On trellis structures for Reed–Muller codes *

Tim Blackmore and Graham Norton

Algebraic Coding Research Group
Centre for Communications Research
University of Bristol

July 13, 1999

Abstract

We study trellises of Reed–Muller codes from first principles. Our approach to local trellis behaviour seems to be new and yields amongst other things another proof of a result of Berger and Be'ery on the state complexity of Reed–Muller codes. We give a general form of a minimal–span generator matrix of the family of Reed–Muller codes with their standard bit–order. We apply this to determining the number of parallel subtrellises in any uniform sectionalisation of a Reed–Muller code and to designing trellises for Reed–Muller codes with more parallel subtrellises than the minimal trellis, but with the same state complexity.

1 Introduction

1.1 Overview

We write $\mathbb{F}_2$ for the field with two elements. By a *code* we mean a linear block code. A trellis T for a code C is a directed graph, the vertices of which are placed at ordered depths. The edges of T join vertices at adjacent depths and are directed according to the order of the depths. Paths through T pass through one vertex at each depth and are in one–to–one correspondence with the codewords of C. The most important application of a trellis for a code is Viterbi decoding (dynamic programming). Trellises with low vertex counts at each depth are of interest, and the state complexity of a trellis measures this. A code has a unique trellis which simultaneously minimises the number of vertices at each depth, its *minimal trellis*, [13].

Here we are interested in trellises and related generator matrices for Reed–Muller ($\mathcal{RM}$)–codes, which have received considerable interest, e. g. in [10, 12] and the articles cited there.

Equivalent codes can have different trellises and so the order of the bits of an $\mathcal{RM}$–code is important. The bit–order of a length 2^m $\mathcal{RM}$–code is determined by the order of $\mathbb{F}_2^m$ (see Section 1.3). The *standard bit–order* of such an $\mathcal{RM}$–code comes from the lex[icographical] order of $\mathbb{F}_2^m$. (This is the natural order from both the 'boolean function' and the '$(u|u+v)$ construction' approaches to $\mathcal{RM}$–codes.) It is known that the standard bit–order of $\mathcal{RM}$–codes is optimal with regard to minimising state complexity, [6], and that the extended cyclic bit–order of $\mathcal{RM}$–codes is worst possible, [7].

In Section 2 we characterise the local trellis behaviour of a length 2^m $\mathcal{RM}$–code whose bit–order is determined by any monomial order of $\mathbb{F}_2^m$. (We note that lex order of $\mathbb{F}_2^m$ is a monomial order

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

but that the extended cyclic bit–order of $\mathcal{RM}$–codes comes from an order of $\mathbb{F}_2^m$ which in general is not monomial.) We use this to show that a total degree order is as bad as the extended cyclic bit–order with regard to state complexity. From Section 2.2 onwards we consider only the standard bit–order. We use our description of the local trellis behaviour to give new (and simpler) proofs of some known results on the state complexity of $\mathcal{RM}$–codes, such as the recurrence relations of [10] and the actual value of the state complexity found in [2]. In the process, we determine a depth at which state complexity is attained, which we use later.

In [9] an algorithm for converting a generator matrix for a code C into a trellis for C is given. A generator matrix that gives the minimal trellis is called a *minimal–span generator matrix.* [1] In Section 3 we give a general form for minimal–span generator matrices for the family of $\mathcal{RM}$–codes with their standard bit–order. (A minimal–span generator matrix can be determined for any given $\mathcal{RM}$–code using an algorithm in [11], however, as far as we are aware, there is no known general form for minimal–span generator matrices for the family of $\mathcal{RM}$–codes.)

A subtrellis of T is a trellis whose paths are a subset of the set of paths through T. Subtrellises are parallel if they have no vertices in common other than the initial vertex and a final vertex. Parallel subtrellises in a trellis can be used for parallel processing, [12]. The number of parallel subtrellises can be increased by dividing a trellis into sections (as described in Section 4). In Sections 4 and 5 we use the general form minimal–span generator matrix and results of [12] to determine the number of parallel subtrellises in uniform sectionalisations of the minimal trellises of $\mathcal{RM}$–codes and to design trellises for $\mathcal{RM}$–codes with more parallel subtrellises than the minimal trellis, but with the same state complexity.

Some of the results on the local trellis behaviour and state complexity of $\mathcal{RM}$–codes first appeared in [3].

1.2 Trellises

For $n \geq 1$, a length n *trellis* T over an alphabet $\mathbb{A}$ is an edge–labelled, directed graph with the following properties:

- its vertex set, V, has an $(n+1)$–way partition, $V = \bigcup_{i=-1}^{n-1} V_i$;
- its edge set, E, has an n–way partition, $E = \bigcup_{i=0}^{n-1} E_i$ such that E_i is a set of triples (v_{i-1}, a, v_i) with $v_{i-1} \in V_{i-1}$, $v_i \in V_i$ and $a \in \mathbb{A}$; such an edge is from vertex v_{i-1} to vertex v_i and has label a.

We also require the following connectivity properties:

- for $1 \leq i \leq n-1$, if $(v_{i-1}, a_i, v_i) \in E_i$ then there exists $v_{i-2} \in V_{i-2}$ and $a_{i-1} \in \mathbb{A}$ such that $(v_{i-2}, a_{i-1}, v_{i-1}) \in E_{i-1}$ and
- for $0 \leq i \leq n-2$, if $(v_{i-1}, a_i, v_i) \in E_i$ then there exists $v_{i+1} \in V_{i+1}$ and $a_{i+1} \in \mathbb{A}$ such that $(v_i, a_{i+1}, v_{i+1}) \in E_{i+1}$.

We note that $\{-1, \ldots, n-1\}$ is refered to as the set of *depths* and that V_i is the set of vertices at depth i etc. Usually the depths are labelled from 0 to n but -1 to $n-1$ will prove to be more natural when considering trellises for $\mathcal{RM}$–codes. Typically the first and last depths each contain a single vertex, the *initial* and *final* vertices, but we deal with trellises with more than one final vertex. A consequence of the 'edge–set property' is that all edges are between vertices at depths $i-1$ and i for some $0 \leq i \leq n-1$ and that distinct edges between the same two vertices must

[1] A minimal–span generator matrix was initially called a trellis–oriented generator matrix in [4], before the advent of deriving non–minimal trellises from generator matrices. We prefer the term minimal–span generator matrix as trellis–oriented generator matrix does not reflect the minimal nature of this generator matrix.

have different labels. We note also that the fact that a trellis is connected ensures that there is an edge from the initial vertex.

When n and $\mathbb{A}$ are understood, T will always denote a length n trellis over $\mathbb{A}$.

For $-1 \leq i < j \leq n-1$, the set of *branches between depth i and depth j* of T consists of those triples

$$(v_i, (a_{i+1}, a_{i+2}, \ldots, a_j), v_j)$$

such that there exists $(v_i, a_{i+1}, v_{i+1}) \in E_{i+1}, (v_{i+1}, a_{i+2}, v_{i+2}) \in E_{i+2}, \ldots, (v_{j-1}, a_j, v_j) \in E_j$. Such a branch is directed from v_i to v_j and $(a_{i+1}, a_{i+2}, \ldots, a_j)$ is the branch label. By convention, the set of *branches between depths i and i* is V_i.

The set of *paths* through T consists of those n–tuples of the form

$$((v_{-1}, a_0, v_0), (v_0, a_1, v_1), \ldots, (v_{n-2}, a_{n-1}, v_{n-1})),$$

such that $(v_{i-1}, a_i, v_i) \in E_i$ for each $0 \leq i \leq n-1$. Such a path has label $(a_0, \ldots, a_{n-1})$. If there is a unique path for each path label then T is called one–to–one. All the trellises that we consider will be one–to–one.

