Linear Programming Upper Bounds on Permutation Code Sizes From Coherent Configurations Related to the Kendall Tau Distance Metric

Fabian Lim\(^1\) and Manabu Hagiwara\(^2\)

\(^1\)Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139
\(^2\)National Institute of Advanced Industrial Science and Technology
Central 2, 1-1-1 Umezono, Tsukuba City, Ibaraki, 305-8568, JAPAN
Email: flim@mit.edu, hagiwara.hagiwara@aist.go.jp

Abstract—Recent interest on permutation rank modulation shows the Kendall tau metric as an important distance metric. This note documents our first efforts to obtain upper bounds on optimal code sizes (for said metric) ala Delsarte’s approach. For the Hamming metric, Delsarte’s seminal work on powerful linear programming (LP) bounds have been extended to permutation codes, via association scheme theory. For the Kendall tau metric, the same extension needs the more general theory of coherent configurations, whereby the optimal code size problem can be formulated as an extremely huge semidefinite programming (SDP) problem. Inspired by recent algebraic techniques for solving SDP’s, we consider the dual problem, and propose an LP to search over a subset of dual feasible solutions. We obtain modest improvement over a recent Singleton bound due to Barg and Mazumdar. We regard this work as a starting point, towards fully exploiting the power of Delsarte’s method, which are known to give some of the best bounds in the context of binary codes.

Index Terms—association schemes, coherent configurations, permutations, linear programming, semidefinite programming

I. INTRODUCTION

A permutation code is designed to only allow certain pairwise distances between any two codewords. These codes have been studied in various contexts, e.g., group codes \([8]\), signal modulation \([2, 3]\), vector quantization \([4]\), rank modulation \([5, 6]\), cost-constrained transpositions \([7]\), etc. This work is motivated by a recent study on a fundamental coding problem. In \([6]\) they looked at optimal code sizes with respect to the Kendall tau distance metric. This metric is important to rank modulation and its applications, e.g., flash memories.

For binary codes, Delsarte’s optimization-based methods \([8]\) give some of the best known bounds \([9]\). For permutation codes, we observe during initial experiments (for very small lengths) that Delsarte-like methods outperform Hamming (sphere packing) bounds \([6, 10]\). Our interest is to investigate, if this improvement carries over for larger codes. Tarnanen extended Delsarte’s work over to permutation codes \([11]\), however only for the Hamming metric (and other metrics with similar symmetries). The novelty here is to consider the Kendall tau metric, and as pointed out in \([6]\), lacks required symmetry to straightforwardly apply Tarnanen’s techniques.

Delsarte’s (and Tarnanen) techniques are based on association schemes, from which linear programming (LP) formulations (of the optimal code size problem) are obtained. For the Kendall tau metric, one needs to consider the more general theory of coherent configurations (CC), which instead deliver semidefinite programming (SDP) formulations. The matrices in these SDP’s turn out to be of unwieldy size, but recent work \([12, 13, 14]\) suggest possible approaches. One may exploit the algebraic structure of the CC’s, to only work with block-diagonalized (and possibly much smaller) versions of these matrices. To our knowledge, such recent techniques are new in the area of permutation codes. However, the solution is not straight-forward. As code lengths increase, the CC’s (related to the Kendall tau metric) become huge quickly, motivating the techniques presented in this preliminary report.

While we believe to be presently unable to fully exploit the power of SDP bounds, we show some initial success. We consider the dual problem (also a SDP), and use an LP to search over a subset of feasible solutions. We obtained modest improvement over a recently published Singleton bound in \([6]\). The reduced complexity allows us to compute up to permutation codes of length 11 (where the matrices were previously of order 11 factorial). Certain bottlenecks, if solved, could allow computation for longer codes. As it stands, our proposed LP bounds perform poorer than known Hamming bounds \([6]\), and it remains to see how far sophisticated SDP-based approaches can ultimately bring us. This note aims to motivate new research to resolve this open question.

II. BACKGROUND

A. Optimal Code Size Problem and Two Metrics

Let \(S_n\) denote the symmetric group on a set \(\{1, 2, \ldots, n\}\) and \(\text{dist}(, )\) be a distance metric on \(S_n\). A subset \(\mathcal{V}\) of \(S_n\) is an \((n, \delta_{\text{min}})\) permutation code (with respect to \(\text{dist}(, )\)), if for any \(g, h \in \mathcal{V}\) such that \(g \neq h\), we have \(\text{dist}(g, h) \geq \delta_{\text{min}}\).

Definition 1 (Optimal code size problem). Let \(\text{dist}(, )\) be a distance metric on the symmetric group \(S_n\). Let \(\delta_{\text{min}} \geq 1\). The following problem is the \textbf{optimal code size problem}.

\[
\max_{\mathcal{V} \subseteq S_n} \#\mathcal{V}
\]

s.t. \(\text{dist}(g, h) \geq \delta_{\text{min}}\) for all \(g, h \in \mathcal{V}\) where \(g \neq h\),

\(\#\mathcal{V}\) denotes the cardinality of the set \(\mathcal{V}\).

Denote \(\mu(n, \delta_{\text{min}})\) to be the maximal cardinality achieved by \((n, \delta_{\text{min}})\) codes, i.e., \(\mu(n, \delta_{\text{min}})\) equals the optimal value of the above problem.

