Asymptotic Improvement of the Gilbert-Varshamov Bound on the Size of Permutation Codes

Michael Tait, Alexander Vardy, and Jacques Verstraëte

Abstract—Given positive integers \( n \) and \( d \), let \( M(n, d) \) denote the maximum size of a permutation code of length \( n \) and minimum Hamming distance \( d \). The Gilbert-Varshamov bound asserts that \( M(n, d) \geq n! / V(n, d-1) \) where \( V(n, d) \) is the volume of a Hamming sphere of radius \( d \) in \( S_n \). Recently, Gao, Yang, and Ge showed that this bound can be improved by a factor \( \Omega(\log n) \), when \( d \) is fixed and \( n \to \infty \). Herein, we consider the situation where the ratio \( d/n \) is fixed and improve the Gilbert-Varshamov bound by a factor that is linear in \( n \). That is, we show that if \( d/n < 0.5 \), then

\[
M(n, d) \geq cn \frac{n!}{V(n, d-1)}
\]

where \( c \) is a positive constant that depends only on \( d/n \). To establish this result, we follow the method of Jiang and Vardy. Namely, we recast the problem of bounding \( M(n, d) \) into a graph-theoretic framework and prove that the resulting graph is locally sparse.

Index Terms—Ajtai-Komlós-Szemeredi bound, Gilbert-Varshamov bound, locally sparse graphs, permutation codes

I. INTRODUCTION

Let \( S_n \) be the symmetric group of permutations on \( n \) elements. A permutation code \( C \) is a subset of \( S_n \). For \( \sigma, \tau \) in \( S_n \), the Hamming distance \( d(\sigma, \tau) \) between them is the number of positions where they differ, namely:

\[
d(\sigma, \tau) = \sum_{i \in [n]} |\sigma(i) - \tau(i)|
\]

We say that a permutation code \( C \subseteq S_n \) has minimum distance \( d \) if each pair of permutations in \( C \) is at Hamming distance at least \( d \). The maximum number of codewords in a permutation code of minimum distance \( d \) will be denoted \( M(n, d) \).

Permutation codes were first investigated in \([4,5,12,13,26]\). A permutation code \( C \subseteq S_n \) has minimum distance \( d \) if each pair of permutations in \( C \) is at Hamming distance at least \( d \). The maximum number of codewords in a permutation code of minimum distance \( d \) will be denoted \( M(n, d) \).

Permutation codes were first investigated in \([4,5,12,13,26]\). A permutation code \( C \subseteq S_n \) has minimum distance \( d \) if each pair of permutations in \( C \) is at Hamming distance at least \( d \). The maximum number of codewords in a permutation code of minimum distance \( d \) will be denoted \( M(n, d) \).

In recent years, they have received considerable renewed attention, due to their application in data transmission over power lines \([9,10,16,24,28]\). Other applications of permutation codes include design of block ciphers \([11]\) and coding for flash memories \([3,7,8]\) (although the latter application involves the Kendall \( T \)-metric in lieu of the Hamming distance).

A key problem in the theory of permutation codes is to determine \( M(n, d) \). Various bounds on \( M(n, d) \) have been proposed in \([10,13,15,17,18,23]\). For small values of \( d \), computer search has been used to find exact values of \( M(n, d) \) in \([23]\). However, in general, the problem is very difficult, and little progress has been made for \( d > 4 \).

For \( x \in \mathbb{R} \), let \( \lceil x \rceil \) denote the operation of rounding to the nearest integer. Then the number of derangements of \( k \) elements is given by \( D_k = \lceil n!/d \rceil \). Under the Hamming distance defined in \([1]\), a sphere of radius \( r \) has volume

\[
V(n, r) = \sum_{k=0}^{r} \binom{n}{k} D_k
\]

The Gilbert-Varshamov bound \([19,27]\) and the sphere-packing bound \([22]\) now give the following.

Theorem 1. \( \frac{n!}{V(n, d-1)} \geq M(n, d) \geq \frac{n!}{V(n, \lfloor \frac{d}{2} \rfloor)} \)

The Gilbert-Varshamov bound is used extensively in coding theory \([22,25]\). As is well known, improving this bound asymptotically is a difficult task \([25]\).

