New upper bounds for binary covering codes

Patric R.J. Östergård\textsuperscript{a,*,1}, Markku K. Kaikkonen\textsuperscript{b}

\textsuperscript{a}Department of Mathematics and Computing Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, Netherlands
\textsuperscript{b}Halkosuontie 65 C 5, 00660 Helsinki, Finland

Received 23 April 1996

Abstract

Improved upper bounds are presented for $K(n, r)$, the minimum cardinality of a binary code of length $n$ and covering radius $r$. The new bounds are obtained by both new and old constructions; in many of these, computer search using simulated annealing and tabu search plays a central role. Some new linear covering codes are also presented. An updated table of upper bounds on $K(n, r)$, $n \leq 64$, $r \leq 12$, is given.

1. Introduction

We consider upper bounds on $K(n, r)$, the minimum cardinality of a (linear or non-linear) binary code of length $n$ and covering radius $r$. Such bounds are obtained by constructing a corresponding covering code. During the last decade — after the milestone papers \cite{3, 5, 12} appeared — many new covering codes have been found. As this paper shows, however, many improvements can still be obtained.

A binary code $C \subseteq F_2^n$ ($F_2 = \{0, 1\}$ is the field of two elements) with covering radius $r$ is said to be an $(n, |C|)r$ code. If it is a linear code with $2^k$ codewords, the notation $[n, k]r$ may be used instead. An $(n, K(n, r))r$ code is called an optimal covering code. Exact values of $K(n, r)$ are known only in the following cases: $K(6, 1) = 12$ \cite{40}, $K(2^k - 1, 1) = 2^{2^k - k - 1}$ (Hamming codes), $K(2^k, 1) = 2^{2^k - k}$ \cite{21}, $K(23, 3) = 4096$ (Golay code), and for all $n \leq 2r + 3$ (see \cite{5}). (The Hamming and Golay codes are perfect codes, that is, any word in the space is at distance at most $r$ from exactly one codeword.) For all other parameters there are gaps between the best known lower and upper bounds. In this paper only upper bounds — which are constructive — are considered. For a table of lower bounds on $K(n, r)$ for $n \leq 33$, $r \leq 10$, see \cite{25}.

\* Correspondence address: Department of Computer Science, Helsinki University of Technology, Otakaari 1, 02150 Espoo, Finland. E-mail: patric.ostergaard@hut.fi.

\^ Supported by the Academy of Finland and the Walter Ahlström Foundation.

0012-365X/98/$19.00$ Copyright © 1998 Elsevier Science B.V. All rights reserved

PII S0012-365X(96)00380-9
Constructions of binary covering codes are considered in Section 2. Note that we do not intend to give an extensive survey of construction methods — for such a survey, the reader is referred to [4] — but only consider methods used to obtain new codes. Methods discussed are both combinatorial and based on computer search. A table of upper bounds on $K(n,r)$ for $n \leq 64$, $r \leq 12$ is displayed in Section 3.

2. Constructions of codes

In this section we consider methods that are used to construct new record-breaking codes. In the theorems only the bounds are given; however, codes can be obtained by following the constructions in the proofs and in the references.

2.1. Computer search for codes

Computer search has come to play a central role in constructions of good covering codes. In particular, the break-through of new heuristics in combinatorial optimization during the last decade has made this approach even more attractive; several new methods can be used effectively to search for covering codes.

In trying to minimize an upper bound for $K(n,r)$, with $n$ and $r$ fixed, we also fix the number of codewords, $|C|$. This value can be set slightly better than the old bound. The task is then to find a code $C \subseteq F_2^n$ such that $d(x,C) \leq r$ for all $2^n$ words in the space. For such a code, the value of

$$f(C) = |\{x \in F_2^n \mid d(x,C) > r\}|$$

is 0. We now have a combinatorial minimization problem with cost function $f(C)$. The first results using this approach were published by Wille in 1987 [43]. He considered ternary codes, and applied an optimization method called simulated annealing [24]. The results were very promising, and inspired research on simulated annealing based searches for covering codes. A recent study by Östergård [35] shows that another optimization method, tabu search [10], performs even better in many cases. For a detailed survey of these results, see [19].

With increasing value of $n$ — and thus increasing number of words in the space $F_2^n$ — it becomes more and more difficult to find good coverings in a direct search. One can then use the so-called matrix method, which was presented in its present form by Blokhuis and Lam [1]. They only treated codes with covering radius $r = 1$, but it is not difficult to show that it works for any $r$ [42].