The image of \(i\) by \(g\) is denoted \(g(i)\). The inverse of \(g\) is denoted \(g^{-1}\). The product of permutations \(g\) and \(h\) is...
denoted \( gh \), whereby \( (gh)(i) = g(h(i)) \). Most literature (e.g., Tarannen [11]) consider the Hamming metric
\[
dist(g, h) \triangleq \# \{ 1 \leq x \leq n : (g^{-1}h)(i) \neq i \},
\]
i.e., the Hamming distance \( \dist(g, h) \) equals the number of moved points of \( g^{-1}h \). For the direct product group \( S_n \times S_n \), define its action on \( S_n \), as \((\alpha, \beta) \cdot g \triangleq \alpha g \beta^{-1} \), where \((\alpha, \beta) \in S_n \times S_n \) and \( g \in S_n \). For any subgroup \( G \) of \( S_n \), a metric dist(.) on \( S_n \) is \( G \)-invariant if for any \( g, h \in S_n \), we have
\[
dist(g, h) = \dist((\alpha, \beta) \cdot g, (\alpha, \beta) \cdot h)
\]
for all \((\alpha, \beta) \in G \). The Hamming metric (4) can be verified to be \((S_n \times S_n)\)-invariant.

The length of a permutation \( g \), denoted \( \text{length}(g) \), equals the minimum integer \( r \) satisfying \( g = \alpha_1 \alpha_2 \cdots \alpha_r \), whereby \( \alpha_i \) are adjacent transpositions in \( S_n \). For rank modulation [5], we consider the Kendall tau metric, given as
\[
dist(g, h) \triangleq \text{length}(g^{-1}h).
\]
There exists a unique element \( w_0 \), termed the longest element, that satisfies \( \text{length}(w_0) = n(n-1)/2 \). Then \( w_0 \) is an involution, i.e., \( w_0^{-1} = w_0 \), and \( \text{dist}(g, h) = \text{dist}(gw_0, hw_0) \), see [15].

Example 1. Consider \( S_3 \) with elements \( e, (12), (23), (13), (123) \), and the Hamming metric. The minimum distance between any two non-equal permutations is 2. For \( \delta_{nm} = 1 \) and 2, the minimum distance satisfies \( \text{length}(w_0) = n(n-1)/2 \). For \( \delta_{nm} = 3 \) the code \( V \) with the optimal size satisfies \( V = \{ e, (123), (123) \} \). Check \( \text{dist}(e, (123)) = \text{dist}(e, (132)) = 3 \), and \( \text{dist}((123), (132)) = \text{dist}(e, (123))^{-1}(132) \) for any \( e, (123), (132) \).

The minimum possible non-zero Kendall tau pairwise distance is 1. For \( \delta_{nm} = 1 \), we have \( \text{dist}(e, (123), (132)) = 3 \), where \( (123) = (123) \). For \( \delta_{nm} = 3 \) the optimal code satisfies \( V = \{ e, (123), (132) \} \), where \( (132) = (123) \).

By Theorem 1, \( A_{S_3,3} = J - I - A_{S_3,2} \). Here \( J \) has all ones.

Example 2. The matrices in \( \mathbb{R}^{S_3 \times S_3} \) corresponding to the conjugacy and length CC (the indexing on \( S_3 \) is done in the same order that appears in Eq. 4, are written as follows. First \( A_{S_3,1} = A_{\Psi_1,1} = I \), where \( I \) is the identity matrix. Next
\[
A_{S_3,2} = \begin{bmatrix}
0 & 1 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 1 & 0
\end{bmatrix}, \quad A_{\Psi_2,2} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 1
\end{bmatrix}.
\]

The focus here is on the length CC, related to the Kendall tau metric. The conjugacy CC (related to the Hamming metric) is actually an association scheme, and is treated in [11]; the recollection is because of connections exploited later.

III. Semidefinite Programming (SDP) BOUNDS

A symmetric matrix \( M \) in \( \mathbb{R}^{S_n \times S_n} \) is positive semidefinite, if all its eigenvalues are non-negative. We now use CC’s to formulate the relaxation of the optimal code size problem. By iii), Theorem 1, a set \( \{ A_1, A_2, \cdots, A_d \} \) of symmetricized adjacency matrices are obtained, whereby \( d \leq d \). For \( A_i \) is not symmetric, then find \( A_i' \) such that \( A_i'^T = A_i' \), and set \( A_{\bar{i}} = A_i + A_i' \). Similarly the symmetricized orbitals \( \Delta_i \) are obtained by setting \( \Delta_{i} = \Delta_i \cup \Delta_i' \) if \( A_i = A_i + A_i' \). Note both \( (g, h) \) and \( (h, g) \) belong to the same \( \Delta_i \), and \( \text{dist}(g, h) = \text{dist}(h, g) \) for any \( (g, h) \in \Delta_i \). The values \( \delta_i \) are called orbit-distances (with respect to a G-invariant metric \( \text{dist}(\cdot) \)). If \( G \) acts transitively on \( S_n \), then by convention \( \Delta_1 = \{ (g, g) : g \in S_n \} \), thus \( \delta_1 \geq 1 \) for all \( j \geq 2 \). The properties of the CC’s can simplify the following optimizations.

Definition 2 (Primal SDP. \( (G, S_n) \) and \( \delta_{min} \)). Let \( G \) be a group which acts on \( S_n \) transitively and \( \text{dist}(\cdot) \) a \( G \)-invariant distance metric on \( S_n \). Let \( \delta_j \) be the orbit-distances w.r.t.
Consider a group which acts on $G$. Let $\tilde{A}_j$ be a corresponding symmetrized adjacency matrix, $J$ is the all-one matrix, and $\text{Tr}$ is the trace function.

**Proposition 1.** Let $G$ be a group which acts on $S_n$ transitively and $\text{dist}()$ a $G$-invariant distance metric on $S_n$. Let $\delta_{\text{min}} \geq 1$. Then, the optimal objective value of (2) upper bounds the optimal objective value of (7) for $\text{dist}(\cdot)$ and $\delta_{\text{min}}$.

