On the covering radius of an unrestricted code as a function of the rate and dual distance

Simon Litsyn\textsuperscript{a}, Patrick Solé\textsuperscript{b}, René Struik\textsuperscript{a}

\textsuperscript{a}Department of Electrical Engineering-Systems. Tel-Aviv University. Ramat-Aviv 69978. Israel
\textsuperscript{b}Lab. d'Informatique, CNRS-URA 1376. ESSI, BP 145. Route des Colles, 06 903 Sophia-Antipolis. France.

Received 31 October 1996; received in revised form 6 August 1997; accepted 29 September 1997

Abstract

We present a uniform approach towards deriving upper bounds on the covering radius of a code as a function of its dual distance structure and its cardinality. We show that the bounds obtained previously by Delsarte, Helleseth et al., Tietäväinen, resp. Solé and Stokes follow as special cases. Moreover, we obtain an asymptotic improvement of these bounds using Chebyshev polynomials. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Coding theory; Dual distance; Covering radius; Asymptotic bounds; Chebyshev polynomials; MacWilliams transform

1. Introduction

It is well-known (see, e.g., [4]) that the covering radius of a code can be related to other code parameters. A classical example is Delsarte’s bound [6], which states that the covering radius of a code is at most the number of nonzero coefficients in the MacWilliams transform of its distance distribution. Helleseth et al. [11] were the first to obtain an upper bound on the covering radius of a code in terms of its dual distance alone. Of a more recent date are upper bounds on the covering radius by Tietäväinen [21], in terms of the dual distance and the location of the zeros of Krawtchouk polynomials, and by Solé and Stokes [19], in terms of the dual distance structure and the code rate. These bounds, and related ones, were obtained using a variety of methods: combinatorial and probabilistic methods [11, 18], algebraic methods [6, 21, 22], and a combination of both algebraic and combinatorial methods [19].

In this paper, we present a general framework for deriving upper bounds on the covering radius of a code in terms of its dual distance structure and its cardinality. We show that this approach allows for uniform and transparent proofs of the previously
reported bounds by Delsarte, Helleseth et al., Tietäväinen, and Solé-Stokes. Moreover, our method supplemented with the use of the extremal properties of Chebyshev polynomials in the spirit of [3] yields a new upper bound on the covering radius, which (partially) improves the asymptotic bounds obtained in [19, 21]. A companion paper [12] is devoted to the more elementary case of linear codes. Here we focus our attention on unrestricted codes. For very recent results in the same vein see [9, 10].

The paper is organized as follows. In Section 2, we review some basic coding theory and mention some simple results regarding Krawtchouk polynomials. In Section 3, we consider in greater detail relations between the weight distribution of a code and its MacWilliams transform and derive some technical results to facilitate the exposition in Section 4. In that section, we show how the relations between the weight distribution of a code and its MacWilliams transform can be used to obtain an upper bound on the covering radius. The method is based upon a variation of the linear programming approach. Using this method, we re-establish the best known upper bounds and obtain a new upper bound, using Chebyshev polynomials. We restrict ourselves to binary codes.

2. Preliminaries

In this section, we review some results from coding theory and state several properties of Krawtchouk polynomials that will be used later on.

We adopt the notations of [13]. The all-zero and all-one vectors are denoted by \( \mathbf{0} \) and \( \mathbf{1} \), respectively.

2.1. Krawtchouk polynomials

Here we mention several properties of Krawtchouk polynomials that will be used in the rest of the paper. For details we refer to [13, Section 1.2].

For \( k = 0, 1, 2, \ldots \) the Krawtchouk polynomial \( K_k(z; n) \) is defined by

\[
K_k(z; n) = \sum_{j=0}^{k} (-1)^j \binom{z}{j} \binom{n-z}{k-j}, \quad \text{where } z \in \mathbb{R}.
\]  

(1)

If the parameter \( n \) is clear from context, we simply write \( K_k(z) \) instead of \( K_k(z; n) \). Notice that \( K_k(n-z; n) = (-1)^k K_k(z; n) \).

We will also need certain orthogonality relations between the Krawtchouk polynomials and some information on the locations of the zeros of these polynomials.

The Krawtchouk polynomials satisfy the following relation:

\[
\sum_{i=0}^{n} \binom{n}{i} K_k(i) K_l(i) = \delta_{kl} \binom{n}{k} 2^n.
\]

(2)
It follows that the polynomials \( \{K_k(z; n)\}_{k=0}^{n} \) form an orthogonal basis of the vector space of all polynomials in \( \mathbb{R}[z] \) of degree at most \( n \) with inner product
\[
\langle f(z), g(z) \rangle_n := \sum_{i=0}^{n} \binom{n}{i} f(i) g(i).
\] (3)

In the rest of this paper we will make extensive use of this orthogonality relation. For later use we mention that a simple calculation shows that
\[
\langle (n - x)f(z), g(z) \rangle_n = n\langle f(z), g(z) \rangle_{n-1}.
\] (4)

The symmetry property for Krawtchouk polynomials
\[
\binom{n}{i} K_i(i) - \binom{n}{l} K_l(l)
\]
yields that, apart from orthogonality relation (2), the Krawtchouk polynomials also satisfy another kind of orthogonality relation, viz.
\[
\sum_{i=0}^{n} K_k(i) K_l(i) = \delta_{k,l} 2^n.
\] (5)

Any polynomial \( \beta(z) \) of degree \( \leq n \) in \( \mathbb{R}[z] \) can be written uniquely as a linear combination of Krawtchouk polynomials, which is called the Krawtchouk expansion of \( \beta(z) \).

