On the state complexity of some long codes

Tim Blackmore and Graham Norton

Abstract. We determine the state complexities of three families of codes that generalise the Reed–Muller codes. Our approach would seem to be new and in particular would seem to provide simplified proofs of known results on trellises for Reed–Muller codes. One of the families is new and its classical code parameters, which compare well with those of the other codes considered, are given. We conclude with a comparison of the asymptotic performance of the codes' parameters.

1. Introduction

1.1. Background. The state complexity (SC) of a code provides a measure of the complexity of the Viterbi decoding algorithm for that code. (We consider linear block codes only.) As such, it is often regarded as the fourth code parameter (the three classical parameters being the length, dimension and minimum distance of the code). Unlike the other three parameters it is dependent on the bit–ordering of the code—i.e. equivalent codes can have different SCs.

It is well–known that cyclic (here we include shortened cyclic and extended cyclic) codes have worst possible SC, [**7**], reaching an upper–bound given by Wolf, [**12**]. There has been considerable work on finding the SCs of short (lengths of up to about 128) $\mathcal{BCH}$–codes under various (non–cyclic) bit–orderings. However, it seems that the only long codes for which SCs under non–cyclic bit–orderings have been considered are the family of binary Reed–Muller ($\mathcal{RM}$–)codes.

In fact, since Forney's defining work, [**5**], in which the SC of a 4–section uniform trellis of a $\mathcal{RM}$–code is determined, there has been considerable interest in the SCs of $\mathcal{RM}$–codes. Most notably, it has been shown that the standard bit–ordering of an $\mathcal{RM}$–code is always optimum with respect to its SC, [**6**], and the SC under this bit–ordering has been determined, [**1**].

Here we determine and compare the SCs of three distinct families of (not necessarily binary) codes, each of which contains the $\mathcal{RM}$–codes as a special case. We show that one of these families of codes can be considered as generalising the $\mathcal{RM}$–codes with respect to SC—a family of codes defined a long time before SCs were first considered but of little interest otherwise. We believe our consideration of the SCs of these codes gives a simplified approach to the determination of the SCs (and other trellis characteristics) of $\mathcal{RM}$–codes.

1991 *Mathematics Subject Classification.* 11T71, 94B99.

Research supported by the U. K. Engineering and Physical Sciences Research Council under grant K27728.

1.2. State complexities. In [**12**] the Viterbi decoding algorithm was applied to block codes. The algorithm takes place along a trellis for a code. A trellis is a directed graph whose vertices are placed at depths. A trellis for a length n code has $n+1$ depths, usually labelled from 0 to n, but here labelled from -1 to $n-1$. The initial and final depth each have only one vertex. Paths through the trellis, passing through a single vertex at each depth, are in one–to–one correspondence with the codewords. It is advantageous for Viterbi decoding that many paths pass through each vertex and hence that there are as few vertices as possible at each depth. A code has a trellis which simultaneously minimises the number of vertices at each depth, called its minimal trellis (e. g. [**8**]). We consider only minimal trellises.

The set of vertices at each depth of a (minimal) trellis forms a vector space. For a length n code C, we write $s_i(C)$ for the dimension of the vertex space at depth i (where $-1 \leq i \leq n-1$). The state complexity (SC) of C is given by

$$s(C) = \max_{-1 \leq i \leq n-1} \{s_i(C)\}.$$

In [**11**] the SC of a code was described as a 'fundamental descriptive characteristic, comparable to the length, size and minimum distance'. A list of more recent publications in which SC plays a central role is given in [**10**].

Calculating the $s_i(C)$ is possible without full knowledge of the trellis, [**5**]. For $-1 \leq i \leq n-1$, the i^{th} past subcode of C, denoted by C_i^-, is the set of codewords of the form $(c_0, c_1, \ldots, c_i, 0, \ldots, 0)$. Similarly the i^{th} future subcode of C, denoted by C_i^+, is the set of codewords of the form $(0, \ldots, 0, c_{i+1}, c_{i+2}, \ldots, c_{n-1})$. If we write $k(D)$ for the dimension of a code D then

$$s_i(C) = k(C) - k(C_i^-) - k(C_i^+). \tag{1.1}$$

Now $k(C_i^-)$ increases in unit steps from 0 to $k(C)$, and $k(C_i^+)$ decreases in unit steps from $k(C)$ to 0. An increase in $k(C_i^-)$ leads to a (possible) decrease in $s_i(C)$ and so we refer to an i where this happens as a point of fall (PofF). Similarly a decrease in $k(C_i^+)$ leads to a (possible) increase in $s_i(C)$ and so we refer to an i where this happens as a point of gain (PofG). It is possible that i is both a PofF and PofG, in which case $s_i(C) = s_{i-1}(C)$. (Such an i can affect other 'trellis complexities', such as branch complexity and edge complexity, not considered here.) We note that if $\gamma_i(C)$ and $\delta_i(C)$ are respectively the number of PsofG and PsofF before and including i then $\gamma_i(C) = k(C) - k(C_i^+)$ and $\delta_i(C) = k(C_i^-)$, so that

$$s_i(C) = \gamma_i(C) - \delta_i(C).$$

It is well–known that the state complexity of a code and its dual are equal—in fact the dimensions of their vertex spaces at each depth are equal, [**5**].