The SDP (4) is a relaxation of the optimal code size problem (1), see appendix for proof. The optimal value of the SDP (4) is at most $\#S_n$, as for any feasible $M$, we have $\text{Tr}(JM) \leq \text{Tr}(J) = \#S_n$. Software like SeDuMi [17] can solve SDP’s.

**Example 3.** Consider $G = S_3 \times \Psi_3$, whereby the Kendall tau metric is $G$-invariant. Let $\Delta_1$ to $\Delta_4$ correspond to $A_{\Psi_3,n}$ to $A_{\Psi_3,4}$ (all symmetric). Using SeDuMi we solve for $\delta_{\text{min}} = 1, 2$ and 3, and get the optimal solutions

$$\frac{1}{7} \cdot J, \quad \frac{1}{6} \cdot (A_{\Psi_3,n} + A_{\Psi_3,3}), \quad \frac{1}{5} \cdot (A_{\Psi_3,n} + A_{\Psi_3,4}),$$

which correspond to optimal objective values 6, 3 and 2.

We need to work with the dual to (2).

**Definition 3 (Dual problem, $(G, S_n)$ and $\delta_{\text{min}}$).** Let $G$ be a group which acts on $S_n$ transitively and $\text{dist}(\cdot)$ a $G$-invariant distance metric on $S_n$. Let $\delta_j$ be the orbit-distances w.r.t. $(G, S_n)$ and $\text{dist}(\cdot)$. Let $\tilde{A}_j$ be a corresponding symmetrized adjacency matrix to $(G, S_n)$. Let $\delta_{\text{min}} \geq 1$. Define the following

$$\min_{(b_1, b_2, \ldots, b_d) \in \mathbb{R}^d} b_1$$

s.t. $b_j \leq 0$ for $2 \leq j \leq d$ with $\delta_j \geq \delta_{\text{min}},$

$$\sum_{j=1}^{d} b_j \tilde{A}_j - J \text{ is positive semidefinite},$$

**IV. LENGTH CC: LINEAR PROGRAMMING (LP) BOUNDS**

Using “duality” we consider the feasible solutions $b$ to (5) (for some $G$-invariant $\text{dist}(\cdot)$ and $\delta_{\text{min}} \geq 1$) that furnish upper estimates $b_1$ to $\mu(n, \delta_{\text{min}})$, see (6) and Proposition 1. While “duality” ideas are not new, the novelty here is to “guess a good subset” of feasible solutions (in the dual program) described by a manageable number of linear equations, and use an LP to optimize over them. For a CC $(G, S_n)$, a feasible solution $b$ corresponds to a positive semidefinite matrix in the following set

$$A_{\tilde{G},S_n} \triangleq \left\{ \sum_{i=1}^{d} b_j \tilde{A}_j : b_j \in \mathbb{R}, \text{ for all } 1 \leq j \leq d \right\}.$$
We claim that the set $A_{S_n} \times \Psi_n \subseteq A_{S_n} \times \Psi_n$, seen by showing each $A_i$ to lie in $A_{S_n} \times \Psi_n$. Observe that $S_n \times \Psi_n$ is a subgroup of $S_n \times S_n$, hence the orbitals of the length $CC$, lie within those of the conjugacy $CC$. In other words, there exists index subsets $\mathcal{T}_{S_n} \subseteq \{1, 2, \ldots, d_{\Psi_n}\}$ such that $A_{S_n} = \bigcup_{i \in \mathcal{T}_{S_n}} A_{\Psi_n,i}$ hold (for all $i$). The claim $A_{S_n} \subseteq A_{S_n} \times \Psi_n$ follows if $A_{S_n}$ is a symmetric matrix, see property i) of the following theorem from [11].

**Theorem 2** (c.f. [11]). Let $(S_n \times S_n, S_n)$ denote the conjugacy $CC$, where $(S_n \times S_n, S_n) = \{A_{S_n,i} : 1 \leq i \leq d_{S_n}\}$, and $A_{S_n,1} = I$. Then all of the following hold for $A_{S_n,i}$:

i) symmetry, i.e., $A_{S_n,i}^T = A_{S_n,i}$ (or $A_{\Psi_n,i} = A_{\Psi_n,i}$).

ii) commutativity, i.e., $A_{S_n,i} A_{S_n,j} = A_{S_n,j} A_{S_n,i}$ for all $i, j$.

iii) diagonalization by an orthonormal matrix $U$ in $\mathbb{R}^{S_n \times S_n}$, i.e., $U^T A_{S_n} U = \sum_{i=1}^{d_{S_n}} p_{i,j} \cdot I_j$ for some $p_{i,j} \in \mathbb{R}$ and 0-1 diagonal matrix $I_j$,

- $I = A_{S_n,1} = \sum_{j=1}^{d_{S_n}} U_1 U_j^T$, therefore $\sum_{j=1}^{d_{S_n}} I_j = I$.
- $\sum_{i=1}^{d_{S_n}} A_{S_n,i} = J$, so $U^T J U = \sum_{i=1}^{d_{S_n}} c_i \cdot I_j$ where $c_j = \sum_{i=1}^{d_{S_n}} p_{i,j}$. By convention $c_1 = \#(S_n)$ (the only non-zero eigenvalue of $J$) and $c_0 = 0$ for $J \geq 2$.