Orthogonality relation (5) can be used to show a relation between the Krawtchouk expansion of polynomial \( \beta(z) \) and that of its Fourier transform, i.e., the polynomial \( \gamma(z) = \sum_{i=0}^{n} \beta(i) K_i(z) \).

**Lemma 1.** Let \( \beta(z) = \sum_{i=0}^{n} \beta_i K_i(z) \) and let \( \gamma(z) = \sum_{i=0}^{n} \gamma_i K_i(z) \). Then \( \gamma_i = \beta(i) \) for all \( i \) iff \( \beta_i = 2^{-n} \gamma_i(j) \) for all \( j \).

**Proof.** Suppose \( \gamma(z) = \sum_{i=0}^{n} \beta(i) K_i(z) \). Then \( \gamma(z) = \sum_{k=0}^{n} \beta_k \sum_{i=0}^{n} K_k(i) K_i(z) \). From Eq. (5) we infer that
\[
\gamma(j) = \sum_{k=0}^{n} \beta_k \sum_{i=0}^{n} K_k(i) K_i(j) = \sum_{k=0}^{n} \beta_k \delta_{k,j} 2^n = \beta_j 2^n \quad \text{for all } j, 0 \leq j \leq n.
\]

The converse statement follows from the fact that \( \gamma(z) \) is uniquely determined by the function values \( \gamma(0), \ldots, \gamma(n) \). \( \square \)

The polynomial \( K_k(z; n) \) has \( k \) distinct zeros, which lie in the interval \( (0, n] \) if \( k \leq n \). If \( z(k,n) \) is the smallest zero of \( K_k(z; n) \) and \( k \leq n \), then
\[
z(k,n-1) < z(k,n) < z(k-1,n-1).
\] (6)

In general, the exact location of the zeros of \( K_k(z; n) \) is not known, but asymptotically it is known [14, p. 563] that if \( 0 < r < \frac{1}{2} \), \( n \to \infty \), and \( k/n \to \tau \) then
\[
z(k,n)/n \to \frac{1}{2} - \sqrt{\tau(1-\tau)}.
\] (7)
The Krawtchouk polynomials of degree up to two are

\[ K_0(z; n) = 1 \] without zeros;

\[ K_1(z; n) = n - 2z \] with zero \( \frac{1}{2}n \);

\[ K_2(z; n) = 2z^2 - 2nz + \binom{n}{2} \] with zeros \( \frac{1}{2}(n \pm \sqrt{n}) \).

2.2. Basic concepts from coding theory

The set of all binary \( n \)-tuples forms a vector space denoted by \( \mathbb{F}_2^n \). The Hamming distance \( d(x, y) \) between two words \( x, y \in \mathbb{F}_2^n \) is defined by \( d(x, y) := |\{ i \mid x_i \neq y_i \}| \). The weight \( wt(x) \) of a word \( x \in \mathbb{F}_2^n \) is defined by \( wt(x) := d(x, 0) \). A binary code \( \mathcal{C} \) of length \( n \) is a nonempty subset of \( \mathbb{F}_2^n \). If \( \mathcal{C} \) has cardinality \( M \), \( \mathcal{C} \) is called an \((n, M)\) code. A linear code of length \( n \) is a linear subspace of \( \mathbb{F}_2^n \). If \( \mathcal{C} \) is a subspace of dimension \( k \), \( \mathcal{C} \) is called an \([n, k]\) code. If \( \mathcal{C} \) is a linear code and \( x \in \mathbb{F}_2^n \), then \( x + \mathcal{C} \) is called a coset of \( \mathcal{C} \). The vectors of minimum weight in such a coset are called coset leaders. If \( \mathcal{C} \) is an \([n, k]\) code, its dual code \( \mathcal{C}^\perp \) is the \([n, n-k]\) code defined by

\[ \mathcal{C}^\perp := \{ y \in \mathbb{F}_2^n \mid \langle x, y \rangle = 0 \text{ for all } x \in \mathcal{C} \}, \]

where \( \langle x, y \rangle \) denotes the standard inner product of \( x \) and \( y \).

The minimum distance \( d \) of a code \( \mathcal{C} \) is the minimum value of \( d(x, y) \) over all pairs of different codewords \( x, y \in \mathcal{C} \). The covering radius \( r \) of a code \( \mathcal{C} \) is the maximum value of \( d(x, \mathcal{C}) \) over all words \( x \in \mathbb{F}_2^n \). Here \( d(x, \mathcal{C}) \) is defined by \( d(x, \mathcal{C}) := \min\{ d(x, c) \mid c \in \mathcal{C} \} \). For linear codes, the covering radius is the highest weight of any coset leader of the code.

3. Weight and distance distributions

In this section we consider relations between the weight and distance distribution of a code and their MacWilliams transforms. These relations will be used in Section 4 to obtain an upper bound on the covering radius.