Let C be a length n code with symbols from $\mathbb{A}$. If the set of path labels of T is equal to the set of codewords of C then T is a *trellis for C*.

EXAMPLE 1.1 *A trellis for a code is not necessarily planar. For example, the reader may verify that the trellis for the $(5,3)$ code in [14, Fig. 2] contains the 'utility' graph $K_{3,3}$ on vertex sets $\{v_1, v_4, v_{10}\}$ and $\{v_2, v_5, v_8\}$, where the root is v_0 and vertices are labelled consecutively within each depth. (We remark that the Viterbi algorithm does not require that the trellis be planar.)*

The complexity of Viterbi decoding is determined by trellis features.

DEFINITION 1.2 *Let T be a length n trellis over $\mathbb{F}_2$. For $-1 \leq i \leq n-1$, we write $s_i(T)$ for $\log_2 |V_i|$, where V_i is the set of vertices of T at depth i. The* state complexity *of T is*

$$s(T) = \max\{s_i(T) \ : -1 \leq i \leq n-1\}.$$

Similarly, for $-1 \leq i \leq j \leq n-1$, we write $b_{i,j}(T)$ for $\log_2 |B_{i,j}|$, where $B_{i,j}$ is the set of branches of T between depths i and j. The branch complexity *of T is*

$$b(T) = \max\{b_{i-1,i}(T) \ : 0 \leq i \leq n-1\}.$$

We note that $s_i(T) = b_{i,i}(T)$ and that $b_{i-1,i}(T) = \log_2 |E_i|$, where E_i is the set of edges of T between depths $i-1$ and i.

There are other measures of trellis complexity, such as the the *edge complexity*, given by $|E| = \sum_{i=0}^{n-1} 2^{b_{i-1,i}(T)}$, and the actual number of computations used in Viterbi decoding with the trellis, given by $2|E| - |V| + 1$, where $|V| = \sum_{i=0}^{n-1} 2^{s_i(T)}$. If T is the minimal trellis for C then we refer to any of the trellis complexities of the minimal trellis of C as a trellis complexity of C. In this case we also write $s_i(C)$ for $s_i(T)$, $b_{i,j}(C)$ for $b_{i,j}(T)$, $s(C)$ for $s(T)$ and $b(C)$ for $b(T)$. We calculate the trellis complexities of an $\mathcal{RM}$–code and its dual in Example 2.3.

1.3 Reed–Muller codes

We work from the definition of Reed–Muller ($\mathcal{R}M$–)codes given in e. g. [1], using variables $X_1, \ldots, X_m$. For $0 \leq r \leq m$, we put

$$\mathrm{Mon}(r,m) = \{X_{i_1} \cdots X_{i_k} : 0 \leq k \leq r \text{ and } 1 \leq i_1 < i_2 < \cdots < i_k \leq m\}$$

and $\mathrm{Poly}(r,m)$ equal to the $\mathbb{F}_2$–linear span of the monomials in $\mathrm{Mon}(r,m)$.

For $\alpha \in \mathbb{F}_2^m$ we write

$$\alpha = (\alpha(1), \ldots, \alpha(m)),$$

where $\alpha(j) \in \mathbb{F}_2$ for $0 \leq j \leq m$. We put $\mathbb{F}_2^m = \{\alpha_0, \ldots, \alpha_{2^m-1}\}$ and assume given a (total) order of $\mathbb{F}_2^m$, $\alpha_0 < \alpha_1 < \cdots < \alpha_{2^m-1}$. The order of $\mathbb{F}_2^m$ is

- a *monomial order* if $\alpha_i(j) \leq \alpha_{i'}(j)$ for all $1 \leq j \leq m$ implies that $\alpha_i \leq \alpha_{i'}$;
- a *total degree order* if $\sum_{j=1}^m \alpha_i(j) \leq \sum_{j=1}^m \alpha_{i'}(j)$ implies that $\alpha_i \leq \alpha_{i'}$;
- *lex[icographical] order* if $i = \sum_{j=1}^m \alpha_i(j) 2^{j-1}$, so that α_i is the *(standard) binary representation* of i.

Lex order and total degree orders are monomial orders. Lex order is the usual counting order of $\mathbb{F}_2^m$.

For a polynomial $f \in \mathrm{Poly}(r,m)$ and the given order of $\mathbb{F}_2^m$, we have the *evaluation* of f,

$$ev(f) = (f(\alpha_0), \ldots, f(\alpha_{2^m-1})),$$

and for $0 \leq r \leq m$ we define $\mathcal{RM}(r,m)$ by

$$\mathcal{RM}(r,m) = ev(\mathrm{Poly}(r,m)).$$

If $\mathbb{F}_2^m$ is ordered by a monomial (respectively total degree) order then we say that $\mathcal{RM}(r,m)$ has a monomial (respectively total degree) bit–order. As in Section 1.1, if $\mathbb{F}_2^m$ has lex order then we say that $\mathcal{RM}(r,m)$ has its standard bit–order. We remark that, although the standard bit–order minimises the state complexity of an $\mathcal{RM}$–code, [6], this does not guarantee that the standard bit-order minimises the other trellis complexities of $\mathcal{RM}$–codes.

We label the columns of a generator matrix of $\mathcal{RM}(r,m)$ from 0 to $2^m - 1$ and write $\dim(r,m)$ for the dimension of $\mathcal{RM}(r,m)$.

2 State complexity of $\mathcal{RM}$–codes

We study local trellis behaviour and introduce points of gain/fall. We then characterise the points of gain/fall of $\mathcal{RM}(r,m)$ with any monomial bit–order. We use this characterisation

(a) to show that the state complexity of $\mathcal{RM}(r,m)$ is worst possible when it has a total degree bit–order,

(b) to give a new proof of the recurrence relations of [10] and

(c) to give a new proof of the value of the state complexity found in [2] when $\mathcal{RM}(r,m)$ has its standard bit–order.

2.1 Points of Gain and Fall

While a minimal–span generator matrix produces a minimal trellis, measures of trellis complexities for a length n code C can also be determined without a minimal–span generator matrix. For $-1 \leq i \leq n-1$ the i^{th} *past subcode of* C, C_i^- is defined as the linear space of codewords of the form $(c_0, \ldots, c_i, 0, \ldots, 0)$ and the i^{th} *future subcode of* C, C_i^+, is the linear space of codewords of the form $(0, \ldots, 0, c_{i+1}, \ldots, c_{n-1})$. In [5] it is shown that for $i \leq j$,

$$b_{i,j}(C) = \dim(C) - \dim(C_i^-) - \dim(C_j^+).$$

In particular, $s_i(C) = \dim(C) - \dim(C_i^-) - \dim(C_i^+)$.

Now $\dim(C_i^-)$ increases from 0 to $\dim(C)$ in unit steps and $\dim(C_i^+)$ decreases from $\dim(C)$ to 0 in unit steps as i goes from -1 to $n-1$. An increase in $\dim(C_i^-)$ leads to a possible decrease in $s_i(C)$ and that a decrease in $\dim(C_i^+)$ leads to a possible increase in $s_i(C)$. Thus we make the following definitions

DEFINITION 2.1 *Let C be a length n code. For $0 \leq i \leq n-1$,*

(i) if $\dim(C_i^+) = \dim(C_{i-1}^+) - 1$ *then i is a* point of gain *of C and*

(ii) if $\dim(C_i^-) = \dim(C_{i-1}^-) + 1$ *then i is a* point of fall *of C.*

We note that if i is both a point of gain and a point of fall of C then $s_i(C) = s_{i-1}(C)$.

