representation theory, they may skip to Section 4.

2.1. Group Representations, G-Modules, and Characters.

Representation theory is the reduction of algebra to linear algebra by representing groups as linear transformations.

Definition 2.1 (Representation)**.** *An n-dimensional representation f of a group G is a homomorphism $f : G \to GL_n(\mathbb{C})$, mapping from the group itself to invertible $n \times n$ matrices.*

While this definition is intuitive, it requires a choice of basis–something we often wish to avoid. We may define a canonical version, a G-module, which skips this step.

Definition 2.2 (G-Module)**.** *An n-dimensional G-module is a vector space V of dimension n equipped with a homomorphism $\rho : G \to GL(V)$, mapping from the group itself to the general linear group of V. For elements $g \in G$, and $v \in V$, we often write $\rho(g)(v)$ simply as gv. Thus V is a G-module in the proper sense, with ρ specifying the group action.*

This latter definition may seem somewhat abstract, but such modules are often familiar objects. Consider, for instance, a group's action on itself–this is called the regular representation.

Example 2.3 (Regular Representation)**.** *Given a group G, the left regular representation is the G-module such that elements of G are basis vectors of V, and $\rho(g)(h) = gh$. Similarly, the right regular representation is given by $\rho(g)(h) = hg$.*

While it may be tempting to compare Fourier coefficients to one-dimensional representations or G-modules, the analogous concept is in fact the trace of a representation, known as a character of G. Character is well-defined for G-modules as trace is independent of basis. Characters also exhibit a property that flies under the radar in abelian groups: they are class functions.

Definition 2.4 (Class Functions)**.** *A function $\chi : G \to \mathbb{C}$ is a class function if $\chi(g) = \chi(g')$, for $g, g' \in G$ that share a conjugacy class.*

Example 2.5. *Given a group G with conjugacy classes $\lambda_1, \ldots, \lambda_n$, the set membership functions*

$$\chi_{\lambda_i}(g) = \begin{cases} 1 & g \in \lambda_i \\ 0 & g \notin \lambda_i \end{cases}$$

form a basis for class functions over G. In the future, we may write χ_g for an element $g \in G$; this refers to the set membership function for the conjugacy class of g.

This is the most intuitive, and certainly the simplest basis for class functions over a group G, but it is far from being the most important. Many important results in representation theory will follow from what is essentially Parseval's theorem, that the inner product of these class functions does not depend on choice of basis. In Fourier analysis, we use this theorem by then examining the Fourier basis. In representation theory, we will examine the analogous basis, the characters of irreducible representations of G.

2.2. **Irreducible Representations and the Maschke Decomposition.** For finite groups, every representation has a decomposition into so-called irreducible representations, called the Maschke Decomposition. Given a G-module V, consider a proper subspace W invariant under the action of G, that is $\forall w \in W,\ gw \in W$. This structure allows V to be expressed (as a G-module), as $W \oplus W^{\perp}$. Further, there exists a basis for which the matrix form of the representation is block diagonal, where the first block is given by W, and the second by $W^{\perp}$. We may apply this technique until each vector space in the decomposition has no proper G-invariant subspaces–these are the irreducible representations. To prove Maschke decomposition, we need a new tool: unitary representations.

Definition 2.6 (Unitary Representations)**.** *Recall that an operator T on a vector space V equipped with a Hermitian product is unitary if $\forall v, w \in V$*

$$\langle Tv, Tw\rangle = \langle v, w\rangle.$$

*Equivalently, if A is the matrix form of the operator T, A is unitary iff $A^{\dagger} = A^{-1}$. Likewise, a **representation** ρ is unitary if there exists a Hermitian product such that $\rho(g)$ is unitary for all $g \in G$.*

Lemma 2.7. *If ρ is a unitary representation of G on a Hermitian space V, and W is a G-invariant subspace, then $\rho = \rho_W \oplus \rho_{W^{\perp}}$*

Proof. If ρ is a unitary representation of G on a Hermitian space V, and W is a G-invariant subspace, then $W^{\perp}$ is G-invariant as well, and $\rho = \rho_W \oplus \rho_{W^{\perp}}$. Basic linear algebra gives $V = W \oplus W^{\perp}$. Further consider $u \in W^{\perp}$

$$\langle u, w\rangle = \langle gu, gw\rangle = \langle gu, w'\rangle = 0$$

which implies $gu \in W^{\perp}$. Then we have that every unitary representation on a Hermitian vector space is the direct sum of irreducible representations. □

Here we have assumed that our representation is unitary, but this is in fact true in general. Recall there always exists a positive definite Hermitian form on V $\{\cdot,\cdot\}$; thus for any ρ we can construct the form required by Definition 2.6 via averaging over G:

$$\langle v, w\rangle = \frac{1}{|G|}\sum_{g\in G}\{gv, gw\}$$

2.3. **The Basis of Irreducible Characters.** Formally, an irreducible character of a group G is the trace of a representation (ρ, V) such that V has no non-trivial G-invariant subspaces. Irreducible characters provide an orthonormal basis for class functions akin to the Fourier basis. We begin by showing orthonormality with respect to our standard Hermitian product.

Lemma 2.8 (Artin 10.4.6)**.** *Given two class functions, $\chi, \chi' : G \to \mathbb{C}$, let the Hermitian product $\langle \chi, \chi'\rangle$ be $\frac{1}{|G|}\sum_{g\in G}\overline{\chi(g)}\chi'(g)$. The irreducible characters of G are orthonormal with respect to $\langle\cdot,\cdot\rangle$.*

Proof. We will say a linear transformation $T : V' \to V$ is G-invariant if $T(gv') = gT(v')$. Note that the kernel and image of G-invariant maps are G-invariant subspaces of V' and V respectively. Thus for irreducible G-modules ρ,ρ' with corresponding vector spaces V, V', any G-invariant map $T : V' \to V$ is either 0 or an isomorphism (this is known as Schur's lemma). Given a general linear transformation on G-modules $T : V' \to V$, we would like to build a G-invariant one. We will do this by taking advantage of translational symmetry and average over G by conjugation:

$$\tilde{T}(v) = \frac{1}{|G|}\sum_{g\in G} g^{-1}Tgv$$

From this point, our analysis will focus around the operator $\Phi(M) = \tilde{M}$, which maps any linear transformation to a G-invariant one. We begin by proving $Tr(\Phi) = \langle \chi, \chi'\rangle$, where χ and χ' are the characters of G-modules V and V' respectively. We will need the following result: for $m \times m$

and $n \times n$ matrices A and B, and a function $F(M) = AMB$, $Tr(F) = Tr(A)Tr(B)$. Now let $F_g(M) = \rho(g)^{-1}M\rho'(g)$, then:

$$Tr(\Phi) = \frac{1}{|G|}\sum_g Tr(F_g) = \frac{1}{|G|}\sum_g Tr(\rho(g)^{-1})Tr(\rho'(g)) = \frac{1}{|G|}\sum_g \overline{\chi(g)}\chi'(g) = \langle \chi, \chi' \rangle$$

Next we show $Tr(\Phi) = Dim(Im(\Phi))$. Consider $M \in Ker(\Phi) \cap Im(\Phi)$. Since it is in the image it is G-invariant, so $\Phi(M) = M = 0$. Since the intersection of image and kernel is trivial, we have $L = Im(\Phi) \oplus Ker(\Phi)$. Using the image and kernel to build a basis for L, the matrix form of Φ is a block form matrix with upper left block given by the identity on $Im(\Phi)$, and the lower right block as identically 0. Thus the trace is the dimension of $Im(\Phi)$ as desired. Then for χ, χ', characters of irreducible G-modules, $\langle \chi, \chi' \rangle$ is the dimension of G-invariant transformations and Schur's lemma finishes the proof–the dimension of this space is zero if χ and χ' are irreducible and non-isomorphic, and one if they are isomorphic. □

It is left to show that the irreducible characters form a basis for class functions.

Theorem 2.9 (Artin 10.8.5)**.** *The irreducible characters of a group G form an orthonormal basis for class functions over G.*

Proof. Let H be the space of class functions and $C \subseteq H$ be the span of the irreducible characters of G. We wish to show $C = H$, which we will do by proving that $Dim(C^\perp) = 0$. Given a class function $\phi \in C^\perp$ and representation ρ, define a G-invariant map $T = \frac{1}{|G|}\sum_g \overline{\phi(g)}\rho(g)$. The trace of T is $\langle \phi, \chi \rangle = 0$, and by Schur's lemma must be scalar multiplication, which implies $T = 0$. Recall the left regular representation V_G, i.e. left group action with basis $e_g, g \in G$, $\rho_g^{reg}(e_h) = e_{gh}$. The basis vectors e_g are linearly independent elements of V_G, but then the $\rho^{reg}(g)$ are linearly independent as well. Putting everything together we have a linear system:

$$\frac{1}{|G|}\sum_g \overline{\phi(g)}\rho^{reg}(g) = 0$$

but then $\forall g, \phi(g) = 0$, and $dim(C^\perp) = 0$ as desired. Thus the irreducible characters span the entire space of class functions, providing an orthonormal basis. □

2.4. **Induced Representations.** On a different note, in this work we will often desire to induce a representation from a subgroup onto the parent group.

Definition 2.10 (Induced Representations)**.** *For $H \leq G$, a left-transversal $t_1, ..., t_l$, and a representation of H, Y, the induced representation on G is $Y \uparrow_H^G (g) = Y(t_i^{-1}gt_j)$, such that $Y(g) = 0$ if $g \notin H$. The representation is independent of the chosen transversal.*

We leave out the proof that this is in fact a representation, but it is detailed in Sagan [32]. For us, the most important case will be when our subgroup representation Y is trivial.

Example 2.11 (Permutation Representation)**.** *Given a subgroup $H \leq G$, inducing from the trivial representation of H gives the Permutation Representation of H, $1 \uparrow_H^G$, a representation we will use heavily in our computations.*

3. The Hoffman Bound for Cayley Graphs

There are two conditions any new independence bound must meet to be of interest: first, it must be tight in certain cases, and second, it must outperform the Hoffman bound. In fact, the KLR technique may be seen as a sort of generalization of the Hoffman bound to take advantage of representation theory and pseudorandmoness. Thus as a baseline, we will first examine how the Hoffman bound performs on Cayley Graphs.

3.1. **The Hoffman Bound.** The Hoffman bound requires two additional graph theoretic definitions: regularity and the adjacency matrix.