Let \( \mathcal{C} \) be a code of length \( n \) and let \( A_i \) be the number of codewords of weight \( i \). The sequence \( \{ A_i \}_{i=0}^n \) is called the weight distribution of code \( \mathcal{C} \). The MacWilliams transform of the weight distribution \( \{ A_i \}_{i=0}^n \) is the sequence \( \{ B_j \}_{j=0}^n \) defined by

\[ B_j = \frac{1}{|\mathcal{C}|} \sum_{i=0}^n A_i K_j(i; n) \quad (j = 0, \ldots, n). \tag{8} \]

Here the numbers \( \{ K_j(i; n) \}_{j=0}^n \) are the Krawtchouk coefficients defined by Eq. (1). If \( \mathcal{C} \) is a linear code with weight distribution \( \{ A_i \}_{i=0}^n \), then the sequence \( \{ B_j \}_{j=0}^n \) defined by Eq. (8) is the weight distribution of the dual code \( \mathcal{C}^\perp \). If \( \mathcal{C} \) is not linear, the MacWilliams transform does not have a natural interpretation. We can still consider this sequence, though.
For all $x \in \mathbb{F}_2^n$, let $A_i(x)$ be the number of codewords at distance $i$ from $x$. It follows that the coset $x + \mathcal{C}$ has weight distribution $\{A_i(x)\}_{i=0}^n$. The MacWilliams transform of $\{A_i(x)\}_{i=0}^n$ is defined by

$$B_i(x) := \frac{1}{|\mathcal{C}|} \sum_{i=0}^n A_i(x)K_j(i; n).$$

The distance distribution $\{A_i(\mathcal{C})\}_{i=0}^n$ of code $\mathcal{C}$ is

$$A_i(\mathcal{C}) := \frac{1}{|\mathcal{C}|} \sum_{x \in \mathcal{C}} A_i(x).$$

Its MacWilliams transform $\{B_i(\mathcal{C})\}_{i=0}^n$ is defined by

$$B_i(\mathcal{C}) := \frac{1}{|\mathcal{C}|} \sum_{i=0}^n A_i(\mathcal{C})K_j(i; n).$$

Let

$$D' := \{1 \leq j \leq n \mid B_j(\mathcal{C}) \neq 0\}.$$ 

The smallest integer in this set is called the dual distance $d'$ of code $\mathcal{C}$.

All numbers $B_j(\mathcal{C})$ are nonnegative, as was proved by Delsarte in the language of association schemes [6, Theorem 3.3].

**Theorem 2.** Let $\mathcal{C}$ be a code of length $n$. Then

1. [14, Ch. 5, Theorem 5] $B_j(\mathcal{C}) = |\mathcal{C}|^{-1} \sum_{y \in \mathcal{C}} \sum_{j=1}^{n-1} (-1)^{(x,y)} \sum_{c \in \mathcal{C}} (-1)^{(c,y)}.$
2. [14, Ch. 5, Theorem 6] $B_j - |\mathcal{C}|^{-2} \sum_{y \in \mathcal{C}} \left( \sum_{c \in \mathcal{C}} (-1)^{(c,y)} \right)^2 \geq 0.$

**Proof.** (1) From Eq. (1) it follows directly that

$$\sum_{y \in \mathcal{C}} (-1)^{(x,y)} = K_k(i; n) \quad \text{if } x \in \mathbb{F}_2^n \text{ has weight } i.$$ (12)

Hence, from the definition of $B_j(x)$ we infer that

$$|\mathcal{C}|B_j(x) = \sum_{i=0}^n A_i(x)K_j(i; n)$$

$$= \sum_{c \in \mathcal{C}} K_j(d(x, c); n)$$

$$= \sum_{c \in \mathcal{C}} \sum_{y \in \mathcal{C}} (-1)^{(x-y)}$$

$$= \sum_{y \in \mathcal{C}} (-1)^{(x,y)} \sum_{c \in \mathcal{C}} (-1)^{(c,y)}.$$
(2) From property (1) of this theorem and the definition of $B_j$ we infer that
\[ |\mathcal{C}|^2 B_j = |\mathcal{C}| \sum_{x \in \mathcal{C}} B_j(x) \]
\[ = \sum_{y: w(y) = j} \left( \sum_{x \in \mathcal{C}} (-1)^{(c, x)} \right)^j \geq 0. \]

The next lemma, a direct consequence of Theorem 2, proves to be useful in the rest of the paper.

**Lemma 3.** Let $\mathcal{C}$ be a code of length $n$. Then
1. $B_j = 0$ if $\sum_{c \in \mathcal{C}} (-1)^{(c, x)} = 0$ for all $y \in \mathbb{F}_2^n$ with weight $j$.
2. $B_j = 0$ if $B_j(x) = 0$ for all $x \in \mathbb{F}_2^n$.

The following lemma, apparently due to Delsarte [6], shows that the weight distribution $\{A_i(x)\}_{i=0}^n$ of code $x + \mathcal{C}$ and its dual distribution, the sequence $\{B_j(x)\}_{j=0}^n$, can be related via a polynomial. The result turns out to be very powerful in our context.

**Lemma 4 (Basic Lemma).** Let $\mathcal{C}$ be a code of length $n$ and let $\beta(z) := \sum_{j=0}^n \beta_j K_j(z; n)$ be a polynomial in $\mathbb{R}[z]$. Then
1. $\sum_{i=0}^n A_i(x) \beta(i) = |\mathcal{C}| \sum_{j=0}^n \beta_j B_j(x)$ for all $x \in \mathbb{F}_2^n$,
2. $\sum_{i=0}^n A_i(x) \beta_j = |\mathcal{C}| 2^{-n} \sum_{j=0}^n \beta(j) B_j(x)$ for all $x \in \mathbb{F}_2^n$.

