J Algebr Comb (2010) 31: 143–158
DOI 10.1007/s10801-009-0191-2

Bounds on permutation codes of distance four

P. Dukes · N. Sawchuck

Received: 30 May 2008 / Accepted: 15 June 2009 / Published online: 26 June 2009

Abstract A *permutation code* of length n and distance d is a set Γ of permutations from some fixed set of n symbols such that the Hamming distance between each distinct $x, y \in \Gamma$ is at least d. In this note, we determine some new results on the maximum size of a permutation code with distance equal to 4, the smallest interesting value. The upper bound is improved for almost all n via an optimization problem on Young diagrams. A new recursive construction improves known lower bounds for small values of n.

Keywords Symmetric group · Permutation code · Permutation array · Characters · Young diagram · Linear programming

1 Introduction and summary

Let n be a positive integer. Two permutations $\sigma, \tau \in \mathcal{S}_n$ are at *distance* d if $\sigma\tau^{-1}$ has exactly $n-d$ fixed points. This is the ordinary Hamming distance when σ and τ are written as words in single-line notation. For example, 14325 and 54123 are at distance three.

A *permutation code* of length n and minimum distance d is a subset Γ of $\mathcal{S}_n$ such that the distance between distinct members of Γ is at least d. The investigation of permutation codes began some time ago with the articles [5, 8]. Little further attention was given to this topic until the past decade. Permutation codes have enjoyed a resurgence due to various applications.

This research is supported by NSERC.

P. Dukes (✉) · N. Sawchuck
Department of Mathematics and Statistics, University of Victoria, Victoria, BC V8W 3R4, Canada
e-mail: dukes@uvic.ca

N. Sawchuck
e-mail: sawchuck@uvic.ca

For instance, consider a common electric power line. While the primary function is delivery of electric power, the frequency can be modulated to produce a family of n 'close' frequencies. At the receiver, as the power itself is received, these small variations in frequency can be decoded as symbols. In order for this information transmission to not interfere with power transmission, it is important that the frequency remain as constant as possible. One means to achieve this is to use block coding with length n, and to insist that each codeword uses each of the n symbols exactly once. See [2] for a survey of constructions and applications of permutation codes.

Let $M(n,d)$ denote the maximum size of a permutation code of length n and minimum distance d. The following are well-known elementary consequences of the definitions.

Lemma 1.1

(a) $M(n,2)=n!$,
(b) $M(n,3)=n!/2$,
(c) $M(n,n)=n$,
(d) $M(n,d)\le nM(n-1,d)$,
(e) $M(n,d)\le n!/(d-1)!$.

Part (a) is clear from the definition. For (b), consider the alternating group $\Gamma = A_n$. The quotient of two permutations in A_n is again in A_n, and thus cannot be a single transposition. The minimum distance is, therefore, equal to three. Permutation codes realizing the bound in (c) are equivalent to Latin squares; see [3] for more on Latin squares and permutation codes. To prove (d), take a permutation code Γ of length n and distance d, and suppose without loss of generality that symbol n appears most often in the last position of words in Γ. Then the code Γ', comprised of the first $n-1$ symbols of all words in Γ ending in n, is a permutation code of length $n-1$ and distance d. We have $|\Gamma'|\ge|\Gamma|/n$. Now (e) follows from (d) by a simple induction.

Various recent papers have investigated permutation codes and their variants. We refer the reader to [2, 6] for related algebraic results, to [10] for a nice probabilistic approach, and to [3, 15] for some combinatorial bounds.

Although nearly all detailed investigations of $M(n,d)$ have considered relatively large distance d, we are presently interested in the smallest undecided distance: $d=4$. By Lemma 1.1, part (e), we have as a starting point $M(n,4)\le n!/6$.

The Gilbert-Varshamov and sphere-packing bounds for permutation codes are well known, and generally outperform other bounds for small values of d.

Lemma 1.2 *Let D_k denote the number of derangements of order k. Then*

$$\frac{n!}{\sum_{k=0}^{d-1} D_k\binom{n}{k}} \le M(n,d) \le \frac{n!}{\sum_{k=0}^{\lfloor\frac{d-1}{2}\rfloor} D_k\binom{n}{k}}.$$

Unfortunately, the sphere-packing upper bound for $d=4$ is simply $n!$. Although distance four has not been explicitly considered on its own, the following improvement for $d=4$ was essentially known to Frankl and Deza in early investigations [8]. A proof is provided here for completeness.

Lemma 1.3 $M(n,4) \le (n-1)!$.

Proof Consider for each $\sigma \in \mathcal{S}_n$ the set of all n words $A_\sigma = \{\sigma\} \cup \{(1i)\sigma : 2 \le i \le n\}$. We have $|A_\sigma| = n$ for any σ. Given a permutation code $\Gamma \subset \mathcal{S}_n$ of distance 4, it suffices to show that if $\sigma \ne \tau$ are both in Γ, then $A_\sigma \cap A_\tau = \emptyset$. Assume without loss of generality that the identity $() = 123\cdots n \in A_\sigma \cap A_\tau$. If either σ or τ equals $()$, then their distance is only two, a contradiction. So $\sigma = (1i)$ and $\tau = (1j)$ for some $1 < i < j \le n$. But then $\sigma\tau^{-1} = (1ji)$ and σ, τ are at distance 3, another contradiction. □

Our main result is an improved upper bound on $M(n,4)$ arising from linear programming and a concrete problem on characters of $\mathcal{S}_n$.