1.3. Outline. In Section 2 we consider a family of binary codes defined by Berman in [**2**]. These codes have defining parameters p, r and m, where p is an odd prime, m and r integers with $m \geq 1$ and $0 \leq r \leq m-1$—we denote such a code $\mathcal{B}(p, r, m)$. Berman used these codes to demonstrate the existence of semisimple abelian codes with better asymptotic performance than any semisimple cyclic codes. Towards this end, he determined their classical code parameters. We show how these codes together with their duals can be considered as a generalisation of $\mathcal{RM}$–codes ($\mathcal{RM}$–codes are the case $p = 2$ and so do not, strictly speaking, belong to the family of $\mathcal{B}$–codes and their duals). We determine the minimum distance of

the dual codes. We also determine the SC of the dual codes and hence the SC of the 'Berman codes'. This SC is greater than might have been expected.

In Section 3 we consider a less well–known family of codes defined in [**4**]. This family generalises $\mathcal{R}M$–codes (this time including them as a special case). These codes are defined over any finite field, $GF(q)$, and have defining parameters n, r and m, where n, m and r are integers with $n \geq 2$, $m \geq 1$ and $0 \leq r \leq m-1$—we denote such a code $\mathcal{D}H_q(n,r,m)$. Thus $\mathcal{D}H$–codes are more numerous than $\mathcal{B}$–codes or their duals. However when both are defined, the $\mathcal{D}H$–codes have poorer classical code parameters than either $\mathcal{B}$–codes or their duals—their poor parameters explaining why $\mathcal{D}H$–codes are less well–known. We determine the SC of $\mathcal{D}H$–codes and hence show that it is these codes that generalise $\mathcal{R}M$–codes with respect to SC.

In Section 4 we introduce a new family codes, defined over any finite field, $GF(q)$, with defining parameters $n \geq 2$, $m \geq 1$, and $0 \leq r \leq m-1$ (as for $\mathcal{D}H$–codes). We denote such a code by $\mathcal{C}_q(n,r,m)$. Their code parameters (including SC) are comparable to those of $\mathcal{B}$–codes, and coincide when both are defined.

When considering the SC of $\mathcal{D}H$– and $\mathcal{C}$–codes, we give a local description of their trellis behaviour, which coincides (and as far as we know was previously unknown) in the case that these codes are $\mathcal{R}M$–codes. From this, we determine recurrence relations for their SCs which generalise those given in [**9**] for $\mathcal{R}M$–codes. From the recurrence relations we determine the SCs. In the case of $\mathcal{R}M$–codes, we believe our derivations to be simpler than those of [**1**] and [**9**].

We summarise the code parameters in Section 5 and in Section 6 we compare the asymptotic performances of the parameters.

2. Berman codes

Berman codes are defined as certain ideals in the ring

$$R_{p,m} = \frac{GF(2)[X_1, \ldots, X_m]}{(X_1^p - 1, \ldots, X_m^p - 1)},$$

where p is an odd prime and m an integer, $m \geq 1$. All such ideals are semisimple, the codes being examples of semisimple abelian codes. For $m = 1$, Berman codes are semisimple cyclic codes. Of course to say that an ideal is a code is to identify a polynomial in the ideal with the codeword of coefficients of monomials in the polynomial (zero coefficients included). Thus Berman codes are binary codes of length p^m. Certainly when considering SC, we need a definite bit–ordering for our code. We take the bit–ordering inherited from the lexicographical ordering of monomials (with $X_1 < \ldots < X_m$) in $R_{p,m}$.

We fix m and for $1 \leq j \leq m$ we put

$$P_p(X_j) = 1 + X_j + X_j^2 + \cdots + X_j^{p-1} \quad \text{and} \quad Q_p(X_j) = X_j + X_j^2 + \cdots + X_j^{p-1}.$$

For $0 \leq r \leq m-1$ we put $G(p,r,m)$ equal to the set of polynomials in $R_{p,m}$ of the form

$$(Q_p(X_{j_1}) \cdots Q_p(X_{j_s})) \cdot \left(P_p(X_{j_{s+1}}) \cdots P_p(X_{j_m})\right),$$

for some $0 \leq s \leq r$ and arrangement, $(j_1, \ldots, j_n)$, of $(1, \ldots, n)$. Berman shows, [**2**, Theorem 2.2], that for each odd prime p, the code (ideal) generated by $G(p,r,m)$

has dimension,

$$K_1(p,r,m) = \sum_{i=0}^{r} \left(\begin{array}{c} m \\ i \end{array} \right) (p-1)^i,$$

and that the check code, which then has dimension,

$$K_2(p,r,m) = p^m - K_1(p,r,m) = \sum_{i=r+1}^{m} \left(\begin{array}{c} m \\ i \end{array} \right) (p-1)^i,$$

has minimum distance 2^{r+1}. It is this code that we denote by $\mathcal{B}(p,r,m)$.

The code generated by $G(p,r,m)$—the check code of $\mathcal{B}(p,r,m)$—is also the dual code of $\mathcal{B}(p,r,m)$, since the permutation $\pi(i) = p^m - i$ is in its automorphism group. We denote this code by $\mathcal{B}^{\perp}(p,r,m)$.

Our initial interest in $\mathcal{B}^{\perp}(p,r,m)$ was that it has the same SC as $\mathcal{B}(p,r,m)$ (this being true of any code and its dual) and that, unlike $\mathcal{B}(p,r,m)$, we have an explicit description of the codewords of $\mathcal{B}^{\perp}(p,r,m)$.