**Proof.** (1) Let $x \in \mathbb{F}_2^n$. By definition of $B_j(x)$ we have
\[ \sum_{i=0}^n A_i(x) \beta(i) = \sum_i A_i(x) \sum_j \beta_j K_j(i; n) \]
\[ = \sum_j \beta_j \sum_i A_i(x) K_j(i; n) \]
\[ = |\mathcal{C}| \sum_j \beta_j B_j(x). \]

(2) The result follows from property (1) of this theorem and Lemma 1. □

If the polynomial $\beta(z)$ has a more restricted form, the weight distribution of a code satisfies certain linear equations.

**Corollary 5.** Let $\mathcal{C}$ be a code of length $n$. Let $\beta(z) := \sum_{j=0}^n \beta_j K_j(z; n)$ be a polynomial in $\mathbb{R}[z]$ for which $\beta_j = 0$ if $j \in D'$. Then
\[ \sum_{i=0}^n A_i(x) \beta(i) = |\mathcal{C}| \beta_0 \quad \text{for all } x \in \mathbb{F}_2^n. \]

**Proof.** Let $x \in \mathbb{F}_2^n$. From Lemma 3 we infer that $B_j(x) = 0$ for all $0 \neq j \notin D'$. Moreover, by definition of $B_j(x)$ we have $B_0(x) = 1$. The result now follows from Lemma 4. □
Delsarte [7, Theorem 2.2] proved that if a code has dual distance \( d' \) and if it has \( s \) nonzero distances, then \( s \geq \lfloor (d' - 1)/2 \rfloor \). This bound was referred to as the dual MacWilliams inequality in [6, Eq. (5.37)]. We will need a slightly stronger result.

**Theorem 6** (Struik [20]). Let \( \mathcal{C} \) be a code of length \( n \) with dual distance \( d' \). Let \( x \in \mathbb{F}_2^n \) and let \( W(x) := \{ i \mid A_i(x) \neq 0 \} \). Then \( |W(x)| \geq \lfloor (d' - 1)/2 \rfloor + 1 \).

**Proof.** Let \( t := \lfloor (d' - 1)/2 \rfloor \). We will show that the only polynomial in \( \mathbb{R}[z] \) of degree at most \( t \) that is zero on \( W(x) \) is the zero polynomial, thus proving that \( |W(x)| \geq t + 1 \).

Let \( x_0, \ldots, x_t \in \mathbb{R} \) and suppose that
\[
\sum_{j=0}^{t} x_j K_j(z) = 0 \quad \text{on} \ W(x). \tag{13}
\]

To prove the theorem we will show that \( x_0 = \cdots = x_t = 0 \). Let \( 0 \leq j \leq t \). From (13) we infer that
\[
\sum_{w=0}^{n} K_j(w)A_{w}(x) \sum_{j=0}^{t} x_j K_j(w) = \sum_{j=0}^{t} x_j \sum_{w=0}^{n} A_{w}(x)K_j(w)K_j(w) = 0. \tag{14}
\]

We now consider the last summation in more detail. For all \( i \), \( 0 \leq i \leq t \), let
\[
S_{ij} = \sum_{w=0}^{n} A_{w}(x)K_i(w)K_j(w). \tag{15}
\]

The polynomial \( \beta(z) := K_i(z)K_j(z) \) has degree \( i + j \leq 2t < d' \). Denote the Krawtchouk expansion of this polynomial by \( \beta(x) = \sum \beta_k K_k(x) \). Using the detailed orthogonality relation (2), we find that \( \langle \beta(z), 1 \rangle_n = \langle \beta(z), K_0(z) \rangle_n = \beta_0 2^n \). Therefore,
\[
\beta_0 2^n = \langle \beta(z), 1 \rangle_n = \langle K_i(z)K_j(z), 1 \rangle_n = \langle K_i(z), K_j(z) \rangle_n = \delta_{ij} \binom{n}{t} 2^n. \tag{16}
\]

Combining Corollary 5 with Eqs. (15) and (16), we infer that
\[
S_{ij} = \delta_{ij} \binom{n}{j} |\mathcal{C}|. \tag{17}
\]

From Eqs. (14), (15), and (17) we now obtain the following result:
\[
\sum_{j=0}^{t} x_j \sum_{w=0}^{n} A_{w}(x)K_i(w)K_j(w) = x_j \binom{n}{j} |\mathcal{C}| = 0. \tag{18}
\]

It follows, that \( x_j = 0 \). \( \square \)

The MacWilliams transform of the weight distribution of code \( x + \mathcal{C} \) can be estimated via the dual distance distribution of code \( \mathcal{C} \).
Lemma 7. Let $C$ be a code of length $n$ and let $x \in \mathbb{F}_2^n$. Then
1. $|B_j(x)| \leq \sqrt{\binom{n}{j}}B_j$.
2. $|B_j(x)| \leq B_j$ if $C$ is a linear code.

Proof. From property (1) of Theorem 2 we infer that $\langle C, B_j(x) \rangle$ can be considered as the inner product $\langle e, 1 \rangle$ over the reals, where $e$ is the vector of length $\binom{n}{j}$ with real components $e(y)$ defined by

$$e(y) := (-1)^{(x,y)} \sum_{c \in C} (-1)^{(c,y)}$$

where $y \in \mathbb{F}_2^n$ has weight $j$.