Recently, Gao, Yang, and Ge \([18]\) showed that the Gilbert-Varshamov bound in \([2]\) can be improved by a factor \( \Omega(\log n) \), when \( d \) is fixed and \( n \to \infty \). In this paper, we complement this work, focusing on the more natural case where \( d/n \) is a fixed ratio and \( n \) tends to infinity. Our main theorem is the following.

Theorem 2. Let \( d/n \) be a fixed ratio with \( 0 < d/n < 0.5 \). Then as \( n \to \infty \), we have

\[
M(n, d) \geq \Omega \left( \frac{n!}{V(n, d-1)} \right)^n
\]

To prove this theorem, we follow the method of \([20]\). Namely, we will construct a graph in which every independent set corresponds to a permutation code of minimum distance \( d \). Using techniques from graph theory, we will then obtain a lower bound on the size of the largest independent set in this graph.

II. GRAPH THEORY

Our graph notation is standard; the reader is referred to \([30]\) for any undefined terms. A graph \( G \) consists of a set of vertices \( V(G) \) and a set of edges \( E(G) \), which are pairs of vertices. We assume throughout that the sets \( V(G) \) and \( E(G) \) are finite. Two vertices \( u \) and \( v \) are adjacent if \( \{u, v\} \in E(G) \). For a subset \( S \) of \( V(G) \), the subgraph induced by \( S \) is the graph that has vertex set \( S \) and edge set \( \{u, v\} \in E(G) : u, v \in S \}. The neighborhood of a vertex \( v \) is the set of all vertices adjacent to \( v \), and the neighborhood graph of \( v \) is the subgraph induced by the neighborhood of \( v \). The degree of a vertex \( v \) is the size of the neighborhood of \( v \). We let \( \Delta(G) \) denote the maximum vertex degree in a graph \( G \). A set \( K \subseteq V(G) \) is a clique if the subgraph induced by \( K \) has every possible edge, and a set \( I \subseteq V(G) \) is
an independent set if the subgraph induced by \( I \) has no edges. A triangle is a clique of size 3. The maximum number of vertices in an independent set of \( G \) is called the independence number and denoted by \( \alpha(G) \).

Consider the graph \( \mathcal{G} \) whose vertex set is \( S_n \) with two permutations \( \sigma, \tau \) being adjacent if and only if \( 1 \leq d(\sigma, \tau) < d \). Then it is clear that any permutation code of length \( n \) and minimum distance \( d \) is an independent set in \( \mathcal{G} \). Conversely, any independent set in \( \mathcal{G} \) is a permutation code of length \( n \) and minimum distance \( d \). This bijection proves the following.

**Lemma 3.** \( M(n, d) = \alpha(\mathcal{G}) \).

Observe that every vertex in \( \mathcal{G} \) has the same degree, given by \( \sum_{k=1}^{d-1} \binom{n}{k}D_k \). It is graph theory folklore that for any graph \( G \),

\[
\alpha(G) \geq \frac{|V(G)|}{\Delta(G) + 1} \tag{4}
\]

It is evident that (4) along with Lemma 3 immediately recovers the Gilbert-Varshamov bound in (2). We note that we have recovered this bound using very little information about \( \mathcal{G} \).

Our strategy to prove Theorem 2 is to use the relative sparseness of the neighborhood graph of every vertex in \( V(\mathcal{G}) \) in order to strengthen (4). This technique has been introduced in [24] to improve the Gilbert-Varshamov bound on the size of binary codes. It was later applied to improve lower bounds on the size of \( q \)-ary codes [28] and sphere packings [21].

We will need some results about locally sparse graphs. Ajtai, Komlós, and Szemerédi [1] showed that one can improve (4) if \( G \) has no triangles. The following lemma is from [1] (but see also [2] p. 272) for a much shorter proof of the same result.

**Lemma 4.** Let \( G \) be a graph with maximum degree \( \Delta \), and suppose that \( G \) has no triangles. Then

\[
\alpha(G) \geq \frac{|V(G)|}{8\Delta} \log_2 \Delta \tag{5}
\]

This result was extended in [6, Lemma 15, p. 296] from graphs with no triangles to graphs with relatively few triangles.