For completeness we have also shown,

PROPOSITION 2.1. *The minimum distance of $\mathcal{B}^{\perp}(p,r,m)$ is p^{m-r}.*

Proposition 2.1 can be proved by induction using the fact that $\mathcal{B}^{\perp}(p,r,m)$ is the direct sum of

$$B_1 = \left\{ \sum_{l=0}^{p-1} f_l(X_1,\ldots,X_{m-1}) \cdot X_m^l : f_0,\ldots,f_{p-1} \in \mathcal{B}^{\perp}(p,r-1,m-1) \right\}$$

and

$$B_2 = \{ g(X_1,\ldots,X_{m-1}) \cdot P_p(X_m) : g \in \mathcal{B}^{\perp}(p,r,m-1) \setminus \mathcal{B}^{\perp}(p,r-1,m-1)^* \}$$

where $\mathcal{B}^{\perp}(p,r-1,m-1)^* = \mathcal{B}^{\perp}(p,r-1,m-1) \setminus \{0\}$. (We adopt the convention that $\mathcal{B}^{\perp}(p,-1,m-1) = \{0\}$ and $\mathcal{B}^{\perp}(p,m-1,m-1) = GF(2)^{p^{m-1}}$.) That $B_1 \oplus B_2 \subseteq \mathcal{B}^{\perp}(p,r,m)$ follows from $X_m^l = P_p(X_m) + Q_p(X_m) \cdot X_m^l$ in $R_{p,m}$ and that $B_1 \oplus B_2$ and $\mathcal{B}^{\perp}(p,r,m)$ have the same dimension is a straightforward counting argument.

2.1. Berman codes and Reed–Muller codes. There is then a strong connection between the classical code parameters of $\mathcal{B}$–codes and their duals (for p an odd prime) and $\mathcal{RM}$–codes (for $p = 2$). We look at another way in which $\mathcal{B}$–, $\mathcal{B}^{\perp}$– and $\mathcal{RM}$–codes can be thought of as being part of the same family.

In [**3**] it is noted that $\mathcal{RM}(r,m)$ is equal to the $(m-r)^{th}$ power of the radical of $R_{2,m}$ (which is not a semisimple ring, unlike $R_{p,m}$ for odd p). It is not hard to see that,

PROPOSITION 2.2. *For $0 \le r \le m-1$, the $(m-r)^{th}$ power of the radical of $R_{2,m}$ is generated by $G(2,r,m)$.*

Thus $\mathcal{RM}(r,m)$ and $\mathcal{B}^{\perp}(p,r,m)$ belong to the same family of codes: those generated by $G(p,r,m)$ for all primes p. (Amongst these codes, the $\mathcal{RM}$–codes are modular abelian codes and the $\mathcal{B}^{\perp}$–codes are semisimple abelian codes.) A code in this family has dimension $K_1(p,r,m)$ and minimum distance p^{m-r} and its dual has dimension $K_2(p,r,m)$ and minimum distance 2^{r+1}.

2.2. State complexity. The ordering implicit in the ideal description of $\mathcal{R}M$–codes is the (non–cyclic) standard bit–ordering, which has been shown to be optimal with respect their SC in [**6**]. Under this bit ordering it is shown in [**1**] that the SC of $\mathcal{R}M(r,m)$ is $S_1(2,r,m)$ where,

$$S_1(p,r,m) = \sum_{i=0}^{\min\{r,m-r-1\}} \left(\begin{array}{c} m-2i-1 \\ r-i \end{array} \right) (p-1)^{r-i}.$$

Thus, in view of the close connection between $\mathcal{B}^\perp$–codes and their duals and $\mathcal{R}M$–codes and their duals, it may have been hoped that the SC of $\mathcal{B}^\perp$–codes, and hence the SC of $\mathcal{B}$–codes, would be $S_1(p,r,m)$. However,

PROPOSITION 2.3. *The SC of* $\mathcal{B}(p,r,m)$ *is,*

$$S_2(p,r,m) = \sum_{i=0}^{r} \left(\begin{array}{c} m-i-1 \\ r-i \end{array} \right) (p-1)^{r-i}.$$

It would seem significant that whilst the dual of an $\mathcal{R}M$–code is an $\mathcal{R}M$–code, the dual of a $\mathcal{B}^\perp$–code is not a $\mathcal{B}^\perp$–code.

The proof of Proposition 2.3 uses an alternative version of Equation (1.1). The i^{th} past *truncated* code of C is $C^i_- = \{(c_0,\ldots,c_i) : (c_0,\ldots,c_n) \in C\}$ and the i^{th} future *truncated* code of C is $C^i_+ = \{(c_{i+1},\ldots,c_n) : (c_0,\ldots,c_n) \in C\}$. Then $k(C^i_-) = k(C) - k(C^+_i)$ and $k(C^i_+) = k(C) - k(C^-_i)$ so that

$$s_i(C) = k(C^i_-) + k(C^i_+) - k(C). \tag{2.1}$$

For codes C_1 and C_2 with $C_1 \cap C_2 = \{0\}$ it is then straightforward to see that

$$s_i(C_1 \oplus C_2) \le s_i(C_1) + s_i(C_2), \tag{2.2}$$

with equality if $k((C_1)^i_- \cap (C_2)^i_-) = k((C_1)^i_+ \cap (C_2)^i_+) = 0$.