Let $A = [a_1 \ a_2 \ \ldots \ a_n]$ be a $k \times n$ binary matrix. A set $S \subseteq F_2^k$ is said to $r$-cover $F_2^k$ using $A$ if for all $x \in F_2^k$, $x = s + At$ has a solution with $s \in S$ and $\text{wt}(t) \leq r$ ($\text{wt}(t)$ denotes the Hamming weight of the vector $t$, that is, the number of nonzero coordinates). The following theorem is proved in [42] for $A = [I_k \ M]$, where $I_k$ is the $k \times k$ identity matrix. It is not difficult to show that the theorem holds also when the rank of $A$ is less than $k$ (see [22, Corollary 1]).
Theorem 1. If \( S \) \( r \)-covers \( F_2^k \) using \( A \), then \( K(n, r) \leq |S|2^{n-k} \).

The codewords of the covering code are explicitly \( \{x \in F_2^n \mid Ax \in S\} \). The matrix method imposes a structure on the code — the code is a union of cosets of a linear code — and (for a given matrix \( A \)) reduces the search space. On the other hand, it is now not so obvious how the search for a code should be done, since we have to find both the elements in \( S \) and the columns of \( A \). For a survey of different approaches to perform this search, see [19]. The new codes found by this method are listed in hexadecimal notation in Table 1. All matrices \( A \) are here in form \( A = [I_k \ M] \). The columns of \( M \) are first given; these are separated from the words in \( S \) by a semi-colon.

The \((13,752)1\) and \((19,320)4\) codes in Table 1 are not best known. They are included since they will be used in constructions later on — the first one already in the next theorem. The result is analogous to a construction for error-correcting codes presented in [37].

Theorem 2. If \( |S| > 1 \) and \( S \) \( 1 \)-covers \( F_2^k \) using \( A \), then \( K(|S| - 1, 1) \leq (n+1)2^{|S|-k-1} \).

**Proof.** Adding a binary vector to all words in \( S \) does not affect the fact that \( S \) is a covering, so we can assume that \( S = \{s_1 = 0, s_2, \ldots, s_{|S|}\} \), where \( 0 \) is the all-zero vector. Furthermore, let \( T = \{a_1, a_2, \ldots, a_n\} \) be the set of column vectors in \( A \). Since \( S \) \( 1 \)-covers \( F_2^k \) using \( A \), all vectors in \( F_2^k \) can be written in one of the following forms: \( 0 \ (s_i) \), \( s_i \), \( a_j \ (s_i + a_j) \), or \( s_i + a_j \), where \( 2 \leq i \leq |S| \), \( 1 \leq j \leq n \).

Now the elements in \( S \) (except the all-zero vector) and \( T \) are swapped, giving \( S' = \{s_1\} \cup T \) and \( T' = S \setminus \{s_1\} = \{t_1', t_2', \ldots, t_{|S|-1}'\} \). It is easily checked that the vectors

<table>
<thead>
<tr>
<th>Code</th>
<th>( k )</th>
<th>Words of ( M ) and ( S )</th>
</tr>
</thead>
<tbody>
<tr>
<td>((13,752)1)</td>
<td>9</td>
<td>(1C0, 1B8, 174, 33; 5, A, 10, 24, 28, 2F, 30, 3F, 56, 59, 73, 7C, 96, 99, B3, BC, C3, CC, DF, E0, E7, EB, F5, FA, 116, 119, 133, 13C, 141, 14E, 155, 157, 167, 168, 169, 172, 182, 18D, 19A, 1A4, 1A6, 1AB, 1B1, 1D6, 1D9, 1F3, 1FC, 1FF.)</td>
</tr>
<tr>
<td>((18,2944)2)</td>
<td>12</td>
<td>(771, 6DC, 56E, 4B7, 7B8, 650; 0, 6E2, 3DA, 45B, AAD, 30E, 5C5, 1DD, 9F4, 418, 47E, 2FA, 1A0, D52, 989, A44, 86A, D0, 8C6, A1C, FCC, AD3, B27, 675, 267, C69, A7F, 7E7, F81, C31, D64, FBB, 59D, 43, 5BB, 978, E1F, ECB, 3FF, CA6, 62D, 235, 9BB, FCE, 716, 7BB.)</td>
</tr>
<tr>
<td>((17,320)3)</td>
<td>11</td>
<td>(27B, 52A, 117, 39D, 5F, 792; 0, 5A9, 5FD, 474, 1C9.)</td>
</tr>
<tr>
<td>((19,320)4)</td>
<td>13</td>
<td>(672, 1952, 15AE, 164E, 894, 1BAB, 0, 5A2, 1DE0, 1BF4, E57.)</td>
</tr>
</tbody>
</table>
$s', s'+t'$, for all possible $s' \in S', t' \in T'$, coincide with those enumerated above, so $S'$
1-covers $F_2^k$ using $A'=[t'_1, t'_2, \ldots, t'_{|S|-1}]$, and the theorem follows from Theorem 1. □