**Lemma 5.** Let \( G \) be a graph with maximum degree \( \Delta \) and suppose that \( G \) has at most \( T \) triangles. Then

\[
\alpha(G) \geq \frac{|V(G)|}{10\Delta} \left( \ln \Delta - \frac{1}{2} \ln \left( \frac{T}{|V(G)|} \right) \right) \tag{6}
\]

A graph has no triangles if and only if the neighborhood of every vertex is an independent set, and a graph has relatively few triangles if the neighborhoods of its vertices are relatively sparse. We make this precise in the following corollary.

**Corollary 6.** Let \( G \) be a graph with maximum degree \( \Delta \) and suppose that for all \( v \in V(\mathcal{G}) \), the subgraph induced by the neighborhood of \( v \) has at most \( E \) edges. Then

\[
\alpha(G) \geq \frac{|V(G)|}{10\Delta} \left( \ln \Delta - \frac{1}{2} \ln \left( \frac{E}{3} \right) \right) \tag{7}
\]

**Proof:** The number of triangles incident with a vertex \( v \) is equal to the number of edges in the subgraph induced by the neighborhood of \( v \). Therefore, for every \( v \in V(G) \), there are at most \( E \) triangles incident with \( v \). Summing over the vertex set of \( G \) and noting that each triangle is incident with exactly 3 vertices gives the result. \( \blacksquare \)

III. **Proof of the Main Theorem**

In order to use Corollary 6 we must count the number of edges in the neighborhood of every vertex in \( V(\mathcal{G}) \). To do so, first note that \( \mathcal{G} \) is vertex transitive because for all \( \sigma, \tau \in S_n \), we have

\[
d(\sigma, \tau) = d(id, \sigma \tau^{-1}) \]

where \( id \) is the identity element of \( S_n \). This implies that the number of edges in the neighborhood of a vertex \( v \in V(\mathcal{G}) \) does not depend on \( v \). Therefore, to simplify the calculation, we will consider only the neighborhood of the identity permutation.

By Corollary 6 in order to prove our main result in (3), it would suffice to show that

\[
\log \left( \frac{\Delta^2}{E} \right) = \Omega(n) \tag{8}
\]

where \( E \) is the number of edges in the subgraph induced by the neighborhood of the identity, and \( \Delta \) is given by

\[
\Delta \overset{\text{def}}{=} \sum_{k=1}^{d} \binom{n}{k}D_k \tag{9}
\]

Note that, for notational convenience, the sum in (9) extends up to \( d \) rather than \( d - 1 \). Since \( d/n \) is fixed, this does not matter.

It follows from (8) that the proof of Theorem 2 reduces to estimating \( E \) and considering the asymptotics of the ratio \( \Delta^2/E \). We count the number of edges in the subgraph induced by the neighborhood of the identity as follows.

1) Vertices in the neighborhood graph of the identity have distance between 1 and \( d - 1 \) from the identity. We will count the edges incident with permutations \( \sigma \) that are at distance \( s \) from \( id \), and then sum over \( s = 1, 2, \ldots, d - 1 \). Note that, given a fixed distance \( s \), there are exactly \( \binom{n}{s} \) \( D_s \) permutations at distance \( s \) from the identity.

2) Let us fix a permutation \( \sigma \) at distance \( s \) from the identity and count how many edges in the neighborhood graph of the identity are incident with \( \sigma \). To do this, we will sum over permutations \( \tau \) that are at distance \( t \) from the identity, for \( t = 1, 2, \ldots, d - 1 \). Note that for \( \tau \) to be incident with \( \sigma \), we must also have \( d(\sigma, \tau) < d \).

3) To count how many permutations \( \tau \) satisfy the above requirements, we let \( m \) be the number of indices where \( \sigma \) and \( \tau \) are deranging the same position. That is, given \( \sigma \) and \( \tau \), let

\[
m \overset{\text{def}}{=} \left| \{ i \in [n] : \sigma(i) \neq i, \tau(i) \neq i \} \right| \tag{10}
\]

By assumption, \( \sigma \) is deranging \( s \) positions while \( \tau \) is deranging \( t \) positions. Thus \( m \leq \min(s,t) \). Also notice that in the \( s - m \) positions where \( \sigma \) is deranging but \( \tau \) is not, the two permutations necessarily differ. Same goes for the \( t - m \) positions where \( \tau \) is deranging but \( \sigma \) is not. Hence

\[
(s - m) + (t - m) \leq d(\sigma, \tau) < d \tag{11}
\]

It follows that the parameter \( m \) defined in (10) is in the range \( \lfloor (s + t - d)/2 \rfloor < m \leq \min(s,t) \), where \( \lfloor x \rfloor \) denotes the smallest nonnegative integer \( k \) with \( k \geq x \).