Writing $\gamma_j(C)$ for the number of points of gain of C *before and including* j and $\delta_i(C)$ for the number of points of fall of C *before and including* i we have that,

$$\gamma_j(C) = \dim(C) - \dim(C_j^+) \quad \text{and} \quad \delta_i(C) = \dim(C_i^-)$$

and hence

$$b_{i,j}(C) = \gamma_j(C) - \delta_i(C). \tag{1}$$

In particular, $s_i(C) = \gamma_i(C) - \delta_i(C)$.

Knowledge of where the points of gain and points of fall of C occur describes how the minimal trellis of C behaves locally. In the terminology of [8],

- if i is neither a point of gain nor a point of fall then there is an 'extension from each vertex at depth $i-1$',
- if i is a point of gain but not a point of fall then there is a 'simple expansion from each vertex at depth $i-1$',
- if i is not a point of gain and a point of fall then there is a 'simple merger into each vertex at depth i',
- if i is a point of gain and a point of fall then there is a 'butterfly between connected vertices at depths $i-1$ and i'.

In particular if all the points of gain and points of fall of C are known all the usual measures of trellis complexity for C can be determined (as in Example 2.3 for $\mathcal{RM}(1,4)$ and $\mathcal{RM}(2,4)$).

Now for $c = (c_0, \ldots, c_{n-1}) \neq (0, \ldots, 0)$ we have the *initial point* of c,

$$\text{initial}(c) = \min\{i : c_i \neq 0\},$$

and the *final point* of c,

$$\text{final}(c) = \max\{i : c_i \neq 0\}.$$

(We note that $0 \leq \text{initial}(c) \leq \text{final}(c) \leq n-1$.) Thus i is a point of gain of C if and only if there exists $c \in C$ with $\text{initial}(c) = i$ and i is a point of fall of C if and only if there exists $c \in C$ with $\text{final}(c) = i$. A set of $\dim(C)$ codewords with distinct initial (respectively final) points can be used to form a generator matrix for C called a *future–oriented generator matrix* (respectively *past–oriented generator matrix*). We extend the notion of initial and final points to polynomials. So for $0 \neq f \in \text{Poly}(r,m)$ we put $\text{initial}(f) = \text{initial}(ev(f)) = \min\{i : f(\alpha_i) \neq 0\}$ and $\text{final}(f) = \text{final}(ev(f)) = \max\{i : f(\alpha_i) \neq 0\}$. Thus the points of gain of $\mathcal{RM}(r,m)$ occur at the initial points of polynomials in $\text{Poly}(r,m)$ and the points of fall of $\mathcal{RM}(r,m)$ occur at the final points of polynomials in $\text{Poly}(r,m)$. In the proof of Proposition 2.2 we give a past–oriented generator matrix and a future–oriented generator matrix for $\mathcal{RM}(r,m)$.

For $\alpha \in \mathbb{F}_2^m$ we write $|\alpha|_0$ for the number of 0's in α and $|\alpha|_1$ for the number of 1's in α.

PROPOSITION 2.2 *If $\mathcal{RM}(r,m)$ has a monomial bit–order then*

(i) i is a point of gain of $\mathcal{RM}(r,m)$ if and only if $|\alpha_i|_1 \leq r$

(ii) i is a point of fall of $\mathcal{RM}(r,m)$ if and only if $|\alpha_i|_0 \leq r$.

PROOF. Fix a monomial order on $\mathbb{F}_2^m$.

Firstly for $0 \leq k \leq r$ and $1 \leq i_1 < i_2 < \cdots < i_k \leq m$ we have that $X_{i_1} \cdots X_{i_k} \in \mathrm{Poly}(r,m)$. Now for $\alpha = (\alpha(1), \ldots, \alpha(m)) \in \mathbb{F}_2^m$, $X_{i_1} \cdots X_{i_k}(\alpha)$ is non–zero if and only if $\alpha(i_1) = \alpha(i_2) = \cdots = \alpha(i_k) = 1$. Since our order is monomial, we have that $\alpha_{\mathrm{initial}(X_{i_1} \cdots X_{i_k})}$ has 1's in positions $i_1, \ldots, i_k$ and 0's elsewhere. Thus each i with $|\alpha_i|_1 \leq r$ is an initial point of an element of $\mathrm{Poly}(r,m)$ and hence a point of gain of $\mathcal{RM}(r,m)$. Moreover there are $\dim(r,m)$ such points, which is the number of points of gain of $\mathcal{RM}(r,m)$.

A similar argument with polynomials of the form $1 + X_{i_1} \cdots X_{i_k}$ for $0 \leq k \leq r$ gives the points of fall of $\mathcal{RM}(r,m)$. □

We recall that $\mathcal{RM}(r,m)^\perp = \mathcal{RM}(m-r-1,m)$. It is straightforward to see from Proposition 2.2 that (i) if $r \leq m-r-1$ then all points of gain and points of fall of $\mathcal{RM}(r,m)$ are respectively points of gain and points of fall of $\mathcal{RM}(m-r-1)$ and (ii) that i is a point of gain of $\mathcal{RM}(m-r-1,m)$ that is not a point of gain of $\mathcal{RM}(r,m)$ if and only if i is a point of fall of $\mathcal{RM}(m-r-1,m)$ that is not a point of fall of $\mathcal{RM}(r,m)$. In particular $s_i(\mathcal{RM}(r,m)) = s_i(\mathcal{RM}(m-r-1,m))$ (as we would expect since $s(C) = s(C^\perp)$ for all codes C, [4]) and the other trellis complexities of $\mathcal{RM}(r,m)$ are typically less than those of $\mathcal{RM}(m-r-1,m)$. This is illustrated in Example 2.3, where we use Proposition 2.2 to determine the trellis complexities of $\mathcal{RM}(1,4)$ and $\mathcal{RM}(2,4)$ with their standard bit–orders.

EXAMPLE 2.3 *Proposition 2.2 yields the following table when $\mathcal{RM}(1,4)$ and $\mathcal{RM}(2,4)$ have their standard bit–orders.*

i	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
α_i		0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
		0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
		0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
		0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
$\gamma_i(\mathcal{RM}(1,4))$	0	1	2	3	3	4	4	4	4	5	5	5	5	5	5	5	5
$\delta_i(\mathcal{RM}(1,4))$	0	0	0	0	0	0	0	0	1	1	1	1	2	2	3	4	5
$s_i(\mathcal{RM}(1,4))$	0	1	2	3	3	4	4	4	3	4	4	4	3	3	2	1	0
$b_{i-1,i}(\mathcal{RM}(1,4))$		1	2	3	3	4	4	4	4	4	4	4	4	3	3	2	1
$\gamma_i(\mathcal{RM}(2,4))$	0	1	2	3	4	5	6	7	7	8	9	10	10	11	11	11	11
$\delta_i(\mathcal{RM}(2,4))$	0	0	0	0	1	1	2	3	4	4	5	6	7	8	9	10	11
$s_i(\mathcal{RM}(2,4))$	0	1	2	3	3	4	4	4	3	4	4	4	3	3	2	1	0
$b_{i-1,i}(\mathcal{RM}(2,4))$		1	2	3	4	4	5	5	4	4	5	5	4	4	3	2	1

Thus

(a) $b(\mathcal{RM}(1,4)) = 4$ and $b(\mathcal{RM}(2,4)) = 5$,

(b) the edge complexities of $\mathcal{RM}(1,4)$ and $\mathcal{RM}(2,4)$ are 172 and 252 respectively and

(c) the total number of computations needed for Viterbi decoding with the minimal trellises of $\mathcal{RM}(1,4)$ and $\mathcal{RM}(2,4)$ are 195 and 355 respectively.

We now show that $\mathcal{RM}(r,m)$ with a total degree bit–order has state complexity reaching the Wolf upper–bound, [14].