Since there are $K(n, 1)$ words that 1-cover $F_2^n$ using $I_n$, we get the following corollary.

**Corollary 1.** $K(K(n, 1) - 1, 1) \leq (n+1)2^{K(n, 1)-n-1}$.

These results imply that we can search for binary codes with covering radius 1 with many words in $S$ and few columns in $M$, or with few words in $S$ and many columns in $M$. The former alternative is generally preferred.

**Example 1.** From $K(5, 1) = 7$, we can deduce $K(6, 1) \leq 12$ using Corollary 1; actually, $K(6, 1) = 12$.

**Example 2.** From $K(9, 1) \leq 62$, we get that $K(61, 1) \leq 5 \cdot 2^{53}$ using Corollary 1.

**Example 3.** The $(13, 752)1$ code in Table 1 is obtained by the matrix method with the following parameters: $n = 13$, $k = 9$, and $|S| = 47$. Theorem 2 then gives that $K(46, 1) \leq 7 \cdot 2^{38}$.

Also the following theorem gives good bounds for codes with covering radius 1; see [5, Theorem 13] for a proof.

**Theorem 3.** $K(2n+1, 1) \leq 2^nK(n, 1)$.

A generalization of this result, presented in [16, Theorem 4; 18], can be used to obtain some best known codes with $r > 1$; acting on a $(n,M)r$ code, the generalized construction gives a $(2n+1,M')r$ code with $M' \approx 2^{n+1}M/r$.

If $S$ contains only the all-zero vector, the codes obtained by the matrix method are linear codes. The new linear codes found by us are listed in Table 2 using hexadecimal notation (only the $M$ part of the parity check matrix is given). An [18,9]3 code is also included. The existence of a code with these parameters was reported in [2], but no such code was listed in that paper. We found several nonequivalent [18,9]3 codes.

The four new codes improve on the best known linear covering codes [27]. All of these, except the [37,25]3 code, have fewer codewords than any previously known nonlinear codes with the same parameters. When the new codes are combined with the Golay code in the ADS construction (see next subsection), three more new linear codes are obtained: [59,36]6, [41,19]7, and [45,22]7 codes.
Table 2
Linear codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Parity check matrix (M part)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[18,9]³</td>
<td>1A0, 174, A5, 173, 17, E8, 9, 18D, 1CF.</td>
</tr>
<tr>
<td>[37,25]³</td>
<td>BFF, 727, 977, 93B, 6B9, 51D, AAE, 9B8, 85B, A2D, B16, 9C5, AE2, 6E2, 371, 5B8, 2DC, 16E, B7, 45B, 62D, 716, 38B, 5C5, 7FF.</td>
</tr>
<tr>
<td>[19,8]³</td>
<td>4EA, 771, 6, 86, 1CD, 3B4, 17E, 7AB.</td>
</tr>
<tr>
<td>[23,11]³</td>
<td>5C5, B9D, 6E2, CF9, 5D4, 275, FB7, DB, 2F5, 753, 166.</td>
</tr>
<tr>
<td>[24,10]5</td>
<td>2D96, 33A6, 1E30, 3551, 35EA, 7F, 3E03, 6BF, 1933, 2BD5.</td>
</tr>
</tbody>
</table>

2.2. Combining codes

Many of the best known codes are constructed by combining other codes. For example, the new codes in Tables 1 and 2 can act as seeds in such constructions. It is easily shown that by combining all words in an \((n_1,M_1)r_1\) code and an \((n_2,M_2)r_2\) code, an \((n_1+n_2,M_1M_2)r_1+r_2\) code is obtained. We can do slightly better than this direct sum construction if the codes involved are normal. The amalgamated direct sum construction (ADS) and the concept of normality were presented for linear codes in [12] and were later generalized to the nonlinear case in [5].