Now we take B_1 and B_2 as above and $i = Qp + R$, where $0 \le Q \le p-1$ and $0 \le R \le p^m - 1$. Then

$$k((B_1)^i_-) = Q \cdot k(\mathcal{B}^\perp(p,r-1,m-1)) + k(\mathcal{B}^\perp(p,r-1,m-1)^R_-)$$

and

$$k((B_1)^i_+) = (p-1-Q) \cdot k(\mathcal{B}^\perp(p,r-1,m-1)) + k(\mathcal{B}^\perp(p,r-1,m-1)^R_+)$$

so that from (2.1), $s_i(B_1) = s_R(\mathcal{B}^\perp(p,r-1,m-1))$. Also, with $B = \mathcal{B}^\perp(p,r,m-1) \setminus \mathcal{B}^\perp(p,r-1,m-1)^*$,

$$k((B_2)^i_-) = \left\{ \begin{array}{cc} k(B^R_-) & \text{if } Q = 0 \\ k(B) & \text{if } Q \ge 1 \end{array} \right. \quad \text{and} \quad k((B_2)^i_+) = \left\{ \begin{array}{cc} k(B^R_+) & \text{if } Q = p-1 \\ k(B) & \text{if } Q \le p-2 \end{array} \right.$$

so that from (2.1), $s_i(B_2) \le k(B)$ with equality if $1 \le Q \le p-2$. Thus noting that if $1 \le Q \le p-2$ then $k((B_1)^i_- \cap (B_2)^i_-) = k((B_1)^i_+ \cap (B_2)^i_+) = 0$, we have from (2.2) that

$$s_i(\mathcal{B}^\perp(p,r,m)) \le s_R(\mathcal{B}^\perp(p,r-1,m-1)) + k(B),$$

with equality if $1 \le Q \le p-2$. Thus

LEMMA 2.4. $s(\mathcal{B}^\perp(p,r,m)) = s(\mathcal{B}^\perp(p,r-1,m-1)) + \left(\begin{array}{c} m-1 \\ r \end{array} \right) (p-1)^r.$

Proposition 2.3 follows from Lemma 2.4 by induction.

Apart from trivial cases, when equality holds, it is straightforward to see that $S_2(p,r,m) > S_1(p,r,m)$. Thus the SC of $\mathcal{B}$–codes would seem disappointing, although in Section 6.2 we see that asymptotically S_2 and S_1 coincide.

The parameters of $\mathcal{B}$– and $\mathcal{B}^\perp$–codes are in Table 1 of Section 5.

3. Dwork–Heller codes

Let n be an integer, $n \geq 1$, $M_{q,n,m}$ the set of monomials in

$$\frac{GF(q)[X_1, \ldots, X_m]}{(X_1^n - 1, \ldots, X_m^n - 1)}$$

and $R_{q,n,m} = \langle M_{q,n,m} \rangle$, the linear span of $M_{q,n,m}$. For $M \in M_{q,n,m}$, we set

$$P_M = \sum_{M|M', M' \in M_{q,n,m}} M'.$$

If $D_{q,n,m}$ is the set of all such polynomials, then $|D_{q,n,m}| = n^m$ and $\langle D_{q,n,m} \rangle = R_{q,n,m}$. For $0 \leq r \leq m-1$, put

$$H_{q,n,m} = \langle M \in M_{q,n,m} : M \text{ is divisible by at most } r \text{ variables} \rangle.$$

Each polynomial in $H_q(n,r,m)$ is a linear combination of elements of $D_q(n,m)$. The coefficients of each such linear combinations (somehow ordered) are the codewords of $\mathcal{DH}_q(n,r,m)$. Thus $\mathcal{DH}_q(n,r,m)$ is a code of length n^m over $GF(q)$.

For example $H_2(3,0,2) = \{0,1\}$. Now $1 = P_1 + P_{X_1} + P_{X_2} + P_{X_1X_2}$, so $\mathcal{DH}_q(3,0,2)$ is $\{(000000000), (110110000)\}$ (not the repetition code we would want).

These codes were defined in [**4**], where it was shown that $\mathcal{DH}_q(n,r,m)$, has dimension $K_1(n,r,m)$ and minimum distance 2^{m-r}, and stated that $\mathcal{DH}_2(2,r,m) = \mathcal{RM}(r,m)$. We note that for p an odd prime, $\mathcal{DH}_2(p,r,m)$ has the same number of codewords as $\mathcal{B}^\perp(p,r,m)$ but inferior minimum distance by Proposition 2.1, and that $\mathcal{DH}_2(p,r,m)$ has the same minimum distance as $\mathcal{B}(p,m-r-1,m)$, but fewer codewords.

3.1. Local behaviour of trellis complexities. To consider the SC of $\mathcal{DH}$–codes we need a bit–ordering. Since $\mathcal{DH}_q(n,m)$ is a length n^m code we label our bit positions from 0 to $n^m - 1$ and our trellis depths from -1 to $n^m - 1$. For $0 \leq i \leq n^m - 1$ we have the n–expansion of i, $(a_1, \ldots, a_m)$, where $i = \sum_{j=1}^m a_j n^{j-1}$ and $0 \leq a_j \leq n-1$. A codeword is the vector of coefficients of an element of $\langle D_q(n,m) \rangle$. The i^{th} symbol of this codeword is the coefficient of $P_M \in D_q(n,m)$, where $M = X_1^{a_1} \ldots X_m^{a_m}$ and $(a_1, \ldots, a_m)$ is the n–expansion of i. In the case of $\mathcal{RM}$–codes this ordering is the standard bit–ordering.

For $0 \leq a \leq n-1$ we write $|i|_a$ for the number of the a_j in the n–expansion of i equal to a. The following result gives a comprehensive local description of the behaviour of the state (or any of the other usual types of trellis) complexity, which as far as we know was previously unknown for $\mathcal{RM}$–codes.