The numbers $d_{S_n}$ tabulated in Table [11] equal the partition number of $n$, see [11]. Consider a matrix $\sum_{i=1}^{d_{\Psi_n}} b_i A_{\Psi_n,i}$ in $A_{S_n} \times \Psi_n$ such that for some $a \in \mathbb{R}^{S_n}$, can be expressed as $\sum_{i=1}^{d_{\Psi_n}} a_i A_{S_n,i}$. Theorem 2 allows us to further express $\sum_{i=1}^{d_{\Psi_n}} b_i A_{\Psi_n,i} = \sum_{j=1}^{d_{S_n}} z_j \cdot (U_1 U_j^T)$ where $z_j = \sum_{i=1}^{d_{\Psi_n}} p_{i,j} a_i$. Then $\sum_{j=1}^{d_{S_n}} b_i A_{\Psi_n,j} = J$ is positive semidefinite (see [5]) if the linear constraints $\sum_{i=1}^{d_{\Psi_n}} p_{i,j} a_i = c_j$ hold for all $j$, for constants $c_j$ in iii). Intuitively, Theorem 2 is an explicit “diagonalization” of all matrices in the subset $A_{S_n} \times \Psi_n$, and facilitates checking of positive semidefined.

A simple extension of the “diagonalization” idea to the following $\text{larger}$ subset of matrices, works reasonably well. Property ii) of Theorem 2 implies iii), as symmetric matrices that commute share common eigenspaces. As such, we desire a subset $B$ of $A_{S_n} \times \Psi_n$, with the property that any $M \in B$, commutes with any $M' \in A_{S_n} \times \Psi_n$. Thus any two matrices in $B$ commute. Such a subset $B$ may be obtained

$$B = \left\{ \sum_{i=1}^{d_{\Psi_n}} (a_i A_{\Psi_n,i}) + \sum_{i=1}^{d_{S_n}} (a_i A_{S_n,i} W) : a \in \mathbb{R}^{2d_{S_n}} \right\},$$

where $W$ is an orthonormal, 0-1 matrix in $\mathbb{R}^{S_n \times S_n}$, that satisfies $(W)_{x,y} = 1$ if and only if $yw_{i-1}^x = x$ for any $x, y \in S_n$. From [8] we see $B$ contains the set $A_{S_n} \times \Psi_n$, considered in Theorem 2. Also by the previous correspondence between $B_j$ and the orbital $\Delta_j$, one can check (see appendix) $W$ commutes with all of $A_{S_n} \times \Psi_n$ (and each $A_{S_n,i}$). Because the longest element satisfies $w_{i-1}^x = w_0$, thus $W^T = W^{-1} = W$. So $A_{S_n,i} W$ are symmetric, and $B$ is a set of symmetric matrices.

One technical lemma, that connects [8] with the dual problem [5], stands in way of finally describing our LP bound. This lemma involves a special subgroup $\Theta_n$ of $S_n$, where $\Theta_n$ is also involved in a few final definitions. Let $\Theta_n = \{ \alpha \in S_n : (\alpha, \alpha) \cdot w_0 = w_0 \}$, where $(\alpha, \alpha) \cdot w_0$ is computed using the action of $S_n \times S_n$ on $S_n$. Let $A_{\Theta_n, \ell}$ denote the symmetrized adjacency matrices belonging to the CC $(S_n \times \Theta_n, S_n)$, where there are $d_{\Theta_n}$ of them. Note $d_{\Theta_n} \leq d_{\Psi_n}$.

**Lemma 1.** Let $A_{S_n,i}$ and $A_{\Theta_n, \ell}$ be the symmetrized adjacency matrices belonging to the conjugacy $CC$ and $(S_n \times \Theta_n, S_n)$, respectively. Let $W$ be defined as before. For $1 \leq \ell \leq d_{\Theta_n}$ and $1 \leq i \leq d_{S_n}$, there exists 0-1 coefficients $t_{\ell,i}$ that satisfy

$$A_{S_n,i} = \sum_{\ell=1}^{d_{\Theta_n}} t_{\ell,i} A_{\Theta_n, \ell}, \quad A_{S_n,i} W = \sum_{\ell=1}^{d_{\Theta_n}} t_{\ell,i} d_{S_n,i} + A_{\Theta_n, \ell} \cdot t_{\ell,i}. \quad (9)$$

See appendix for the proof of Lemma 1. The coefficients $t_{\ell,i}$ satisfying (9) are used to state the following main theorem. For $\Theta_n \subseteq S_n$, let index subsets $\mathcal{T}_{\Theta_n, \ell}$ satisfy $A_{\Theta_n, \ell} = \sum_{\ell=1}^{d_{\Theta_n}} A_{\Theta_n, \ell}$. Using orbit-distances $\delta_j \text{ w.r.t. } (S_n \times \Psi_n, S_n)$ and the Kendall tau metric dist$(\cdot, \cdot)$, define constants $\gamma_\ell$ that satisfy $\gamma_\ell = \max \{ \delta_j : j \in \mathcal{T}_{\Theta_n, \ell} \}$.

**Theorem 3 (LP Bound on $(S_n \times \Psi_n, S_n)$ and $\delta_{min}$).** Let $W$ be the $01$ orthonormal matrix defined as before.