Theorem 1.4 *If $k^2 \le n \le k^2 + k - 2$ for some integer $k \ge 2$, then*

$$\frac{n!}{M(n,4)} \ge 1 + \frac{(n+1)n(n-1)}{n(n-1) - (n-k^2)((k+1)^2 - n)((k+2)(k-1) - n)}.$$

The next two sections are devoted to the proof of this result. Specifically, Section 2 introduces various background necessary for the proof, and Section 3 handles the details through a certain optimization problem.

Some interesting special cases of Theorem 1.4 are now given.

Corollary 1.5 *If $n = k^2$, or if $n = (k+2)(k-1)$, where $k \ge 2$ is an integer, then*

$$M(n,4) \le \frac{n!}{n+2}. \tag{1.1}$$

Corollary 1.6 *There exists a positive constant ϵ such that for infinitely many values of n,*

$$(n-1)! - M(n,4) \ge \epsilon\sqrt{n}(n-2)!.$$

Proof In Theorem 1.4, take $n = k^2 + k/2$ for k an even integer. Then

$$\frac{n!}{M(n,4)} \ge 1 + \frac{(n+1)n(n-1)}{n(n-1) - (k/2)(3k/2+1)(k/2-2)}.$$

Multiplying both numerator and denominator of the fraction on the right by $(n-3)!$, and neglecting insignificant terms, we have $(n-1)! - M(n,4) \sim \frac{3}{8}k^3(n-3)!$. □

In investigating linear programming bounds for permutation codes, Tarnanen [13] gives the explicit bound $n!/M(n,4) \ge 12$ for $n \ge 10$. This is one instance of Corollary 1.5 above. Indeed, the method in that article inspired Theorem 1.4, whose proof is given in Section 3.

A similar expression to Theorem 1.4 (though unpleasant) holds as well for $k^2 + k + 1 \le n \le k^2 + 2k$. See the end of Section 3 for details. A table of upper bounds for small values of n is also provided.

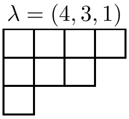
On the other hand, the Gilbert-Varshamov lower bound, specialized to $d = 4$, is

$$M(n,4) \geq \frac{6n!}{2n^3 - 3n^2 + n + 6}. \tag{1.2}$$

It is possible to construct, through recursive methods in [2], permutation codes of minimum distance 4 which come close to or improve (1.2) for small values of n. This is the content of Section 4.

2 Partitions, characters, and LP bounds

For n a positive integer, a *partition* λ of n, denoted $\lambda \vdash n$, is an unordered list of positive integers which sum to n. Equivalently, λ may be written as an ordered t-tuple $(\lambda_1, \lambda_2, \ldots, \lambda_t)$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_t$. A partition is often identified with its *Young diagram*, in which λ_i boxes occupy the ith row, left-justified. For instance, the partition $1 + 3 + 4$ of $n = 8$ is written as the triple $\lambda = (4, 3, 1)$ and has Young diagram as shown.



The *conjugate* λ^* of a partition λ is the partition whose Young diagram is the transpose of that for λ. Specifically, the ith part of the conjugate is

$$\lambda_i^* = |\{j : \lambda_j \geq i\}|.$$

The conjugate of the partition shown above is $(3, 2, 2, 1)$. The number of ones in λ, which is simply $\lambda_1^* - \lambda_2^*$, is denoted by $\varphi(\lambda)$.

In what follows, we shall use without definition terms such as ‘main diagonal’, ‘outside corner box’, and ‘hook’, which are standard for Young diagrams and Young tableaux. A comprehensive reference is [7]. When convenient, we may use exponential notation $1^{t_1}2^{t_2}\ldots$ to denote a partition with t_1 ones, t_2 twos, etc.

We now summarize some terminology and basic facts on the representation theory of the symmetric group. See [7] for further detail.

A *representation* of a group G is a homomorphism $h : G \to GL(N, \mathbb{C})$. The *character* associated with h is $\chi_h = \mathrm{Trace} \circ h$, mapping G to $\mathbb{C}$. Its *dimension* (or *degree*) is $N = \chi_h(1_G)$. The dimension is also abbreviated $\dim \chi_h$. Since h is a homomorphism, a character χ_h is clearly constant on any conjugacy class of G.

Both the irreducible representations of the symmetric group $\mathcal{S}_n$ and the conjugacy classes of $\mathcal{S}_n$ are in one-to-one correspondence with the set of all partitions of n. Each irreducible character of $\mathcal{S}_n$ is an integer-valued function on the conjugacy classes of $\mathcal{S}_n$. Here, we represent the character corresponding to partition λ by χ^λ, and the conjugacy class corresponding to μ simply by μ. So the (λ, μ)-entry of the character table of $\mathcal{S}_n$ is $\chi^\lambda(\mu)$.

We also have $\chi^\lambda(1^n) = \dim \chi^\lambda$, and this is often written $\dim \lambda$. The explicit value of $\dim \lambda$ is easily obtained from the hook length formula, [7]. More generally, the so-called *Frobenius character formulas* (see [11] for details) give $\chi^\lambda(1^{n-t}t^1)/\dim\lambda$, for small values of t. These are

$$\frac{\chi^\lambda(1^{n-2}2^1)}{\dim\lambda}=\frac{\sum_i\beta_i(\beta_i+1)-\sum_i\alpha_i(\alpha_i+1)}{n(n-1)},\quad\text{and}$$
$$\frac{\chi^\lambda(1^{n-3}3^1)}{\dim\lambda}=\frac{\sum_i\alpha_i(\alpha_i+1)(2\alpha_i+1)+\sum_i\beta_i(\beta_i+1)(2\beta_i+1)-3n(n-1)}{2n(n-1)(n-2)},$$

where $\lambda\vdash n$ has

- exactly s boxes on its main diagonal,
- $\alpha_1>\alpha_2>\cdots>\alpha_s$ boxes below the diagonal in columns 1 through s, and
- $\beta_1>\beta_2>\cdots>\beta_s$ boxes right of the diagonal in rows 1 through s.