1. By the Cauchy–Schwartz inequality we have $|\langle C, B_j(x) \rangle|^2 = \langle e, 1 \rangle^2 \leq \langle e, e \rangle \langle 1, 1 \rangle$. Using property 2 of Theorem 2, we find that $|B_j(x)| \leq \sqrt{\binom{n}{j}}B_j$.

2. Suppose $C$ is a linear code. If $y \in C^\perp$, then the inner product $\langle c, y \rangle$ always assumes the value 0. If $y \not\in C^\perp$, then the inner product $\langle c, y \rangle$ assumes the values 0 and 1 equally often, since $C$ is a linear code. It follows, that $e(y)$ has value $\pm|C|$ if $y \in C^\perp$ and value 0 if $y \not\in C^\perp$. Hence we find that $|B_j(x)| \leq B_j$. \qed

Lemma 8 (Solé and Stokes \cite{Sol}). Let $C$ be a code of length $n$ and let $x \in \mathbb{F}_2^n$. Then
1. $\sum_{j=0}^n |B_j(x)| \leq 2^n/\sqrt{|C|}$.
2. $\sum_{j=0}^n |B_j(x)| \leq 2^n/|C|$ if $C$ is a linear code.

Proof. (1) It follows from property 1 of Lemma 7 and the Cauchy–Schwartz inequality that

$$\sum_{j=0}^n |B_j(x)| \leq \sum_{j=0}^n \sqrt{\binom{n}{j}} \cdot B_j \leq \sqrt{2^n \cdot \left( \sum_{j=0}^n B_j \right)^{1/2}} = 2^n/\sqrt{|C|}.$$

(2) If $C$ is a linear code, then it follows from property 2 of Lemma 7 that

$$\sum_{j=0}^n |B_j(x)| \leq \sum_{j=0}^n B_j = 2^n/|C|.$$

\qed

4. Upper bounds on the covering radius

Linear programming has proved to be a highly successful technique to obtain upper bounds on the size of error-correcting codes. In this section we use a variation of this linear programming approach to derive upper bounds on the covering radius of a code as a function of its dual distance structure and its cardinality. The main idea is to use Lemma 4, which shows that the weight distribution of a code and its dual distribution can be related via a polynomial. By a proper choice of this polynomial one obtains an upper bound on the covering radius of the code. We show that all results follow from one theorem (Theorem 9), thus offering a uniform approach.
Theorem 9 (Main Theorem). Let $\mathcal{C}$ be a code of length $n$. Let $\beta(z) := \sum_{i=0}^{n} \beta_i K_i(z; n)$ be a polynomial in $\mathbb{R}[z]$. For all $x \in \mathbb{F}_2^n$ let

$$S(x) := \sum_{i=0}^{n} A_i(x) \beta(i). \quad (19)$$

Suppose $\beta(z) \leq 0$ for all integers in the interval $(0, n]$. Then $d(x, \mathcal{C}) \leq 0$ in each of the following two cases:

1. $S(x) > 0$.
2. $S(x) = 0$ and $\beta(z)$ has at most $\lfloor (d' - 1)/2 \rfloor$ integral zeros in the interval $(0, n]$.

If $\mathcal{C}$ is a self-complementary code (i.e. $\mathcal{C}$ is invariant under the translation $x \rightarrow x + 1$), then the bound on the covering radius remains valid if one replaces the constraints on the interval $(0, n]$ by constraints on the smaller interval $(0, n - 1)$.

Proof. Let $x \in \mathbb{F}_2^n$.

1. If $S(x) > 0$, then it follows from Eq. (19) that not all the numbers $A_i(x)$ with $i \leq 0$ can be zero, hence $d(x, \mathcal{C}) \leq 0$.
2. Let $S(x) = 0$ and suppose that $d(x, \mathcal{C}) > 0$. It follows from Eq. (19) that $A_i(x) \beta(i) = 0$ for all $i > 0$, i.e. $\beta(z)$ is zero on the set $W(x) := \{i \mid A_i(x) \neq 0\}$. By Theorem 6 this set has cardinality $|W(x)| \geq \lfloor (d' - 1)/2 \rfloor + 1$ and hence $\beta(z)$ has more than $\lfloor (d' - 1)/2 \rfloor$ integral zeros on the interval $(0, n]$. This proves the statement.

If $\mathcal{C}$ is a self-complementary code, then $A_i(x) = A_{n-i}(x)$. Therefore, the result remains valid if we replace the constraints on the interval $(0, n]$ by constraints on the smaller interval $(0, n - 1) = (0, n] \cap [0, n - 1)$. \hfill $\square$

Remark 10. It might be interesting to point out that the set of admissible polynomials $\beta(z)$, relative to a given $\theta$ and a given $x$, is a convex cone (without its vertex $0$).

Remark 11. The theorem can also be formulated in terms of the Fourier transform of polynomial $\beta(z)$, but, for reasons of space, it is omitted.

As an application of this theorem, we derive a number of upper bounds on the covering radius of a code, including all best known bounds reported. We consider two cases separately, viz., the case where in Theorem 9 $S(x)$ does not depend on the actual choice of vector $x$ (Type I) and the general case (Type II).