Definition 3.1 (d-regular)**.** *An undirected graph $G(V, E)$ is d-regular if every vertex v has exactly d edges.*

Recalling our definition from the introduction, Cayley graphs then are $|S|$-regular. One often wishes to consider a graph in matrix form–indeed much of graph theory is done via linear algebra in this manner.

Definition 3.2. *Let G(V,E) be a graph. The adjacency matrix A of G is defined to be:*

$$A_{ij} = \begin{cases} 1, & (i,j) \in E \\ 0, & else \end{cases}$$

For an undirected graph, A is always symmetric, which makes spectral analysis a particularly useful tool. Indeed, in 1970, Hoffman proved an independence bound based on the spectrum of A:

Theorem 3.3 (Hoffman Bound)**.** *Let $G(V, E)$ be an undirected d-regular graph, and let A be the adjacency matrix of G. Let $|V| = n$, and $d = \lambda_1 \geq \ldots \geq \lambda_n$ be the eigenvalues of A corresponding to an orthonormal eigenbasis $\mathbf{1}/n = v_1, v_2, \ldots, v_n$ with respect to the inner product defined in Lemma 2.8. Any independent set $S \subseteq V$ must then sastisfy:*

$$|S| \leq n \frac{-\lambda_n}{d - \lambda_n}$$

Proof. Let S be an independent set of G, and let $f : V \to \mathbb{F}_2$ be the indicator function of S. Consider the inner product $\langle f, Af \rangle$. What does the application of the adjacency matrix A do to the indicator function f? One can think about it as spreading f out across its neighbors. In particular, $Af_i > 0$ exactly when vertex i shares an edge with an element in S. Since S is independent, $\langle f, Af \rangle = 0$. Let $\{a_i\}_{i=1}^n$ be the coefficients of f in the eigenbasis. We first note two important properties of the a_i: $a_1 = \langle f, v_1 \rangle = \frac{|S|}{n}$, and $\langle f, f \rangle = \sum_{i=1}^n a_i^2 = \frac{|S|}{n}$. Now we can prove the theorem:

$$\begin{aligned} \langle f, Af \rangle &= \sum_{i=1}^{n} \lambda_i a_i^2 \\ &= d a_1^2 + \sum_{i=2}^{n} \lambda_i a_i^2 \\ &\geq d \frac{|S|^2}{n^2} + \lambda_n \left(\frac{|S|}{n} - \frac{|S|^2}{n^2} \right) \end{aligned}$$

Substituting in 0 for $\langle f, Af \rangle$ then gives the desired result. □

3.2. **Cayley Graphs.** The question remains, how well does the Hoffman bound perform on Cayley graphs? To answer this, let's first recall the formal definition of a Cayley graph from the introduction:

Definition 3.4 (Cayley Graph)**.** *Given a group G and generator set $S \subseteq G$, Cay(G,S) is the graph (V, E) where $V = G$, and $(g, g') \in E$ iff $g = sg'$ for some $s \in S$.*

Example 3.5. *Given a finite group $|G| = n$, Cay(G,G) is the complete graph K_n.*

To apply the Hoffman bound, we must first analyze the eigenvalues of Cayley graphs. In general, the computation of these spectra is difficult, but we can gain some insight by formulating them in terms of the representation theory of the underlying group G [13].

Theorem 3.6 (Diaconis & Shahshahani)**.** *Given a group G, let S be a generating set containing no incomplete conjugacy classes. Further, let these classes be indexed by i, and χ^i be the ith irreducible representation of G. Then the eigenvalues of the adjacency matrix A are exactly:*

$$\lambda_i = \frac{1}{\chi^i(e)} \sum_{g \in S} \chi^i(g)$$

An easier treatment of this theorem can be found in [25].

3.3. **The Abelian Case: $A_q(n,d)$.** In general, it is difficult to compute these characters, but for a special set of abelian Cayley graphs, the eigenvalues have an easily computable form. Recall from the introduction that $\mathbb{S}_q(n,d)$ is the set of elements of $\mathbb{F}_q^n$ of hamming weight less than d. We will examine $G = \mathrm{Cay}(\mathbb{Z}_q^n, \mathbb{S}_q(n,d))$, where Max-Ind-Set corresponds to the famous problem of finding the largest q-ary (n,d)-code, whose size is denoted $A_q(n,d)$. Computing $A_q(n,d)$ has been studied in great depth since Hamming's seminal paper on error correcting codes, in which he proved the first major upper bound for q-ary codes commonly seen via sphere-packing [18]:

Theorem 3.7 (Hamming Bound)**.** *Let $t = \lfloor \frac{d-1}{2} \rfloor$, then:*

$$A_q(n,d) \leq \frac{q^n}{\sum_{k=1}^{t} \binom{n}{k}(q-1)^k}$$

In the late 1970s, Delsarte provided a linear program which gave a uniform improvement over the Hamming bound [12]. The bound stood unchallenged for decades, and it is only recently that semi-definite programming methods have provided slight improvements [28]. How does the Hoffman bound compare to Hamming and Delsarte? Let's look at some explicit computations. First, we will need to compute the eigenvalues of G. To do so, we will need to compute the irreducible characters of $\mathbb{F}_q^n$ (for a full treatment of the representation theory of $\mathbb{F}_q^n$, see [3]). Consider the set of one-dimensional representations indexed by $g \in \mathbb{F}_q^n$:

$$\rho_g(g') = e^{\frac{2\pi i \langle g, g' \rangle}{q}}$$

Here, $\langle g, g' \rangle = \left(\sum_{i=1}^{n} g_i * g_i' \right) \pmod q$. This is clearly a homomorphism, as

$$\rho_g(g_1 + g_2) = e^{\frac{2\pi i \langle g, g_1 + g_2 \rangle}{q}} = e^{\frac{2\pi i \langle g, g_1 \rangle}{q}} e^{\frac{2\pi i \langle g, g_2 \rangle}{q}} = \rho_g(g_1)\rho_g(g_2)$$

Further, these representations are distinct for each group element, and irreducible due to being degree one (they have no non-zero proper subspaces). Since there are $|G|$ of these representations, they form the irreducible basis for class functions $f : \mathbb{F}_q^n \to \mathbb{C}$.

Now we are ready to compute the eigenvalues of G.

Lemma 3.8 (Prop 25 in [22])**.** *Let $K_{n,k}^q(x)$ be the Krawtchouk polynomials, defined by:*

$$K_{n,k}^q(x) = \sum_{i=0}^{k} (-1)^i (q-1)^{k-i} \binom{x}{i} \binom{n-x}{k-i}$$

Given $g \in \mathbb{Z}_q^n$ the corresponding eigenvalue λ_g of eigenvector χ^g is the Krawtchouk polynomial:

$$\lambda_g = K_{n-1,d-1}^q(|g| - 1) - 1$$

Proof. To begin, applying Theorem 3.6 gives

$$\lambda_g = \sum_{a \in \mathbb{S}_q(n,d)} \chi^g(a).$$

Our computation will revolve around these irreducible characters. Without loss of generality, we may assume that g is non-zero in its first $|g| = i$ elements–permuting coordinates makes no computational difference. Let us begin by considering the case of the sum just over weight k. Select j non-zero values to be placed in the first i slots, and let their positions be denoted by $p(m)$. Recall that the irreducible character associated to an element $g \in G$ is given by the product of the characters on each component, and that χ^k for $k \in \mathbb{Z}_n$ is the homomorphism given by $\chi^k(1) = e^{2k\pi i/n}$. Letting ω be a primitive qth root of unity, our computation then becomes:

$$\sum_{|x|=k} \chi^g(x) = \sum_{|x|=k} \prod_{m=1}^{j} \chi^{g_{p(m)}}(x_{p(m)}) = \sum_{|x|=k} \prod_{m=1}^{j} \omega^{g_{p(m)} x_{p(m)}} = (q-1)^{k-j} \prod_{l=1}^{j} \sum_{y \in \mathbb{Z}_q \setminus 0} \omega^{x_{p(l)} y}$$

The roots of unity sum to 0, thus in dropping trivial root this last summation becomes -1. Further the $k-j$ non selected values have $q-1$ possible values. There are $\binom{i}{j}$ methods to select our j non-zero values, and $\binom{n-i}{k-j}$ ways to arrange the other values. Summing over all $j \leq k$ gives us exactly the Krawtchouk polynomial $K^q_{n,k}(i)$. At this point, we have shown:

$$\lambda_g = \sum_{k=1}^{d-1} K^q_k(|g|).$$

The Krawtchouk polynomials admit a number of interesting recurrence relations which in turn provide a formula for exactly this sum [10]. It is, in fact, the case that

$$\sum_{k=1}^{d-1} K^q_k(|g|) = K^q_{n-1,d-1}(|g|-1) - 1.$$

This concludes the proof. □

In fact, the reader familiar with coding theory, and in particular the history of $A_q(n,d)$, may not be surprised by the form of the above expression. The Krawtchouk polynomials were used by Delsarte to derive his Linear Programming bound [12], as well as in later analysis of closed-form solutions [31]. With this in mind, let us examine how Hoffman compares to Delsarte and Hamming for the case of single error correcting codes, i.e. distance 3.

Corollary 3.9 ($A_q(n,3)$)**.** *Let $R[x]$ round x to the nearest integer.*

$$A_q(n,3) \leq \frac{q^{n-1}\left(q\left(qR\left[\frac{n-2}{q}+\frac{1}{2}\right]+n(q-2)-q+4\right)R\left[\frac{2nq-2n-q+4}{2q}\right]-n(q-1)(n(q-1)-q+3)\right)}{\left(qR\left[\frac{n-2}{q}+\frac{1}{2}\right]+n(q-2)-q+4\right)R\left[\frac{2nq-2n-q+4}{2q}\right]}$$

Proof. We need only compute the smallest eigenvalue, which will occur at the closest integer to the minimum of the quadratic:

$$K^q_{n,k}(m-1) = \frac{1}{2}\left(-n(q-1)(2mq+q-3)+mq(mq+q-4)+n^2(q-1)^2\right)$$

Our minimum value occurs at $\frac{2nq-2n-q+4}{2q}$ and takes the value:

$$\frac{1}{2}\left(-q\mathrm{R}\left[\frac{2nq-2n-q+4}{2q}\right]\left(q\mathrm{R}\left[\frac{2nq-2n-q+4}{2q}\right]-2n(q-1)+q-4\right)-n(q-1)(n(q-1)-q+3)\right)$$

Plugging this value into Theorem 3.3 gives the desired bound. □

Corollary 3.10. *For $n \equiv 1,2 \pmod q$, Theorem 1.5 gives*

$$A_q(aq+r,3) = q^n \frac{(q-1)n-r(q-r)}{((q-1)n+r)((q-1)n+r-q)}$$

Which is exactly the closed-form of Delsarte's Linear Program for non-binary codes proved in [31].