A (*symmetric*) *k-class association scheme* on a set X consists of $k+1$ nonempty symmetric binary relations $R_0,\dots,R_k$ which partition $X\times X$, where R_0 is the identity relation $\{(x,x):x\in X\}$, and such that for any $x,y\in X$ with $(x,y)\in R_h$, the number of $z\in X$ such that $(x,z)\in R_i$ and $(y,z)\in R_j$ depends only on the indices h,i,j. For $J\subset\{1,\dots,k\}$, a *J-clique* is a subset W of X such that for any distinct $w_1,w_2\in W$, $(w_1,w_2)\in R_j$ for some $j\in J$.

The symmetric group defines an association scheme, called the *conjugacy scheme*, where $X=\mathcal{S}_n$ are the points, relations are indexed by partitions $\lambda\vdash n$, and $(\sigma,\tau)\in R_\mu$ if and only if $\sigma\tau^{-1}$ belongs to conjugacy class μ. Of course, σ and τ are at distance d if and only if $(\sigma,\tau)\in R_\mu$, where $\varphi(\mu)=n-d$.

This motivates a generalized form of permutation codes. We say $\Gamma\subset\mathcal{S}_n$ is a D-permutation code if any two distinct permutations in Γ are at some distance in D. The maximum size of such a set Γ is denoted $M(n,D)$. It is an easy observation that

$$M(n,D)M(n,D^c)\le n!,\tag{2.1}$$

where $D^c=\{1,\dots,n\}\setminus D$. It is not hard to see that (2.1) implies Lemma 1.3; see Section 5 for more details.

Tarnanen [13] considered the following specialization of Delsarte's inequality (see [4]) to cliques in the conjugacy scheme. Notation has been changed slightly for convenience.

Theorem 2.1 ([13]) *Subject to $a_\mu\ge 0$ for all $\mu\vdash n$, $a_{(1,\dots,1)}=1$, $a_\mu=0$ for all $\mu\vdash n$ having $n-\varphi(\mu)\notin D$, and*

$$\sum_{\mu\vdash n}a_\mu\chi^\lambda(\mu)\ge 0$$

for all $\lambda\vdash n$, put

$$M_{\mathrm{LP}}(n,D)=\max\sum_{\mu\vdash n}a_\mu.$$

Then

$$M(n,D)\le M_{\mathrm{LP}}(n,D).\tag{2.2}$$

Delsarte in fact proved that (2.1) holds analogously for LP bounds. In our context,

$$M_{\mathrm{LP}}(n, D)M_{\mathrm{LP}}(n, D^c) \le n!. \tag{2.3}$$

The preceding algebraic tools set the stage for a proof of Theorem 1.4 and some additional observations in the next section.

3 Proof of the upper bound on $M(n, 4)$

3.1 Outline of the proof

By (2.2) and (2.3), it follows that

$$M(n, 4) = M(n, \{4, \dots, n\}) \le \frac{n!}{M_{\mathrm{LP}}(n, \{2, 3\})}. \tag{3.1}$$

In this way, our results are obtained from lower bounds on $M_{\mathrm{LP}}(n, \{2, 3\})$. The convenient choice of $D = \{2, 3\}$ above offers a nice simplification of Theorem 2.1.

Proposition 3.1 *Let $n \ge 4$. Then $M_{\mathrm{LP}}(n, \{2, 3\})$ is given by*

$$\begin{aligned} \max\{1 + a + b : a, b \ge 0 \text{ and } \\ \forall \lambda \vdash n,\ \dim(\chi^\lambda) + a\chi^\lambda(1^{n-2}2) + b\chi^\lambda(1^{n-3}3) \ge 0\}. \end{aligned} \tag{3.2}$$

Each feasible point for the LP in Proposition 3.1 leads to a lower bound on M_{LP}. For the proof of Theorem 1.4, we will consider feasible points (a, b) which are multiples of $(3, n - 2)$. As we shall see in Sect. 3.2, such feasible points lead to the optimum LP value for n in the relevant range.

A necessary and sufficient condition for the point $(a, b) = (3C, (n - 2)C)$ to be feasible is that, for all $\lambda \vdash n$,

$$\dim(\lambda) + 3C\chi^\lambda(1^{n-2}2) + (n - 2)C\chi^\lambda(1^{n-3}3) \ge 0,$$

or equivalently, using the Frobenius character formulas,

$$\frac{\sum_{i=1}^{s}[\alpha_i(\alpha_i + 1)(2\alpha_i - 5) + \beta_i(\beta_i + 1)(2\beta_i + 7)]}{n(n - 1)} \ge 3 - \frac{2}{C}. \tag{3.3}$$

Therefore, we obtain the largest possible C by minimizing, over all $\lambda \vdash n$, the numerator on the left of (3.3). (Recall that n is fixed.)

Define the polynomials $f(x) = x(x + 1)(2x - 5)$ and $g(x) = x(x + 1)(2x + 7)$, and put

$$\Phi(\lambda) = \sum_{i=1}^{s}[f(\alpha_i) + g(\beta_i)], \tag{3.4}$$

where as before λ has $\alpha_1 > \cdots > \alpha_s$ boxes below the diagonal and $\beta_1 > \cdots > \beta_s$ boxes right of the diagonal. Again, s is the number of boxes on the main diagonal.

Proposition 3.2 *Let $n = k^2 + l$, $0 \le l \le k-2$. With Φ as defined in* (3.4),

$$\Phi(\lambda) \ge n(n-1) + 2l(k-l-2)(2k-l+1) \text{ for all } \lambda \vdash n,$$

with equality if $\lambda = l^1 k^k$ or $\lambda = (l+1)^1(k-1)^{k+1}$.

