Title
Long nonbinary codes exceeding the Gilbert-Varshamov bound for any fixed distance

Permalink
https://escholarship.org/uc/item/3cd9654v

Journal
IEEE Transactions on Information Theory, 50(10)

ISSN
0018-9448

Authors
Yekhanin, S
Dumer, I

Publication Date
2004-10-01

Peer reviewed
Long Nonbinary Codes Exceeding the Gilbert–Varshamov Bound for any Fixed Distance

Sergey Yekhanin and Ilya Dumer, Senior Member, IEEE

Abstract—Let $A(q,n,d)$ denote the maximum size of a $q$-ary code of length $n$ and distance $d$. We study the minimum asymptotic redundancy

$$\rho(q,n,d) = n - \log_q A(q,n,d)$$

as $n$ grows while $q$ and $d$ are fixed. For any $d$ and $q \geq d - 1$, long algebraic codes are designed that improve on the Bose–Chaudhuri–Hocquenghem (BCH) codes and have the lowest asymptotic redundancy

$$\rho(q,n,d) \leq \left( (d - 3) + 1/(d - 2) \right) \log_q n$$

known to date. Prior to this work, codes of fixed distance that asymptotically surpass BCH codes and the Gilbert–Varshamov bound were designed only for distances 4, 5, and 6.

Index Terms—Affine lines, Bose–Chaudhuri–Hocquenghem (BCH) code, Bezout’s theorem, norm.

I. INTRODUCTION

Let $A(q,n,d)$ denote the maximum size of a $q$-ary code of length $n$ and distance $d$. We study the asymptotic size $A(q,n,d)$ if $q$ and $d$ are fixed as $n \to \infty$, and introduce a related quantity

$$c(q,d) = \lim_{n \to \infty} \frac{n - \log_q A(q,n,d)}{\log_q n}$$

which we call the redundancy coefficient.

The Hamming upper bound

$$A(q,n,d) \leq q^n / \sum_{i=0}^{\lfloor (d-1)/2 \rfloor} (q-1)^i \binom{n}{i}$$

leads to the lower bound

$$c(q,d) \geq \left\lfloor (d-1)/2 \right\rfloor$$

which is the best bound on $c(q,d)$ known to date for arbitrary values of $q$ and $d$. On the other hand, the Varshamov existence bound admits any linear $[n,k,d]_q$ code of dimension

$$k \leq n - 1 - \log_q \sum_{i=0}^{d-2} (q-1)^i \binom{n-1}{i}$$

This leads to the redundancy coefficient

$$c(q,d) \leq d - 2.$$  \hfill (2)

(Note that the Gilbert bound results in a weaker inequality $c(q,d) \leq d - 1$.)

Let $e$ be a primitive element of the Galois field $F_{q^n}$. Consider (see [20]) the narrow-sense Bose–Chaudhuri–Hocquenghem (BCH) code defined by the generator polynomial with zeros $e^1, \ldots, e^{d-2}$. Let $C_q^{q^n}(d)$ denote the extended BCH code obtained by adding the overall parity check. Code $C_q^{q^n}(d)$ has length $q^n$, constructive distance $d$, and redundancy coefficient

$$c(q,d) \leq \left[ (d-2)(q-1)/q \right].$$  \hfill (3)

Note that the above BCH bound (3) is better than the Varshamov bound (2) for $q < d - 1$ and coincides with (2) for $q \geq d - 1$. Note also that (3) meets the Hamming bound (1) if $q = 2$ or $d = 3$. Therefore,

$$c(2,d) = \left\lfloor (d-1)/2 \right\rfloor \quad \text{and} \quad c(q,3) = 1.$$

For distances 4, 5, and 6, infinite families of nonbinary linear codes are constructed in [5] and [6] that reduce asymptotic redundancy (3). Open Problem 2 from [6] also raises the question if the BCH bound (3) can be improved for larger values of $d$. Our main result is an algebraic construction of codes that gives an affirmative answer to this problem for all $q \geq d - 1$. In terms of redundancy, the new bound is expressed by the following.

