



HAL
open science

Improving the Gilbert-Varshamov bound for q -ary codes

van H. Vu, Lei Wu

► **To cite this version:**

van H. Vu, Lei Wu. Improving the Gilbert-Varshamov bound for q -ary codes. 2005 European Conference on Combinatorics, Graph Theory and Applications (EuroComb '05), 2005, Berlin, Germany. pp.285-288, 10.46298/dmtcs.3456 . hal-01184445

HAL Id: hal-01184445

<https://inria.hal.science/hal-01184445>

Submitted on 14 Aug 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Improving the Gilbert-Varshamov bound for q -ary codes

Van H. Vu^{1†} and Lei Wu^{1‡}

¹Department of Mathematics, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0112, USA.

Given positive integers q , n and d , denote by $A_q(n, d)$ the maximum size of a q -ary code of length n and minimum distance d . The famous Gilbert-Varshamov bound asserts that

$$A_q(n, d + 1) \geq q^n / V_q(n, d),$$

where $V_q(n, d) = \sum_{i=0}^d \binom{n}{i} (q-1)^i$ is the volume of a q -ary sphere of radius d .

Extending a recent work of Jiang and Vardy on binary codes, we show that for any positive constant α less than $(q-1)/q$ there is a positive constant c such that for $d \leq \alpha n$, $A_q(n, d + 1) \geq c \frac{q^n}{V_q(n, d)}$. This confirms a conjecture by Jiang and Vardy.

1 Introduction

Given a set Ω of q symbols, without loss of generality, let $\Omega = \{0, 1, \dots, q-1\}$. A q -ary word of length n is a sequence $x = (x_1, \dots, x_n)$, where $x_i \in \Omega$. The number of non-zero symbols in a word x is the weight of x . Given two words x and y , the (Hamming) distance between x and y is the number of coordinates i in which x_i and y_i are different. A set \mathcal{C} of words is called a code with minimum distance d if any two codewords in \mathcal{C} have distance at least d . For a word x , the Hamming sphere of radius d centered at x has volume

$$V_q(n, d) = \sum_{i=0}^d \binom{n}{i} (q-1)^i.$$

Thanks to symmetry, the volume of the sphere does not depend on x .

For integers q , n and d , let $A_q(n, d)$ denote the maximum size of a q -ary code of length n and minimum distance d . Estimating $A_q(n, d)$ is one of the most important problems in coding theory. The famous Gilbert-Varshamov bound [4, 11] asserts that

$$A_q(n, d + 1) \geq \frac{q^n}{V_q(n, d)}.$$

[†]Email address: vanvu@ucsd.edu; webpage: <http://www.math.ucsd.edu/~vanvu/> Research supported in part by an NSF CAREER award and by an Alfred P. Sloan fellowship.

[‡]Email: lwu@math.ucsd.edu; <http://math.ucsd.edu/~lwu>

This bound is used extensively in numerous contexts and has been generalized in many different settings [7, 8, 6]. Improving upon the Gilbert-Varshamov bound asymptotically is a notoriously difficult task [8]. Tsfasman, Vlăduț, and Zink [10] made a breakthrough for the case when $q \geq 49$. More recently, Jiang and Vardy [6] improved the Gilbert-Varshamov bound, for the case $q = 2$, for certain range of d :

Theorem 1.1 *Let α be a constant satisfying $0 < \alpha \leq .4994$. Then there is a positive constant c depending on α such that the following holds. For $d \leq \alpha n$,*

$$A_2(n, d + 1) \geq c \frac{2^n}{V_2(n, d)} \log_2 V_2(n, d) \quad (1)$$

If $d \geq \alpha' n$ for some constant $\alpha' > 0$, then $V_2(n, d)$ is exponential in n . Thus, Theorem 1.1 improved Gilbert-Varshamov bound by a factor linear in n . We can rewrite (1) in the following more pleasant form (the constant c here, of course, would be different):

$$A_2(n, d + 1) \geq c \frac{2^n}{V_2(n, d)} n. \quad (2)$$

Jiang and Vardy asked if one can get to $\alpha < 0.5$ using a different method than computer simulations as they did (the strange constant .4994 resulted from these simulations). They also conjectured that an improvement similar to (2) can be achieved for q -ary codes, for any $q \geq 3$.

The main result of this paper resolves both of these issues. For the binary case, our main theorem (Theorem 1.2) extends the assumption $\alpha < 0.4994$ in [6] to its natural limit $\alpha < 0.5$. The proof of Theorem 1.2 does not rely on computers, and reflects, in a clean way, the necessity of the assumption $\alpha < (q - 1)/q$.

Throughout the paper, asymptotic notations are used under the assumption that n goes to infinity. We also emphasize the case when d is proportional to n , namely, $d = \alpha n$ for some positive constant α . This case is of special interest in coding theory.

Theorem 1.2 *Let q be a fixed positive integer and α be a constant satisfying $0 < \alpha < \frac{q-1}{q}$. There is a positive constant c depending on q and α such that for $d = \alpha n$,*

$$A_q(n, d + 1) \geq c \frac{q^n}{V_q(n, d)} n \quad (3)$$

In general, the constant α can take any value less than or equal to one. However, it is well known and easy to show that for $\alpha \geq (q - 1)/q$, the volume $V_q(n, d)$ is close to q^n , namely, $q^n \leq 2V_q(n, d)$. In this case, the Gilbert-Varshamov bound gives no useful information. Thus, the value $(q - 1)/q$ serves as a natural threshold and we will assume $\alpha < (q - 1)/q$.

2 Graph theoretic frame work

We recall a folklore in graph theory.