We defer a proof of Proposition 3.2 until Sect. 3.3. From this, (3.1) and (3.3), it follows that

$$\begin{aligned}\frac{n!}{M(n,4)} &\ge M_{\mathrm{LP}}(n,\{2,3\})\\ &\ge 1 + (n+1)C\\ &\ge 1 + \frac{2(n+1)n(n-1)}{3n(n-1) - \Phi(l^1k^k)}\\ &= 1 + \frac{(n+1)n(n-1)}{n(n-1) - l(2k-l+1)(k-l-2)}.\end{aligned}$$

This completes the proof of Theorem 1.4.

3.2 Optimality of $(3C, (n-2)C)$

We saw in the last section that each feasible point (a, b) for the LP in (3.2) results in a lower bound on $M_{\mathrm{LP}}(n, \{2, 3\})$. Therefore, it is important to note that our choice $a = 3C$, $b = (n-2)C$ is indeed best-possible. Since $M_{\mathrm{LP}}(n, \{2, 3\}) = 1 + a + b$, the optimality will follow provided we show that

- $(3C, (n-2)C)$ lies on at least two constraints of the form $t_i a + u_i b \le 1$, $i = 1, 2$, such that $t_1 < u_1$ while $t_2 > u_2$.

It is a routine calculation that

$$\sum_{i=0}^{k}[f(i) + g(i)] = k(k+1)^2(k+2). \tag{3.5}$$

Actually, from this identity, it is straightforward to prove the second statement of Proposition 3.2. In particular, for $\lambda^{(1)} = l^1k^k$ one has sequences

$$\begin{aligned}&\alpha_i : k, k-1, \dots, k-l+1, k-l-1, \dots, 1, 0,\\ \text{and}\quad &\beta_i : k-1, k-2, \dots, 1, 0;\end{aligned}$$

similarly for $\lambda^{(2)} = (l+1)^1(k-1)^{k+1}$,

$$\begin{aligned}&\alpha_i : k+1, k, \dots, k-l+1, k-l-1, \dots, 3, 2,\\ \text{and}\quad &\beta_i : k-2, k-3, \dots, 1, 0.\end{aligned}$$

After simplifying Φ, we see that $(3C, (n-2)C)$ lies on the two constraints in (3.2) which correspond to these two $\lambda^{(i)}$.

Now, the constraints are of the form $t_i a + u_i b \leq 1$, where

$$t_i = -\frac{\chi^{\lambda^{(i)}}(1^{n-2}2)}{\dim(\lambda^{(i)})} \quad \text{and} \quad u_i = -\frac{\chi^{\lambda^{(i)}}(1^{n-3}3)}{\dim(\lambda^{(i)})}.$$

Using the Frobenius character formulas and (3.5),

$$u_1 - t_1 = \frac{k(k+1)(k-l)(k-l-1)}{n(n-1)(n-2)} > 0$$

and

$$t_2 - u_2 = \frac{k(k+1)(k^2+k+2+l+2kl-l^2)}{n(n-1)(n-2)} > 0.$$

It follows that, if C is chosen according to Proposition 3.2, $(3C, (n-2)C)$ is indeed the optimum for (3.2).

3.3 Minimizing Φ over Young diagrams

The purpose here is to prove Proposition 3.2 using some neat local changes to Young diagrams. Recall $\Phi(\lambda) = \sum_{i=1}^{s} [f(\alpha_i) + g(\beta_i)]$, where f and g are the cubic polynomials defined in Sect. 3.1. The following simple properties of f and g are easily verified.

Lemma 3.3

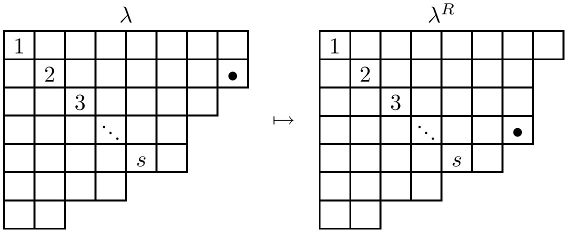
$$\boxed{\begin{array}{rl}
f(y) \geq f(x) & \textit{for integers } \ 0 < x < y \\
g(y) > g(x) & \textit{for } \ 0 \leq x < y \\
g(x) \geq f(x) & \textit{for all integers } \ x \\
f(x) + f(y) \geq f(x+1) + f(y-1) & \textit{for integers } \ 0 \leq x < y \\
g(x) + g(y) \geq g(x+1) + g(y-1) & \textit{for integers } \ 0 \leq x < y \\
f(x) + g(y) \geq f(y) + g(x) & \textit{for integers } \ 0 \leq x \leq y \\
f(y) + g(x) \geq f(y-1) + g(x+1) & \textit{for integers } \ 0 \leq x \leq y-3
\end{array}}$$

Using Lemma 3.3, we show that diagram operations 1 through 5 below do not increase Φ. The diagrams for which these operations cannot be performed are then characterized and compared relative to Φ.