COROLLARY 2.4 *If $\mathcal{RM}(r,m)$ has a total degree bit–order, then*

$$s(\mathcal{RM}(r,m)) = \min\{\dim(r,m), \dim(m-r-1,m)\}.$$

PROOF. Since the state complexity of a code is equal to the state complexity of its dual (e. g. [4]), it is sufficient to show that for $r \leq m-r-1$, the state complexity of $\mathcal{RM}(r,m)$ (with a total degree bit–order) is $\dim(r,m)$.

Since there are $\dim(r,m)$ points of gain of $\mathcal{RM}(r,m)$ it is sufficient (from Equation (1) with $j=i$) to show that all points of gain come before all points of fall i. e. that if i is a point of gain and j a point of fall then $\alpha_i < \alpha_j$. Now from Proposition 2.2 if i is a point of gain then $|\alpha_i|_1 \leq r$ and if j is a point of fall then $|\alpha_j|_0 \leq r$ so that $|\alpha_j|_1 \geq m-r > r \geq |\alpha_i|_1$. By the definition of a total degree order if $|\alpha_i|_1 < |\alpha_j|_1$ then $\alpha_i < \alpha_j$. □

2.2 Recurrence relations for $\mathcal{RM}(r,m)$ with its standard bit–order

For the rest of the paper we take $\mathcal{RM}(r,m)$ to have its standard bit–order. Thus for $0 \leq i \leq 2^m-1$, α_i is the binary representation of i:

$$\alpha_i = (\alpha_i(1), \ldots, \alpha_i(m)) \text{ if and only if } i = \sum_{j=1}^{m} \alpha_i(j) 2^{j-1}.$$

Where there may be ambiguity regarding the value of m we write $\alpha_i^{(m)}$ for α_i. Thus if $0 \leq i = \sum_{j=1}^{n} \alpha_i(j)2^{j-1} \leq 2^{n-1}-1$ and $m \geq n$ then

$$\alpha_i^{(m)} = (\alpha_i(1), \ldots, \alpha_i(n), \underbrace{0, \ldots, 0}_{m-n}).$$

We will use the following abbreviations:

$\gamma_i(r,m)$	$\gamma_i(\mathcal{RM}(r,m))$
$\delta_i(r,m)$	$\delta_i(\mathcal{RM}(r,m))$
$b_{i,j}(r,m)$	$b_{i,j}(\mathcal{RM}(r,m))$
$s_i(r,m)$	$s_i(\mathcal{RM}(r,m))$
$s(r,m)$	$s(\mathcal{RM}(r,m))$.

We will make considerable use of a special case (Corollary 2.6) of the next result. While Proposition 2.5 may be known, we include a proof based on Proposition 2.2 for completeness.

PROPOSITION 2.5 *For $-1 \leq i \leq j \leq 2^m-1$,*

$$b_{i,j}(r,m) = b_{2^m-j-2, 2^m-i-2}(r,m).$$

PROOF. We write $\gamma_j^+(r,m)$ for the number of points of gain *after and including* j and $\delta_i^+(r,m)$ for the number of points of fall *after and including* i. We note that

$$\gamma_j^+(r,m) = \dim(r,m) - \gamma_{j-1}(r,m) \quad \text{and} \quad \delta_i^+(r,m) = \dim(r,m) - \delta_{i-1}(r,m).$$

Now if i has binary representation $(\alpha_i(1), \ldots, \alpha_i(m))$ then $2^m-1-i = \sum_{k=1}^{m} 2^{k-1} - \sum_{k=1}^{m} \alpha_i(k)2^{k-1}$ has binary representation $(1-\alpha_i(1), \ldots, 1-\alpha_i(m))$. Thus with lex order of $\mathbb{F}_2^m$, $|\alpha_i|_0 = |\alpha_{2^m-i-1}|_1$ and $|\alpha_i|_1 = |\alpha_{2^m-i-1}|_0$. In particular, from Proposition 2.2, i is a point of gain (respectively point

of fall) if and only if $2^m - i - 1$ is a point of fall (respectively point of gain). Also, with lex order of $\mathbb{F}_2^m$, $\alpha_i \leq \alpha_j$ if and only $\alpha_{2^m-i-1} \geq \alpha_{2^m-j-1}$. Thus

$$\gamma_j(r,m) = \delta^+_{2^m-j-1}(r,m) \quad \text{and} \quad \delta_i(r,m) = \gamma^+_{2^m-i-1}(r,m)$$

and from (1),

$$\begin{aligned} b_{i,j}(r,m) &= \gamma_j(r,m) - \delta_i(r,m) = \delta^+_{2^m-j-1}(r,m) - \gamma^+_{2^m-i-1}(r,m) \\ &= (\dim(r,m) - \delta_{2^m-j-2}(r,m)) - (\dim(r,m) - \gamma_{2^m-i-2}(r,m)) \\ &= \gamma_{2^m-i-2}(r,m) - \delta_{2^m-j-2}(r,m) = b_{2^m-j-2,2^m-i-2}(r,m). \end{aligned}$$

□

Putting $i = j$ in Proposition 2.5 gives,

COROLLARY 2.6 *For* $-1 \leq i \leq 2^m - 1$,

$$s_i(r,m) = s_{2^m-i-2}(r,m).$$

Next we use Proposition 2.2 to give a new proof of the recurrence relations of [10]. We begin with

PROPOSITION 2.7 *For* $0 \leq n \leq m$,

$$s_{2^n-1}(r,m) = \sum_{j=r-m+n+1}^{r} \binom{n}{j}.$$

PROOF. Since $\alpha^{(m)}_{2^n-1} = (\underbrace{1,\ldots,1}_{n},\underbrace{0,\ldots,0}_{m-n})$, if $i \leq 2^n - 1$ then $\alpha^{(m)}_i = (\alpha_i(1),\ldots,\alpha_i(n),\underbrace{0,\ldots,0}_{m-n})$ for some $\alpha_i(1),\ldots,\alpha_i(n) \in \mathbb{F}_2$. From Proposition 2.2 i is a point of gain if and only if $|\alpha^{(n)}_i|_1 \leq r$ and i is a point of fall if and only if $|\alpha^{(n)}_i|_0 \leq r - m + n$. Thus

$$\gamma_{2^n-1}(r,m) = \sum_{j=0}^{r} \binom{n}{j} \quad \text{and} \quad \delta_{2^n-1}(r,m) = \sum_{j=0}^{r-m+n} \binom{n}{j}$$

from which the result follows. □

In [7, Example 1], the values of $s_i(r,m)$ are calculated for $i = 2^{m-3} - 1, 2^{m-2} - 1, 2^{m-2} + 2^{m-3} - 1, 2^{m-1} - 1$. In view of Corollary 2.6 these give the values of $s_i(r,m)$ for $i = 2^{m-1} + 2^{m-3} - 1, 2^{m-1} + 2^{m-2} - 1, 2^{m-1} + 2^{m-2} + 2^{m-3} - 1$ also. Thus effectively these are the values of s_i for the '8–way uniform sectionalisation' of the minimal trellis of $\mathcal{RM}(r,m)$. (Sectionalisations are described in Section 4.) In Example 2.8 we illustrate how Propositions 2.2 and 2.7 can be used to calculate the values of s_i for uniform sectionalisations of the minimal trellis of $\mathcal{RM}(r,m)$ by recalculating the values of $s_i(r,m)$ given in [7, Example 1].