We have seen that Hoffman (and thus the KLR-technique as we will show in Section 4) matches Delsarte's bound, but we would like to understand why this is the case only for $n \equiv 1,2 \pmod q$. In part, it has to do with the relative location of the true minimum of the Krawtchouk polynomial to integer values. We noted that the minimum occurs exactly at $n - \frac{1}{2} + \frac{2-n}{q}$. Thus, when $\frac{2-n}{q} = 0$, ($n \equiv 2 \pmod q$) the closest integer to our minimum is exactly $\frac{1}{2}$ away, the furthest it can possibly be. $\frac{2-n}{q} = 1/q$, ($n \equiv 1,3 \pmod q$) is the second furthest away it can be (for $n \equiv 3 \pmod q$, the bound matches the Hamming bound but not Delsarte's bound). For $n > 3 \pmod q$, Hamming's bound outperforms Hoffman's bound.

We would further like to understand any connection between Delsarte's and Hoffman's bounds. In brief, Delsarte's Linear Program relies on Hamming schemes, commutative association schemes whose eigenvalues correspond to the sums of characters of a given Hamming weight [12]. In other words, the connection between the two methods lies in the irreducible characters of the underlying group.

Unfortunately, for larger distance, the Hoffman bound fails to recover even the Hamming bound. It is possible one could use the KLR technique to get around this method. In Section 6, we will see how KLR extend the Hoffman bound by using representation theoretic structure and pseudorandomness to bound certain troublesome characters. Unfortunately, this actually appears to be more difficult in the abelian case due to the lack of induced representations with significant structure.

4. The KLR Technique

Recently, Kane, Lovett, and Rao (KLR) introduced a novel representation-theoretic technique that gave an exponentially improved upper bound on the independence number of the Birkhoff graph, a special Cayley graph of S_n. In this section, we will introduce the basic setup of the KLR technique for a general Cayley graph Cay(G,S), and show how it relates to the Hoffman bound.

4.1. **The KLR Technique.** Consider a Cayley graph Cay(G,S), and a subset $A \subseteq G$ which we would like to test for independence. Kane et al. begin by introducing two class functions $\phi, \psi : G \to \mathbb{C}$. In particular, ϕ is defined to be a combination of all pairs of elements in A, i.e. a convolution of the indicator set of A with itself, and ψ as some subset of generators with no incomplete conjugacy classes. When we take the inner product of ϕ and ψ, we are testing whether any two elements differ by a generator. Formally, let our generator subset be $C \subseteq S$. Kane et al. define $\phi, \psi \in \mathbb{C}[G]$ as follows:

$$\phi = \frac{1}{|G||A|^2} \sum_{\sigma \in G, a, a' \in A} \sigma a a'^{-1} \sigma^{-1}$$

$$\psi = \frac{1}{|C|} \sum_{\tau \in C} \tau$$

At first glance it may be unclear why ϕ is conjugated by all of G. The sole reason for this inclusion is that it forces ϕ to be a class function by spreading out the value of each pair evenly across its conjugacy class. Let's examine the inner product of ϕ and ψ:

$$\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\phi(g)} \psi(g) = \frac{1}{|C||G|} \sum_{g \in C} \phi(g).$$

Here, we have used the fact that ψ is 0 on $g \in G - C$. Note that in the above ϕ and ψ have taken on the form of functions, in particular, $\phi(g) = \chi_g(\phi)$, expanding the definition of χ_g linearly over the group ring. Then it is clear that the above equation is greater than 0 if and only if ϕ is non-zero on some $g \in C$. In particular, this occurs exactly when two elements in A differ by an element of C:

$$g = \sigma a a'^{-1} \sigma^{-1}$$

$$a = \sigma^{-1} g \sigma a'$$

$$a = g' a'$$

Here $g' \in C$ as C has no incomplete conjugacy classes. One might think you could be clever, and, by removing the conjugation by σ, allow C to have incomplete conjugacy classes. However, one way or another, ϕ and ψ must be class functions, as their inner product is only sensitive up to difference in conjugacy class.

On its own, this inner product is not particularly useful, as we have no information in the standard basis. However, given that ϕ is a convolution, it will become a product in the basis of irreducible characters. This will be more tractable. In this case:

$$\langle \phi, \chi^\lambda \rangle = \frac{1}{|G|} \sum_{g \in G} \phi(g) \chi^\lambda(g)$$

$$= \chi^\lambda(\phi)$$

where we have expanded the definition of χ^λ linearly–for an abelian group, this is simply the Fourier coefficient of ϕ evaluated at λ. Finally this implies that

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

4.2. **KLR vs. Hoffman.** The difference between the KLR technique and Hoffman's bound lies in the use of representation theory rather than spectral analysis. Hoffman uses the orthonormal eigenbasis for his inner product, whereas KLR instead use the irreducible characters, a basis for class functions. Let's begin by examining the naive case of the KLR technique, where we do not take advantage of extra representation-theoretic structure, and use a strategy similar to Hoffman.

Proposition 4.1 (KLR-Hoffman)**.** *Let G be a group of size n, and S a generating set with no incomplete conjugacy classes. Further, let the conjugacy classes of G be indexed by $[k]$ s.t. $|S| = \lambda_1 \geq \lambda_2 \ldots \geq \lambda_k$, where for simplicity we have let $\lambda_i = \chi^i(S)$. If $A \subseteq Cay(G, S)$ is independent, then:*

$$|A| \leq n\frac{-\lambda_k}{|S| - \lambda_k}$$

Proof. The proof follows much the same as Theorem 3.3. Define ϕ, ψ as in Section 4.1, with $C = S$. First, let's examine $\chi^g(\phi)$ (Claim 3.2 in [23]):

$$\begin{aligned}
\chi^g(\phi) &= \frac{1}{|A|^2}\sum_{a,a' \in A}\chi^g(aa'^{-1}) = \frac{1}{|A|^2}\sum_{a,a' \in A}\chi^g(a)\overline{\chi^g(a')} \\
&= \frac{1}{|A|^2}\left(\sum_{a\in A}\chi^g(a)\right)\left(\sum_{a\in A}\overline{\chi^g(a)}\right) = \frac{1}{|A|^2}\left(\sum_{a\in A}\chi^g(a)\right)\overline{\left(\sum_{a\in A}\chi^g(a)\right)} \\
&\geq 0
\end{aligned}$$

We see that we have decomposed $\chi^g(\phi)$ into the Fourier weights of the indicator function of $\mathbf{1}_A = \sum_{a\in A} a$. Because our basis is orthonormal, we may use the same two pieces of information regarding the character decomposition as Hoffman, namely that: $\langle \mathbf{1}_A, \chi^1\rangle = \frac{|A|}{n}$, and $\langle \mathbf{1}_A, \mathbf{1}_A\rangle = \frac{|A|}{n} = \sum_{i=1}^{k} \hat{\mathbf{1}_A}(i)^2$. This latter equality is known as Parseval's theorem. This allows us to use the same trick: pull out the trivial coefficient, and bound the remaining coefficients knowing the entire sum.

$$\begin{aligned}
\langle \phi, \psi\rangle &\propto \sum_{i=1}^{k} |\mathbf{1}_A(i)|^2\lambda_i \\
&= |S|\frac{|A|^2}{n^2} + \sum_{i=2}^{k} |\mathbf{1}_A(i)|^2\lambda_i \\
&\geq |S|\frac{|A|^2}{n^2} + \lambda_n(\frac{|A|}{n} - \frac{|A|^2}{n^2})
\end{aligned}$$

Substituting 0 for $\langle \phi, \psi\rangle$ concludes the proof. □

Now for abelian base groups, this is exactly the Hoffman bound. However, for non-abelian base groups, the eigenvalues differ from the character sums by the degree of the character–making the Hoffman bound potentially stronger. On the other hand, this is far outweighed by the ability to take further advantage of representation theory, as we will see in Section 6.

5. Representation Theory of the Symmetric Group

We have shown that even a naive application of the KLR technique is as good as the Hoffman bound for abelian groups, but for the method to truly shine we will need a group with significantly more structure. With such structure, however, comes the complication of group representations. This section aims to provide requisite background in the representation theory of the symmetric group. If the reader is already familiar with this topic, and in particular its combinatorial foundations (Young tableaux, tabloids, etc.), they should skip to Section 6. If, on the other hand, the reader would like more information on the representation theory of S_n, we refer them to Sagan [32].

5.1. **Young Tableaux and Tabloids.** We will express partitions of n as $\lambda \vdash n$, where $\lambda = \{\lambda_1, ..., \lambda_l\}$ is a partition with decreasing integer values that sum to n. The representation theory of the symmetric group shares a deep connection to certain combinatorial objects known as tableaux. We will use tableaux to construct irreducible representations, as well as for computations such as Young's rule which will be essential for the KLR technique.

Definition 5.1 (Ferrers Diagram)**.** *Given a partition $\lambda \vdash n = (\lambda_1, \ldots, \lambda_n)$, its Ferrers diagram is a diagram of dots with left justified rows such that the ith row has λ_i dots. We will use empty boxes instead of dots, as we intend to fill the boxes with numbers.*

Ferrers diagrams are the basis for Young diagrams, which replace the dots with integers.

Definition 5.2 (Young Tableaux)**.** *Given the Ferrers diagram F of a partition $\lambda \vdash n$, a Young tableaux is a bijective map $f : F \to [n]$, and is displayed by placing the number i in box $f^{-1}(i)$.*

Young tableaux are useful combinatorial objects in their own right, but we may wish to generalize them to allow for non bijective functions. These are called generalized Young tableaux.

Definition 5.3 (Generalized Young Tableaux)**.** *Given the Ferrers diagram F of a partition $\lambda \vdash n$, a generalized Young tableaux is a map $f : F \to \mathbb{Z}$. The cardinality of the pre-image of i is denoted as μ_i, and f is said to have content μ.*

Finally, there are many specific types of Tableaux which will be important for certain representations or computations. One particularly important such type are the semi-standard Tableaux.