Operation 1 *Flattening right of the diagonal.* Here, assume that there is an integer $t \geq s$ with $\lambda_j = t$ and $\lambda_i \geq t+2$ for some $i < j \leq s$. Without loss of generality, suppose i is the greatest such index and j is the least such index. So the box in position (i, λ_i) is an outside corner. By choice of j, moving this box into position $(j, \lambda_j + 1)$ yields a valid diagram λ^R with

$$\Phi(\lambda) - \Phi(\lambda^R) = g(\beta_i) + g(\beta_j) - g(\beta_i - 1) - g(\beta_j + 1) \geq 0.$$

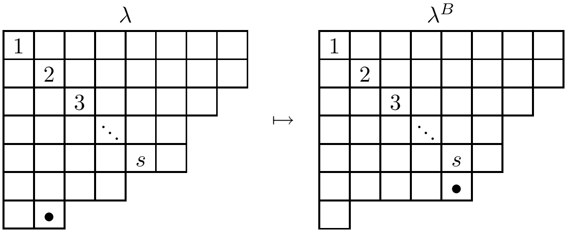
We illustrate this diagram operation with an example. The diagonal cells are numbered and the box which moves is indicated.



Operation 2 *Flattening below the diagonal.* Assume that there is an integer t with $\lambda_j^* = t$ and $\lambda_i^* \geq t + 2$ for some $i < j \leq s$. Without loss of generality, suppose i is the greatest such index and j is the least such index. Moving the lowermost box of column i into column j yields a valid diagram λ^B with

$$\Phi(\lambda) - \Phi(\lambda^B) = f(\alpha_i) + f(\alpha_j) - f(\alpha_i - 1) - f(\alpha_j + 1) \geq 0.$$

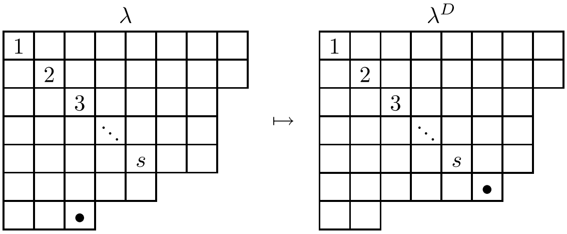
Operation 2 is illustrated below.



Operation 3 *Adding to the diagonal.* Assume now that both $\lambda_s, \lambda_s^* > s$, so that both α_s and β_s are positive, but that there is an outside corner box in some position other than $(s, s + 1)$ or $(s + 1, s)$. Moving any such outside corner box (say from row i) into position $(s + 1, s + 1)$ creates a new valid diagram λ^D. As expected,

$$\Phi(\lambda) - \Phi(\lambda^D) = g(\beta_i) - g(\beta_i - 1) \geq 0.$$

A similar inequality holds if an outside corner box is selected below the diagonal. This operation is illustrated below.



After some combination of these three operations, we are left with a diagram approximating a rectangle. Define a *near-rectangle* to be a Young diagram obtained by removing a (possibly empty) hook from a rectangle. Then Φ attains its minimum on near-rectangles, i.e., on partitions of the form $\lambda = (k+1, \dots, k+1, k, \dots, k, r)$, where the k and $r < k$ terms may not be present. Near-rectangles enjoy the property that the sequences (α_i) and (β_i) are each an interval of consecutive integers, possibly minus a single integer.

We now consider two further operations which do not increase Φ.

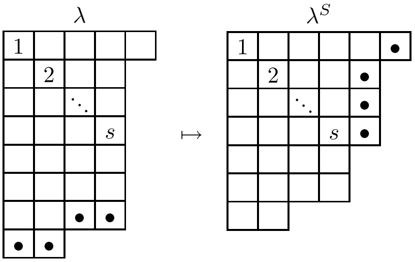
Operation 4 *Transposing.* Suppose that $\lambda_1 > \lambda_1^*$ for a near-rectangle λ. Then $\beta_i \geq \alpha_i$ for all $i = 1, \dots, s$. Taking the conjugate of λ simply interchanges α_i with β_i. Estimating term-by-term and using Lemma 3.3, we have

$$\Phi(\lambda) - \Phi(\lambda^*) = \sum_{i=1}^{s} [f(\alpha_i) + g(\beta_i) - f(\beta_i) - g(\alpha_i)] \geq 0.$$

For near-rectangles, this leads to the simplifying assumption that the bounding rectangle be square or tall (not wide).

Operation 5 *Squaring.* Suppose λ is a near-rectangle with $\alpha_i \geq \beta_i + 3$ for all $i = 1, \dots, s$. Move all lowermost boxes in columns 1 through s onto the right of rows 1 through s, one per row. Call the resulting diagram λ^S and refer to the illustration below. Estimating term-by-term,

$$\Phi(\lambda) - \Phi(\lambda^S) = \sum_{i=1}^{s} [f(\alpha_i) - f(\alpha_i - 1) + g(\beta_i) - g(\beta_i + 1)] \geq 0.$$



After possibly several iterations of these five operations, we arrive at Young diagrams of the following structure.

Lemma 3.4 *Function Φ attains its minimum at a near-rectangle λ with $\beta_i \leq \alpha_i \leq \beta_i + 3$ for each $i = 1, \dots, s$.*

In other words, the Young diagram of λ is obtained from an $(m+j) \times m$ rectangle, for some $j \in \{0, 1, 2, 3\}$, by adding an extra column of $Y \geq 0$ boxes appended on the right, and an extra row of $X \geq 0$ boxes appended below. If there are n boxes in total, we have $X + Y = n - m(m+j)$.

We now invoke the assumption of Theorem 1.4 that $n = k^2 + l$ for some integers $k \geq 2$ and $0 \leq l \leq k - 2$.

Case 1: $j = 0$. This forces $m = k$ and $X + Y = l$. Using (3.5), we have

$$\Phi(\lambda) = k(k+1)^2(k+2) - f(k-X) - g(k-Y).$$

Since $Y = l - X$, further calculation shows that $\Phi(\lambda)$ has a negative coefficient $6(n - (k+1)^2)$ of X^2. Therefore, Φ is minimized at either $X = 0$ or $X = l$. By Lemma 3.3, it is easily seen that $(X, Y) = (l, 0)$ minimizes Φ in this case. This results in partition $\lambda = l^1 k^k$.