EXAMPLE 2.8 *Let* $m \geq 3$. *From Proposition 2.7 we get*

$$s_{2^{m-2}-1}(r,m) = \binom{m-2}{r-1} + \binom{m-2}{r} = \binom{m-1}{r} = s_{2^{m-1}-1}(r,m)$$

and

$$s_{2^{m-3}-1}(r,m) = \binom{m-3}{r-2} + \binom{m-3}{r-1} + \binom{m-3}{r},$$

in agreement with [7]. To proceed from $s_{2^{m-2}-1}(r,m)$ to $s_{2^{m-2}+2^{m-3}-1}(r,m)$, we need to count the number of i in the range $2^{m-2} \le i < 2^{m-2}+2^{m-3}-1$ that are points of gain and the number that are points of fall. For $2^{m-2} \le i < 2^{m-2}+2^{m-3}-1$ we have $\alpha_i^{(m)} = (\alpha_i(1),\ldots,\alpha_i(m-3),0,1,0)$ for some $\alpha_i(1),\ldots,\alpha_i(m-3) \in \mathbb{F}_2$. From Proposition 2.2 such an i is a point of gain if and only if $|\alpha_i^{(m-3)}|_1 \le r-1$ and is a point of fall if and only if $|\alpha_i^{(m-3)}|_0 \le r-2$. Thus

$$\begin{aligned} s_{2^{m-2}+2^{m-3}-1}(r,m) &= s_{2^{m-2}-1}(r,m) + \sum_{j=0}^{r-1}\binom{m-3}{j} - \sum_{j=0}^{r-2}\binom{m-3}{j} \\ &= \binom{m-1}{r} + \binom{m-3}{r-1}. \end{aligned}$$

The value of $s_{2^{m-2}+2^{m-3}-1}(r,m)$ given in [7] is $\sum_{j=0}^{r}\binom{m-1}{j} - \sum_{j=0}^{r-2}\binom{m-2}{j} - 2\sum_{j=0}^{r-2}\binom{m-3}{j}$, which agrees with the value above after a little rearrangement.

We now prove the recurrence relations of [10] using points of gain/fall.

THEOREM 2.9 *Let $1 \le r \le m-1$. For $0 \le n \le m-2$ and $2^n \le i \le 2^{n+1}-1$,*

$$s_i(r,m) = s_{i-2^n}(r-1,m-2) + \sum_{j=r-m+n+1}^{r}\binom{n}{j}.$$

PROOF. We take $2^n \le i \le 2^{n+1}-1$ and count the number of points of gain and points of fall between 2^n and i. We treat the cases $n \le m-3$ and $n = m-2$ separately.

Firstly if $2^n \le i \le 2^{n+1}-1$ for some $0 \le n \le m-3$ and $2^n \le j \le i$ then

$$\alpha_j^{(m)} = (\alpha_j(1),\ldots,\alpha_j(n),1,\underbrace{0,\ldots,0}_{m-n-1})$$

for some $\alpha_j(1),\ldots,\alpha_j(n) \in \mathbb{F}_2$, where $m-n-1 \ge 2$. Now $0 \le j-2^n \le i-2^n \le 2^{m-2}-1$ and using Proposition 2.2

1. j is a point of gain of $\mathcal{RM}(r,m)$ if and only if $|\alpha_j^{(m-2)}|_1 \le r$ if and only if $|\alpha_{j-2^n}^{(m-2)}|_1 \le r-1$ if and only if $j-2^n$ is a point of gain of $\mathcal{RM}(r-1,m-2)$ and
2. j is a point of fall of $\mathcal{RM}(r,m)$ if and only if $|\alpha_j^{(m-2)}|_0 \le r-2$ if and only if $|\alpha_{j-2^n}^{(m-2)}|_0 \le r-1$ if and only if $j-2^n$ is a point of fall of $\mathcal{RM}(r,m)$.

Secondly if $2^{m-2} \le i \le 2^{m-1}-1$ and $2^{m-2} \le j \le i$ then $\alpha_j^{(m)} = (\alpha_j(1),\ldots,\alpha_j(m-2),1,0)$ for some $\alpha_j(1),\ldots,\alpha_j(m-2) \in \mathbb{F}_2$. Again $0 \le j-2^{m-2} \le i-2^{m-2} \le 2^{m-2}-1$ and using Proposition 2.2

1. j is a point of gain of $\mathcal{RM}(r,m)$ if and only if $|\alpha_j^{(m-2)}|_1 \le r-1$ if and only $j-2^{m-2}$ is a point of gain of $\mathcal{RM}(r-1,m-2)$ and similarly
2. j is a point of fall of $\mathcal{RM}(r,m)$ if and only if $j-2^{m-2}$ is a point of fall of $\mathcal{RM}(r-1,m-2)$.

In both cases, the number of points of gain between 2^n and i (inclusive) is equal to $\gamma_{i-2^n}(r-1,m-2)$ and the number of points of fall between 2^n and i is $\delta_{i-2^n}(r-1,m-2)$. Thus

$$s_i(r,m) = s_{2^n-1}(r,m) + \gamma_{i-2^n}(r-1,m-2) - \delta_{i-2^n}(r-1,m-2) = s_{2^n-1}(r,m) + s_{i-2^n}(r-1,m-2)$$

and the result follows from Proposition 2.7. □

We note that in view of Corollary 2.6 it does not really matter that Theorem 2.9 does not give recurrence relations for $2^{m-1} \le i \le 2^m-1$.

2.3 A result of Berger and Be'ery revisited

In [2], a rather technical proof of the formula

$$s(r,m) = \sum_{j=0}^{\min\{r,m-r-1\}} \binom{m-2j-1}{r-j} \tag{2}$$

is given. We give a simple inductive proof of Equation (2), based on Theorem 2.9. We show at the same time that state complexity is attained at $i(r,m)$ with

$$\alpha_{i(r,m)} = (0,\ldots,0,\underbrace{1,0,1,0,\ldots,1,0}_{2\min\{r,m-r-1\}})$$

i. e. that $s_{i(r,m)}(r,m) = s(r,m)$. We will require the following lemma:

LEMMA 2.10 *Let ϱ, μ, ν and ν' be integers. If $0 \le \nu < \nu'$ and $\varrho \ge \nu' - \mu$ then*

$$\sum_{j=\nu-\mu}^{\varrho} \binom{\nu}{j} \le \sum_{j=\nu'-\mu}^{\varrho} \binom{\nu'}{j}.$$

PROOF. It is sufficient to show that the result holds when $\nu' = \nu + 1$. Then

$$\sum_{j=\nu+1-\mu}^{\varrho} \binom{\nu+1}{j} - \sum_{j=\nu-\mu}^{\varrho} \binom{\nu}{j} = \left(\sum_{j=\nu+1-\mu}^{\varrho} \binom{\nu}{j-1}\right) - \binom{\nu}{\nu-\mu},$$

which is non–negative since $\varrho \ge \nu + 1 - \mu$. □

The notation

$$s_I(r,m) = \max\{s_i(r,m) : i \in I\}$$

for $I \subseteq \{0,\ldots,2^m-1\}$ will be useful. Often I will be of the form $[i,j] \stackrel{\text{def}}{=} \{i,i+1,\ldots,j\}$.

It is straightforward to see that (2) holds and that state complexity is attained at $i(r,m)$ for $r = 0$ and $r = m$. In particular this is true for $m = 1$ and $0 \le r \le m$. Thus we assume inductively that, for all $m' < m$ and $0 \le r' \le m'$,

$$s(r',m') = \sum_{j=0}^{\min\{r',m'-r'-1\}} \binom{m'-2j-1}{r'-j} \tag{3}$$

and that $s_{i(r',m')}(r',m') = s(r',m')$. We then show that (2) holds and that $s_{i(r,m)}(r,m) = s(r,m)$ for all $0 \le r \le m$. Since we know that the latter are true for $r = 0$ and $r = m$, we take $1 \le r \le m-1$.