Let \(C\) be an \((n,M)r\) code with \(C_a^{(i)} = \{(c_1, c_2, \ldots, c_n) \in C | c_i = a\}\) and \(C_a^{(i)} = \{(c_1, \ldots, c_{i-1}, c_i, \ldots, c_n) | (c_1, \ldots, c_{i-1}, a, c_{i+1}, \ldots, c_n) \in C\}\). \(C_a^{(i)}\) and \(C_b^{(i)}\) are assumed to be nonempty for all \(i\).

**Definition 1.** \(C\) has norm \(N^{(i)}\) with respect to the \(i\)th coordinate if \(d(x, C_a^{(i)}) + d(x, C_b^{(i)}) \leq N^{(i)}\) for all \(x \in F_2^n\). If \(N^{(i)} \leq N\) for at least one coordinate \(i\), \(C\) is said to have norm \(N\), and coordinates for which this equation holds are called acceptable. A code with norm \(2r+1\) is said to be normal.

Let \(A\) be an \((n_A,M_A)r_A\) normal code and let \(B\) be an \((n_B,M_B)r_B\) normal code. Before taking the ADS of two normal codes, we assume that the coordinates have been permuted so that \(A\) and \(B\) are normal with respect to the last and the first coordinate, respectively. The ADS of these codes is \(C = A \oplus B = A_0^{(n_A)} \oplus B_0^{(n_B)} \cup A_1^{(n_A)} \oplus B_1^{(n_B)}\).

**Theorem 4** (Cohen et al. [5, Theorem 11]). \(C\) is an \((n_A + n_B - 1, |A_0^{(n_A)}||B_0^{(n_B)}| + |A_1^{(n_A)}||B_1^{(n_B)}|)r_C\) code, where \(r_C \leq r_A + r_B\), and \(C\) is normal if equality holds.

These results are important, since codes obtained by the ADS construction are sometimes used once more in this construction. We thus want the new codes to be normal, so that we only have to check normality of the seeds. Although Theorem 4 does not assure that the new code is normal (due to the possibility that \(r_C < r_A + r_B\)), we know...
that \( C \) has norm \( 2(r_A + r_B) + 1 \) \([5, \text{Eq. (43)}]\), so \( C \) can be used in constructions as a normal code with covering radius \( r_A + r_B \).

We have checked the normality of all new codes in this paper that are used in the ADS construction. Normality of new linear codes with minimum distance at most 4 follows from a result in [20]. Old codes have been proved normal in, for example, [5, 13].

We now define some other useful concepts. The codes in these definitions are not necessarily binary.

**Definition 2.** A code \( C \) with covering radius \( r \) has \((k, t)\)-subnorm \( S \) if there is a partition \( C = C_0 \cup C_1 \cup \cdots \cup C_{k-1} (|C_i| > 0) \) such that \( \min_i d(x, C_i) + \max_j d(x, C_j) \leq S \) whenever \( r - t \leq d(x, C) \leq r \). Such a partition is called acceptable. If \( C \) has \((k, t)\)-subnorm \( 2r + 1 \), it is called \((k, t)\)-subnormal. Furthermore, if \( C \) is \((k, 0)\)-subnormal, it is called \( k \)-seminormal.

**Definition 3.** A code \( C \) is called strongly \( k \)-seminormal if there is a partition \( C = C_0 \cup C_1 \cup \cdots \cup C_{k-1} \) such that all \( C_i \) have covering radius \( r + 1 \). Again, such a partition is called acceptable.

The concept of subnormality was introduced by Honkala in [15]. A lot of research has been done on the conjectures (see, for example, [5]) that among all optimal binary codes with given parameters there are normal and subnormal (that is, \((2, r)\)-subnormal) codes. Despite this effort, the conjectures have only been partially settled.

The first result on normality of codes is apparently the following theorem, which was proved already in 1964 by Golomb and Posner.

**Theorem 5** (Golomb and Posner [11, Theorem 23]). A code \( C = F_2^{2^k - 1} \), which has covering radius 0, is strongly \( 2^k \)-seminormal (and thus has \((2^k, 0)\)-subnorm 1).