4) Now let us suppose that \( s, t, m \) are fixed, and count those positions where \( \sigma \) and \( \tau \) are both deranging, but still map to the same value. That is, let

\[
r \overset{\text{def}}{=} \left| \{ i \in [n] : \sigma(i) \neq i, \tau(i) \neq i, \sigma(i) = \tau(i) \} \right| \tag{12}
\]
Notice that \( d(\sigma, \tau) = s + t - m - r \). It follows that the parameter \( r \) defined in (12) satisfies \( r > [s + t - m - d]^+ \).

Our strategy for counting \( E \) is to fix \( s, t, m, r \) and count the number of pairs of permutations \( \sigma \) and \( \tau \) that have these parameters. We will then sum over \( s, t, m, r \) in the appropriate ranges.

As already noted above, there are \( \binom{m}{s} \binom{m}{t} \) permutations \( \sigma \) at a given distance \( s \) from the identity. Given \( \sigma \), there are \((s)\) ways to select the \( m \) positions where both \( \sigma \) and \( \tau \) are deranging. Given these \( m \) positions, there are \((r)\) ways to pick the \( r \) positions where both permutations are deranging but have the same image, as in (12). We have now chosen our permutation \( \sigma \), and part of our permutation \( \tau \). In particular, we have chosen \( m \) positions where \( \tau \) is deranging. To specify the rest of \( \tau \), we first choose the other \( t - m \) positions where \( \tau \) is deranging. This can be done in exactly \((t-m)\) ways. Now that we have chosen the \( t \) positions where \( \tau \) is deranging, we must p...
Since \( \lim_{c \to 0} 3T_2(3c) = 0 \), we can conclude from (8) and (27) that it remains to show
\[
\log_2 \left( \frac{(\delta n)^2 D_2^2}{d^4 g(d, d, d, 0)} \right) = \log_2 \left( \frac{(\delta n)^2 D_2^2}{d^4 (\delta n) D_2 d!} \right) = \Omega(n)
\]
This is easily accomplished as follows:
\[
\log_2 \left( \frac{(\delta n)^2 D_2^2}{d^4 (\delta n) D_2 d!} \right) \geq d \log_2 \left( \frac{n}{d} \right) - 4 \log_2(d) - \log_2 3
\]
\[
= \delta n \log_2 \left( \frac{1}{\delta} \right) - 4 \log_2(\delta n) - \log_2 3
\]
Since \( \delta = d/n \) is a positive constant strictly less than 0.5, we see that \( \delta \log_2(1/\delta) \) is positive and, hence, the expression above is \( \Omega(n) \). This completes the proof of Theorem 3.