Firstly we note that from Corollary 2.6, $s(r,m)$ is attained at some $0 \le i \le 2^{m-1}-1$. Next from Theorem 2.9 we have that for $0 \le n \le m-2$,

$$s_{[2^n,2^{n+1}-1]}(r,m) = s_{[0,2^n-1]}(r-1,m-2) + \sum_{j=r-m+n+1}^{r} \binom{n}{j}.$$

We note that $s_{[0,2^n-1]}(r-1,m-2)$ is non–decreasing as n increases, so that it is sufficient to show that $\sum_{j=r-m+n+1}^{r} \binom{n}{j}$ is also non–decreasing as n increases to deduce that $s(r,m)$ is attained at some $i \in [2^{m-2}, 2^{m-1}-1]$. This then implies that

$$s(r,m) = s(r-1,m-2) + \binom{m-2}{r-1} + \binom{m-2}{r} = s(r-1,m-2) + \binom{m-1}{r}. \tag{4}$$

Applying Lemma 2.10 with $\varrho = r$, $\mu = m - r - 1$ and $0 \leq n = \nu < n' = \nu' \leq m - 2$ (since $r \geq n' - (m - r - 1)$) we have that $\sum_{j=r-m+n+1}^{r} \binom{n}{j}$ is indeed non–decreasing as n increases so that $s(r,m)$ is attained at some $i \in [2^{m-2}, 2^{m-1} - 1]$ and (4) holds. Thus from (3) with $m' = m-2$ and $r' = r - 1 \leq m - 2$ we have that

$$s(r,m) = \sum_{j=0}^{\min\{r-1, m-r-2\}} \binom{m-2-2j-1}{r-1-j} + \binom{m-1}{r} = \sum_{j=1}^{\min\{r, m-r-1\}} \binom{m-2j-1}{r-j} + \binom{m-1}{r}$$

and so (2) holds.

Finally since we know that $s(r,m) = s_{[2^{m-2}, 2^{m-1}-1]}(r,m)$ we have from Theorem 2.9 that if $s(r-1, m-2)$ is attained at j then $s(r,m)$ is attained at $i = j + 2^{m-2}$. From the inductive hypothesis we can take $j = i(r-1, m-2)$ which gives $i = i(r,m)$. Thus we have

THEOREM 2.11 *For $m \geq 1$ and $0 \leq r \leq m$,*

$$s(r,m) = s_{i(r,m)}(r,m) = \sum_{j=0}^{\min\{r, m-r-1\}} \binom{m-2j-1}{r-j}.$$

3 Minimal–span generator matrix for $\mathcal{RM}(r,m)$

Recall that a minimal–span generator matrix for a code C is a generator matrix that gives the minimal trellis of C using the algorithm of [9]. Equivalently a minimal–span generator matrix is a generator matrix which is simultaneously a past–oriented generator matrix and a future–oriented generator matrix, e. g. [9]. (Past–oriented generator matrices and future–oriented generator matrices were defined in Section 2.1).

The two generator matrices for $\mathcal{RM}(r,m)$ implicit in the proof of Proposition 2.2 are well-known. A generator matrix can be converted into a minimal–span generator matrix, e. g. [9]. Thus it is possible to determine a minimal–span generator matrix for a given $\mathcal{RM}$–code (with any bit–order). In this section we determine a general form for minimal–span generator matrices for the family of $\mathcal{RM}$–codes when they have their standard bit–order.

The $2^m \times 2^m$ identity matrix is a minimal–span generator matrix for $\mathcal{RM}(m,m)$. *For the rest of the paper we assume that* $0 \leq r \leq m-1$.

For $0 \leq k \leq r$ we define $\mathcal{U}(k,m)$ to be

$$\left\{ \prod_{j=m-k+1}^{m} (X_j + 1 + \beta(j)) : \beta(j) \in \mathbb{F}_2 \text{ for } m-k+1 \leq j \leq m \right\}$$

and $\mathcal{V}(k,r,m)$ to be

$$\left\{ \prod_{j=1}^{r-k} X_{i_j}(X_{m-k}+1) + \prod_{j=1}^{r-k} (X_{i_j}+1) X_{m-k} : 1 \leq i_1 < \cdots < i_{r-k} \leq m-k-1 \right\}.$$

We note that, by convention $\prod_{\emptyset} = 1$, so that $\mathcal{U}(0,m) = \{1\} = \mathcal{V}(r,r,m-r-1)$. The sets $\mathcal{U}(k,4)$ and $\mathcal{V}(k,2,4)$ for $0 \leq k \leq 2$ are given in Example 3.7.

For sets of polynomials P, Q we write $P \cdot Q = \{p \cdot q : p \in P, q \in Q\}$ and $[ev(P)]$ for a matrix whose rows are the elements of $ev(P)$.

THEOREM 3.1 *For $m \geq 1$ and $0 \leq r \leq m-1$,*

$$G(r,m) = \left[ev\left(\bigcup_{k=0}^{r} \mathcal{U}(k,m) \cdot \mathcal{V}(k,r,m) \right) \right]$$

is a minimal–span generator matrix for $\mathcal{RM}(r,m)$.

PROOF. We take $m \geq 1$ and $0 \leq r \leq m-1$. We prove the theorem by showing that

(i) the rows of $G(r,m)$ are in $\mathcal{RM}(r,m)$ (Lemma 3.2)

(ii) the initial points of the rows of $G(r,m)$ are distinct (Lemma 3.4)

(iii) the final points of the rows of $G(r,m)$ are distinct (Lemma 3.5)

(iv) each point of gain of $\mathcal{RM}(r,m)$ is an initial point of $G(r,m)$ (Lemma 3.6).

We note that (i)—(iii) imply that $G(r,m)$ is a minimal–span generator matrix for a subcode of $\mathcal{RM}(r,m)$ and that (iv) ensures that this subcode is $\mathcal{RM}(r,m)$.

Throughout the proof we take $0 \leq k \leq r$, $p \in \mathcal{U}(k,m)$, given by

$$p = \prod_{j=m-k+1}^{m} (X_j + 1 + \beta(j))$$

for some $\beta(m-k+1), \ldots, \beta(m) \in \mathbb{F}_2$ and $q = q_0 + q_1 \in \mathcal{V}(k,r,m)$, where

$$q_0 = \prod_{j=1}^{r-k} X_{i_j}(X_{m-k}+1) \quad \text{and} \quad q_1 = \prod_{j=1}^{r-k} (X_{i_j}+1) X_{m-k},$$

for some $1 \leq i_1 < \cdots < i_{r-k} \leq m-k-1$. We remark that p is determined by k and $\beta(m-k+1), \ldots, \beta(m)$ and that q is determined by k and $i_1, \ldots, i_{r-k}$.

Part (i) of the proof of Theorem 3.1 follows directly from

LEMMA 3.2 *For $0 \leq k \leq r$, $\mathcal{U}(k,m) \cdot \mathcal{V}(k,r,m) \subseteq \mathrm{Poly}(r,m)$.*

PROOF. Both q_0 and q_1 are in $\mathrm{Poly}(r-k+1,m)$ and have $X_{i_1} \cdots X_{i_{r-k}} X_{m-k}$ as their only monomial in $\mathrm{Mon}(r-k+1,m)$. Thus $q \in \mathrm{Poly}(r-k,m)$ and $p \cdot q \in \mathrm{Poly}(r,m)$. □

For parts (ii) and (iii), we use

LEMMA 3.3 *The initial point of $p \cdot q$ is*

$$\mathrm{initial}(p \cdot q) = 2^{i_1 - 1} + 2^{i_2 - 1} + \cdots + 2^{i_{r-k} - 1} + \sum_{j=m-k+1}^{m} \beta(j) 2^{j-1}$$

and the final point of $p \cdot q$ is

$$\mathrm{final}(p \cdot q) = 2^{m-k} - 1 - \mathrm{initial}(p \cdot q)$$

i. e.