For $1 \leq i, j \leq d_{S_n}$, let constants $p_{i,j}$ and matrices $U, I_j$ be obtained from Theorem 2. Let matrices $M_{1,j}$ and $M_{2,j}$ satisfy $M_{1,j} = \frac{1}{2}(U_1 U_j^T)(I + W)$ and $M_{2,j} = \frac{1}{2}(U_1 U_j^T)(I - W)$. Then $\ell \leq \delta_{\Theta_n}$, let the constants $\gamma_\ell$ be defined as above. For $1 \leq \ell \leq 2d_{S_n}$, let the coefficients $t_{\ell,i}$ satisfy (9). Let $\alpha^* \in \mathbb{R}^{2d_{S_n}}$ solve the following LP problem

$$\min_{(a_1, a_2, \ldots, a_{2d_{S_n}}) \in \mathbb{R}^{2d_{S_n}}} \sum_{i=1}^{2d_{S_n}} t_{\ell,i} \cdot a_i$$

s.t. $\sum_{i=1}^{2d_{S_n}} t_{\ell,i} \cdot a_i \leq 0$ for $2 \leq \ell \leq \delta_{\Theta_n} \cdot \gamma_\ell \leq \delta_{min}$.

$$\sum_{i=1}^{d_{S_n}} (a_i + a_{d_{S_n} + i}) \cdot p_{i,j} \geq c_j \text{ for } 1 \leq j \leq d_{S_n} \text{ with } M_{1,j} \neq 0$$

$$\sum_{i=1}^{d_{S_n}} (a_i - a_{d_{S_n} + i}) \cdot p_{i,j} \geq c_j \text{ for } 1 \leq j \leq d_{S_n} \text{ with } M_{2,j} \neq 0$$

Let $b_1^\text{LP}$ and $\mu(n, \delta_{min})$ respectively denote the optimal objective values of the dual problem [5] and the optimal code size problem [11], for $G = S_n \times \Psi_n$ and the Kendall tau metric dist$(\cdot, \cdot)$ and $\delta_{min}$. Then we have the following inequalities

$$\mu(n, \delta_{min}) \leq b_1^\text{LP} \leq \sum_{i=1}^{\delta_{min}} t_{\ell,i} \cdot a_i^*.$$
One issue: no known efficient method to compute “max distances” \( \gamma_\ell \), where \( \gamma_\ell = \max \{ \delta_j : j \in I_{\tilde{S}_n, \ell} \} \). If one replaces \( \tilde{S}_n \) by \( S_n \) in the expression for \( \gamma_\ell \), then [18] has closed-forms for \( \gamma_\ell \). Also its is unclear how large the number \( \# \{ 1 \leq \delta_k \leq \tilde{d}_{\tilde{S}_n} : \delta_k \geq \delta_{\min} \} \) of non-positive constraints could be. No rigorous analysis is done here, but see [19] for a characterization of \( \tilde{d}_{\tilde{S}_n} \).

V. CONCLUSION & FUTURE DIRECTIONS

Motivated by recent work on solving SDP’s with algebraic structure, we formulated the optimal code size problem w.r.t. Kendall tau metric as a SDP, and propose using LP to search for solutions. The problem seems difficult, but we report modest improvement over a recent Singleton bound.

The interest is to progress toward (possibly) beating known Hamming bounds, for the cases \( n \geq 6 \) (other than those shown here). We offer some future directions. As previously mentioned, it would be nice to analyze the subsets that should be searched (for \( \delta_{\min} < n \)). Next, one might generalize to larger subsets where a manageable SDP (not a LP as here) is used for searching. Finally, one might seek a similar Fourier-type analysis as [9], using representation-theoretic techniques.

REFERENCES

[1] D. Slepian, “Group codes for the Gaussian channel (Abstr.),” IEEE Trans. Inform. Theory, vol. 14, no. 2, p. 242, 1968.
[2] ——, “Permutation modulation,” Proc. IEEE, vol. 53, no. 3, pp. 228–236, 1965.
[3] T. Mittelholzer and J. Lahtonen, “Group codes generated by finite reflection groups,” IEEE Trans. Inform. Theory, vol. 42, no. 2, pp. 519–528, Mar. 1996.
[4] V. K. Goyal, S. A. Savari, and W. Wang, “On optimal permutation codes,” IEEE Trans. Inform. Theory, vol. 47, no. 7, pp. 2961–2971, 2001.
[5] A. A. Jiang, R. Mateescu, M. Schwartz, and J. Bruck, “Rank modulation for flash memories,” IEEE Trans. Inform. Theory, vol. 55, no. 6, pp. 2659–2673, 2009.
[6] A. Barg and A. Mazumdar, “Codes in permutations and error correction for rank modulation,” IEEE Trans. Inform. Theory, vol. 56, no. 7, pp. 854–858, 2009.
[7] F. Farnoud and O. Milenkovic, “Sorting of permutations by cost-constrained transpositions” IEEE Trans. Inform. Theory, vol. 58, no. 1, pp. 3–23, 2012.
[8] P. Delcourt, An algebraic approach to the association schemes of coding theory, 1st ed. Philips Research Reports, Supplement No. 10, 1973.
[9] M. Navon and A. Samorodnitsky, “Linear programming bounds for codes via a covering argument,” Discrete Computational Geometry, vol. 41, no. 2, pp. 199–207, 2007.
[10] B. H. Margolius, “Permutations with inversions,” J Integer Seq, vol. 4, no. 1–13, 2001.
[11] H. Tarnanen, “Upper bounds on permutation codes via linear programming,” Europ. J. Combinatorics, vol. 20, pp. 101–114, 1999.
[12] C. Bachoc, D. C. Gijswijt, A. Schrijver, and F. Vallentin, Invariant semidefinite programs. Online: http://arxiv.org/abs/1007.2005, 2010.
[13] A. Schrijver, “New code upper bounds from the Terwilliger algebra and SDP,” IEEE Trans. Inform. Theory, vol. 51, no. 8, pp. 2859–2866, 2005.
[14] M. X. Goemans and F. Rendl, “Semidefinite programs and association schemes,” Computing, vol. 63, pp. 331–340, 1999.
[15] J. E. Humphreys, Reflection Groups and Coxeter Groups, 4th ed. Cambridge University Press, 1990.
[16] E. Bannai and T. Ito, Algebraic Combinatorics I, 1st ed. Benjamin/Cummings Publishing Co., 1984.
[17] J. F. Sturm, “Using SeDuMi 1.02, a MATLAB toolbox for opt. over symm. cones,” Optim. Meth. and Soft., vol. 11, no. 1, pp. 625–653, 1999.
[18] M. Geck, S. Kim, and G. Pfeiffer, “Minimal length elements in twisted reflection groups,” Journal of Algebra, vol. 229, no. 2, pp. 570–600, Jul. 2000.
[19] B. Brink and R. B. Howlett, “Normalizers of parabolic subgroups in Coxeter groups,” Inventiones Mathematicae, vol. 136, no. 2, pp. 323–351, Apr. 1999.
APPENDIX