Definition 5.4 (Semi-standard Tableaux)**.** *A Tableaux is semi-standard if its rows are weakly increasing, and its columns are strictly increasing*

Example 5.5. *Consider the hook partition $\lambda \vdash n = (3, 1, 1)$. Figure 1 shows, from left to right, the Ferrers diagram, a Young tableaux, a generalized Young tableaux of content $(2, 2, 0, 1)$, and a semi-standard tableaux. Hook partitions will become particularly important for character computations*

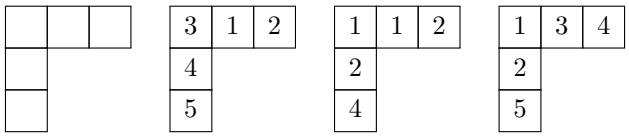


FIGURE 1. Young diagrams

later on.

Definition 5.6 (Young Tabloid)**.** *Given the Ferrers diagram F of a partition $\lambda \vdash n$, a Young tabloid is an equivalence class of Young tableaux with relation given by equality of the set of integers of each row. In other words, a Young tabloid is a Young tableaux with un-ordered rows.*

Example 5.7. *Figure 2 shows an equality and inequality of tabloids.*

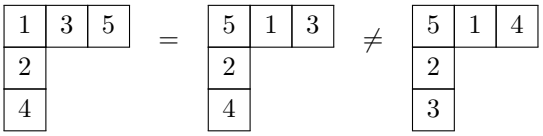


FIGURE 2. Tabloid Equivalence

5.2. **Permutation Modules.** Permutation modules are large representations with significant structure induced from Young subgroups.

Definition 5.8. *Given a partition $\lambda = (\lambda_1, \ldots, \lambda_l)$, the Young subgroup $S_\lambda \subset S_n$ is*

$$S_{1,\ldots,\lambda_1} \times S_{\lambda_1+1,\ldots,\lambda_1+\lambda_2} \times \cdots \times S_{n-\lambda_l+1,\ldots,n}$$

which is isomorphic to

$$S_{\lambda_1} \times S_{\lambda_2} \times \cdots \times S_{\lambda_l}$$

These subgroups are the base of our irreducible representations, and have very important induced structure.

Definition 5.9 (Permutation Modules)**.** *For a partition $\lambda \vdash n$, the permutation module M^λ is the induced module of the Young subgroup $S_\lambda, 1 \uparrow_{S_\lambda}^{S_n} = \mathbb{C}\{\pi_1 S_\lambda, \ldots, \pi_k S_\lambda\}$ for $\{\pi_1, \ldots, \pi_k\}$ a transversal of S_λ.*

Permutation modules have a number important properties. First, tabloids provide a combinatorial basis for permutation modules, that is: $M^\lambda = \mathbb{C}\{\{t_1\}, \ldots, \{t_k\}\}$, where $\{t_1\}, \ldots, \{t_k\}$ are the lambda-tabloids and group action is simply given by $\pi\{t\} = \{\pi(t)\}$. Equally importantly for our purposes, permutation modules decompose into irreducible representations in an often computable way, via Young's rule.

Theorem 5.10 (Young's Rule)**.** $M^\lambda = \bigoplus_{\mu \vdash n} K_{\mu,\lambda} S^\mu$

Here, $K_{\mu,\lambda}$ is the number of semi-standard tableaux of shape μ and content λ, and together are known as the Kostka numbers.

5.3. **Specht Modules.** Now we can introduce the irreducible representations of S_n, which are known as the Specht modules. There are two primary ways these are defined, we will note both definitions here but primarily use the second later on. We first define the row and column stabilizers of a tableaux t with rows $\{R_i\}$ and columns $\{C_i\}$:

$$R_t = S_{R_1} \times \ldots \times S_{R_l}$$
$$C_t = S_{C_1} \times \ldots \times S_{C_k}$$

Further, we define two useful group algebra sums:

$$\mathbb{R}_t^+ = \sum_{\pi \in C_t} \pi, \qquad \mathbb{C}_t^- = \sum_{\pi \in C_t} sgn(\pi)\pi$$

This allows us to define an important extension of a tabloid.

Definition 5.11 (Polytabloid)**.** *Given a tabloid $\{t\}$, its polytabloid is*

$$e_t = \mathbb{C}_t^- \{t\}$$

The irreducible representations of S_n, the Specht modules, may be derived from polytabloids.

Definition 5.12 (Specht Modules)**.** *Given a partition $\lambda \vdash n$, the Specht Module S^λ is the sub-module of M^λ spanned by all polytabloids of shape λ*

Perhaps a more common but somewhat less intuitive equivalent definition for the Specht Modules uses the Young symmetrizer s_λ, which takes advantage of the fact that the tabloid $\{t\} = \mathbb{R}_t t$.

$$s_\lambda = \mathbb{R}_t^+ \mathbb{C}_t^- = \sum_{r \in R_t, c \in C_t} (-1)^c rc$$

Then the Specht module S^λ is $\mathbb{C}[S_n]s_\lambda$. For proof that the Specht modules are irreducible and distinct, we refer the reader to Sagan [32].

6. The KLR Technique for the Symmetric Group

Now that we have seen an abelian application of the KLR technique, and have the requisite background in the representation theory of the symmetric group, let us examine the specifics of its use on a Cayley graph of the symmetric group, the Birkhoff Graph.

Definition 6.1 (Birkhoff Graph)**.** *The Birkhoff Graph Cay(S_n, C) is the Cayley graph on the symmetric group with generating set C consisting of all pure cycles.*

The setup of the technique is the same as above, but the Birkhoff graph will require a more subtle condition on A to get a tight bound. Before, our condition merely regulated the trivial representation. Here, it will bound the norm of the permutation modules evaluated on the indicator function for A, effectively smoothing out its Fourier coefficients. If, on the other hand, the condition is violated and any coefficient is too large, we will be able to restrict to a smaller case and use induction. This technique and the simplification of our inner product requires a number of computational lemmas. Further, since it is partly these computations which informs our decision for a condition on A, we will examine them before proving the independence bound.

6.1. Computational Lemmas and the Test Set Condition.

In general, computing the characters on a given conjugacy class of the symmetric group is difficult, and the result is not necessarily clean. When we applied the KLR Technique to abelian groups, our function ψ summed over all generators, but here, the characters are simply to complicated to make this a reasonable computation. There is, however, a combinatorial rule for computing these values called the Murnaghan-Nakayama rule. While this rule is complicated in general, it simplifies nicely for n-cycles on S_n, so our ψ will only test for these generators.

Lemma 6.2 (Murnaghan-Nayakama Rule for (n)-Cycles)**.** *Given a partition $\lambda \vdash n$,*

$$\chi^{\lambda}((n)) = \begin{cases} (-1)^m, \lambda = h_m \\ 0, \ else \end{cases}$$

Here, h_k are the aptly named "hook partitions" of shape $(n-k, 1, \ldots, 1)$.

This computation motivates us to focus on hook partitions, as it will kill all other terms in our inner product. In particular, we will examine the decomposition of permutation modules indexed by hook partitions.

Lemma 6.3 (Young's Rule for Hook Partitions)**.** *For a hook partition $h_k \vdash n$:*

$$M^{h_k} = \bigoplus_{m=0}^{k} \binom{k}{m} S^{h_m}$$

According to this lemma, if we can bound M^{h_k}, we can bound our irreducible characters and thus our inner product. Yet, the above lemmas barely give any hint of how to relate this representation theory back to the graph itself–what conditions on the graph will bound the permutation module? The idea brought forward by Kane et al. was to pick an explicit basis for M^{h_k} and examine its matrix form.

Remark 6.4. *The permutation module M^{λ} with basis given by λ-tabloids is of the form:*

$$M^{\lambda}_{I,J}(\pi) = \begin{cases} 1, \ \pi(I) = J \\ 0, \ else \end{cases}$$

Further, for $A \subset S_n, \zeta = \frac{1}{|A|} \sum_{\tau \in A} \pi$, extending the definition of M^{λ} linearly gives:

$$M^{\lambda}_{I,J}(\zeta) = Prob_A[\pi(I) = (J)]$$

Here, in other words, is a connection to our graph: we bound the probability that a random draw from A sends tabloid I to J. This allows us to take advantage of the structure vs. randomness paradigm. In other words, if A is sufficiently random, we will be able to bound M and our computation successfully. If, on the other hand, A is structured, we will be able to use that structure for an inductive argument.

6.2. **Independence Bound for the Birkhoff Graph.** With the above lemmas out of the way, we are ready to prove that a set A, with special conditions to allow for induction and bounded characters, cannot be independent.

Proposition 6.5 (Prop 3.1 in [23]). *Given odd n, c sufficiently small, and $A \subset S_n$ s.t.*

(1) *All elements of A are of the same sign*

(2) *For any even $k < n$, and any I, J h_k-tabloids, $Pr_{\pi \in A}[\pi(I) = J] < \frac{c^k (n-k)!}{n!}$*

two elements in A differ by an n-cycle, and thus A is not an independent set of the Birkhoff Graph. If a set A satisfies condition (2), we call it c-pseudorandom.

Proof. The proof of this statement mainly requires the simplification of $\langle \phi, \psi \rangle$ by the lemmas presented in Section 6.1. While ϕ is defined equivalently to the abelian case, ψ is summed over n-cycles rather than the entire generating set, which allows us to compute the irreducible characters directly. Applying Lemma 6.2 reduces our inner product to a sum over hook partitions:

$$\begin{aligned} \langle \phi, \psi \rangle &= \sum_{\lambda \vdash n} \chi^\lambda(\phi) \chi^\lambda(\psi) \\ &= \sum_{m=0}^{n-1} (-1)^m \chi^{h_m}(\phi) \end{aligned}$$

Now we will show how our condition from Section 6.1 bounds the characters of ϕ. In particular, our condition on A allows us to bound its Frobenius norm:

$$M^{h_k}(\zeta)_{I,J} = Pr_{\pi \in A}[\pi(I) = J] \leq c^k/(n)_k$$

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

On its own, this is not of much help, but expanding out ϕ, we can see that:

$$Tr(M^{h_k})(\phi) = ||M^{h_k}(\zeta)||_F^2$$

and combining this with Lemma 6.3 gives a bound on $\chi^{h_m}(\phi)$!

$$\begin{aligned} 1 + \binom{k}{m} \chi^{h_m}(\phi) &\leq \sum_\lambda K_{\lambda, h_k} \chi^\lambda(\phi) \leq c^k \\ \implies \chi^{h_m}(\phi) &\leq \frac{c^k - 1}{\binom{k}{m}} \end{aligned}$$

From here, the argument essentially devolves intro tricks for picking the correct k to bound our sum. For $m \leq n/2$, pick $k = 2m$ for the following final bound:

$$\chi^{h_m}(\phi) \leq \frac{c^{2m} - 1}{\binom{2m}{m}}$$

This works for $m \leq n/2$, but we need a bit more work for $m > n/2$. Given a partition $\lambda \vdash n$, we receive its **conjugate partition** by flipping its rows and columns. Further, the Specht modules of conjugate partitions λ, λ^* differ exactly by the sign representation, i.e. $S^\lambda = S^{\lambda^*} \otimes Sign$. Without loss of generality we may assume our set A is entirely even, which gives $\chi^{h_m}(\phi) = \chi^{h_{n-1-m}}(\phi)$ and bounds the remaining characters.