4.1. Upper bounds: Type I

In this subsection, we re-establish the upper bounds on the covering radius obtained by Delsarte [6], Helleseth et al. [11], and Tietäväinen [21].
Theorem 12. Let \( \mathcal{C} \) be a code of length \( n \). Let \( \beta(z) := \sum_{j=0}^{n} \beta_j K_1(z; n) \) be a polynomial in \( \mathbb{F}_2[z] \) for which \( \beta_j = 0 \) if \( j \in D' \). Suppose \( \beta(z) < 0 \) for all integers in the interval \( (0, n] \). Then \( \mathcal{C} \) has covering radius at most 0 in each of the following two cases:

1. \( \beta_0 > 0 \),
2. \( \beta_0 = 0 \) and \( \beta(x) \) has at most \( \lfloor (d' - 1)/2 \rfloor \) integral zeros in the interval \( (0, n] \).

If \( \mathcal{C} \) is a self-complementary code (i.e. \( \mathcal{C} \) is invariant under the translation \( x \to x + 1 \)), then the bound on the covering radius remains valid if we replace the constraints on the interval \( (0, n] \) by constraints on the smaller interval \( (0, n-1) \).

Proof. Let \( x \in \mathbb{F}_2^n \). From Corollary 5 we infer that

\[
S(x) := \sum_{i=0}^{n} A_i(x) \beta(i) = |\mathcal{C}| \beta_0. \tag{20}
\]

The result now follows from Theorem 9.

Remark 13. In [21] the same result was proved, but only for polynomials \( \beta(z) \) of degree at most \( d' - 1 \) for which \( \beta_0 > 0 \).

Any polynomial \( \beta(z) \) that satisfies the conditions of Theorem 12 yields an upper bound on the covering radius of a code \( \mathcal{C} \) as a function of the set \( D' \).

Example 14. Let \( \mathcal{C} \) be a self-complementary code of length \( n \) with dual distance \( d' \geq 3 \). We want to find an upper bound on the covering radius of this code by considering the polynomial \( \beta(z) = \beta_0 K_0(z) + \beta_1 K_1(z) + K_2(z) = \beta_0 + \beta_1 (n-2z) + (2z^2 - 2nz + \frac{1}{2} n(n-1)) \). Choose \( \beta_0 \geq 0 \) and \( \beta_1 \) in such a way that \( \beta(0) = \beta(n-1) = 0 \), where \( \theta < \frac{1}{2} n \). Since \( \beta(z) < 0 \) on the interval \( (\theta, n-\theta) \), the assumptions of Theorem 12 are satisfied and hence \( \theta \) is an upper bound on the covering radius of \( \mathcal{C} \). For each \( \beta_0 \) in the range \([0, \frac{1}{2} n]\) we find that \( \beta_1 = 0 \) and \( \theta(n-\theta) = \frac{1}{2} \beta_0 + \frac{1}{4} n(n-1) \), i.e. \( \theta_1 = \frac{1}{2} (n \pm \sqrt{n-2\beta_0}) \). Clearly, one obtains the best bound if \( \beta_0 = 0 \), i.e. if \( \beta(z) = K_2(z) \). It follows, that \( \mathcal{C} \) has covering radius \( r(\mathcal{C}) \leq \frac{1}{2} (n - \sqrt{n}) \).

Remark 15. This result was originally proved by Helleseth et al. [11] and was referred to as the Norse bound in [5]. In fact, the proof of [11, Theorem 3] already uses the polynomial \( \beta(z) = K_2(z; n) \) implicitly. The bound in the above example is tight, since the first-order Reed–Muller code \( R(1, m) \) of length \( n = 2^m \), \( m \) even, has dual distance \( d' = 4 \) and covering radius \( r = \frac{1}{2} (n-\sqrt{n}) \). If \( x \) has maximal distance to this Reed–Muller code, then the coset \( x + R(1, m) \) only contains words of weights \( w_{1,2} = \frac{1}{2} (n \pm \sqrt{n}) \).

Theorem 12 yields an upper bound on the covering radius of a code \( \mathcal{C} \) as a function of the set \( D' \). The best upper bounds known on the covering radius were obtained by Tietäväinen [21] and Delsarte [6] and depend on the dual distance \( d' = \min D' \), resp. on the number \( s' = |D'| \). First we give Tietäväinen’s bound.
Proof. Our proof is based upon an application of Theorem 12. For any polynomial \( \beta(z) \) we denote its Krawtchouk expansion by \( \beta(z) = \sum \beta_i K_i(z; n) \). In order to apply Theorem 12, we need to know \( \beta_0 \). Recall from Section 2.1 that the Krawtchouk polynomials \( \{K_i(z; n)\}_{i=0}^n \) form an orthogonal basis of the vector space of all polynomials in \( \mathbb{R}[z] \) of degree at most \( n \) with inner product \( \langle f(z), g(z) \rangle_n \). Using the detailed orthogonality relation (2), we find that \( \langle \beta(z), 1 \rangle_n = \langle \beta(z), K_0(z) \rangle_n = \beta_0 2^n \).

1. Recall that \( z(t,n) \) stands for the smallest zero of \( K_i(z) := K_i(z; n) \). Let \( \beta(z) := -K_i^2(z)/(z - z(t,n)) \). Since \( K_i(z) \) has degree \( t \) and \( K_i(z)/(z - z(t,n)) \) is a polynomial of degree \( t - 1 \), we find that

\[
\beta_0 2^n = \langle \beta(z), 1 \rangle_n = \langle \beta(z), K_0(z)/(z - z(t,n)) \rangle_n = 0.
\]

Notice that \( \beta(z) \leq 0 \) on \( (z(t,n), n] \) and has \( t - 1 \) distinct zeros on this interval. It follows that if \( \mathcal{C} \) has dual distance \( d' \geq 2t \), then \( \beta(z) \) satisfies the assumptions of Theorem 12 with \( \theta = z(t,n) \). Therefore code \( \mathcal{C} \) has covering radius \( r \leq z(t,n) \).