1) SDP relaxation of optimal code size problem

First we prove (4) is a relaxation of (1). In the following for a subset $\mathcal{V}$ of $S_n$, let $\mathcal{V}^2$ denote the product set $\mathcal{V} \times \mathcal{V}$. Let $\mathbb{R}^{S_n}$ denote the set of vectors with real number entries with index set $S_n$.

**Proof of Proposition 7:** Let $\mathcal{V}$ denote a solution of (1), i.e., let $#\mathcal{V} = \mu(n, \delta_{min})$. Identify the subset $\mathcal{V}$ of $S_n$ with an 0-1 vector $b \in \mathbb{R}^{S_n}$, where $b_\gamma = 1$ if and only if $\gamma \in \mathcal{V}$. We construct a matrix $M$ whose objective value in (4) equals $#\mathcal{V}$, i.e., $TrJM = #\mathcal{V}$. Let $M = \frac{1}{\mathcal{V}} bb^T$, i.e., $M = \frac{1}{\mathcal{V}} \mathbb{1} bb^T$, and let $1$ denote the all-ones vector. Observe that $TrJM = Tr(1^TM) = Tr(1^T) = \frac{1}{\mathcal{V}} Tr(\mathcal{V}^2) = #\mathcal{V}$. Next we show that matrix $M$ is a feasible solution to (4).

Because $M = \frac{1}{\mathcal{V}} bb^T$, therefore $M$ is positive semidefinite and $Tr(M) = 1$ is satisfied. Next observe $Tr(AI,M) = #((\Delta) \cap \mathcal{V})$, so $Tr(AI,M) \geq 0$ is satisfied. Now consider any $(x, y) \in S_n$ where $x \neq y$. If $x, y \in \mathcal{V}$ then $dist(x, y) \geq \delta_{min}$. By contraposition, if $dist(x, y) < \delta_{min}$ then $(x, y) \notin \mathcal{V}^2$. Let $(x, y) \notin \Delta$, for some $j \geq 2$, then $\delta_j < \delta_{min}$ also implies $(x, y) \notin \mathcal{V}^2$, which in turn implies $Tr(AI,M) = 0$.

2) Matrix $W$ and set $AS_n \times \Psi \times S_n$

Next we prove the orthonormal, 0-1 matrix $W$ commutes with all matrices in the set $AS_n \times \Psi \times S_n$. Recall $W$ is related to the longest element $w_0$, where for any $x, y \in S_n$, we have $W_{x,y} = 1$ if and only if $y w_0^{-1} x = x$.

**Proof:** It suffices to show that $W$ commutes with any adjacency matrix $A_{\Psi \times S_n}$. For any $A_{\Psi \times S_n}$, observe that $(W^TA_{\Psi \times S_n})_{x,y} = (A_{\Psi \times S_n})_{xw_0^{-1}y, w_0^{-1}y}$. Recall $(A_{\Psi \times S_n})_{x,y} = 1$ if and only if $(x, y) \in \Delta$, whereby $\Delta$ is an orbital of the induced action of $S_n \times \Psi$ on $S_n \times S_n$. Hence $(A_{\Psi \times S_n})_{xw_0^{-1}y, w_0^{-1}y} = (A_{\Psi \times S_n})_{x,y}$ because $(x, y)$ and $(xw_0^{-1}y, w_0^{-1}y)$ both belong to same orbital. Hence for any $x$ we have $A_{\Psi \times S_n}W = WA_{\Psi \times S_n}$, which implies $W$ commutes with all of $AS_n \times \Psi \times S_n$.

3) Technical Lemma 7

To show Lemma 7 we need to first establish a relationship between the adjacency matrices $A_{Z,i}$ of the CC $(S_n \times Z, S_n)$, where $Z$ is a subgroup of $S_n$, with orbits on $S_n$, of a subgroup of $S_n \times S_n$ that is related to $Z$. Recall our definition of the action of any $(\alpha, \beta)$ in $S_n \times S_n$ on any $x \in S_n$, given as $(\alpha, \beta)x = \alpha x \beta^{-1}$. For any subgroup $Z$ of $S_n$, denote the subgroup $((\beta, \beta) : \beta \in Z)$ of $S_n \times S_n$ by $H_Z$. Let $C_{Z,1}, C_{Z,2}, \ldots, C_{Z,r}$ denote the $r$ orbits, obtained from the action of $H_Z$ on $S_n$. Each orbit $C_{Z,i}$ is called a conjugacy class (or the action of $H_Z$ on $S_n$). Let $(\rho(\beta))$ denote the 0-1 matrix where $(\rho(\beta))_{x,y} = 1$ if and only if $y \beta^{-1} = x$ (i.e., by our previous definition, $W = (\rho(w_0))$. We claim a one-to-one correspondence between some conjugacy class $C_{Z,i}$ and some adjacency matrix $A_{Z,i}$ of the CC $(S_n \times Z, S_n)$, given as

$$A_{Z,i} = \sum_{\beta \in C_{Z,i}} \rho(\beta).$$

By this claim the number $r$ of conjugacy classes $C_{Z,i}$, equals the number $d$ of adjacency matrices $A_{Z,i}$. To show (11), consider the following.