Putting everything together, we first pull out the trivial character and its conjugate (similar to the abelian case!), then bound the remaining negative terms (odd m).

$$\frac{1}{2} \langle \phi, \psi \rangle \geq 1 - \sum_{m \geq 1, m \text{ odd}}^{\frac{n-1}{2}} \frac{c^{2m} - 1}{\binom{2m}{m}}$$

which happens to be positive for $c > 1$ small enough [23]. Kane et al's specific constant of $c = \sqrt{2}$ requires a slight modification to k. □

In the abelian case, our condition immediately gave us an independence bound. Because of the more stringent conditions on A, in this case we need to apply induction to finish our bound.

Theorem 6.6 (Theorem 1.8 in [23]). *For an independent set A in the Birkhoff graph B_n, $|A| \leq \frac{n!}{2^{\frac{n-4}{2}}}$*

Proof. We will first prove the theorem assuming all permutations of A have the same sign and that n is odd. In this case, our bound will be better, $|A| \leq \frac{n!}{2^{\frac{n-1}{2}}}$. If n = 1, then the bound is $|A| \leq 1$, which is trivially true as $|S_1| = 1$. This serves as the base case for our induction. Let n be odd, if A is $\sqrt{2}$-pseudorandom, we are done. Otherwise, there exists even $m < n$ and h_m-tabloids I, J s.t. $Pr_{\pi \in A}[\pi(I) = J] \geq 2^{m/2}/(n)_m$. Consider the set:

$$A' = \{\pi \in A : \pi(I) = J\}$$

Note that these are now permutations on S_{n-m} (with some relabeling), since they fix m coordinates. $|A'|$, therefore, may be viewed as an independent set in B_{n-m}. Since this set is of smaller dimension, we may apply the inductive hypothesis:

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

Now we reduce to the case above in general. To restrict to the case where A has only permutations of the same sign, simply pick the larger set of permutations in A with the same sign, A_2, losing only a factor of 2. If n is odd we are done. If n is even, we will take a similar strategy to the above and reduce to B_{n-1}. Let j be the most common value of $\pi(n)$. We have $Pr_{\pi \in A}[\pi(n) = j] \geq 1/n$. For simplicity, with relabeling let $j = n$. Consider the set $A_3 = \{\pi \in A_2 | \pi(n) = n\}$. A_3 is an independent set in B_{n-1}, and $n-1$ is odd. Thus we have

$$|A| \leq 2|A_2| \leq 2n|A_3| \leq 2n(n-1)!/2^{(n-2)/2} = n!/2^{(n-4)/2}$$

□

7. The Generalized Symmetric Group

For a given Cayley graph $\mathrm{Cay}(G, S)$ and test set A, the key to the KLR technique lies in the pseudorandomness condition on A and its relation to the representation theory of G. For the abelian case, there was no obvious condition, which brings up the question of whether the technique extends to other groups. An obvious choice for further exploration are the generalized symmetric groups, the wreath product of $\mathbb{Z}_k$ and S_n, due to their similar representation theory to S_n.

7.1. The Generalized Symmetric Group.

To define the Generalized Symmetric Group, we will need the wreath product, which in turn relies on the semi-direct product.

Definition 7.1 (Semi-direct Product). *Given groups G and H along with a homomorphism $\phi : H \to Aut(G)$, the (external) semi-direct product of groups G and H, $G \rtimes H$, is the group $(G \times H, \circ)$ where*

$$(g, h) \circ (g', h') = (g(\phi(h)g'), hh').$$

ϕ defines a group action of H on G. In the future, we may drop ϕ when notation is clear.

Example 7.2 (Dihedral Group). *The semi-direct product $C_n \rtimes C_2$ with $\phi(0)(g) = g, \phi(1)(g) = g^{-1}$ is the dihedral group D_{2n}. In this case, take $(1, 0) = r$, the rotation, and $(1, 1) = f$, the reflection. Clearly these generate the group and $r^n = e$. Further, $f^2 = (1 - 1, 1 + 1) = e$, and $frf = (-1, 0) = (1, 0)^{-1}$. This gives the classic presentation of the dihedral group.*

Definition 7.3 (Wreath Product). *Given groups G and H, and a set Ω with H acting on it, let $K = \prod_{w \in \Omega} G_w$. In other words, K is the direct product of copies of G indexed by the set Ω. Further, we extend the action of H on Ω to an action over K by changing the value at coordinate ω to that at $h^{-1}\omega$. Formally, elements $k \in K$, $h \in H$, and a coordinate ω of k,*

$$h(k)_\omega = k_{h^{-1}\omega}$$

Here, k_ω denotes the value of k at the coordinate index ω. The (unrestricted) wreath product $G \wr H$ is then the semi-direct product $K \rtimes H$, where the action of H on K is as above.

We call a wreath product regular when $\Omega := H$, and the action is given by left multiplication.

Example 7.4. *The simplest regular wreath product, $\mathbb{Z}_2 \wr \mathbb{Z}_2$, is the dihedral group D_8. We will work through this example in more detail to get a feeling for wreath products. To begin, note that $\mathbb{Z}_2 \wr \mathbb{Z}_2 = ((\mathbb{Z}_2)_0 \times (\mathbb{Z}_2)_1) \rtimes \mathbb{Z}_2$. In fact, this is a particularly nice example to showcase the importance of the group action for a semi-direct product. We are about to show that $\mathbb{Z}_4 \rtimes \mathbb{Z}_2 \cong (\mathbb{Z}_2 \times \mathbb{Z}_2) \rtimes \mathbb{Z}_2$ (albeit with different group actions)–an equivalence that should look odd to even a new student of abstract algebra! Let's choose $r = (1, 0, 1)$. Then we have:*

$$r = (1, 0, 1), r^2 = (1, 1, 0), r^3 = (0, 1, 1), r^4 = (0, 0, 0)$$

Further let:

$$f = (1, 0, 0), f^2 = (0, 0, 0)$$

Then:

$$frf = ((1, 0, 0) \circ (1, 0, 1)) \circ (1, 0, 0) = (0, 0, 1) \circ (1, 0, 0) = (0, 1, 1) = r^{-1}$$

Finally, we show r and f generate the group. $fr = (0, 0, 1)$, then $fr \circ (1, 0, 1) = (0, 1, 0) = f'$. Together, f and f' clearly generate $(\mathbb{Z}_2 \times \mathbb{Z}_2) \rtimes 0$, and to get one in the last coordinate, we can simply add $fr = (0, 0, 1)$!

Finally we are ready to define the generalized symmetric group

Definition 7.5 (Generalized Symmetric Group)**.** *The generalized symmetric group S(m,n) is the wreath product $\mathbb{Z}_m \wr S_n$ with $\Omega = \{1, \ldots, n\}$ with the natural group action of permutation.*

Many wreath products may be quite hard to visualize, but the generalized symmetric group is delightfully simple. Elements may be viewed as pairs of permutations and elements in $\mathbb{Z}_m^n$, and composition is a simple matter of applying permutations to the coordinates of $\mathbb{Z}_m^n$ with some addition and composition of permutations! To show this, let's do a quick example of element composition:

Example 7.6. *Let $G = S(3, 3)$, $\pi = (12)$, $\sigma = (23)$, $g = (0, 1, 2)$, and $g' = (1, 0, 0)$. Then*

$$\begin{aligned}(g, \pi) \circ (g', \sigma) &= (g + (\pi(g')), \pi\sigma) \\ &= ((0, 1, 2) + (0, 1, 0), (123)) \\ &= ((0, 2, 2), (123))\end{aligned}$$

Just like the symmetric group, the irreducible representations of the generalized symmetric group may be constructed via Specht modules, and rely heavily on Young tableaux.

7.2. **The Hyperoctahedral Group.** As is often the case in mathematics, we will examine the case $m = 2$:

Definition 7.7 (Hyperoctahedral Group)**.** *The hyperoctahedral group is $S(2, n) = \mathbb{Z}_2 \wr S_n$*

The hyperoctahedral group gives the symmetries of the hypercube (we showed this for the case $n = 2$ in Example 7.4). Due to the common use of the hypercube, especially with the ubiquity of boolean function analysis, one could imagine such a group of symmetries and its representations to be useful throughout much of computer science. Indeed, the hyperoctahedral group has found use in cryptography [30, 7], as well as a number of other areas in discrete math and combinatorics. Yet, despite its stark similarity to the symmetric group, there is a dearth of usage of its representation theory. This section provides a novel use of the representation theory of the hyperoctahedral group and shows that the KLR technique may be expanded to non-abelian groups beyond S_n.