**Proof.** Let \( C \) be the union of the \([2^k - 1, 2^k - k - 1]_1 \) Hamming code and its cosets. \( \square \)

The general problem of finding good partitions of codes is very difficult. For small codes, however, a computer-based search is feasible; this approach will be further discussed in Section 2.4. See also [34].

**Example 4.** The following four sets of codewords (in hexadecimal form) constitute an acceptable partition of an \((11, 192)_1 \) code with \((4, 1)\)-subnorm 3:

\[
C_1: \text{C1D, 53, 91, A8, E6, F7, F9, 106, 13F, 141, 169, 16E, 192, 1BA, 1C4, 230, 270, 28B, 2CB, 327, 35E, 3A5, 3DC, 423, 45A, 4A1, 4D8, 534, 574, 58F, 5CF, 602, 63B, 645, 66A, 66D, 696, 6BE, 6C0, 708, 719, 757, 795, 7AC, 7E2, 7F3, 7FD;}
\]
Among several nonequivalent \((11,192)_1\) codes, we were only able to prove that this one is \((4,1)\)-subnormal.

**Example 5.** The \((19,320)_4\) code in Table 1 is strongly 5-seminormal, since its matrix \(A\) is a parity check matrix for a \([19,6]_5\) code.

The blockwise direct sum (BDS) is an important construction involving codes with properties defined in Definitions 2 and 3.

**Theorem 6** (Honkala [17, Theorem 1]). Let \(A\) be a code with covering radius \(r_A\), \((k,r_A)\)-subnorm \(S_A\), and the partition \(A = A_0 \cup A_1 \cup \cdots \cup A_{k-1}\) acceptable, and let \(B\) be a code with covering radius \(r_B\), \((k,|r_B - (S_B - S_A + 1)/2|)\)-subnorm \(S_B\), and the partition \(B = B_0 \cup B_1 \cup \cdots \cup B_{k-1}\) acceptable. Then the covering radius of the blockwise direct sum of \(A\) and \(B\) is \(r \leq \lfloor (S_A + S_B)/2 \rfloor\), where

\[
\text{BDS}(A,B) = \bigcup_{i=0}^{k-1} (A_i \oplus B_i).
\]

Theorem 6 can be used to obtain good codes when the subnorms of \(A\) and \(B\) have different parity. If both codes have odd subnorm, one of the codes can be extended before the BDS construction is applied.

**Theorem 7** (Struik [41, Lemma 4.15]). If a code with covering radius \(r\) and odd \((k,r)\)-subnorm \(S\) is extended by adding a parity check bit, a code with covering radius \(r'\) and even \((k,r')\)-subnorm \(S + 1\) is obtained.

**Example 6.** By taking the BDS of a code obtained by extending a code from Theorem 5 for \(k = 2\) — the extended code has \((4,1)\)-subnorm 2 — and the code in Example 4 with \((4,1)\)-subnorm 3, we get a record-breaking \((15,384)_2\) code. We further conjecture that there is a \((4,1)\)-subnormal \((14,1408)_1\) code; if this conjecture holds, an \((18,2816)_2\) code can be constructed in the same way.

**Example 7.** The code in Example 5 is strongly 5-seminormal. It is then also strongly 4-seminormal and has \((4,4)\)-subnorm 9. By taking the BDS of this code
and an extended code from Theorem 5 for \( k = 2 \), a \((23,640)5\) code is obtained.

Codes with nonbinary coordinates are useful in some cases. The discussion of constructions that combine codes will thus be continued in the next subsection.

### 2.3. Constructions acting on nonbinary codes

In the following we are going to consider nonbinary codes and codes over mixed alphabets, that is, codes in a space \( H = F_{q_1}^{n_1} \cdots F_{q_m}^{n_m} \), with \( m \geq 1, n_i \geq 0, \sum_{i=1}^{m} n_i \geq 1, \) and \( q_i \geq 2 \). Such a code with \( M \) words and covering radius \( r \) is denoted by \((q_1, q_2, \ldots, q_m; n_1, n_2, \ldots, n_m; M; r)\), and \( K_{q_1, q_2, \ldots, q_m}(n_1, n_2, \ldots, n_m; r) = \min \{M \mid \text{there is a } (q_1, q_2, \ldots, q_m; n_1, n_2, \ldots, n_m; M; r) \text{ code}\} \).

A strongly \( k \)-seminormal code and a code with a \( k \)-ary coordinate can be effectively combined.