Case 2: $j = 2$. This leads to $m = k - 1$ and $X + Y = l + 1$. Since $X \leq k - 1$ in this case, it follows that $Y \geq l + 2 - k$. Calculating with (3.5), one has

$$\Phi(\lambda) = k(k+1)^2(k+2) - 6k(k+2) - f(k+1-X) - g(k-1-Y).$$

Working as in Case 1, this function is minimized at one or both endpoints. Since $l \leq k - 2$, the endpoints are $(0, l+1)$, $(l+1, 0)$, with the minimizing partition being $\lambda = (l+1)^1(k-1)^{k+1}$.

Case 3: $j = 1$. Since $l \leq k - 2$, we have $m = k - 1$ and $X + Y = k + l$. As before, Φ can only be minimized at either endpoint $(X, Y) = (k-1, l+1)$ or (l, k). The two relevant values agree with that calculated in Case 1. We have

$$\Phi((k-1)^{k-l}k^{l+1}) = \Phi(l^1 k^k) = n(n-1) + 2l(k-l-2)(2k-l+1).$$

Case 4: $j = 3$. A $(k+2) \times (k-1)$ rectangle minimizes Φ for $l = k - 2$, in addition to previous shapes. We have

$$\Phi((k-1)^{k+2}) = n(n-1).$$

3.4 Other values of n and d

It should be stressed that the case $n = k^2 + l$, $k + 1 \leq l \leq 2k$ is not qualitatively different. Using feasible points of the form $(3C/2, (n-2)C)$, a slightly modified function Φ, and working mostly with the $(m+1) \times m$ and $(m+3) \times m$ rectangles, one arrives at a similar bound. For $n = k^2 + k - 1$ and $k^2 + k$, the bounds obtained from this method are worse, the latter bound agreeing with Lemma 1.3. We omit details, but state the results for the interested reader.

Theorem 3.5 *If $n = k^2 + l$, with $k + 1 \leq l \leq 2k$ for some integer $k \geq 2$, then*

$$\frac{n!}{M(n,4)} \geq 1 + \frac{(n+1)n(2n-1)}{2n(n-1) - ((k+1)^2 - n)(l(5-2l) + k(4l-1))}.$$

For $l = k - 1$, $n!/M(n,4) \geq n + 1$ and for $l = k$, $n!/M(n,4) \geq n$.

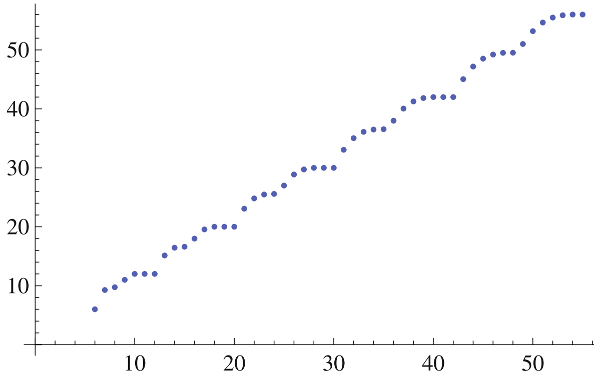


Fig. 1 Lower bounds on $M_{\mathrm{LP}}(n, \{2, 3\})$ versus n

A plot of the small LP bounds we obtain by this method is given in Figure 1.

To guess feasible points of the form $(3C, (n-2)C)$, we initially computed explicit values of M_{LP} for small values of n. Using Operations 1 through 5 above, the number of constraints is drastically reduced, and the LP becomes computationally efficient.

A preliminary look at the case $d = 5$ shows that near-rectangles are not necessarily the optimizing diagrams. This is essentially due to the 'nonlinear' $\chi^{\lambda}(1^{n-4}2^2)$ term. Therefore, the technique may fail to have much success for upper bounds on permutation codes of higher distances.

4 Constructions for small *n*

A permutation code of length n and distance d is here denoted by $\mathrm{PC}(n, d)$. It is well-known that $M(n, 4) = n!/6$ for $n = 4, 5, 6$. In fact, more is true. We will later make use of the fact that for these values of n, $\mathcal{S}_n$ can be partitioned into six disjoint $\mathrm{PC}(n, 4)$. See [2] and earlier references for details of the constructions.

Our starting point is a construction found in [2]. This is actually analogous to the 'partitioning construction' [14] for constant weight binary codes.

Lemma 4.1 ([2]) *Suppose there are disjoint* $PC(n_0, 4)$ *of sizes* $s_1, \ldots, s_p$ *and disjoint* $PC(n_1, 4)$ *of sizes* $t_1, \ldots, t_p$. *Suppose further that there is a constant weight binary code of length* $n = n_0 + n_1$, *weight* n_1, *and distance* 4 *of size* c. *Then there is a* $PC(n, 4)$ *of size* $c\sum_{j=1}^{p} s_j t_j$.

The idea here is to consider each word $\mathbf{x}$ of the constant weight code in turn. On the positions in which $\mathbf{x}$ has a 0, place any word from the ith $\mathrm{PC}(n_0, 4)$. Likewise, on the positions in which $\mathbf{x}$ has a 1, place any word (symbols shifted to $\{n_0+1, \ldots, n_0+n_1\}$) from the ith $\mathrm{PC}(n_1, 4)$. The distance between permutations resulting from different constant weight binary words $\mathbf{x} \neq \mathbf{y}$ is at least 4 by virtue of the constant weight code. The distance between permutations arising from the same constant weight binary word $\mathbf{x}$ is at least 4, either because a given $\mathrm{PC}(n_i, 4)$ alone carries the distance, or because for each $j = 0, 1$, the $\mathrm{PC}(n_j, 4)$ are disjoint.