Theorem 1: For all $q$ and $d \geq 3$

$$c(q,d) \leq (d-3) + 1/(d-2).$$  \hfill (4)

Combining (3) and (4), we obtain

$$c(q,d) \leq \min \left( \left[ (d-2)(q-1)/q \right], (d-3) + 1/(d-2) \right).$$

Note that the above bound is better than the Varshamov existence bound for arbitrary $q$ and $d \geq 4$. 

Manuscript received February 28, 2004; revised June 23, 2004. The work of S. Yekhanin was supported in part by NTT Award MIT 2001-04 and by the National Science Foundation under Grant CCR-0219218. The work of I. Dumer was supported by the National Science Foundation under Grant CCR-00097125.

S. Yekhanin is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: yekhanin@mit.edu).

I. Dumer is with the College of Engineering, University of California, Riverside, CA 92521, USA (e-mail: dumer@ee.ucr.edu).

Communicated by C. Carlet, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2004.834744
that surpass the BCH
cannot increase when the alphabet size is
Several families of linear codes con-
Linear, there exist better bounds for small
Namely, 
has been of special interest, and general
codes are equivalent to
and 
the construction of [14] also improves (5). Namely,
[10].
The constructions are rather complex and the resulting
, and
We start with the bounds
for all
The Hamming bound yields
for small values of 
we review the upper bounds for
we make some concluding remarks in Section V.
Finally, we review Theorem 4, which is proven in Section IV.
Now we proceed to the bounds for 
. The Hamming bound yields
for all values of 
for arbitrary value of distance
The following Lemma 2 due to Gevorkyan [13] shows that
redundancy cannot increase when the alphabet size is
reduced.

**Lemma 2:** For arbitrary value of distance
, we have

**Proof:** Given a code $V$ of length $n$ over the $q_2$-ary alphabet we prove the existence of a code $V'$ of the same length over $q_1$-ary alphabet with the same redundancy coefficient. Let $q_2$-ary alphabet be an additive group $E_{q_2}$, and $q_1$-ary alphabet form a subset $E_{q_1} \subseteq E_{q_2}$. Define the componentwise shift $V' = V + V$ of code $V$ by an arbitrary vector $v \in E_{q_1}$. Note that any vector $f \in E_{q_1}$ belongs to exactly $|V|$ codes among all $q_2$ codes $V_{q_2}$ as $v$ runs through $E_{q_1}$. Hence, codes $V_{q_2}$ include on average $q_2 |V|/q_1$ vectors of the subset $E_{q_1} \subseteq E_{q_2}$. Therefore, some set $V_{q_1} \cap E_{q_2}$ has at least this average size. Denote this set by $V'$. Clearly, $V'$ is a $q_1$-ary code with the same distance as code $V$. It remains to note that
\[
\frac{n - \log_{q_1} (q_2 |V|/q_1^2)}{\log_{q_1} n} = \frac{n - \log_{q_2} |V|}{\log_{q_2} n}.
\]
The proof is completed. \qed

**Corollary 3:** Let $\{q_i\}$ be an infinite sequence of growing alphabet sizes. Assume there exist $c^*$ and $d$ such that for all $i$, $c(q_i, d) \leq c^*$. Then $c(q, d) \leq c^*$ for all values of $q$.

**Proof:** This follows trivially from Lemma 2. \qed

### III. Code Construction

In the sequel, the elements of the field $F_q$ are denoted by Greek letters, while the elements of extension fields $F_{q^m}$ are denoted by Latin letters.

We start with an extended BCH code $C = C_{q^m}(d - 1)$ of length $n = q^m$ and constructive distance $d - 1$. Here for any position $j \in [1, q^m]$, we define its locator $e_j$, where $e_j = e^j$ for $j < n$ and $e_n = 0$. Then the parity-check matrix of code $C$ has the form

\[
H_{q^m}(d - 1) = \begin{pmatrix}
1 & \cdots & 1 \\
1 & \cdots & e_{n-1} & 0 \\
\vdots & \ddots & \vdots & \vdots \\
e_{d-3} & \cdots & e_{d-3} & 0 \\
\end{pmatrix}.
\]

Here the powers of locators $e_j$ are represented with respect to some basis of $F_{q^m}$ over $F_q$. Note that the redundancy of $C$ is at most $(d-3)n+1$. Also, we assume in the sequel that $q$ does not divide $d - 2$, since code $C$ has constructive distance $d$ instead of $d - 1$ otherwise.