7.3. **The Signed Symmetric Group.** For a full treatment of the signed symmetric group, see Richard Bayley's dissertation [8]. As a permutation group, $S(2, n)$ is a subgroup of S_{2n} known as the signed symmetric group, and has a very simple formulation. In particular, let us relabel the normal $\{1, \ldots, 2n\}$ to $\Omega = \{-n, -n + 1, \ldots, -1, 1, \ldots, n\}$. Then

$$S(2, n) = \{\pi \in S_\Omega \mid \pi(-i) = -\pi(i)\}$$

To understand the representation theory of $S(2, n)$, we will first need to understand its structure as a group. Being a permutation group, every element $\pi \in S(2, n)$ has a unique cycle decomposition

in S_{2n}. Cycle notation does not work well with negative numbers, so from here on out we will write $-k = \bar{k}$, and $\bar{\bar{k}} = k$–notation introduced by Bayley. Given a cycle $c = (a_1 a_2 \dots a_n)$, with $a_i \in \{\bar{n}, \dots, n\}$, let $\bar{c}$ denote the "negation" of the cycle, $\bar{c} = (\bar{a_1}\bar{a_2}\dots\bar{a_n})$. The representation theory of the signed symmetric group will rely on two different types of cycles:

Definition 7.8 (Stable and Anti-Stable Cycles)**.** *Let $\Omega = \{\bar{n}, \dots, \bar{1}, 1, \dots, n\}$. Given a cycle $c \in S_\Omega \cong S_{2n}$:*

(1) *if $c = \bar{c}$, we call c a stable cycle*
(2) *if $c \neq \bar{c}$, we call c an anti-stable cycle*

Example 7.9. *Let's consider the simplest non-trivial signed symmetric group, $S(2,2)$. In general, the group has $2^n n!$ elements, so we need only list 8 elements in our case. They are:*

$$\{(), (1\bar{1}), (2\bar{2}), (1\bar{1})(2\bar{2}), (12)^*(\bar{1}\bar{2})^*, (1\bar{2})^*(2\bar{1})^*, (12\bar{1}\bar{2}), (1\bar{2}\bar{1}2)\}$$

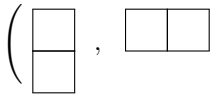
In the above, cycles with a star denote anti-stable cycles, those without are stable. As one can see, anti-stable cycles come in pairs. This is true in general, as any element of $S(2,n)$ must be stable as a whole.

Recall that the conjugacy classes of the symmetric group correspond to certain cycle decompositions. In particular, two elements are in the same conjugacy class if they have the same number of cycles of each length in their decomposition. Further, these decompositions then correspond to partitions of n. The signed symmetric group mirrors this structure, but instead of using only the length of cycles to distinguish them, it uses length and stability.

Remark 7.10. *$\pi, \sigma \in S(2,n)$ are conjugate if their decomposition has the same number of stable cycles of every size, and same number of anti-stable cycles of every size* [8].

Just as in the case of S_n, these are related to partitions of n. In fact, the conjugacy classes are indexed by bi-partitions, i.e. two partitions λ, μ s.t. $|\lambda| + |\mu| = n$, denoted $(\lambda, \mu) \vdash n$ This relation is given by the following rules laid out in Bayley's work [8]. The first partition of the pair corresponds to the stable cycles in the permutation. Each row of the partitions Ferrers diagram corresponds to a single stable cycle, with length equivalent to half the length of the cycle. The second partition corresponds to anti-stable cycles, where each row corresponds to half a pair of anti-stable cycles, with length equivalent to this single cycle.

Example 7.11. *The element $(1\bar{1})(2\bar{2})(34)(\bar{3}\bar{4}) \in S(2,4)$ corresponds to the bi-partition*



In fact, these bi-partitions were originally used to index the conjugacy classes of the hyperoctahedral group thought of as a wreath product.

7.4. **The Representation Theory of the Hyperoctahedral Group.** For a full treatment of the representation theory of the hyperoctahedral group, refer to [15]–we will only provide an overview of what is necessary for the KLR technique. In many ways, the representation theory of the hyperoctahedral group is the generalization of the representation theory of the symmetric group from partitions to bi-partitions of n. In fact, for right-handed degenerate bi-partitions, i.e. when the left partition is the empty-set, the theory often simply reduces to that of S_n. The irreducible representations, just as in the symmetric group, are Specht modules indexed by bi-partitions which we will write as $S^{\lambda,\mu}$. Further, these representations may be constructed as sub-modules of permutation modules similarly indexed by bi-partitions $M^{\lambda,\mu}$.

Definition 7.12 (Permutation Module for S(2,n))**.** *Given a bi-partition $\lambda, \mu \vdash n$, the permutation module $M^{\lambda,\mu}$ is defined to be the induced representation $1 \uparrow_{S_\lambda \times B_\mu}^{S(2,n)}$. Here S_λ is the Young Subgroup corresponding to the partition λ, and B_μ, for $\mu \vdash m$, is the semi-direct product of $S_\mu \rtimes E(m)$, the subgroup of diagonal $m \times m$ matrices with ± 1 values.*

This mirrors the original definition of the permutation module for S_n, but in general, $M^{\lambda,\mu}$ has no known corresponding basis of combinatorial objects. Around the turn of the century, Halicioglu and Morris developed a combinatorial method to build the Permutation and Specht modules using Dynkin diagrams, which they expand to a generalized version of Young tableaux [17]. However, the combinatorial method does not seem to lend itself as well to decomposition into irreducibles. Geissinger and Kinch's method, on the other hand, leads to an analogous result to Young's rule:

Theorem 7.13 (Theorem III.5 in [15])**.** *Given a bi-partition $\lambda, \mu \vdash n$, the permutation module decomposes into the irreducible Specht modules by:*

$$M^{\lambda,\mu} = \bigoplus_{\alpha,\beta} \left(\sum_{m \subset \lambda} K_{\alpha,m} K_{\beta,(\lambda - m) \cup \mu} \right) S^{\alpha,\beta}$$

Here, a partition $m \subset \lambda$ if $\forall i, m_i \leq \lambda_i$.

As in the previous use of the KLR technique, it is this decomposition which is key to bounding our computation.

7.5. **Symmetric Functions.** A polynomial f is said to be symmetric if $\forall \pi \in S_n, f(x_1, \ldots, x_n) = f(x_{\pi(1)}, \ldots, x_{\pi(n)})$. Symmetric functions play a pivotal roll in both the representation theory of the symmetric and hyperoctahedral groups. While we were able to define the irreducible representations using equivalent ideas over Young tableaux, we will need some results on symmetric functions for our computations. Perhaps the most obvious symmetric polynomial is the power sum:

Definition 7.14 (Power sum)**.** *The power sum of x at k is given by:*

$$p_k(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^k$$

Amazingly, despite their simplicity, power sums end up having great importance. Often we will wish to decompose symmetric functions into some basis. In fact, there is a basis of symmetric functions closely tied to Young tableaux known as Schur polynomials.

Definition 7.15 (Schur polynomials)**.** *Given a partition $\lambda \vdash n$, let the matrix $(M_\lambda)_{ij} = x_i^{\lambda_i + n - j}$. The associated Schur polynomial is then:*

$$s_\lambda(x) = \frac{det(M_\lambda)}{det(M_\emptyset)}$$

Example 7.16. *The Schur polynomial associated to the empty-set, $s_\emptyset$, is* 1*.*

This was the initial definition for Schur polynomials offered by Cauchy. Almost a century later, Littlewood proved a connection between Schur polynomials and Young tableaux [35]:

Theorem 7.17. *Given a partition $\lambda \vdash n$:*

$$s_\lambda(x) = \sum_T \prod_i x^{c(T)_i}$$

where the sum is over semi-standard tableaux T of size λ and content $c(T)$.

This theorem provides some intuition for why Schur polynomials show up in our computations.

8. KLR Technique for the Hyperoctahedral Group

8.1. **Picking a Cayley Graph.** With our base group chosen, we are left to pick a generating set. Recall that chosen for the Birkhoff graph: the set of all pure cycles C. Perhaps the most natural extension, then, would be the set of pure cycles of the signed symmetric group. However, upon closer examination, one notices that this graph to have very large independent sets based on parity, similar to $Cay(\mathbb{F}_2^n, \{e_i\})$. What about pairs of anti-stable cycles? This generator set leads to a richer structure with no large independent sets. In fact, due to the inclusion of the elements $\{(0, c) \mid c \in C\}$, the graph contains a partition of 2^n subgraphs which contain the Birkhoff graph. Then, to be robust

to different groups, we would like to show that we may substitute the hyperoctahedral group for the symmetric group in our analysis to recover a similar bound.

We note as well that the anti-stable cycles do not generate the hyperoctahedral group; they generate the subgroup of index 2 called the coxeter group D_n. Contrary to the definition given previously, Cayley graphs actually do not need a full generator set S. Let C now be the set of anti-stable pairs. Our analysis below will be on Cay($S(2,n)$, C), which is two disjoint copies of a Cayley graph with a full generator set, Cay(D_n, C).

8.2. **Computational Lemmas.** As for the symmetric group, there are a number of computational results we will need to apply the KLR technique. These computations inform our generator subset and condition on A, so we present them first here while explaining their use in more detail in Section 8.5. Our first lemma will regard the computation of $\chi^{\lambda,\mu}(\psi)$, and will lead us to define a set of partitions analogous to hooks for bi-partitions.

Lemma 8.1 (Murnaghan-Nakayama for $((n),\emptyset)$)**.** *Given a bipartition $\lambda,\mu \vdash n$, let $\chi^{\lambda,\mu}$ be the irreducible character corresponding to $S^{\lambda,\mu}$. Then:*

$$\chi^{\lambda,\mu}((n),\emptyset) = \begin{cases} (-1)^m, & \lambda,\mu = h_m,\emptyset \\ (-1)^m, & \lambda,\mu = \emptyset,h_m \\ 0, & \textit{else} \end{cases}$$

Proof. We will use a well known analog of the Murnaghan-Nakayama Rule [34, 33]. Given a bi-partition $\lambda,\mu \vdash n$, the rule states:

$$\prod_i (p_{\lambda_i}(x) + p_{\lambda_i}(y)) \prod_i (p_{\mu_j}(x) - p_{\mu_j}(y)) = \sum_{\alpha,\beta\vdash n} \chi^{\alpha,\beta}(\lambda,\mu) s_\alpha(x) s_\beta(y)$$

Of course, we only wish to examine a single bi-partition, $(n),\emptyset \vdash n$, which simplifies the above to:

$$p_n(x) + p_n(y) = \sum_{\alpha,\beta\vdash n} \chi^{\alpha,\beta}(n,\emptyset) s_\alpha(x) s_\beta(y)$$

The left-hand side may be simplified to hook partitions using the original Murnughan-Nakayama rule:

$$p_n(x) + p_n(y) = \sum_{i=0}^{n-1} (-1)^i s_{h_i}(x) + \sum_{i=0}^{n-1} (-1)^i s_{h_i}(y)$$

This leaves us with an equality of Schur functions:

$$\sum_{i=0}^{n-1} (-1)^i s_{h_i}(x) + \sum_{i=0}^{n-1} (-1)^i s_{h_i}(y) = \sum_{\alpha,\beta\vdash n} \chi^{\alpha,\beta}(\lambda,\mu) s_\alpha(x) s_\beta(y)$$

Schur functions are a basis of symmetric functions, thus matching the coefficients gives the desired result. □

This lemma reduces our computation down to degenerate hook bi-partitions, which we call uni-hooks.