The difficult ingredient in Lemma 4.1 is a good set of disjoint $\mathrm{PC}(n, 4)$. In this construction, the set of positions in which 'high' symbols are placed is different for distinct constant weight codewords. So it follows that the construction results in disjoint PC$(n,4)$ when the ingredient constant weight codes are disjoint. For this purpose, we cite an easy but helpful fact from coding theory. The proof is rather well known, but provided here for completeness.

Lemma 4.2 ([9]) *Let $\mathbb{U}_w^n$ denote the set of all binary words of length n and weight w. Then $\mathbb{U}_w^n$ partitions into n codes of minimum distance* 4.

Proof Define a mapping $T:\mathbb{U}_w^n\to\mathbb{Z}/(n)$ as follows. For a binary word $\mathbf{x}=x_0x_1\cdots x_{n-1}$ of weight w, put $T(\mathbf{x})=\sum_i ix_i \pmod n$. Consider $C_j=T^{-1}(j)$, where $j\in\mathbb{Z}/(n)$. If $\mathbf{x},\mathbf{y}\in C_j$ were to disagree in exactly two positions, say positions h and k, then

$$0=j-j=T(\mathbf{x})-T(\mathbf{y})=h(x_h-y_h)+k(x_k-y_k)=\pm(h-k).$$

It follows that each C_j has minimum distance at least 4. □

In Lemma 4.1, suppose our disjoint PC$(n_j,4)$ are $\Gamma_1^{(j)},\dots,\Gamma_p^{(j)}$, $j=0,1$. For any i, we could equally as well have paired up $\Gamma_1^{(0)}$ with $\Gamma_i^{(1)}$, $\Gamma_2^{(0)}$ with $\Gamma_{i+1}^{(1)}$, and so on. This results in p disjoint PC$(n,4)$ of sizes $s_1t_i+s_2t_{i+1}+\cdots+s_pt_{i-1}$, where indices wrap (mod p) and $i=1,\dots,p$.

These are essentially our new observations which make the partitioning construction for permutation codes now recursive.

Theorem 4.3 *Suppose there exist disjoint PC$(n_0,4)$ of sizes $s_1,\dots,s_p$ and disjoint PC$(n_1,4)$ of sizes $t_1,\dots,t_p$. Suppose further that there are disjoint constant weight binary codes of length $n=n_0+n_1$, weight n_1, and distance* 4 *of sizes $c_1,\dots,c_q$. Put $u_i=s_1t_i+s_2t_{i+1}+\cdots+s_pt_{i-1}$, $i=1,\dots,p$, where indices are read* (mod p). *Then there are disjoint PC$(n,4)$ of sizes u_ic_j, $i=1,\dots,p$, $j=1,\dots,q$.*

In either the binary constant weight setting or the permutation setting, consider any set of disjoint codes of sizes $a_1,\dots,a_N$. By including singletons, these codes may be assumed to partition the relevant space of words. Following [1], define the *norm* of this partition to be $\sum_{i=1}^N a_i^2$. Let $A^2(n,4,w)$ denote the maximum norm of a partition into constant weight binary codes of length n, weight w, and distance 4. Let $M^2(n,4)$ denote the maximum norm of a partition into PC$(n,4)$. It is clear that, in both cases, the norm is bounded above in terms of the maximum possible code size. This is stated below for permutation codes.

Lemma 4.4

$$\frac{n!}{M(n,4)}\le\frac{(n!)^2}{M^2(n,4)}.$$

Define the *binorm* of $a_1,\dots,a_N$ to be $\sum_{i=1}^N(a_1a_i+a_2a_{i+1}+\cdots+a_Na_{i-1})^2$, where indices are read mod N. Observe that if all $a_i=n!/N$, then the binorm is simply $n!^4/N$. Let $M^4(n,4)$ be the maximum binorm of a partition of $\mathcal{S}_n$ into PC$(n,4)$.

Table 1 Improved lower bounds on $M^2(2n,4)$ and $M(n,4)$ from the partitioning construction

$2n$	$\frac{\binom{2n}{n}^2}{A^2(2n,4,n)} \le$	$\frac{n!^4}{M^4(n,4)} \le$	$\frac{(2n)!^2}{M^2(2n,4)} \le$	$\frac{(2n)!}{M(2n,4)} \le$	$(2n)!/\mathrm{GV}$
10	7.89458, [1]	6, [2]	47.3675	42, [2]	286
12	8.54186, [1]	6, [2]	51.251	42, [2]	507
14	12.8985, [1]	15, [2]	193.48	158.4, [2]	820
16	15.9995, [9]	15, [2]	239.99	239.99	1241
20	20	47.9975	959.95	959.95	2471
24	24	59.692	1432.6	1432.6	4325
28	28	209.85	5875.7	5875.7	6931
32	32	240	7680	7680	10417

Specializing to $n_0 = n_1 = n$ in Theorem 4.3, and interpreting in terms of norms and binorms, we arrive at the following inequality.

Corollary 4.5 $M^2(2n,4) \ge A^2(2n,4,n)M^4(n,4)$.