LEMMA 3.1. *For $0 \leq i \leq n^m - 1$,*

1. *i is a PofG of $\mathcal{DH}_q(n,r,m)$ if and only if $|i|_0 \geq m - r$ and*
2. *i is a PofF of $\mathcal{DH}_q(n,r,m)$ if and only if $|i|_1 \geq m - r$.*

We note that for $r \leq m-r-1$ no i is a PofG and PofF (since e. g. if $|i|_0 \geq m-r \geq m-(m-r-1) = r+1$ then $|i|_1 \leq m-r-1$). Also for $n=2$, for $r \geq m-r$, i is a PofG of $\mathcal{D}H_q(2,r,m)$ which is not a PofG of $\mathcal{D}H_q(2,m-r-1,m)$ if and only if $m-(m-r-1)-1 = r \geq |i|_0 \geq m-r$ if and only if $m-r \leq |i|_1 \leq r$ if and only if i is a PofF of $\mathcal{D}H_q(2,r,m)$ which is not a PofF of $\mathcal{D}H_q(2,m-r-1,m)$. This of course ties in with fact that the vertex dimension at each depth of an $\mathcal{R}M(r,m)$ is the same as for its dual, $\mathcal{R}M(m-r-1,m)$. It also shows that the edge complexity of a low–rate $\mathcal{R}M$–code will be less than that of its high–rate dual.

REMARK 3.2. The bit–ordering of $\mathcal{D}H_q(n,r,m)$ described corresponds to the lexicographic ordering of $D_q(n,m)$—in the sense that the coefficient of P_{M_1} comes before that of P_{M_2} in a codeword if and only if $M_1 < M_2$ where $<$ is the lexicographic ordering of monomials. In fact Lemma 3.1 holds for any monomial ordering of $D_q(n,m)$. This is not true of the results in the rest of the section.

Lemma 3.1 makes calculating the dimension of a vertex space at a given depth quite straightforward—we just have to count the i with $|i|_0 \geq m-r$ and $|i|_1 \geq m-r$ that occur before our given depth and subtract the latter from the former. For example, the results of [**5**] and [**7**] on the 4–section and 8–section uniform trellises of $\mathcal{R}M$–codes follow quite easily. However to determine at which depth the difference between these counts is maximised requires some more work.

3.2. Recurrence relations. We derive recurrence relations that generalise those for $\mathcal{R}M$–codes given in [**9**]. In the case of $\mathcal{R}M$–codes our approach would seem to simplify that of [**9**]. For $\mathcal{R}M$–codes we have that

$$s_i(\mathcal{R}M(r,m)) = s_{2^m-2-i}(\mathcal{R}M(r,m))$$

for $-1 \leq i \leq 2^m-1$, (a known fact that is easily deducible from Lemma 3.1) so that it is only necessary to calculate recurrence relations for $-1 \leq i \leq 2^{m-1}-1$. No such identity holds for $\mathcal{D}H$–codes in general. It is the recurrence relations for $n^{m-1} \leq i \leq 2n^{m-1}-1$ that cause difficulty. The cases $i=-1,0$ are trivial. For $i \geq 1$ not in the range $n^{m-1} \leq i \leq 2n^{m-1}-1$ either

1. there is a j, $1 \leq j \leq m-1$, such that $n^{j-1} \leq i \leq 2n^{j-1}-1$ (this being the only case for $\mathcal{R}M$–codes), in which case we get the recurrence relation

$$s_i\left(\mathcal{D}H_q(n,r,m)\right) = s_{i-n^{j-1}}\left(\mathcal{D}H_q(n,r-1,m-2)\right) + s_{n^{j-1}-1}\left(\mathcal{D}H_q(n,r,m)\right),$$

   or

2. there is a j, $1 \leq j \leq m$ and an a, $2 \leq a \leq n-1$, such that $an^{j-1} \leq i \leq (a+1)n^{j-1}-1$, in which case we get the recurrence relation

$$s_i\left(\mathcal{D}H_q(n,r,m)\right) = s_{i-an^{j-1}}\left(\mathcal{D}H_q(n,r-1,m-1)\right) + s_{an^{j-1}-1}\left(\mathcal{D}H_q(n,r,m)\right).$$

(We note that calculating the second terms in the right–hand sides of the recurrence relations is straightforward from Lemma 3.1.)

For $n^{m-1} \leq i \leq 2n^{m-1}-1$ we need some notation. We write $u(l) = \sum_{j=l+1}^{m} n^{j-1}$ and $v(l) = u(m-2) - n^{l-1}$ and $u(l,a) = u(l) + an^{l-1}$ and $v(l,a) = u(m-1) + (a-1)n^{l-1}$ $(v(l,0) \neq v(l))$. Then for i in this range, except $i = \sum_{j=1}^{m} q^{j-1}$ which is always a PofF (and so cannot be a depth at which SC is attained), either

1. there is an l, $1 \leq l \leq m-1$, such that $u(l) \leq i \leq u(l)+n^{l-1}$, in which case

$$s_i\left(\mathcal{D}H_q(n,r,m)\right) - s_{u(l)-1}\left(\mathcal{D}H_q(n,r,m)\right) =$$
$$s_{i-v(l)}\left(\mathcal{D}H_q(n,r-1,m-2)\right) - s_{u(l)-v(l)-1}\left(\mathcal{D}H_q(n,r-1,m-2)\right),$$

or

2. there is an l, $1 \leq l \leq m-1$, and an a, $2 \leq a \leq n-1$, such that $u(l,a) \leq i \leq u(l,a) + n^{l-1}$, in which case

$$s_i\left(\mathcal{DH}_q(n,r,m)\right) - s_{u(l,a)-1}\left(\mathcal{DH}_q(n,r,m)\right) =$$
$$s_{i-v(l,a)}\left(\mathcal{DH}_q(n,r-1,m-1)\right) - s_{u(l,a)-v(l,a)-1}\left(\mathcal{DH}_q(n,r-1,m-1)\right).$$

(Again the terms not dependent on i are quite straightforward to calculate.)