Definition 8.2 (Unihooks)**.** *We call a bi-partition $\lambda,\mu \vdash n$ a left-unihook if it is of the form $(h_m,\emptyset)$. Similarly, we call a bi-partition $\lambda,\mu \vdash n$ a right-unihook if it is of the form $(\emptyset,h_m)$. We call the union of left and right unihooks the set of unihooks.*

Let's examine how permutation modules indexed by such partitions decompose.

Lemma 8.3. *The decomposition of the permutation module $M^{\emptyset,h_k}$ is given by:*

$$M^{\emptyset,h_k} = \bigoplus_{m=0}^{k} \binom{k}{m} S^{\emptyset,h_m}$$

Proof. This is a direct application of Geissenger and Kinch's Rule, Theorem 7.13:

$$\begin{aligned} M^{\emptyset,h_k} &= \bigoplus_{\alpha,\beta} \left(\sum_{m \subset \emptyset} K_{\alpha,m} K_{\beta,(\emptyset - m)\cup h_k} \right) S^{\alpha,\beta} \\ &= \bigoplus_{\alpha,\beta} K_{\alpha,\emptyset} K_{\beta,h_k} S^{\alpha,\beta} \\ &= \bigoplus_{\beta} K_{\beta,h_k} S^{\emptyset,\beta} \\ &= \bigoplus_{m=0}^{k} \binom{k}{m} S^{\emptyset,h_m} \end{aligned}$$

□

Unfortunately, the permutation modules of left unihooks are a great deal more complicated than right unihooks, and the corresponding Geissenger and Kinch decomposition is not useful. With this in mind, we will need the following lemma concerning the relationship between conjugate and transpose bi-partitions.

Lemma 8.4 (Proposition II.I in [15])**.**

1. $\chi^{\lambda,\mu} = \chi^{\lambda^*,\mu^*}\chi^{(1^n),\emptyset}$
2. $\chi^{\lambda,\mu} = \chi^{\mu,\lambda}\chi^{(n),\emptyset}$

Further $S^{(1^n),\emptyset}$ is the alternating representation on S_n with trivial action on $E(n)$, and $S^{(n),\emptyset}$ acts trivially on S^n while counting parity on $E(n)$.

To get equivalence between unihooks, we will want to define the idea of sign and parity for $S(2,n)$.

Definition 8.5 (Sign and Parity)**.** *An element $x = (a_1, \ldots, a_n, \pi) \in \mathbb{Z}_2 \wr S_n$ shares its sign with π, and parity with $(a_1, \ldots, a_n)$*

Now let's explicitly write out equivalences between unihooks where we have assumed even sign and parity.

Corollary 8.6. *Given $x \in S(2,n)$ with even sign and parity, then:*

(1) $\chi^{\emptyset,h_k}(x) = \chi^{\emptyset,h_{n-k-1}}(x)$
(2) $\chi^{\emptyset,h_k}(x) = \chi^{h_k,\emptyset}(x)$

8.3. **A Basis for Permutation Modules.** In general, the permutation modules $M^{\lambda,\mu}$ do not seem to correspond to any nice basis structure as M^{λ} did with λ-tabloids. However, unihooks provide a special case in which the permutation modules are significantly simplified:

$$\begin{aligned} M^{h_k,\emptyset,} &= 1 \uparrow_{S_{h_k}}^{S(2,n)} = 1 \uparrow_{S_{n-k}}^{S(2,n)} \\ M^{\emptyset,h_k} &= 1 \uparrow_{B_{h_k}}^{S(2,n)} = 1 \uparrow_{S_{n-k} \rtimes E(n)}^{S(2,n)} \end{aligned}$$

Both right and left unihooks have nice combinatorial bases, generated by versions of tabloids based off of naturally defined signed Young tableaux. However, we will focus on the simpler right unihook permutation modules, whose basis remains the Young h_k-tabloids. We rename these Unsigned Young tabloids, as they are equivalent to Young tabloids with the sole difference being action by $\pi \in S(2,n)$ as $\pi(\{t\}) = \{|\pi t|\}$.

Lemma 8.7. *The unsigned Young tabloids provide a basis for the permutation module $M^{\emptyset,h_k} = 1 \uparrow_{S_{h_k} \rtimes E(n)}^{S(2,n)}$*

Proof. Recall that

$$1\uparrow_{S_{n-k}\rtimes E(n)}^{S(2,n)} = \mathbb{C}S(2,n)\{\pi_1 B_{h_k},\ldots,\pi_m B_{h_k}\}$$

for a left-transversal $\{\pi_1,\ldots,\pi_m\}$ of B_{h_k} [32]. Let T be the module spanned by the unsigned tabloids: $\mathbb{C}S(2,n)\{\{t^{h_k}\}\}$. We will construct an isomorphism:

$$\theta : M^{\emptyset,h_k} \to T$$

by $\theta(\pi_i B_{h_k}) = \{\pi_i t^{h_k}\}$. This is clearly bijective, as it sends a basis to a basis (note that the number of Unsigned h_k-tabloids $\frac{n!}{(n-k)!} = [S(2,n) : B_{h_k}]$). The additive homomorphic property follows immediately from definition. Multiplication is only slightly more subtle, and follows due to the fact that the subgroup fixes $\{t^{h_k}\}$ (the tableaux filled in order (left to right, up to down) by $[n]$). In particular, given $\sigma \in B_{h_k}$, $\exists j$ s.t. $\sigma\pi_i B_{h_k} = \pi_j B_{h_k}$. In this case, $\theta(\sigma\pi_i B_{h_k}) = \{\pi_j t^{h_k}\}$. Further, we may write $\sigma = \pi_j s \pi_i^{-1}$, for some $s \in B_{h_k}$. Then

$$\sigma\theta(\pi_i B_{h_k}) = \sigma\{\pi_i t^{h_k}\} = \{\pi_j s\pi_i^{-1}\pi_i t^{h_k}\} = \{\pi_j t^{h_k}\}.$$

□

8.4. **Permutation Modules and the Test Set Condition.** Recall that for the symmetric group, the main condition on our test set A stemmed from the matrix representation of M^λ under the tabloid basis. Because we have a decomposition rule mirroring Young's rule, the same approach will allow us to bound the irreducible characters of the hyperoctahedral group. Given unsigned Young h_k-tabloids I, J, then:

$$M_{I,J}^{\emptyset,h_k}(\pi) = \begin{cases} 1, \pi(I) = J \\ 0, \text{ else} \end{cases}$$

Now in creating our condition, we have to hold two considerations in mind: bounding our characters, and allowing for induction. For the symmetric group, we only really had one choice, the tabloid condition, and it worked for both. This is no longer the case: attempting to use a condition on left unihooks will result in failure of induction. Consider the most natural condition on A: $\forall I, J$, signed h_k-tabloids,

$$Prob_{\pi\in A}[\pi(I) = J] \leq \frac{c^k}{|S|}, |S| = \frac{2^n n!}{(n-k)!}.$$

On the surface, this condition seems great, and even implies the natural condition on unsigned tabloids:

$$Prob_{\pi\in A}[\pi(I) = J] \leq \frac{c^k}{|U|}, |U| = \frac{n!}{(n-k)!}.$$

Together these give what seem like appropriate bounds: $Tr(M^{h_k,\emptyset}(\phi)) \leq c^k$ and $Tr(M^{\emptyset,h_k}(\phi)) \leq c^k$. Yet, even without doing any computation we can immediately tell that induction will fail under this property. Our independence bound $T(n)$ requires the following recursion:

$$T(n) \leq T(n-m)\frac{n!2^n}{(n-m)!c^k}$$

Even setting $T(n)$ to be the size of the entire group doesn't work, giving us

$$(n-m)!2^{n-m}\frac{n!2^n}{(n-m)!c^m} = n!2^n\frac{2^{n-m}}{c^m}$$

Since c, m are constant, c^m cannot possibly compete with 2^{n-m} asymptotically, so our condition cannot be used inductively. Thus informed, we will instead use only the condition on unsigned tabloids, and use Corollary 8.6 to bound left unihooks. Recall our normalized indicator function $\zeta = \frac{1}{|A|}\sum_{\tau\in A}\tau$. Then analogous to the S_n case, we have:

$$M^{\emptyset,h_k}(\zeta)_{I,J} = Pr_{\pi\in A}[\pi(I) = J] \leq \frac{c^k}{|U|}$$
$$\implies ||M^{\emptyset,h_k}(\zeta)||_F^2 \leq c^k$$

These bounds will be instrumental to applying the KLR technique.

8.5. **Applying the KLR Technique.** We are finally ready to apply the KLR Technique to the hyperoctahedral group.

Proposition 8.8. *Given a set of vertices A in $Cay(S(2,n),C)$, n odd, with the properties:*

(1) *All elements in A are of the same sign and parity*

(2) *For unsigned tabloids I,J, $Prob_{\pi\in A}[\pi(I)=J] \leq \frac{c^k(n-m)!}{n!}$*

then two elements $\pi,\sigma \in A$ differ by an anti-stable pair of length $2n$ for small enough $c>1$, e.g. $c=\sqrt{2}$.

Proof. As a reminder, we will redefine ϕ and ψ.

$$\phi = \frac{1}{|S(2,n)||A|^2}\sum_{\sigma\in S(2,n),a,a'\in A} \sigma a a'^{-1}\sigma^{-1}$$

$$\psi = \frac{1}{|C_{2n}|}\sum_{\tau\in C_{2n}} \tau$$

Here ψ is summed over C_{2n}, pure stable cycles of length exactly $2n$. As with the Birkhoff graph, we begin with:

$$\langle\phi,\psi\rangle = \sum_{(\lambda,\mu)\vdash n} \chi^{\lambda,\mu}(\phi)\chi^{\lambda,\mu}(\psi)$$

Our next step is to compute the irreducible characters evaluated on ψ, but since we have chosen ψ to be purely cycles of length $2n$, we may apply Lemma 8.1, which reduces our problem to a sum over the characters of unihooks:

$$\begin{aligned}\langle\phi,\psi\rangle &= \sum_{m=0}^{n-1}(-1)^m\left(\chi^{h_m,\emptyset}(\phi)+\chi^{\emptyset,h_m}(\phi)\right)\\ &= 2\sum_{m=0}^{n-1}(-1)^m\chi^{\emptyset,h_m}(\phi)\end{aligned}$$

In the second step, we have applied Corollary 8.6 which condition (1) allows us to use (shifting by an odd sign and/or parity if necessary) to further the simplification. Recall from Section 6.2 that $Tr(M^{\emptyset,h_k}(\phi)) = ||M^{\emptyset,h_k}(\zeta)||_F^2$. Geissenger and Kinch's rule then gives:

$$Tr(M^{\emptyset,h_k}(\phi)) = \sum_{m=0}^{n-1}\binom{k}{m}\chi^{\emptyset,h_m} \leq c^k$$

This clearly mirrors the process for S_n and gives:

$$\chi^{\emptyset,h_m}(\phi) \leq \frac{c^k-1}{\binom{k}{m}}$$

We are now in the same position as in Proposition 6.5, and the desired bound follows for $c=\sqrt{2}$. □

With this in hand, we can prove our independence bound.