Consider any nonzero codeword $c \in C$ of weight $w$ with nonzero symbols in positions $j_1, \ldots, j_w$. Let $X(c) = \{x_1, \ldots, x_w\}$ denote its locator set, where we use notation $x_i = e_{j_i}$ for all $i = 1, \ldots, w$. We say that $X(c)$ lies on an affine line $L(a, b)$ over $F_q$ if there exist $a, b \in F_{q^m}$ such that

\[
x_i = a + \lambda_i b
\]

where $\lambda_i \in F_q$ for all values of $i = 1, \ldots, w$.

The key observation underlying our code construction is that under some restrictions on extension $m$ and characteristic char $F_q$ of the field $F_q$, any code vector $c \in C$ of weight $d - 1$ has its locator set $X(c)$ lying on some affine line. Formally, this is expressed by the following.

**Theorem 4:** Let $m$ be a prime, $m > (d - 3)!$ and char $F_q > d - 3$. Consider the extended BCH code $C_{q^m}(d - 1)$ of constructive distance $d - 1$. Then, any codeword $c$ of minimum weight $d - 1$ has its locator set $X(c)$ lying on some affine line $L(a, b)$ over $F_q$.

We defer the proof of Theorem 4 till Section IV and proceed with the code construction. Let

\[
s = \lceil m/(d - 2) \rceil, \quad \mu = s(d - 2).
\]

Consider the field $F_{q^s}$ and its subfield $F_{q^n}$. Let $g = \{g_1, \ldots, g_n\}$ be the basis of $F_{q^n}$ over $F_q$ such that $F_{q^n}$ is spanned by $\{g_1, \ldots, g_n\}$. Let $h = \{h_1, \ldots, h_m\}$ be an arbitrary basis of $F_{q^m}$ over $F_{q^n}$ in what follows, we map each element $x = \sum_{i=1}^{m} a_i h_i$ of the field $F_{q^m}$ onto the element

\[
\hat{x} = \sum_{i=1}^{m} \alpha_i g_i
\]

of the field $F_{q^n}$. It is readily seen that for arbitrary $a, b \in F_{q^n}$ and $\lambda \in F_q$

\[
a + \lambda b = \bar{a} + \lambda \bar{b}
\]

Recall that the norm [15] of $\hat{x} \in F_{q^n}$

\[
N_{F_{q^n}/F_{q^s}}(\hat{x}) = N_{d=2}(\hat{x}) = \hat{q}^{d-3} \hat{x}_1^{1} + \cdots + \hat{q}^{d-1} \hat{x}_d^{1}
\]

is a classical mapping from $F_{q^n}$ to $F_{q^s}$.

Now we are ready to present our code construction. Consider the $q$-ary code $C'(n, k', d')$ of length $n = q^m$ with the parity-check matrix

\[
\hat{H}_{q^n}^{m} = \begin{pmatrix}
1 & \cdots & 1 & 1 \\
e_1 & \cdots & e_{n-1} & 0 \\
\vdots & \ddots & \vdots & \vdots \\
e_{d-3} & \cdots & e_{d-3} & 0 \\
N_{d=2}(\hat{e}_1) & \cdots & N_{d=2}(\hat{e}_{n-1}) & 0
\end{pmatrix}
\]

where the locators $e_j$ and their powers are represented in $F_{q^n}$ with respect to the basis $h$ and values of $N_{d=2}$ are represented in $F_q$ with respect to $g$. Recall that $N_{d=2}(\hat{x})$ takes values in $F_{q^n}$. Therefore, the redundancy of $C'$ does not exceed $(d - 3)n + s + 1$.