3.3. State complexity. It is possible to inductively determine from the recurrence relations at which depth the SC is attained and the value of the SC. For some $\mathcal{DH}$–codes the depth where the SC is attained is unique and in this case will be in the problem area, $n^{m-1} \leq i \leq 2n^{m-1} - 1$.

We put $[r,m] = m - 2\min\{r, m-r-1\} - 1$. Then,

PROPOSITION 3.3. *The state complexity of $\mathcal{DH}_q(n,r,m)$ is attained at depth with n–expansion,*

$$(\underbrace{n-1,\ldots,n-1}_{[r,m]},\underbrace{0,0,1,0,1,0,1,\ldots,0,1}_{m-[r,m]}).$$

Its value is $S_1(n,r,m)$.

Thus $\mathcal{DH}$–codes can be considered to generalise $\mathcal{RM}$–codes with respect to state complexity.

Again the parameters of $\mathcal{DH}$–codes are in Table 1 of Section 5.

4. A new family of codes

These codes are defined similarly to $\mathcal{DH}$–codes. Again we work in the vector space $R_{q,n,m}$, but with the basis $E_q(n,m)$ of polynomials of the form

$$Q_M = \sum_{M' \mid M, M' \in M_{q,n,m}} M',$$

and for $0 \leq r \leq m-1$ we put $I_q(n,r,m)$ equal to the linear span of all monomials divisible by at least $r+1$ variables. Then again each polynomial in $I_q(n,r,m)$ is a linear combination of elements of $E_q(n,m)$ and the codewords of the code which we denote $\mathcal{C}_q(n,r,m)$, are the coefficients of such linear combinations.

PROPOSITION 4.1. *$\mathcal{C}_q(n,r,m)$ is an $[n^m, K_2(n,r,m), 2^{r+1}]$ code.*

Thus $\mathcal{C}$–codes are defined as often as $\mathcal{DH}$–codes but have classical code parameters comparable to the superior, but less often defined, $\mathcal{B}$–codes. For $q = n = 2$ we again get the $\mathcal{RM}$–codes, but here $\mathcal{C}_2(2,r,m) = \mathcal{RM}(m-r-1,m)$, the dual of $\mathcal{RM}(r,m)$.

4.1. Local behaviour of trellis complexities. For $\mathcal{C}$–codes, the i^{th} symbol position is the coefficient of Q_M, where $M = X_1^{a_1} \ldots X_m^{a_m}$ and $(a_1,\ldots,a_m)$ is the n–expansion of i $(0 \leq i \leq n^m - 1)$. Again when $\mathcal{C}$–codes are $\mathcal{RM}$–codes we have their standard bit–ordering. We have the following analogue of Lemma 3.1,

LEMMA 4.2. *For $0 \leq i \leq n^m - 1$,*

1. *i is a PofG of $\mathcal{C}_q(n,r,m)$ if and only if $|i|_{n-1} \leq m-r-1$ and*
2. *i is a PofF of $\mathcal{C}_q(n,r,m)$ if and only if $|i|_0 \leq m-r-1$.*

We note that for $\mathcal{RM}(r,m) = \mathcal{C}_2(2, m-r-1, m)$, Lemma 4.2 implies that i is a PofG if and only if $|i|_1 \leq r$ if and only if $|i|_0 \geq m-r$, and that i is a PofG if and only if $|i|_0 \leq r$ if and only if $|i|_1 \geq m-r$, both of which are in agreement with Lemma 3.1. We note also that Lemma 4.2 holds for any monomial ordering of $E_q(n,m)$—c.f. Lemma 3.1 and Remark 3.2.

4.2. Recurrence relations. We quickly get from Lemma 4.2 that, for $-1 \leq i \leq n^m - 1$,

$$s_i\left(\mathcal{C}_q(n,r,m)\right) = s_{n^m-2-i}\left(\mathcal{C}_q(n,r,m)\right),$$

a property of $\mathcal{RM}$–codes not shared in general by $\mathcal{DH}$–codes, as noted in the previous section.

Thus for $\mathcal{C}$–codes, as for $\mathcal{RM}$–codes, we do not need to find recurrence relations for $(n-1)n^{m-1} \leq i \leq n^m - 1$ (the vertex dimensions for these depths being deducible from those for $-1 \leq i \leq n^{m-1} - 2$). We do need to divide all other $i \geq 1$ into two sets though.

For $1 \leq i \leq (n-1)n^{m-1} - 1$ either,

1. there is a j, $1 \leq j \leq m$, and an a, $1 \leq a \leq n-2$, such that $an^{j-1} \leq i \leq (a+1)n^{j-1} - 1$, in which case

$$s_i\left(\mathcal{C}_q(n,r,m)\right) = s_{i-an^{j-1}}\left(\mathcal{C}_q(n,r-1,m-1)\right) + s_{an^{j-1}-1}\left(\mathcal{C}_q(n,r,m)\right),$$

   or

2. there is a j, $1 \leq j \leq m-1$, such that $(n-1)n^{j-1} \leq i \leq n^j - 1$ (this being the only case for $\mathcal{RM}$–codes), in which case

$$s_i\left(\mathcal{C}_q(n,r,m)\right) = s_{i-(n-1)n^{j-1}}\left(\mathcal{C}_q(n,r-1,m-2)\right) + s_{(n-1)n^{j-1}-1}\left(\mathcal{C}_q(n,r,m)\right).$$

(Again the second terms on the right–hand side of these recurrences can be easily calculated from Lemma 4.2.)