Theorem 8.9. *For any independent set $A\subset Cay(S(2,n),C)$, $|A| \leq \frac{n!2^n}{2^{\frac{n-6}{2}}}$*

Proof. We begin by assuming that n is odd and A consists entirely of one sign and parity. Here we will prove $|A| \leq \frac{n!2^n}{2^{\frac{n-1}{2}}}$. Our base case $n=1$ is clear, since simplifying gives $|A|\leq 2$ which is clearly true as $|S(2,1)|=2$. If A satisfies the unsigned tabloid conditions, then A is not independent and we are done. Assume A does satisfy this condition, then there exist unsigned h_k-tabloids I,J s.t. $Prob_{\pi\in A}[\pi(I)=J] \geq \frac{2^{k/2}(n-k)!}{n!}$. Now let $A'=\{\pi\in A \mid \pi(I)=J\}$. In the case of the symmetric group, A' was an independent set in the lower dimensional space, but here because the tableaux are unsigned, this is no longer true. However, we can reduce to the lower dimensional graph by restricting to the most common group of k signs on the lower rows of the tabloids, losing a factor of

at most 2^k. With some relabeling, the restricted A' is then an independent set in $Cay(S(2, n-k), C)$, so we may apply our inductive hypothesis:

$$|A| \leq \frac{2^k n!}{(n-k)!2^{k/2}}|A'| \leq \frac{n!2^n}{2^{(n-1)/2}}$$

Now let's generalize this to even n and A of any sign. Our base case still clearly holds for our new bound $\frac{2^n n!}{2^{(n-6)/2}}$. First, we simply reduce to the more common sign and parity in A, reducing its size by a factor of 4 at most. Second, if n is even, we will restrict to the most common value of $\pi(n)$, which reduces the size of A by $\frac{1}{2n}$ at most. This gives a new set A' which is an independent set in $Cay(S(2, n-1), C)$. Then inductively:

$$|A| \leq 8n|A'| \leq \frac{n!2^n}{2^{(n-6)/2}}$$

□

9. Further Directions

9.1. **Abelian Base Groups: $A_q(n, d)$.** Perhaps the clearest direction of further research is an extension of our new result on $A_q(n, d)$. One possibility lies in examining the Fourier coefficients of our indicator function more closely. Our bound at the moment is extremely crude, as it simply assumes the Fourier weight is entirely concentrated at the minimum coefficient–this is, of course, true in the case of some perfect codes. Understanding the layout of Fourier coefficients in large independent sets could allow not only for an improved closed-form bound for single error correcting codes (perhaps one could match the new semi-definite bounds), but might also open up the possibility of expanding to larger distances.

9.2. **Non-Abelian Base Groups and Coding Theory.** Another obvious direction is to continue the work this paper presents on non-abelian groups. For instance, one could likely generalize the technique to $S(m, n)$ for any m, rather than picking $m = 2$ as we have done above. The representation theory of the generalized symmetric group is well-studied, and the results necessary for application of the KLR technique are likely known as well.

While studying the generalized symmetric group would help show the robustness of the method, novel applications to other families such as affine groups could have implications in coding theory. The main application of the KLR Technique in [23] is to provide an exponential lower bound on alphabet size of maximally recoverable codes over grid topologies, a popular layout for error correcting codes. Studying other non-abelian base groups with well understood representation theory could shine light on lower bounds for other code topologies as well.

9.3. **Lower Bounds for $S(2, n)$.** In order to further examine whether the KLR technique is the "correct" way to view independence number, we should prove the tightness of our bounds. We know for both the abelian and symmetric case our bound is often tight. However, we have provided no lower bound for the hyperoctahedral group, which forced the addition of extra generators and possibly weakened the tightness of our bound.

10. Acknowledgements

I would like to thank my research advisor Professor Madhu Sudan who initially directed me towards the work of Kane, Lovett, and Rao–all of whom I would also like to thank for speaking with me about their paper. Further, I would like to thank Professor Richard Stanley who both took the time to ask about my research and directed me towards the generalized symmetric group. Finally, I'd like to thank Sam Hopkins for his insightful comments in the writing phase.

References

[1] Ahmadi, A. (2016). ORF 523 Lecture 11 [pdf]. Retrieved from `http://www.princeton.edu/~amirali/Public/Teaching/ORF523/S16/ORF523_S16_Lec11_gh.pdf`

[2] Alon, N. (2013). The chromatic number of random Cayley graphs. European Journal of Combinatorics, 34(8), 1232-1243.

[3] Alon, N., Dinur, I., Friedgut, E., & Sudakov, B. (2004). Graph Products, Fourier Analysis and Spectral Techniques. Geometric & Functional Analysis GAFA, 14(5), 913-940.

[4] Ambainis, A., Magnin, L., Roetteler, M., & Roland, J. (2011). Symmetry-Assisted Adversaries for Quantum State Generation. Computational Complexity (CCC), 2011 IEEE 26th Annual Conference on, 167-177.

[5] Artin, M. (2011). Algebra (2nd ed.). Boston: Pearson Prentice Hall.

[6] Babai, L. (1979). Spectra of Cayley graphs. Journal of Combinatorial Theory, Series B, 27(2), 180-189.

[7] Bauer, F. (1997). Decrypted secrets : Methods and maxims of cryptology. Berlin ; New York: Springer.

[8] Bayley, R. (2007). Relative Character Theory and the Hyperoctahedral Group, PQDT - UK & Ireland.

[9] Burgisser, P., & Ikenmeyer, C. (2013). Explicit lower bounds via geometric complexity theory. Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing, 141-150.

[10] Coleman, R. (2011). On Krawtchouk Polynomials. arXiv:1101.1798 [math.CA]

[11] Cooperman, G., Finkelstein, L., & Sarawagi, N. (1990). Applications of Cayley graphs. In International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes (pp. 367-378). Springer, Berlin, Heidelberg.

[12] Delsarte, P., & Levenshtein, V. (1998). Association schemes and coding theory. Information Theory, IEEE Transactions on, 44(6), 2477-2504.

[13] Diaconis, P., & Shahshahani, M. (1981). Generating a random permutation with random transpositions. Zeitschrift Fur Wahrscheinlichkeitstheorie Und Verwandte Gebiete, 57(2), 159-179.

[14] Dinur, I., Friedgut, E., & Regev, O. (2008). Independent Sets in Graph Powers are Almost Contained in Juntas. Geometric and Functional Analysis, 18(1), 77-97.

[15] Geissinger, & Kinch. (1978). Representations of the hyperoctahedral groups. Journal of Algebra, 53(1), 1-20.

[16] Gopalan, P., Hu, G., Kopparty, S., Saraf, S., Wang, C., & Yekhanin, S. (2016). Maximally Recoverable Codes for Grid-like Topologies.

[17] Halicioglu, S., & Morris, A. (2003). Specht Modules for Weyl Groups. Beitrage Algebra Geom. 34 (1993), no. 2, 257–276.

[18] Hamming, R. (1950). Error Detecting and Error Correcting Codes. Bell System Technical Journal, 29(2), 147-160.

[19] Hastad, J. (1999). Clique is hard to approximate within $\frac{n}{1?\epsilon}$. Acta Mathematica, 182(1), 105-142.

[20] Huang, J., Guestrin, C., & Guibas, L. (2009). Fourier theoretic probabilistic inference over permutations. Journal of Machine Learning Research, 10, Article ID 997-1070.

[21] Kahn, J., Kalai, G., & Linial, N. (1988). The influence of variables on Boolean functions. Foundations of Computer Science, 1988., 29th Annual Symposium on, 68-80.

[22] Kaplan, N. (2011). Coding Theory Lecture Notes. Unpublished.

[23] Kane, D., Lovett, S., & Rao, S. (2017). The independence number of the Birkhoff polytope graph, and applications to maximally recoverable codes.

[24] Karp, R. M. (1972). Reducibility among combinatorial problems. In Complexity of computer computations (pp. 85-103). Springer, Boston, MA.

[25] Kaski, P. (2002). Eigenvalues and Spectra of Cayley Graphs

[26] Konstantinova, E. (2008). Some problems on Cayley graphs. Linear Algebra and Its Applications, 429(11), 2754-2769.

[27] Lovasz, L. (1979). On the Shannon capacity of a graph. Information Theory, IEEE Transactions on, 25(1), 1-7.

[28] Musin, O. (2009). Bounds for codes via semidefinite programming. Information Theory and Applications Workshop, 2009, 237-239.

[29] Navon, M., & Samorodnitsky, A. (2005). On Delsarte's linear programming bounds for binary codes. Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on, 327-336.

[30] Raharinirina, I. (2016). Use of Signed Permutations in Cryptography. arXiv:1612.05605 [C.CR]

[31] C. Roos, C. de Vroedt (1978). Some upper bounds for codes derived from Delsartes inequalities for Hamming schemes, Delft Progress Rep., 3, pp. 127-138

[32] Sagan, B. (2001). The symmetric group : Representations, combinatorial algorithms, and symmetric functions (2nd ed., Graduate texts in mathematiC ; 203). New York: Springer.

[33] R. Stanley, Evaluation of irreducible representations of the hyperoctahedral group at bipartition $(\lambda, \mu) = ([n], \emptyset)$, URL (version: 2017-11-26): `https://mathoverflow.net/q/287018`

[34] Stembridge, J. (1987). Unpublished Notes.

[35] Tamvakis, H. (2012). The theory of Schur polynomials revisited. L'Enseignement Mathematique, 58(1), 147-163.