$$\alpha_{\mathrm{initial}(p \cdot q)} = (0\text{—}0, \underset{i_1}{1}, 0\text{—}0, \underset{i_2}{1}, 0\text{—}0, \ldots, 0\text{—}0, \underset{i_{r-k}}{1}, 0\text{—}0, \underset{m-k}{0}, \beta(m-k+1), \ldots, \beta(m)) \qquad (5)$$

and

$$\alpha_{\mathrm{final}(p \cdot q)} = (1\text{—}1, \underset{i_1}{0}, 1\text{—}1, \underset{i_2}{0}, 1\text{—}1, \ldots, 1\text{—}1, \underset{i_{r-k}}{0}, 1\text{—}1, \underset{m-k}{1}, \beta(m-k+1), \ldots, \beta(m)). \qquad (6)$$

PROOF. For $\alpha = (\alpha(1), \ldots, \alpha(m)) \in \mathbb{F}_2^m$ we have that $q_0(\alpha) \neq 0$ if and only if $\alpha(i_1) = \alpha(i_2) = \cdots = \alpha(i_{r-k}) = 1$ and $\alpha(m-k) = 0$. Likewise, $q_1(\alpha) \neq 0$ if and only if $\alpha(i_1) = \alpha(i_2) = \cdots = \alpha(i_{r-k}) = 0$ and $\alpha(m-k) = 1$. Also, $p(\alpha) \neq 0$ if and only if $\alpha(m-k+1) = \beta(m-k+1), \ldots, \alpha(m) = \beta(m)$.

Thus $\alpha_{\text{initial}(p \cdot q_0)}$ is the right–hand side of (5) and

$$\alpha_{\text{final}(p \cdot q_0)} = (\underbrace{1, \ldots, 1}_{m-k-1}, \underset{m-k}{0}, \beta(m-k+1), \ldots, \beta(m)) \tag{7}$$

and

$$\alpha_{\text{initial}(p \cdot q_1)} = (\underbrace{0, \ldots, 0}_{m-k-1}, \underset{m-k}{1}, \beta(m-k+1), \ldots, \beta(m)) \tag{8}$$

and $\alpha_{\text{final}(p \cdot q_1)}$ is the right–hand side of (6). Since (7) is less than (8) in lex order of $\mathbb{F}_2^m$, all non–zero points of $p \cdot q_0$ come before all non–zero points of $p \cdot q_1$ and in particular $\text{initial}(p \cdot q) = \text{initial}(p \cdot q_0)$ and $\text{final}(p \cdot q) = \text{final}(p \cdot q_1)$. □

For the proofs of parts (ii) and (iii) we take $0 \leq k' \leq r$, $p' \in \mathcal{U}(k', m)$, given by

$$p' = \prod_{j=m-k'+1}^{m} (X_j + 1 + \beta'(j))$$

for some $\beta'(m-k'+1), \ldots, \beta'(m) \in \mathbb{F}_2$ and $q' = q'_0 + q'_1 \in \mathcal{V}(k', r, m)$, where

$$q'_0 = \prod_{j=1}^{r-k'} X_{i'_j}(X_{m-k'} + 1) \quad \text{and} \quad q'_1 = \prod_{j=1}^{r-k'} (X_{i'_j} + 1) X_{m-k'},$$

for some $1 \leq i'_1 < \ldots < i'_{r-k'} \leq m-k'-1$. As for p and q, p' is determined by k' and $\beta'(m-k'+1), \ldots, \beta'(m)$ and q' is determined by k' and $i'_1, \ldots, i'_{r-k'}$.

LEMMA 3.4 *The initial points of the rows of $G(r, m)$ are distinct.*

PROOF. From Lemma 3.3,

$$\alpha_{\text{initial}(p' \cdot q')} = (0\text{—}0, \underset{i'_1}{1}, 0\text{—}0, \underset{i'_2}{1}, 0\text{—}0, \ldots, 0\text{—}0, \underset{i'_{r-k'}}{1}, 0\text{—}0, \underset{m-k'}{0}, \beta'(m-k'+1), \ldots, \beta'(m)).$$

We assume that $\alpha_{\text{initial}(p \cdot q)} = \alpha_{\text{initial}(p' \cdot q')}$ and show that $p \cdot q = p' \cdot q'$.

Firstly if $k = k'$ then $i_1 = i'_1, \ldots, i_k = i'_k$ so that $q = q'$ and $\beta(m-k+1) = \beta'(m-k+1), \ldots, \beta(m) = \beta'(m)$ so that $p = p'$. Thus it suffices to show that $k = k'$.

We put $(\alpha(1), \ldots, \alpha(m)) = \alpha_{\text{initial}(p \cdot q)} = \alpha_{\text{initial}(p' \cdot q')}$. Thus $\sum_{j=1}^{m-k-1} \alpha(j) = r-k$, $\alpha(m-k) = 0$ and $\sum_{j=1}^{m-k'-1} \alpha(j) = r-k'$, $\alpha(m-k') = 0$. Now if $k' < k$ then $m-k'-1 \geq m-k$ and

$$r-k' = \sum_{j=1}^{m-k-1} \alpha(j) + 0 + \sum_{j=m-k+1}^{m-k'-1} \alpha(j) \leq (r-k) + [(m-k'-1) - (m-k+1) + 1] = r-k'-1.$$

Similarly $k < k'$ implies that $r-k \leq r-k-1$. □

LEMMA 3.5 *The final points of the rows of $G(r, m)$ are distinct.*

PROOF. From Lemma 3.3,

$$\alpha_{\mathrm{final}(p'\cdot q')} = (1\text{—}1, \underset{i'_1}{0}, 1\text{—}1, \underset{i'_2}{0}, 1\text{—}1, \ldots, 1\text{—}1, \underset{i'_{r-k'}}{0}, 1\text{—}1, \underset{m-k'}{1}, \beta'(m-k'+1), \ldots, \beta'(m)).$$

We assume that $\alpha_{\mathrm{final}(p\cdot q)} = \alpha_{\mathrm{final}(p'\cdot q')}$ and show that $p \cdot q = p' \cdot q'$.

Again if $k = k'$ it is straightforward to see that $p = p'$ and $q = q'$ so it is sufficient to show that $k = k'$. We put $(\alpha'(1), \ldots, \alpha'(m)) = \alpha_{\mathrm{final}(p\cdot q)} = \alpha_{\mathrm{final}(p'\cdot q')}$. Thus

$$\sum_{j=1}^{m-k-1} \alpha'(j) = (m-k-1) - (r-k) = m-r-1 = \sum_{j=1}^{m-k'-1} \alpha'(j)$$

and $\alpha'(m-k) = \alpha'(m-k') = 1$. If $k' < k$ then $m-k'-1 \geq m-k$ and $m-r-1 = \sum_{j=1}^{m-k'-1} \alpha'(j) \geq \sum_{j=1}^{m-k} \alpha'(j) = m-r$. Similarly $k < k'$ implies that $m-r-1 \geq m-r$. □

Finally we prove part (iv):

LEMMA 3.6 *Each point of gain of $\mathcal{RM}(r,m)$ is an initial point of $G(r,m)$.*

PROOF. Let i be a point of gain with $\alpha_i = (\alpha_i(1), \ldots, \alpha_i(m))$. From Lemma 3.3 it suffices to show that there exists a k, $0 \leq k \leq r$, such that $\sum_{j=1}^{m-k-1} \alpha_i(j) = r-k$ and $\alpha_i(m-k) = 0$.