**Theorem 5:** Suppose $m > (d - 3)!$ is a prime, and char $F_q > d - 3$; then code $C'(n, k', d')$ defined by (13) has parameters

\[
[g^m, k' \geq q^m - (d - 3)m - \lceil m/(d - 2) \rceil - 1, d' \geq d]_{q^s}.
\]

**Proof:** Note that $d' \geq d - 1$, since $C'$ is a subcode of the extended BCH code $C$ defined in (8). Let $C'_{d=1} \subseteq C$ be the set of all codewords of weight exactly $d - 1$. It remains to prove that $C' \cap C_{d=1} = \emptyset$.

Assume the converse. Let $c \in C'$ be a codeword of weight $d - 1$ with locator set $X(c) = \{x_1, \ldots, x_{d-1}\}$. This implies that for some nonzero symbols $\xi_1, \xi_{d-1} \in F_q$

\[
\sum_{i=1}^{d-1} \xi_i x_i = 0, \quad t = 0, \ldots, d - 3
\]

\[
\sum_{i=1}^{d-1} \xi_i N_{d=2}(\hat{x}_i) = 0.
\]
Note that $c \in C_d$. Therefore according to Theorem 4, there exist $a, b \neq 0$ from $F_q$ and pairwise distinct $\{\lambda_i\} \in F_q$ such that
\[ x_i = a + \lambda_i b. \]
Consider the affine permutation $\pi(x) = A + Bx$ of the entire locator set $F_q$, where $A = -ab^{-1}$ and $B = b^{-1}$. Clearly, $\pi$ maps each $x_i$ onto $\lambda_i$, i.e.,
\[ \lambda_i = A + Bx_i. \]

It is well known ([1], [20]) that the extended BCH code $C$ is invariant under any affine permutation of the locators, so that $\{\lambda_i\}$ is also a locator set in $C_d$. Indeed, for any $t \in [0, d-3]$, we have an equality
\[
\sum_{i=1}^{d-1} \xi_i \lambda_i^t = \sum_{i=1}^{d-1} \xi_i (A + Bx_i)^t = \sum_{j=0}^t A^{t-j} B^j \left( \sum_{i=1}^{d-1} \xi_i x_i^j \right) = 0.
\]

The proof is completed.

\[ \square \]

The proof is completed.

Lemma 6: Suppose $\text{char } F_q > d - 3$; then
\[ c(q, d) \leq (d - 3) + 1/(d - 2). \]

Proof: We estimate the asymptotic redundancy of the family of codes presented in Theorem 5. Here $q$ and $d$ are fixed, while $m > (d - 3)!$ runs to infinity over primes. Then
\[
c(q, d) \leq \lim_{m \to \infty} \left( \frac{(d - 3)m + \left[ m/(d - 2) \right] + 1}{m} \right).
\]

The proof is completed.

IV. AFFINE LINES

Before we proceed to the proof of Theorem 4, let us introduce some standard concepts and theorems of algebraic geometry. Let $F$ be an algebraically closed field and $r, t$ be two positive integers. Let $f_1, \ldots, f_r \in F[x_1, \ldots, x_t]$. For any $x = (a_1, \ldots, a_t) \in F^t$, the matrix
\[
J_x(f_1, \ldots, f_r) = \begin{pmatrix}
\frac{\partial f_1}{\partial x_1} | x \\
\vdots \\
\frac{\partial f_r}{\partial x_t} | x
\end{pmatrix}
\]
is called the Jacobian of functions $f_i$ at point $x$.

The set $V$ of common roots to the system of equations
\[
\begin{align*}
    f_1(x_1, \ldots, x_t) &= 0 \\
    \vdots \\
    f_r(x_1, \ldots, x_t) &= 0
\end{align*}
\]
is called an affine variety. The ideal $I(V)$ is the set of all polynomials $f \in F[x_1, \ldots, x_t]$ such that $f(x) = 0$ for all $x \in V$. One important characteristic of a variety is its dimension $\dim V$. Dimension of a nonempty variety is a nonnegative integer. Let $x = (a_1, \ldots, a_t) \in V$ be an arbitrary point on $V$. The dimension of a variety $V$ at a point $x$, denoted $\dim_x V$, is the maximum dimension of an irreducible component of $V$ containing $x$. A point $x \in V$ such that $\dim_x V = 0$ is called an isolated point.