4.3. State complexity. It follows quickly by induction from the recurrence relations that

PROPOSITION 4.3. *For $n = 2$ the state complexity of $\mathcal{C}_q(n,r,m)$ is attained at depth with n–expansion*

$$(\underbrace{1,\ldots,1}_{[r,m]},\underbrace{0,0,1,0,1,0,1,\ldots,0,1}_{m-[r,m]}).$$

Its value is $S_1(n,r,m)$.

For $n \geq 3$, the state complexity of $\mathcal{C}_q(n,r,m)$ is attained at all depths with n–expansion $(a_1,\ldots,a_m)$, where $1 \leq a_1,\ldots,a_m \leq n-2$. Its value is $S_2(n,r,m)$.

The parameters of $\mathcal{C}$–codes are in Table 1 of Section 5.

5. Code Parameters

The parameters of the codes discussed are summarised in Table 1. A defining parameter for each code is $m \geq 1$. For the other defining parameters (DPs) we have n, r integers, $n \geq 2$, and $0 \leq r \leq m-1$, p an odd prime and q a power of a prime. The codes are of length p^m, n^m or 2^m according to whether p, n or neither appear as a defining parameter; K_1, K_2, S_1 and S_2 are defined in Section 2.

TABLE 1. Code parameters

Code	DPs	Field	Dimension	Minimum Distance	State Complexity
$\mathcal{RM}$	r	$GF(2)$	$K_1(2,r,m)$	2^{m-r}	$S_1(2,r,m)$
	$m-r-1$	$GF(2)$	$K_2(2,r,m)$	2^{r+1}	$S_1(2,r,m)$
$\mathcal{B}^\perp$	p,r	$GF(2)$	$K_1(p,r,m)$	p^{m-r}	$S_2(p,r,m)$
$\mathcal{B}$	p,r	$GF(2)$	$K_2(p,r,m)$	2^{r+1}	$S_2(p,r,m)$
$\mathcal{DH}$	q,n,r	$GF(q)$	$K_1(n,r,m)$	2^{m-r}	$S_1(n,r,m)$
$\mathcal{C}$	q,n,r	$GF(q)$	$K_2(n,r,m)$	2^{r+1}	$S_1(2,r,m)$ if $n=2$ $S_2(n,r,m)$ if $n\geq 3$

6. Asymptotic analysis

The general theory of asymptotic behaviour of SCs has received some attention (e. g. [**10**] and references given there). However the SCs of few long codes are known and hence little is known about the asymptotic behaviour of SCs for specific families of codes. Here we look at the behaviour of $S_1(n,r,m)$ and $S_2(n,r,m)$ as $m\to\infty$. For r fixed the behaviour is trivial so we want r to increase with m. We fix λ, $0<\lambda<1$, and put $r=\lfloor\lambda m\rfloor$. Our results of course apply to $\mathcal{RM}(\lfloor\lambda m\rfloor,m)$ as a special case.

6.1. Asymptotic comparison of classical code parameters. We know that, for an odd prime p, $\mathcal{B}^\perp(p,\lfloor\lambda m\rfloor,m)$ and $\mathcal{DH}_2(p,\lfloor\lambda m\rfloor,m)$ both have the same number of codewords, but that the former has minimum distance $p^{m-\lfloor\lambda m\rfloor}$, compared with the latter's $2^{m-\lfloor\lambda m\rfloor}$. A non–trivial asymptotic comparison of these minimum distances (one for which both do not either tend to 0 or ∞) is given by

$$\lim_{m\to\infty}\frac{\log_p p^{m-\lfloor\lambda m\rfloor}}{m}=(1-\lambda)>(1-\lambda)\log_p 2=\lim_{m\to\infty}\frac{\log_p 2^{m-\lfloor\lambda m\rfloor}}{m}. \tag{6.1}$$

Thus asymptotically the minimum distance of $\mathcal{B}^\perp$–codes remains superior.

For convenience we introduce the following notation,

$$\begin{array}{ll}\mathcal{DH}_q^\lambda(n,m)=\mathcal{DH}_q(n,\lfloor\lambda m\rfloor,m) & \mathcal{C}_q^\lambda(n,m)=\mathcal{C}_q(n,m-\lfloor\lambda m\rfloor-1,m)\\ K_1^\lambda(n,m)=K_1(n,\lfloor\lambda m\rfloor,m) & K_2^\lambda(n,m)=K_2(n,m-\lfloor\lambda m\rfloor-1,m)\\ R_1^\lambda(n,m)=K_1^\lambda(n,m)/q^m & R_2^\lambda(n,m)=K_2^\lambda(n,m)/q^m.\end{array}$$

Thus $\mathcal{DH}_q^\lambda(n,m)$ and $\mathcal{C}_q^\lambda(n,m)$ both have minimum distance $2^{m-\lfloor\lambda m\rfloor}$.

We also put

$$R_i^\lambda(n,\infty)=\lim_{m\to\infty}R_i^\lambda(n,m)$$

for $i=1,2$.

PROPOSITION 6.1. *With the above notation, we have*

$$R_1^\lambda(n,\infty)=\begin{cases}0 & \text{for } 0<\lambda<(n-1)/n\\ 1 & \text{for } (n-1)/n<\lambda<1\end{cases}$$

and

$$R_2^\lambda(n,\infty)=\begin{cases}0 & \text{for } 0<\lambda<1/n\\ 1 & \text{for } 1/n<\lambda<1.\end{cases}$$

Thus for $1/n<\lambda<(n-1)/n$, $\mathcal{C}_q^\lambda(n,m)$ has asymptotic rate 1 whereas $\mathcal{DH}_q^\lambda(n,m)$ has asymptotic rate 0.