First, we establish the one-to-one correspondence. By the definition of the orbital $\Delta_i$, for any $(x, y), (\tilde{x}, \tilde{y}) \in \Delta_i$, there exists some $\alpha \in S_n$ and $\beta \in Z$ such that $\tilde{x} = \alpha x \beta^{-1}$ and $\tilde{y} = \alpha y \beta^{-1}$. Equivalently for any $(x, y), (\tilde{x}, \tilde{y}) \in \Delta_i$, there exists some $\beta \in Z$ that satisfies $\tilde{x}^{-1} \tilde{y} = \beta x^{-1} y^{-1}$, which means that $\tilde{x}^{-1} \tilde{y}$ and $x^{-1}y$ are both in $C_{Z,i}$. Note that $\sum_{\beta \in C_{Z,i}} \rho(\beta)$ is a 0-1 matrix, and by the definition of $\rho(\beta)$ we conclude

$$\left( \sum_{\beta \in C_{Z,i}} \rho(\beta) \right)_{x,y} = 1,$$

if and only if $x^{-1}y \in C_{Z,i}$, if and only if $(x, y) \in \Delta_i$. This establishes (11) by referring to the original definition of $A_{Z,i}$ from $\Delta_i$.

**Proof of Lemma 7:** Denote a set $\{\beta w_0 : \beta \in C_{S_n,i}\}$ of elements in $S_n$ by $P_{S_n,i}$. Hence $P_{S_n,i}$ is obtained using the conjugacy class $C_{S_n,i}$ and the longest element $w_0$. Let $A_{S_n,i}$ be an adjacency matrix of the conjugacy CC, and $W = (\rho(w_0))$. It follows $A_{S_n,i}W = \sum_{\beta \in C_{S_n,i}} \rho(\beta w_0) = \sum_{\beta \in P_{S_n,i}} \rho(\beta)$. Because $\Psi_n$ is a subgroup of $\Theta_n$, so for each conjugacy class $C_{\Theta_n,\ell}$ there exists index sets $I_{\ell} \subset \Psi_n$ satisfying $C_{\Theta_n,\ell} = \cup_{j \in I_{\ell}} C_{\Psi_n, j}$. The sets $I_{\ell} \subset \Psi_n$ partition $\{1, 2, \ldots, d_{\Psi_n}\}$. For $1 \leq i \leq d_{\Psi_n}$, we claim there exists new index sets $\bar{I}_i \subset I_i$ that satisfy

$$C_{S_n,i} = \cup_{\ell \in \bar{I}_i} C_{\Theta_n,\ell}, \quad (12)$$

$$P_{S_n,i} = \cup_{\ell \in \bar{I}_i} C_{\Theta_n,\ell}. \quad (13)$$

If the claim holds, Lemma 7 is easily proved as follows. By the previously established (11), we can write $A_{\Theta_n,\ell} = \sum_{\beta \in C_{\Theta_n,\ell}} \rho(\beta)$, where $A_{\Theta_n,\ell}$ is an adjacency matrix of the CC $(S_n \times \Theta_n, S_n)$. By (11), again, an adjacency matrix $A_{S_n,i}$ of the conjugacy CC satisfies $A_{S_n,i} = \sum_{\beta \in C_{S_n,i}} \rho(\beta)$, so (12) implies $A_{S_n,i} = \sum_{\ell \in \bar{I}_i} A_{\Theta_n,\ell}$. Also because $A_{S_n,i}W = \sum_{\beta \in C_{S_n,i}} \rho(\beta w_0)$, by definition of $P_{S_n,i}$ then (13) implies $A_{S_n,i}W = \sum_{\ell \in \bar{I}_i} A_{\Theta_n,\ell}W$. But because both $A_{S_n,i}$ and $A_{\Theta_n,\ell}$ are symmetric 0-1 matrices, there must exist sets $\bar{I}_i \subset I_i$ to express $A_{S_n,i}$ and $A_{\Theta_n,\ell}$ in terms of symmetrized adjacency matrices $A_{\Theta_n,\ell}$, i.e.

$$A_{S_n,i} = \sum_{\ell \in \bar{I}_i} A_{\Theta_n,\ell}, \quad (14)$$

$$A_{S_n,i}W = \sum_{\ell \in \bar{I}_i} A_{\Theta_n,\ell}W,$$

where $t_{\ell,i}$ are coefficients appearing in the lemma statement.

We end by showing the previous claims. The first identity (12) follows easily from the fact $\Theta_n \subset S_n$. The second identity (13) follows by arguing if $c_{\Theta_n,\ell} \cap \bar{P}_{S_n,i} \neq \emptyset$ then $C_{\Theta_n,\ell} \subset P_{S_n,i}$. Consider some $x w_0 \in C_{\Theta_n,\ell} \cap \bar{P}_{S_n,i}$, where $x \in C_{S_n,i}$. By definition of the conjugacy class $C_{\Theta_n,\ell} = \{(\alpha, \alpha) : w_0 \}$, so $x w_0 = \alpha x w_0 \alpha^{-1} = (\alpha x w_0 \alpha^{-1}) (\alpha w_0 \alpha^{-1}) = (\alpha x w_0 \alpha^{-1}) w_0$, and it follows $(\alpha x w_0 \alpha^{-1}) w_0$ is also in $P_{S_n,i}$ as $\alpha x w_0 \alpha^{-1} \in C_{S_n,i}$. Hence (15) is shown.
4) Main theorem

Finally, the following verifies that our proposed LP bound \([10]\) indeed provides an upper bound to the optimal value of the dual problem \([5]\).