**Theorem 8** (Östergård [30, Theorem 4]). Let \( A \) be a strongly \( k \)-seminormal code with covering radius \( r_A \) and \( A = A_0 \cup A_1 \cup \cdots \cup A_{k-1} \) an acceptable partition. Let \( B \) be a code with covering radius \( r_B \) whose first coordinate is \( k \)-ary. Then the code \( C = \bigcup_{i=0}^{k-1} A_i \oplus B_i^{(1)} \) has covering radius \( r_C \leq r_A + r_B \).

The following construction of binary covering codes was first used by Kabatyanskii and Panchenko [22] in their proof that (among other things) the density of an optimal binary covering code with covering radius one tends to 1 as \( n \) tends to infinity. Consider all \( q^k \) cosets of a \((q; n = (q^k - 1)/(q - 1); q^{n-k})1\) Hamming code, where \( q \) is a prime or a prime power. Transform these codes into \( q^k \)-ary codes (\( q^k > q \)) in the following way: Let \( a_0 + a_1 + \cdots + a_{q^k-1} = a' \), with \( a_i \geq 1 \). Encode the words in the codes in all possible ways using the scheme \( 0 \rightarrow 0_1, 0_2, \ldots, 0_{q^k}; 1 \rightarrow 1_1, 1_2, \ldots, 1_{q^k} \). It is easily seen that the new codes have covering radius 1 and partition the space \( F_q^n \). The cardinality of at least one of these \( q^k \) codes is then

\[
|C| \leq n^{q^k}/q^k. \tag{1}
\]

It turns out, however, that it is often possible to (slightly) improve this bound by choosing a proper representation of the original code. Stanton et al. studied such constructions in [23, 39]; for example, \( K_8(8,1) \leq 342272 \) follows from [23, Theorem 4], whereas (1) gives \( K_8(8,1) \leq 342392 \).

**Example 8.** By combining a \((8; 8; 342272)1\) code (8 times to get rid of all 8-ary coordinates) in Theorem 8 with a code \( A \) from Theorem 5 \((k = 3)\), a \((56,1337 \cdot 2^{40})1\) binary code is obtained.

Some of the codes obtained in the previous subsection can be constructed from a mixed code. In the next theorem, the original code must have even subnorm; if its
subnorm is odd, the code can be extended (cf. Theorem 7). The proof is trivial and is omitted.

**Theorem 9.** If a code $C$ has $(k,0)$-subnorm $2S$ and the partition $C = C_0 \cup C_1 \cup \cdots \cup C_{k-1}$ acceptable, then the code $\{(i,c) : c \in C_i\}$ has covering radius $S$.

**Example 9.** By using the code from Example 4 in Theorem 9, we get a $(4,2;1,12;192)_2$ code. This can further be used in Theorem 8 to get an alternative proof of $K(15,2) \leq 384$.

Theorem 9 also works in the other direction. By puncturing a code with covering radius $r$ by deleting a $k$-ary coordinate, we get a code that has $(k,0)$-subnorm $2r$ [31, Lemma 2]. Moreover, if the original code is a nontrivial perfect code with covering radius $r$, then the punctured code has $(k,r)$-subnorm $2r$ [17, Lemma 1].

**Example 10.** We can get one more proof of $K(15,2) \leq 384$ by taking the BDS of the code in Example 4 and a punctured perfect $(4,2;1,4;8)_1$ code [26].

2.4. **Partitioning codes using tabu search**

We mentioned earlier that the problem of finding an acceptable partition of a code with $(k,t)$-subnorm $S$ is very difficult in general. Computer search for such partitions using simulated annealing was considered in [31]. We will now briefly discuss a method based on tabu search. This approach has turned out to outperform the earlier approach based on simulated annealing.

As in [31], we set $t = 0$. If we find an acceptable partition of a code with $(k,0)$-subnorm $S$, we can check (using the same partition) whether it also has $(k,t)$-subnorm $S$ for $t > 0$ — that is usually the case.