2. Let \( \beta(z) = (z - n)K_i^2(z)/(z - z(t,n - 1)) \). Since \( K_i(z) \) has degree \( t \) and \( K_i(z)/(z - z(t,n - 1)) \) is a polynomial of degree \( t - 1 \), we find that

\[
\beta_0 2^n = \langle \beta(z), 1 \rangle_n = \langle (z - n)K_i(z), K_i(z)/(z - z(t,n - 1)) \rangle_n = 0.
\]

Here we used Eq. (4). Notice that \( \beta(z) \leq 0 \) on \( (z(t,n - 1), n] \) and has \( t \) distinct zeros on this interval. It follows, that if \( \mathcal{C} \) has dual distance \( d' \geq 2t + 1 \), then \( \beta(z) \) satisfies the assumptions of Theorem 12 with \( \theta = z(t,n - 1) \). Therefore, code \( \mathcal{C} \) has covering radius \( r \leq z(t,n - 1) \).

Notice that the upper bounds on the covering radius can only be attained if all the zeros of the Krawtchouk polynomial \( K_i(z) \) in the proof are integers. When applying Theorem 16, the exact value of \( z(k,n) \) is usually not known. In that case one can use the estimate for \( z(k,n) \) given in Section 2.1.

Now we prove Delsarte’s bound.

**Theorem 17** (Delsarte [6]). Let \( \mathcal{C} \) be a code of length \( n \). Then \( \mathcal{C} \) has covering radius at most \( s' := |D'| \).
Proof. Our proof is based upon an application of Theorem 12 and Lemma 5. Let \( \sigma(z) \) be the annihilator polynomial of code \( \mathcal{C} \), i.e.,

\[
\sigma(z) = \frac{2^n}{|\mathcal{C}|} \prod_{j \in D'} \left(1 - \frac{z}{j}\right). \tag{21}
\]

Let \( \beta(z) = \sum \beta_j K_j(z; n) \) and let \( \gamma(z) = \sum \gamma_j K_j(z; n) \) be the Fourier transform of \( \beta(z) \).

If we choose \( \gamma(z) = \sigma(z) \), then \( \gamma(z) \) has degree \( s' \), \( \gamma(j) = 0 \) for all \( j \in D' \), and \( \gamma(0) = 2^n |\mathcal{C}|^{-1} \). By Lemma 1 we now have \( \beta_j = |\mathcal{C}|^{-1} > 0 \), \( \beta_j = 0 \) for all \( j \in D' \), and \( \beta(s' + 1) = \cdots = \beta(n) = 0 \). It follows that \( \beta(z) \) satisfies the conditions of Theorem 12 with \( \theta = s' \). Therefore, code \( \mathcal{C} \) has covering radius at most \( s' \). \( \square \)

4.2. Upper bounds: Type II

In this subsection, we investigate general criteria under which Theorem 9 can be applied. As an application, we derive the upper bound on the covering radius originally obtained by Solé and Stokes [19]. Moreover, we obtain an absolute and asymptotic improvement of their result, using Chebyshev polynomials.

Theorem 18. Let \( \mathcal{C} \) be a code of length \( n \). Let \( \beta(z) = \sum_{j=0}^{n} \beta_j K_j(z; n) \) be a polynomial in \( \mathbb{R}[z] \). Suppose \( \beta_j \leq 0 \) for all integers \( j \) in the interval \((0, n]\). Then \( \mathcal{C} \) has covering radius at most \( \theta \) in each of the following two cases:

1. \( \beta(0) > \max_{k \in D'} |\beta(k)|(2^n/|\mathcal{C}| - 1) \),
2. \( \beta(0) > \max_{k \in D'} |\beta(k)|(2^n/|\mathcal{C}| - 1) \) and \( \mathcal{C} \) is a linear code.

Proof. Let \( x \in \mathbb{F}_2^n \). By Lemma 4 we have

\[
S(x) := \sum_{i=0}^{n} A_i(x) \beta(i) - |\mathcal{C}| \sum_{j=0}^{n} \beta_j B_j(x). \tag{22}
\]

We claim that in each of the two cases considered in the theorem we have \( S(x) > 0 \). If so, then the result follows from Theorem 9, after application of the Fourier transform (Lemma 1). To prove this claim, observe that quantity \( S(x) \) can be bounded from below as follows:

\[
S(x) = |\mathcal{C}| \sum_{j=0}^{n} \beta_j B_j(x) \geq |\mathcal{C}| \left( \beta_0 - \sum_{j \in D'} \beta_j B_j(x) \right). \tag{23}
\]

We have

\[
\left| \sum_{j \in D'} \beta_j B_j(x) \right| \leq \max_{k \in D'} |\beta_k| \cdot \sum_{j \in D'} |B_j(x)| \leq \max_{k \in D'} |\beta_k| \left( \sum_{j=0}^{n} |B_j(x)| - 1 \right). \tag{24}
\]
Combining Lemma 8 with Eqs. (23) and (24), we find that $S(x) > 0$ in each of the two cases of the theorem. Hence the claim follows. □

**Remark 19.** Notice that Theorem 18 generalizes Case I of Theorem 12 (after application of the Fourier transform).

Any polynomial that satisfies the conditions of Theorem 18 yields an upper bound on the covering radius of code $\mathcal{C}$ as a function of the set $D'$ and the cardinality of $\mathcal{C}$.