Proof of Theorem \([7]\): For some \(b \in \mathbb{R}^{d_{\Phi}}, z \in \mathbb{R}^{d_{\omega}}\) and \(a \in \mathbb{R}^{d_{\Phi}}\), we have the following chain of equalities

\[
\sum_{j=1}^{d_{\Phi}} b_j \cdot \tilde{A}_{\Phi,j} = \sum_{t=1}^{d_{\omega}} z_t \cdot \tilde{A}_{\omega,t}
\]

\[
\leq \sum_{i=1}^{d_{\Phi}} a_i \cdot A_{s,i} + \sum_{i=1}^{d_{\Phi}} a_{d_{\Phi}+i} \cdot A_{s,i} W_i
\]

where \((a)\) follows by setting \(b_j = z_t\) if \(j \in \tilde{I}_{\omega,t}\), and \((b)\) follows by setting \(z_t = \sum_{i=1}^{d_{\omega}} t_{t,i} \cdot a_i\). The theorem will be proved by showing for any feasible \(a\) in \(\mathbb{R}^{d_{\Phi}}\) to \([10]\), there corresponds some feasible \(b\) in \(\mathbb{R}^{d_{\omega}}\) to \([5]\) by the above relationship \([14]\).

Firstly the objectives of \([5]\) and \([10]\) are equal because \(b_1 = z_1 = \sum_{i=1}^{d_{\Phi}} t_{1,i} a_i\). Let \(a\) satisfy the second constraint of \([10]\) and let \(b\) satisfy \([14]\). By \(z_t = \sum_{i=1}^{d_{\omega}} t_{t,i} a_i\), if \(\gamma_t \geq \delta_{\min}\) we have \(z_t \leq 0\). Because \(\gamma_t = \max\{\delta_j : j \in \tilde{I}_{\omega,t}\}\), then for any \(j \in \tilde{I}_{\omega,t}\) such that \(\delta_j \geq \delta_{\min}\), we must have \(b_j \leq 0\). Finally if \(\tilde{d}_{\Phi} \{\delta_j : j \in \tilde{I}_{\omega,t}\} = \{\delta_2, \delta_3, \ldots, \delta_{\Phi}\}\), implying that \(b_j \leq 0\) for all \(j \geq 2\) whereby \(\delta_j \geq \delta_{\min}\), therefore \(b\) satisfies the non-positive constraint of \([5]\).

Next consider the matrices \(M_{1,j}\) and \(M_{2,j}\) given in the theorem statement. Note \(M_{1,j} + M_{2,j} = U I_j U^T\) and \(M_{1,j} - M_{2,j} = (U I_j U)^T W\). Using \(A_{s,i} = \sum_{j=1}^{d_{\Phi}} p_{i,j} \cdot (U I_j U)^T\) in Theorem \([2]\) we express

\[
A_{s,i} = \sum_{j=1}^{d_{\Phi}} p_{i,j} \cdot M_{1,j} + \sum_{j=1}^{d_{\Phi}} p_{i,j} \cdot M_{2,j},
\]

\[
A_{s,i} W = \sum_{j=1}^{d_{\Phi}} p_{i,j} \cdot M_{1,j} - \sum_{j=1}^{d_{\Phi}} p_{i,j} \cdot M_{2,j},
\]

\[
J = \sum_{j=1}^{d_{\Phi}} c_j \cdot M_{1,j} + \sum_{j=1}^{d_{\Phi}} c_j \cdot M_{2,j},
\]

where we claim (shown below) that the matrices \(M_{1,j}\) and \(M_{2,j}\) are i) all symmetric, and ii) have eigenvalues only 0 or 1, and iii) \(M_{1,j} + M_{2,j} = 0\) and iv) \(\sum_{j=1}^{d_{\Phi}} (M_{1,j} + M_{2,j}) = \sum_{j=1}^{d_{\Phi}} (U I_j U)^T = I\). For any \(a\) satisfying the last two constraints of \([10]\), then \([15]\) implies \(\sum_{i=1}^{d_{\Phi}} a_i \cdot A_{s,i} + \sum_{i=1}^{d_{\Phi}} a_{d_{\Phi}+i} \cdot A_{s,i} W - J\) is positive semidefinite. Then for \(b\) that corresponds by \([14]\) to such an \(a\), we will have \(\sum_{j=1}^{d_{\Phi}} b_j \cdot \tilde{A}_{\Phi,j} - J\) satisfying the positive semidefinite constraint in \([5]\).

To finish the proof we address the above claims i) and ii), whereby iii) and iv) will then follow from similar arguments. Claim i) follows because all matrices \(U I_j U^T\) commute with all matrices \(A_{s,i}\), see Theorem \([2]\). Recall \(W\) commutes with all matrices in \(B\), therefore \(W\) commutes with all matrices \(A_{s,i}\), which implies \(W\) commute with all \(U I_j U^T\). This implies \(M_{1,j}\) and \(M_{2,j}\) are symmetric, since both \(W\) and \(U I_j U^T\) are symmetric.

Claim ii) follows because \(w_0^{-1} = w_0\), and it can be verified that \(W = w_0 I\), which implies that the possible eigenvalues of \(W\) are \(-1\) and \(1\). Thus the possible eigenvalues of matrices \(M_{1,j}\) and \(M_{2,j}\) are 0 or 1.