Proposition 2.1 *Let G be a D -regular graph on n vertices. Then G contains an independent set of size $n/(D + 1)$.*

Given q, n and d , we follow [6] and define a graph \mathcal{G} whose vertices are the q -ary words of length n and two words are adjacent if their Hamming distance is at most d . It's easy to see that \mathcal{G} has q^n vertices, the degree of every vertex is $D = V_q(n, d) - 1$, and $A_q(n, d + 1)$ is the independence number of \mathcal{G} , denoted by $I(\mathcal{G})$. The Gilbert-Varshamov bound is simply the realization of Proposition 2.1 on this graph.

For a D -regular graph, each neighborhood has at most $\binom{D}{2}$ edges. We say that such a graph is *locally sparse* if in every neighborhood the number of edges is much less than $\binom{D}{2}$. In the extreme case when the graph is triangle-free, i.e., when the number of edges in each neighborhood is zero, Proposition 2.1 was improved by a logarithmic factor by Ajtai, Komlós and Szemerédi in [1]. Namely, they obtained $I(G) \geq cn \log D/D$. This result has been extended to locally sparse graphs (i.e. with few triangles) by Shearer [9].

Lemma 2.2 (Shearer) *For any positive constant $\epsilon \leq 2$ there is a positive constant c such that the following holds. Let G be a D -regular graph on N vertices. Assume that each neighborhood in G contains at most $D^{2-\epsilon}$ edges. Then the independence number of G , denoted by $I(G)$, satisfies:*

$$I(G) \geq c \frac{N}{D} \ln D.$$

In order to prove Theorems 1.1 and 1.2, one needs to verify the hypothesis of Lemma 2.2 for \mathcal{G} . Due to symmetry, every neighborhood in \mathcal{G} has the same number of edges. Thus, for convenience, we can consider the neighborhood of the word consisting of only zeros. Let T be the number of edges in this neighborhood and \mathcal{G}_0 be the graph spanned by these edges. Our goal is to show that there is a positive constant ϵ such that

$$T \leq D^{2-\epsilon}. \tag{4}$$

It is not hard to give explicit formulae for T and D . Fixed $q \geq 2$, we have

$$D = V_q(n, d) - 1 = \sum_{i=1}^d \binom{n}{i} (q-1)^i,$$

$$T = \Theta \left(\sum_{w=1}^d \binom{n}{w} (q-1)^w \sum_{\{i,j,k\} \in N} \binom{w}{i} \binom{w-i}{k} \binom{n-w}{j} (q-2)^k (q-1)^j \right),$$

where N is the set of all triples $\{i, j, k\}$ that satisfies:

$$i + k \leq w, \quad j \leq n - w, \quad w - i + j \leq d, \quad \text{and} \quad d(x, y) = i + j + k \leq d$$

One can easily see the difficulty of dealing with these two variables directly, especially T . In fact, this was the main hurdle for further improvement of [6].

Our approach is to translate (4) into simpler inequalities which we are able to prove using the following notion. Let X and Y be two functions in n . We call X and Y polynomially equivalent and write $X \sim Y$ if there are positive constants c_1, c_2 such that

$$n^{-c_1} X \leq Y \leq n^{c_2} X.$$

We find new parameters $T' \sim T$, $D' \sim D$ where both T' and D' are relatively simple. Since both T and D are exponential functions in n , if we can show

$$T' \leq D'^{2-\delta}, \quad (5)$$

for a positive constant δ , then it follows that for all sufficiently large n , $T \leq D^{2-\epsilon}$, where, say, $\epsilon = .999\delta$.

Finding D' is easy. For T' , we will apply a technique which can be viewed as a discrete analogue of Lagrange's multiplier. Once D' and T' are determined, (5) becomes equivalent to a reasonable inequality concerning entropy functions, which serve as good estimates of binomial coefficients. This inequality is not obvious, but can be proved using the assumption $\alpha < (q-1)/q$ and an analytic argument. The readers are invited to check the full version of the paper for the (rather technical) details.

References

- [1] M. Ajtai, J. Komlós and E. Szemerédi, A note on Ramsey numbers, *J. Combinatorial Theory (A)*, 29, (1980), 354-360.
- [2] N. Alon and J. Spencer, *Probabilistic Method*, John Wiley & Sons, Inc., 2000.
- [3] B. Bollobas, *Random Graphs*, 2nd ed., Cambridge University Press, 2001.
- [4] E. N. Gilbert, A comparison of signalling alphabets, *Bell Syst. Tech. J.*, 31, (1952), 504-522.
- [5] R. J. Graham, M. Grötschel and L. Lovász, *Handbook of Combinatorics*, MIT Press, North-Holland, 1995.
- [6] T. Jiang and A. Vardy, Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes, *IEEE Transactions on Information Theory*, to appear.
- [7] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes*, Amsterdam: North-Holland/Elsevier, 1977.
- [8] V. S. Pless, W. C. Huffman (editors), *Handbook of Coding Theory*, Amsterdam: North-Holland/Elsevier, 1998.
- [9] J. B. Shearer, A note on the independence number of triangle-free graphs, *Discrete Mathematics*, 46, (1983), 83-87.
- [10] M. A. Tsfasman, S. G. Vlăduț, T. Zink, Modular curves, Shimura curves, and Goppa codes better than the Varshamov-Gilbert bound, *Mathematische Nachrichten*, 104, (1982), 13-28.
- [11] R. R. Varshamov, Estimate of the number of signals in error correcting codes, *Dokl. Acad. Nauk*, 117, (1957), 739-741, (in Russian).