REFERENCES

[1] M. Ajtai, J. Komlós, and E. Szemerédi, “A note on Ramsey numbers,” J. Combinatorial Theory Ser. A, vol. 29, pp. 354–260, 1980.
[2] N. Alon and J. Spencer, The Probabilistic Method, 2nd edition, New York: John Wiley and Sons, 2000.
[3] A. Barg and A. Mazumdar, “Codes in permutations and error correction for rank modulation,” IEEE Trans. Inform. Theory, vol. 56, pp. 3158–3165, July 2010.
[4] I.F. Blake, “Permutation codes for discrete channels,” IEEE Trans. Inform. Theory, vol. 20, pp. 138–140, 1974.
[5] I.F. Blake, G. Cohen, and M. Deza, “Coding with permutations,” Inf. Control, vol. 43, no. 1, pp. 1–19, 1979.
[6] B. Bollobás, Random Graphs, 1st edition, London: Academic Press, 1985.
[7] J. Bruck, A. Jiang, R. Mateescu, and M. Schwartz, “Rank modulation for flash memories,” Proc. IEEE Int. Symp. Information Theory, pp. 1731–1735, Toronto, Canada, 2008.
[8] J. Bruck, A. Jiang, and M. Schwartz, “Error-correcting codes for rank modulation,” Proc. IEEE Int. Symp. Information Theory, pp. 1736–1740, Toronto, Canada, 2008.
[9] W. Chu, C.J. Colbourn, and P. Dukes, “Constructions for permutation codes in paperware communications,” Des. Codes Cryptogr., vol. 32, pp. 51–64, 2004.
[10] C.J. Colbourn, T. Kløve, and A.C.H. Ling, “Permutation arrays for paperware communications and mutually orthogonal Latin squares,” IEEE Trans. Inform. Theory, vol. 50, no. 6, pp. 1289–1291, June 2004.
[11] C.J. Colbourn, A.C.H. Ling, D.R. de la Torre, “An application of permutation arrays to block ciphers,” Proc. Southeastern International Conference on Combinatorics, Graph theory and Computing, vol. 145, pp. 5–7, Boca Raton, FL, 2000.
[12] M. Deza and P. Frankl, “On the maximum number of permutations with given maximal or minimal distance,” J. Combinatorial Theory Ser. A, vol. 22, no. 3, pp. 352–360, 1977.
[13] M. Deza and P. Frankl, “Bounds for permutation arrays,” J. Statist. Plann. Inference, vol. 2, no. 2, pp. 197–209, 1978.
[14] C. Ding, F.-W. Fu, T. Kløve, and V.K.-W. Wei, “Construction of permutation arrays,” IEEE Trans. Inform. Theory, vol. 48, no. 4, pp. 977–980, 2002.
[15] P. Dukes and N. Savchuck, “Bounds on permutation codes of distance four,” J. Algebraic Combin., vol. 31, no. 1, pp. 143–158, 2010.
[16] H.C. Ferreira and A.J.H. Vinck, “Inference cancellation with permutation trellis arrays,” Proc. IEEE Vehicular Technology Conf., pp. 2401-2407, Boston, MA, September 2000.
[17] F.-W. Fu and T. Kløve, “Two constructions of permutations arrays,” IEEE Trans. Inform. Theory, vol. 50, no. 5, pp. 881–883, 2004.
[18] F. Gao, Y. Yang, and G. Ge, “An improvement to the Gilbert-Varshamov bound for permutation codes,” IEEE Trans. Inform. Theory, vol. 59, pp. 3059–3063, May 2013.
[19] E.N. Gilbert, “A comparison of signalling alphabets,” Bell Syst. Tech. J., vol. 31, pp. 504–522, 1952.
[20] T. Jiang and A. Vardy, “Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes,” IEEE Trans. Inform. Theory, vol. 50, pp. 1655–1664, August 2004.
[21] M. Krivelevich, S. Litsyn, and A. Vardy, “A lower bound on the density of sphere packings via graph theory,” International Math. Research Notices, vol. 43, pp. 2271–2279, June 2004.
[22] F.J. MacWilliams and N.J.A. Sloane, The Theory of Error Correcting Codes, Amsterdam: North-Holland/Elsevier, 1977.
[23] R. Montermini and D.H. Smith, “A new table of permutation codes,” Des. Codes Cryptogr., vol. 63, no. 2, pp. 241–253, 2012.
[24] N. Pavlidou, A.J.H. Vinck, J. Yazdani, and B. Honary, “Power line communications: State of the art and future trends,” IEEE Commun. Mag., vol. 41, pp. 34–40, 2003.
[25] V.S. Pless and W.C. Huffman (Editors), Handbook of Coding Theory, Amsteram: North-Holland/Elsevier, 1998.
[26] D. Slepian, “Permutation modulation,” Proc. IEEE, vol. 53, pp. 228–236, March 1965.
[27] R.R. Varshamov, “Estimate of the number of signals in error correcting codes,” Dokl. Acad. Nauk, vol. 117, pp. 739–741, 1957 (in Russian).
[28] A.J.H. Vinck, “Coded modulation for powerline communications,” A.E.U. Int. J. Electron. Commun., vol. 54, no. 9, pp. 3200–3208, 2005.
[29] V.H. Vu and L. Wu, “Improving the Gilbert-Varshamov bound for q-ary codes,” IEEE Trans. Inform. Theory, vol. 51, no. 9 pp. 3200–3208, 2005.
[30] D.B. West, Introduction to Graph Theory, Prentice Hall, 1996.