**Example 20.** Let $\mathcal{C}$ be a code of length $n$ with $D' \subset \{a, \ldots, b\}$. We want to find an upper bound on the covering radius of this code by considering a suitable power of the polynomial $\gamma(z) = ((a + b - 2z)/(b - a))$. We have $\gamma(a) = 1 = -\gamma(b)$ and $\gamma(0) = (a + b)/(b - a) > 1$. Now choose integer $t$ such that $t > 2^n/\sqrt{|\mathcal{C}|}$ if $\mathcal{C}$ is a linear code, and $t > 2^n/|\mathcal{C}|$ if $\mathcal{C}$ is a nonlinear code. In both cases it is clear that $\gamma(t)$ satisfies the conditions of Theorem 18 with $\theta = t$. Therefore code $\mathcal{C}$ has covering radius

$$r \leq \begin{cases} \left\lfloor \frac{\log_2(2^n/|\mathcal{C}|)}{\log_2(\frac{a+b}{b-a})} \right\rfloor & \text{if } \mathcal{C} \text{ is a linear code}, \\
\left\lfloor \frac{n}{2} + \frac{1}{2} \frac{\log_2(2^n/|\mathcal{C}|)}{\log_2(\frac{a+b}{b-a})} \right\rfloor & \text{if } \mathcal{C} \text{ is a nonlinear code}. \end{cases}$$

**Remark 21.** This result was originally proved by Sole and Stokes [10], for the case $a + b = n$. Interestingly, their result is based upon a combinatorial interpretation of a generalization of the Camion-Courteau-Delsarte formula [1]. For details we refer to that paper.

From Theorem 18 it is clear that if $\beta(z)$ has fixed degree and $\beta(0) = 1$, then an optimum choice for polynomial $\beta(z)$ satisfying the conditions of the theorem will be obtained if one chooses $\beta(z)$ in such a way that $\max_{k \in D'} |\beta(k)|$ is minimal. This strongly suggests the use of Chebyshev polynomials.

**Theorem 22.** Let $\mathcal{C}$ be a code of length $n$ with $D' \subset \{a, \ldots, b\}$. Then $\mathcal{C}$ has covering radius $r$ with

$$r \leq \begin{cases} \frac{\cosh^{-1}(2^n/|\mathcal{C}|)}{\cosh^{-1}(\frac{a+b}{b-a})} & \text{if } \mathcal{C} \text{ is a linear code}, \\
\frac{\cosh^{-1}(2^n/\sqrt{|\mathcal{C}|})}{\cosh^{-1}(\frac{a+b}{b-a})} & \text{if } \mathcal{C} \text{ is a nonlinear code}. \end{cases}$$
Proof. We use the Chebyshev polynomials $T_\ell(z)$ (see e.g. [2, 16]) written in the following explicit form:

$$T_\ell(z) = \begin{cases} 
\cosh(\ell \cosh^{-1}(z)) & \text{if } z \geq 1, \\
\cos(\ell \arccos(z)) & \text{if } -1 \leq x \leq 1, \\
(1)\ell T_\ell(z) & \text{if } z \leq 1.
\end{cases}$$

(Recall that $\cosh^{-1}(z) = \ln(z + \sqrt{z^2 - 1}).$)

We want to find an upper bound on the covering radius of the code by considering the polynomial $\gamma_\ell(z)$ defined by

$$\gamma_\ell(z) = \frac{T_\ell\left(\frac{a+b-z}{b-a}\right)}{T_\ell\left(\frac{a+b}{b-a}\right)}.$$

We have $\gamma_\ell(a) = (T_\ell((a+b)/(b-a)))^{-1} = (-1)^\ell \gamma_\ell(b)$ and $\gamma_\ell(0) = 1$. Moreover, we have

$$\gamma_\ell(a) = \max \{|\gamma_\ell(x)| : a \leq x \leq b\}.$$ 

Now choose integer $\ell$ in such a way that

$$\frac{\gamma_\ell(0)}{\gamma_\ell(a)} = T_\ell\left(\frac{a+b}{b-a}\right) > 2^n / |\mathcal{C}|$$

if $\mathcal{C}$ is a linear code, and such that

$$\frac{\gamma_\ell(0)}{\gamma_\ell(a)} = T_\ell\left(\frac{a+b}{b-a}\right) > 2^n / \sqrt{|\mathcal{C}|}$$

if $\mathcal{C}$ is a nonlinear code. In both cases it is clear that $\gamma_\ell(z)$ satisfies the conditions of Theorem 18 with $\theta = \ell$. This proves the theorem. □

To compare asymptotically the bounds of Theorems 16 and 22, one should use (7) in the first one and the best known upper bounds on the size of codes with given minimum distance (e.g. the linear programming bound [15]) in the second one. As it is shown in [12], the linear case of the bound of Theorem 22 improves on the bound of Theorem 16 for codes of growing length $n$ with $d' > 0.298n + o(n)$.

Acknowledgements

We are grateful to the referee for many useful comments which led to a significant improvement of the style of the paper. The first author would like to thank I. Honkala, T. Laihonen and A. Tietäväinen for helpful discussions.
References