Let $C \subseteq H$ be the original code. Initially we partition the code $C$ in some way into $k$ subcodes: $C = C_0 \cup C_1 \cup \cdots \cup C_{k-1}$. (One difference from the approach in [31] is that here the sizes of the subcodes $C_i$ are neither fixed in advance nor kept constant.) A cost function is related to such a partition:

$$E = \sum_{i=0}^{k-1} |\{x \in H : d(x,C_i) > S - r, \ d(x,C) = r\}|.$$

We now try to find a partition with $E = 0$, which then is a global minimum of this optimization problem. Tabu search is a local search method, that is, starting from the initial partition, a new partition that only slightly differs from the previous partition is chosen at each step of the search process. The new partition must belong to the neighborhood of the old one. We here define the neighborhood as the set of all partitions obtained by moving any one codeword to another subcode (the size of the neighborhood is thus $|C|(k - 1)$).

A feature of tabu search is that inverses of the $L$ (length of tabu list) recent moves are not allowed. We here add a codeword to the tabu list when it is moved to another
subcode. We can now express our heuristic in one sentence: At each step, taking the tabu list into account, choose the neighbor with smallest cost \( E \). If there are many neighbors with smallest cost, one is chosen at random. As \( L \) codewords are tabu, the number of moves that are considered at each step is \((|C| - L)(k - 1)\).

Strongly seminormal codes can be found using the same procedure and the following cost function:

\[
E = \sum_{i=0}^{k-1} |\{x \in H \mid d(x, C_i) > r + 1\}|.
\]

2.5. Codes with a nontrivial permutation group

The concept of piecewise constant codes was introduced in [5]. It is pointed out in [32] that a piecewise constant code can be seen as a union of orbits under a nontrivial permutation group. In the following examples, two new record-breaking codes with nontrivial permutation groups are given. These codes are all quasi-cyclic (some have fixed points). The coordinates in the examples are numbered from left to right, starting from 1.

Example 11. By applying the permutation group generated by \((2,3,4,5,6,7,8)(9,10,11,12,13,14,15)\) to the words in \{0000000000000000, 0000000111111111, 011000101000, 01110100010101\} and their complements, we get a \((15,32)4\) code. This code can also be obtained from a Hadamard matrix of order 16 by taking the rows and their complements as codewords and deleting one coordinate (such a code is called \(B_{16}\) in [28, p. 49]). Interestingly, two out of the five nonequivalent Hadamard matrices of order 16 give (nonequivalent) codes with the aforementioned parameters; the other three matrices lead to \((15,32)5\) codes.

Example 12. By applying the permutation group generated by \((1,2,\ldots,13)(14,15,\ldots,26)\) to the words in \{00000000000000000000000000, 000000000000011111111111, 11111111111111111111111111, 01010011001010010010011001, 001100100000111111111111\}, we obtain a \((26,56)9\) code.

Also the code in Example 12 can be obtained from a Hadamard matrix (of order 28). It would be interesting to know whether codes obtained by generalizing these constructions lead to infinite series of good covering codes. The following old bound can also be explained in the same manner.

Example 13 (G. Exoo [8]). By applying the permutation group generated by \((3,4,5,6,7)(8,9,10,11,12)\) to the words in \{00000000000000000000111111, 001111111110110000, 001111111110110000, 001100100101000101, 0100100010010101\} and their complements, we get a \((12,28)3\) code.
3. Table of upper bounds

The most extensive published tables of nonlinear covering codes have been up to $n = 33$ and $r = 10$. Our table, Table 3, is extended up to $n = 64$ and $r = 12$. There are several reasons for this extension. First, we have here presented new effective constructions for codes with $r = 1$ that give good codes only for long codes ($n > 33$). Second, also other researchers have recently presented constructions for long covering codes [7, 41]; without a proper reference table it is very hard to evaluate such results. Third, we want to pay the readers' attention to long codes with large covering radius. Such codes, which are here almost without exception obtained by the so-called ADS construction, are certainly for many parameters far from optimal.

Entries in Table 3 without keys follow from one of the following results (see, for example, [5, Eqs. (48), (53), (54)]).

**Theorem 10.** (a) $K(n, r) = 1$, for $n \leq r$;
(b) $K(n, r) \leq K(n + 1, r)$;
(c) $K(n + 1, r) \leq 2K(n, r)$.

Only one construction is given for each code, the one that in our opinion is the nicest. It is our intention to make it possible for the reader to construct codes corresponding to all bounds in the table. We have checked by computer that application of the ADS construction to any two codes in Table 3 does not render any improvements.