Proof Suppose we have partitions achieving $A^2(2n,4,n)$ and $M^4(n,4)$. With notation as in Theorem 4.3, there exists a partition of $\mathcal{S}_{2n}$ into PC$(2n,4)$ of sizes $(s_1s_i + s_2s_{i+1} + \cdots + s_ps_{i-1})c_j$ for all relevant i, j. Computing the norm of this partition,

$$\begin{aligned} M^2(2n,4) &\ge \sum_{i,j} c_j^2(s_1s_i + s_2s_{i+1} + \cdots + s_ps_{i-1})^2 \\ &= A^2(2n,4,n)M^4(n,4). \end{aligned}$$ □

Writing Corollary 4.5 more conveniently, we have

$$\frac{(2n)!^2}{M^2(2n,4)} \le \frac{\binom{2n}{n}^2}{A^2(2n,4,n)} \cdot \frac{n!^4}{M^4(n,4)}. \tag{4.1}$$

In Table 1, partitions from [2] and [1, 9] are used to obtain lower bounds on $M(2n,4)$ which improve the Gilbert-Varshamov bound, here denoted by GV. Lemma 4.4 and (4.1) are used.

5 Conclusion

Our main upper bound on $M(n,4)$ is actually obtained through $M_{\mathrm{LP}}(n,\{2,3\})$, yet we have so far said nothing about $M(n,\{2,3\})$. It is not hard to show that $M(n,\{2,3\}) = n$ for $n \ge 6$. For consider a set $\Delta \subset \mathcal{S}_n$, $n \ge 6$, with $\sigma\tau^{-1}$ either a transposition or 3-cycle for distinct $\sigma, \tau \in \Delta$. Without loss of generality, $() \in \Delta$. Any transpositions in Δ must intersect. So if Δ has only transpositions, then $|\Delta| \le 1 + (n-1) = n$. On the other hand, any two 3-cycles, as well as any transposition and any 3-cycle, must pairwise intersect in two points. Checking various cases completes the argument.

It is curious that $M_{\mathrm{LP}}(n,\{2,3\})$ being a poor upper bound on $M(n,\{2,3\})$ is actually advantageous for bounding $M(n,4)$. Along these lines, we are very interested in deciding whether

$$M_{\mathrm{LP}}(n,\{2,3\})M_{\mathrm{LP}}(n,\{4,\dots,n\})=n!.$$

A variation on this identity does hold. Consider the graph G with vertex set $\mathcal{S}_n$, and where pair $\{\sigma,\tau\}$ is an edge of G if and only if $d(\sigma,\tau)\notin D$. The *Lovász θ-function* $\theta(n,D)$ of G is an upper bound on the size of an independent set in G; thus it is also an upper bound on $M(n,D)$. Furthermore,

$$\theta(n,D)\theta(n,D^c)=n!.$$

The interested reader is referred to [12] for more details on the Lovász bound in relation to Delsarte's bound.

It is not hard to see that $M_{\mathrm{LP}}(n,\{2,3\})=\theta(n,\{2,3\})$. In fact, the latter quantity amounts to dropping the nonnegativity condition on the variables in Proposition 3.1, and the maximum is still attained in the first quadrant $a,b\geq 0$. Actually, one of the referees has observed more generally that semidefinite programming may be a fruitful technique for bounding permutation codes. We leave this to possible future work.

On the lower bound side, we have for $n\leq 32$ reported new lower bounds substantially better than the Gilbert-Varshamov bound. It is shown that the partitioning construction is recursive and accounts for all permutations in $\mathcal{S}_{2n}$. However, it appears difficult to control the number of disjoint arrays, and hence $n!/M(n,4)$, as n increases.

It is our hope that LP bounds and the partitioning construction will enjoy further application to permutation codes, and to constant composition codes in general.

Acknowledgement The authors are thankful for the careful reading of the anonymous referees, whose suggestions greatly improved the manuscript.

References

1. Brouwer, A.E., Shearer, J.B., Sloane, N.J.A., Smith, W.D.: A new table of constant weight codes. IEEE Trans. Inform. Theory **36**, 1334–1380 (1990)
2. Chu, W., Colbourn, C.J., Dukes, P.J.: Permutation codes for powerline communication. Des. Codes Cryptography **32**, 51–64 (2004)
3. Colbourn, C.J., Kløve, T., Ling, A.C.H.: Permutation arrays for powerline communication and mutually orthogonal Latin squares. IEEE Trans. Inform. Theory **50**, 1289–1291 (2004)
4. Delsarte, P.: An algebraic approach to the association schemes of coding theory. Philips Res. Rep. Suppl. **10** (1973)
5. Deza, M., Vanstone, S.A.: Bounds for permutation arrays. J. Statist. Plann. Inference **2**, 197–209 (1978)
6. Ellis, D., Friedgut, E., Pilpel, H.: Intersecting families of permutations. Preprint
7. Fulton, W.: Young Tableaux. London Mathematical Society Student Texts, vol. 35. Cambridge University Press, Cambridge (1997)
8. Frankl, P., Deza, M.: On the maximum number of permutations with given maximal or minimal distance. J. Combin. Theory Ser. A **22**, 352–360 (1977)

9. Graham, R.L., Sloane, N.J.A.: Lower bounds for constant weight codes. IEEE Trans. Inform. Theory **26**, 37–43 (1980)
10. Keevash, P., Ku, C.Y.: A random construction for permutation codes and the covering radius. Des. Codes Cryptogr. **41**, 79–86 (2006)
11. Murnaghan, F.D.: The Theory of Group Representations. Johns Hopkins Press, Baltimore (1938)
12. Schrijver, A.: A comparison of the Delsarte and Lovász bounds. IEEE Trans. Inform. Theory **25**, 425–429 (1979)
13. Tarnanen, H.: Upper bounds on permutation codes via linear programming. European J. Combin. **20**, 101–114 (1999)
14. Van Pul, C.L.M., Etzion, T.: New lower bounds for constant weight codes. IEEE Trans. Inform. Theory **35**, 1324–1329 (1989)
15. Yang, L., Chen, K.: New lower bounds on sizes of permutation arrays. Preprint