We shall need the following lemma ([19, p. 166]).

Lemma 7: Let $V$ be an affine variety with the ideal
\[ I(V) \subset F[x_1, \ldots, x_t]. \]

Then for any $x = (a_1, \ldots, a_t) \in V$ and $f_1, \ldots, f_r \in I(V)$
\[ \text{rank } J_x(f) \leq t - \dim_x V. \]

The next lemma is a corollary to the classical Bezout’s theorem ([16, p. 53]).
Lemma 8: Let $V$ be an affine variety defined by (20). Then the number of isolated points on $V$ does not exceed
$$
\prod_{i=1}^r (\deg f_i).
$$
Let $\xi_1, \ldots, \xi_{t+1}$ be fixed nonzero elements of some finite field $F_q$. Consider a variety $V$ in the algebraic closure of $F_q$ defined by the following system of equations:
$$
\begin{cases}
\xi_1 x_1 + \cdots + \xi_t x_t + \xi_{t+1} = 0 \\
\xi_1 x_1^2 + \cdots + \xi_t x_t^2 + \xi_{t+1} = 0 \\
\vdots \\
\xi_1 x_1^{d_i} + \cdots + \xi_t x_t^{d_i} + \xi_{t+1} = 0
\end{cases}
$$
(21)
Let $x = (a_1, \ldots, a_t)$ be an arbitrary point on $V$. We say that $x$ is an interesting point if $a_i \neq a_j$ for all $i \neq j$.

Lemma 9: Let $V$ be the variety defined by (21). Suppose $\text{char } F_q > t$; then every interesting point on $V$ is isolated.

Proof: Let $x = (a_1, \ldots, a_t)$ be an arbitrary interesting point on $V$. Let $f_i(x_1, \ldots, x_t)$ denote the left-hand side of the $i$th equation of (21). Consider the Jacobian of $\{f_i\}$ at point $x$:
$$
J_x(f_1, \ldots, f_t) = \begin{pmatrix}
\xi_1 & \cdots & \xi_t \\
2\xi_1 a_1 & \cdots & 2\xi_t a_t \\
\vdots & & \vdots \\
t\xi_1 a_1^{d_i-1} & \cdots & t\xi_t a_t^{d_i-1}
\end{pmatrix}
$$
Thus, we have
$$
\det J_x(f_1, \ldots, f_t) = t! \prod_{i=1}^t \xi_i \begin{vmatrix}
1 & \cdots & 1 \\
a_1 & \cdots & a_t \\
\frac{a_1^{d_i-1}}{} & \cdots & \frac{a_t^{d_i-1}}{}
\end{vmatrix}.
$$
Using standard properties of the Vandermonde determinant and the facts that $\xi_i$ are nonzero and $\text{char } F_q > t$, we get
$$
\text{rank } J_x(f_1, \ldots, f_t) = t.
$$
(22)
It is easy to see that $f_1, \ldots, f_t \in I(V)$. Combining (22) with Lemma 7, we obtain $\text{dim}_q V = 0$. The proof is completed.

Lemma 10: Let $m$ be a prime $m > t!$. Assume $\text{char } F_q > t$. Let $V$ be the variety defined by (21). Suppose $x \in F_q^m$ is an interesting point on $V$; then $x \in F_q^t$. In other words, every interesting point on $V$ that is rational over $F_q^m$ is rational over $F_q$.

Proof: Assume the converse. Let $x = (a_1, \ldots, a_t)$ be an interesting point on $V$ such that $x \in F_q^m \setminus F_q^t$. Consider the following $m$ conjugate points:
$$
p_k = (a_1^q, \ldots, a_t^q), \quad \text{for all } 0 \leq i \leq m - 1.
$$
Each of the above points is interesting. Since $m$ is a prime, the points are pairwise distinct. However, according to Lemma 9, every interesting point on $V$ is isolated. Thus, we have $m > t!$ isolated point on $V$. This contradicts Lemma 8.