The keys are as follows:

- **ai** Amalgamated direct sum (Theorem 4), where one of the codes has the following parameters: (1) [3, 1] $^1$, (2) [7, 4] $^1$, (3) (13, 704) $^1$, (4) [15, 11] $^1$, (5) (23, 2$^{15}$) $^2$, (6) [39, 30] $^2$, (7) [23, 12] $^3$, (8) (31, 2$^{19}$) $^3$, (9) [19, 6] $^5$, (10) [24, 10] $^5$, (11) [31, 11] $^7$.

- **bi** BDS construction in one of the following references: (1) this paper (Examples 6 and 7), (2) [7, 41].

- **ci** Combinatorial construction in one of the following references: (1) this paper (Examples 11 and 12), (2) [16, 18], (3) [5], (4) [38].

- **li** Linear code listed in one of the following references: (1) this paper (Table 2), (2) [12], (3) [9], (4) [6].

- **mi** Code found by the matrix method and listed in one of the following references: (1) this paper (Table 1, Examples 2 and 3), (2) [36], (3) [33], (4) [13], (5) [14].

- **ni** Code with no known structure listed in one of the following references: (1) [44], (2) [36].

- **p** Perfect code (Hamming or Golay code).

- **si** Combination (Theorem 8) of strongly seminormal code(s) from Theorem 5 and a nonbinary code with the following parameters: (1) (4, 2; 1, 7; 60) $^1$ [29]; (2) (8; 8; 3427272) $^1$ [23] (see text).

$u$ $K(2n + 1, 1) \leq 2^n \cdot K(n, 1)$ (Theorem 3).
Table 3
Upper bounds on $K(n,r)$, $n \leq 64$, $r \leq 12$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$r$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>$7^{c3}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>$12^{c3}$</td>
<td>4</td>
<td>2</td>
<td></td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>$16^{p}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>$32$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>$64^{c1}$</td>
<td>$16^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>$120^{c1}$</td>
<td>$30^{a2}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>$192^{c3}$</td>
<td>$44^{a2}$</td>
<td>$16^{a1}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>$380^{m2}$</td>
<td>$78^{m3}$</td>
<td>$28^{m4}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>$704^{a4}$</td>
<td>$128^{a2}$</td>
<td>$42^{m4}$</td>
<td>$16^{a1}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>$1408$</td>
<td>$256$</td>
<td>$64^{a2}$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>$2048^{p}$</td>
<td>$384^{a1}$</td>
<td>$112^{c2}$</td>
<td>$32^{c1}$</td>
<td>$16^{a1}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>$4096$</td>
<td>$768$</td>
<td>$192^{m5}$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>$8192$</td>
<td>$1536$</td>
<td>$320^{m1}$</td>
<td>$112^{c1}$</td>
<td>$32^{c1}$</td>
<td>$16^{a1}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>$16384$</td>
<td>$2944^{m1}$</td>
<td>$512^{c1}$</td>
<td>$192^{a1}$</td>
<td>$16^{a1}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>$32768$</td>
<td>$4096^{c1}$</td>
<td>$1024$</td>
<td>$256^{a1}$</td>
<td>$64^{a1}$</td>
<td>$16^{a1}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>$65536$</td>
<td>$61440$</td>
<td>$1024^{c1}$</td>
<td>$256^{a1}$</td>
<td>$64^{a1}$</td>
<td>$16^{a1}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>$256^{a1}$</td>
<td>$32768$</td>
<td>$1024^{c1}$</td>
<td>$256^{a1}$</td>
<td>$64^{a1}$</td>
<td>$16^{a1}$</td>
<td>$7^{a1}$</td>
<td>$2^{a1}$</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>22</td>
<td>$2^21^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>$2^32^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>$2^42^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>$2^52^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>26</td>
<td>$2^62^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>27</td>
<td>$2^72^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>28</td>
<td>$2^82^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>29</td>
<td>$2^92^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>30</td>
<td>$2^{10}2^{a1}$</td>
<td>$3^2$</td>
<td>$4096$</td>
<td>$1536$</td>
<td>$512$</td>
<td>$128$</td>
<td>$64$</td>
<td>$28^{a1}$</td>
<td>$12^{a1}$</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Acknowledgements

The authors thank Professor Geoffrey Exoo for allowing us to include his construction in Example 13 and Heikki Hämäläinen and Iiro Honkala for rewarding discussions.

References


[38] P.R.J. Östergård and W.D. Weakley, Constructing covering codes with given automorphisms, submitted for publication.