CONJECTURE 6.2. *Both* $R_1^{(n-1)/n}(n,\infty)$ *and* $R_2^{1/n}(n,\infty)$ *are equal to* $1/2$.

In fact, for $n \geq 3$, we can also distinguish between the asymptotic performance of $K_1^\lambda(n,m)$ and $K_2^\lambda(n,m)$ for $0 < \lambda < 1/n$ using a $\frac{\log_n}{m}$ comparison similar to Equation (6.1). Explicitly, using the entropy function $H_n(\lambda) = \lambda \log_n(n-1) - \lambda \log_n \lambda - (1-\lambda)\log_n(1-\lambda)$, we have

PROPOSITION 6.3. *For* $0 < \lambda < 1/n < 1/2$,

$$\lim_{m\to\infty} \frac{\log_n K_1^\lambda(n,m)}{m} = H_n(\lambda) < H_n(1-\lambda) = \lim_{m\to\infty} \frac{\log_n K_2^\lambda(n,m)}{m},$$

6.2. Asymptotic SC performance. It is quite straightforward to see that

$$\frac{S_2(n,\lfloor \lambda m \rfloor, m)}{m} \geq \frac{S_1(n,\lfloor \lambda m \rfloor, m)}{m} \longrightarrow \infty \text{ as } m \longrightarrow \infty$$

and that

$$0 \leq \frac{S_1(n,\lfloor \lambda m \rfloor, m)}{n^m} \leq \frac{S_2(n,\lfloor \lambda m \rfloor, m)}{n^m} \longrightarrow 0 \text{ as } m \longrightarrow \infty.$$

Thus neither of these provides a substantial comparison of the asymptotic performance of the SCs and so we look at the more subtle $\frac{\log_n}{m}$ comparison, used above.

PROPOSITION 6.4. *For* $0 < \lambda < 1$,

$$\lim_{m\to\infty} \frac{\log_n S_1(n,\lfloor \lambda m \rfloor, m)}{m} = H_n(\lambda) = \lim_{m\to\infty} \frac{\log_n S_2(n,\lfloor \lambda m \rfloor, m)}{m}.$$

Thus the $\frac{\log_n}{m}$ comparison fails to distinguish between the asymptotic performances of the SCs of $\mathcal{D}H_2(p,r,m)$ and the superior $\mathcal{B}^\perp(p,r,m)$.

Writing $S_1^\lambda(n,m)$ and $S_2^\lambda(n,m)$ for the SCs of $\mathcal{D}H_q^\lambda(n,m)$ and $\mathcal{C}_q^\lambda(n,m)$ respectively, we have

COROLLARY 6.5. *For* $0 < \lambda < 1$,

$$\lim_{m\to\infty} \frac{\log_n S_1^\lambda(n,m)}{m} = H_n(\lambda) \quad \textit{and} \quad \lim_{m\to\infty} \frac{\log_n S_2^\lambda(n,m)}{m} = H_n(1-\lambda).$$

Thus for $n \geq 3$, $\mathcal{D}H_q^\lambda(n,m)$ has asymptotically lower SC for $\lambda < 1/2$ but the superior $\mathcal{C}_q^\lambda(n,m)$ has asymptotically lower SC for $\lambda > 1/2$.

Acknowledgements. The authors gratefully acknowledge financial support from the U. K. Engineering and Physical Sciences Research Council. The first author was supported by the EPSRC.

References

[1] Y. Berger, Y. Be'ery, *Bounds on the trellis size of linear block codes,* IEEE Trans. Information Theory **39**, 203–209.

[2] S. D. Berman, *Semisimple cyclic and abelian codes II,* Kibernetika **3**, 21–30.

[3] ———, *On the theory of group codes,* Kibernetika **3**, 31–39.

[4] B. M. Dwork, R. M. Heller, *Results of a geometric approach to the theory and construction of non-binary multiple error and failure correcting codes,* IRE Nat. Conv. Rec., 123–129.

[5] G. D. Forney, *Coset codes—Part II: Binary lattices and related codes,* IEEE Trans. Information Theory **34** (1988), 1152–1187.

[6] T. Kasami, T. Takata, T. Fujiwara, S. Lin, *On the optimum bit orders with respect to state complexity of trellis diagrams for binary linear codes,* IEEE Trans. Information Theory **39** (1993), 242–245.

[7] ———, *On complexity of trellis structure of linear block codes,* IEEE Trans. Information Theory **39** (1993), 1057–1064.

[8] A. B. Kiely, S. J. Dolinar, R. J. McEliece, L. L. Ekroot, W. Lin, *Trellis decoding complexity of linear block codes,* IEEE Trans. Information Theory **42** (1996), 1687–1697.
[9] C. Lu, S. Huang, *On bit–level trellis complexity of Reed–Muller codes,* IEEE Trans. Information Theory **41** (1995), 2061–2064.
[10] A. Lafourcade, A. Vardy, *Lower bounds on trellis complexity of block codes,* IEEE Trans. Information Theory **41** (1995), 1938–1954.
[11] D. J. Muder, *Minimal trellises for block codes,* IEEE Trans. Information Theory **34** (1988), 1049–1053.
[12] J. K. Wolf, *Efficient maximum likelihood decoding of linear block codes using a trellis,* IEEE Trans. Information Theory **24**(1978), 76–80.

Algebraic Coding Research Group, Centre for Communications Research, University of Bristol, England
E-mail address: Tim.Blackmore@Bristol.ac.uk

Algebraic Coding Research Group, Centre for Communications Research, University of Bristol, England
E-mail address: Graham.Norton@Bristol.ac.uk