Remark 11: Note that we can slightly weaken the condition of Lemma 10 replacing $m > t!$ with the following condition: $\forall s \neq 1, s|m$ implies $s > t!$.

Now we are ready to prove Theorem 4.

Proof: Assume $C_{d-1}$ is nonempty (this fact will be proven later) and consider the locator set $X(c) = (x_1, \ldots, x_{d-1})$ for any $c \in C_{d-1}$. Recall that $X(c)$ satisfies the first $(d-2)$ equations in (14) where $\xi_i \neq 0$ for all $i$. Assume an affine permutation $\pi(x) = a + bx$ of the locator set $F_q^m$ of the code $C$. Let $a, b \neq 0 \in F_q^m$ be such that
$$
\pi(x_{d-2}) = 1 \quad \text{and} \quad \pi(x_{d-1}) = 0.
$$
(23)
Let $y_c$ denote $\pi(x_i)$. Now we again use the fact that code $C_8$ is invariant under affine permutations. Therefore, the new locator set $y(c) = (y_1, \ldots, y_{d-3}, 1)$ satisfies similar equations
$$
\begin{cases}
\xi_1 + \cdots + \xi_{d-3} + \xi_{d-2} = -\xi_{d-1} \\
\xi_1 y_1 + \cdots + \xi_{d-3} y_{d-3} + \xi_{d-2} = 0 \\
\xi_1 y_2 + \cdots + \xi_{d-3} y_{d-3} + \xi_{d-2} = 0 \\
\vdots \\
\xi_1 y_{d-3} + \cdots + \xi_{d-3} y_{d-3} + \xi_{d-2} = 0
\end{cases}
$$
(24)
Now we remove the first equation (which does not include variables $y_c$) from (24), and obtain the system of equations, which is identical to system (21) for $t = d-3$. Recall that $x_1, \ldots, x_{d-1}$ are pairwise distinct elements of $F_q^m$. Therefore, $y_1, \ldots, y_{d-3}, 1, 0$ are also pairwise distinct. Thus, $y_1, \ldots, y_{d-3}$ is an interesting solution to the above system.

It is straightforward to verify that all the conditions of Lemma 10 hold. This yields
$$
y_c = a + bx = \lambda_i \in F_q^t, \quad \forall i \in [1, d-1].
$$
Thus, we obtain all locators $x_i$ on the affine line
$$
x_i = -\frac{a}{b} + \frac{\lambda_i}{b}, \quad \lambda_i \in F_q^t.
$$
Finally, we prove that $C_{d-1}$ is nonempty. Note that $\text{char } F_q \geq d-2$. Also, recall that we consider codes $C_{d-1}$ with constructive distance $d-1$, in which case $q$ does not divide $d-2$. Thus, we now assume that $q \geq d-1$. Then we consider (24) taking $\xi_{d-1} = 1$ and arbitrarily choosing $d-3$ different locators $y_1, \ldots, y_{d-3}$ from $F_q^t \setminus \{0, 1\}$. Obviously, the resulting system of linear equations has nonzero solution $\xi_1, \ldots, \xi_{d-2}$. This gives the codeword of weight $d-1$ and completes the proof of Theorem 4.

V. CONCLUSION

We have constructed an infinite family of nonbinary codes that reduce the asymptotic redundancy of BCH codes for any given alphabet size $q$ and distance $d$ if $q \geq d-1$. Families with such a property were earlier known only for distances 4, 5, and 6 [6]. Even the shortest codes in our family have very big length $n \approx q^{(d-2)!}$, therefore, the construction is of theoretical interest.

The main question (i.e., the determination of the exact values of $c(q,d)$) remains open.

ACKNOWLEDGMENT

S. Yekhanin would like to express his deep gratitude to M. Sudan for introducing the problem to him and many helpful
discussions during this work. He would also like to thank J. Kelner for valuable advice.

REFERENCES