We know from Proposition 2.2 there exists a z, $0 \leq z \leq r$, such that $\sum_{j=1}^{m} \alpha_i(j) = r-z$. Put $K = \{w : \sum_{j=m-w}^{m} \alpha_i(j) \leq w-z\}$. Now $\sum_{j=m-r}^{m} \alpha_i(j) \leq r-z$ so that $r \in K$ and K is non–empty. Let k be the least element of K. Then $0 \leq k \leq r$ and $\sum_{j=m-k}^{m} \alpha_i(j) \leq k-z$. Also $k-1 \notin K$, so that $\sum_{j=m-k+1}^{m} \alpha_i(j) \geq k-z$. Thus we have

$$k-z \leq \sum_{j=m-k+1}^{m} \alpha_i(j) \leq \sum_{j=m-k}^{m} \alpha_i(j) \leq k-z,$$

so that there must be equality throughout. From the central equality we have $\alpha_i(m-k) = 0$ and using the right–hand equality we have

$$\sum_{j=1}^{m-k-1} \alpha_i(j) = \sum_{j=1}^{m} \alpha_i(j) - \sum_{j=m-k}^{m} \alpha_i(j) = (r-z) - (k-z) = r-k.$$

□

This completes the proof of Theorem 3.1. □

EXAMPLE 3.7 *We take $m = 4$ and $r = 2$ and work from the statement of Theorem 3.1. With $k = 0$, $\mathcal{U}(0,4) = \{1\}$ and $\mathcal{V}(0,2,4) = \{X_1X_2 + X_1X_4 + X_2X_4 + X_4, X_1X_3 + X_1X_4 + X_3X_4 + X_4, X_2X_3 + X_2X_4 + X_3X_4 + X_4\}$. With $k = 1$, $\mathcal{U}(1,4) = \{X_4, X_4 + 1\}$ and $\mathcal{V}(1,2,4) = \{X_1 + X_3, X_2 + X_3\}$. With $k = 2$, $\mathcal{U}(2,4) = \{X_3X_4, X_3X_4 + X_4, X_3 + X_3X_4, 1 + X_3 + X_4 + X_3X_4\}$ and $\mathcal{V}(2,2,4) = \{1\}$.*

This gives

$$G(2,4) = \begin{bmatrix} 0001000110001000 \\ 0000010110100000 \\ 0000001111000000 \\ 0000000001011010 \\ 0000000000111100 \\ 0101101000000000 \\ 0011110000000000 \\ 0000000000001111 \\ 0000000011110000 \\ 0000111100000000 \\ 1111000000000000 \end{bmatrix}.$$

Finally for this section, we note that, while the generator matrices implicit in the proof of Proposition 2.2 are a past–oriented generator matrix and a future–oriented generator matrix for $\mathcal{RM}(r,m)$ with any monomial bit–order the generator matrix of Theorem 3.1 is not in general a minimal–span generator matrix for $\mathcal{RM}(r,m)$ with a non–lexicographical monomial bit–order.

4 Parallel subtrellises in uniform sectionalisations of the minimal trellis for $\mathcal{RM}(r,m)$

We start with a trellis T over an alphabet $\mathbb{A}$, with vertex set $V = \bigcup_{i=-1}^{n-1} V_i$ and edge set $E = \bigcup_{i=0}^{n-1} E_i$, as in Section 1.2. A sectionalisation of T is another edge–labelled, directed, connected graph. So given $h_{-1} = -1 < h_1 < \cdots < h_{\nu-1} = n-1$, the *sectionalisation* $T_{h_{-1},\ldots,h_{\nu-1}}$ *of* T *with section boundaries* $h_{-1},\ldots,h_{\nu-1}$ consists of

1. the vertex set $V_{h_{-1},\ldots,h_{\nu-1}} = \bigcup_{j=-1}^{\nu-1} V_{h_j}$ and
2. the edge set $B_{h_{-1},\ldots,h_{\nu-1}} = \bigcup_{j=0}^{\nu-1} B_{h_{j-1},h_j}$ where B_{h_{j-1},h_j} is the set of branches between depths h_{j-1} and h_j.

We refer to the edges of $T_{h_{-1},\ldots,h_{\nu-1}}$ as branches. Clearly $T_{h_{-1},\ldots,h_{\nu-1}}$ can be regarded as having ν sections and so we call it a ν–way sectionalisation of T. For $0 \le j \le \nu-1$, $h_j - h_{j-1}$ is the *length* of section of j. If all the sections have the same length then the sectionalisation is said to be *uniform*. We note that a ν–way uniform sectionalisation of T necessarily has section boundaries $-1, \frac{n}{\nu}-1, \frac{2n}{\nu}-1, \ldots, \frac{(\nu-1)n}{\nu}-1, n-1$. Also for $\nu = n$ we identify $T_{-1,0,\ldots,n-1}$ with T in the obvious way (i. e. by identifying the branch $(v_{i-1},(a_i),v_i) \in B_{i-1,i}$ with the edge $(v_{i-1},a_i,v_i) \in E_i$).

The set of paths through $T_{h_{-1},\ldots,h_{\nu-1}}$ are those ν–tuples of the form

$$((v_{-1},(a_0,\ldots,a_{h_0}),v_{h_0}),(v_{h_0},(a_{h_0+1},\ldots,a_{h_1}),v_{h_1}),\ldots,(v_{h_{\nu-2}},(a_{h_{\nu-2}+1},\ldots,a_{n-1}),v_{n-1})),$$

such that $(v_{h_{j-1}},(a_{h_{j-1}+1},\ldots,a_{h_j}),v_{h_j}) \in B_{h_{j-1},h_j}$ for $0 \le j \le \nu-1$. Such a path has label $(a_0,\ldots,a_{n-1})$ and vertex set $\{v_{-1},v_{h_0},\ldots,v_{n-1}\}$. Paths with vertex sets $\{v_{-1},v_{h_0},\ldots,v_{h_{\nu-2}},v_{n-1}\}$ and $\{v_{-1},v'_{h_0},\ldots,v'_{h_{\nu-2}},v'_{n-1}\}$ are *parallel* if $v_{h_j} \ne v'_{h_j}$ for $0 \le j \le \nu-2$.

A subtrellis of $T_{h_{-1},\ldots,h_{\nu-1}}$ is a trellis whose set of paths are a non–empty set of paths through $T_{h_{-1},\ldots,h_{\nu-1}}$. Two subtrellises $\{P_1,\ldots,P_\mu\}$ and $\{P'_1,\ldots,P'_{\mu'}\}$ are *parallel* if P_i and P'_j are parallel for all $1 \le i \le \mu$ and $1 \le j \le \mu'$. We note that parallel subtrellises are necessarily disjoint. Subtrellises $S_1,\ldots,S_\lambda$ are *parallel* if they are pairwise parallel. We are interested in the largest number of parallel subtrellises that form a partition of $T_{h_{-1},\ldots,h_{\nu-1}}$, which we write $||T_{h_{-1},\ldots,h_{\nu-1}}||$. Thus

$$||T_{h_{-1},\ldots,h_{\nu-1}}|| = \max\{|\{S_1,\ldots,S_\lambda\}| : T = \bigcup_{k=1}^{\lambda} S_k \text{ and } S_1,\ldots,S_\lambda \text{ are parallel}\}.$$

We refer to $||T_{h_{-1},\ldots,h_{\nu-1}}||$ as the number of parallel subtrellises in $T_{h_{-1},\ldots,h_{\nu-1}}$. We note that $||T_{h_{-1},\ldots,h_{\nu-1}}||$ is not the number of sets of parallel subtrellises.

EXAMPLE 4.1 *Let T be the trellis over $\mathbb{A}$ with vertex set $V = V_{-1} \cup V_0 \cup V_1$ and edge set $E = E_0 \cup E_1$ given by*

- $V_{-1} = \{v_{-1}\}$, $V_0 = \{v_0, v_0'\}$ *and* $V_1 = \{v_1\}$;
- $E_0 = \{e_{00} = (v_{-1}, a_{00}, v_0), e_{01} = (v_{-1}, a_{01}, v_0), e_{00}' = (v_{-1}, a_{00}', v_0'), e_{01}' = (v_{-1}, a_{01}', v_0')\}$ *and* $E_1 = \{e_{10} = (v_0, a_{10}, v_1), e_{11} = (v_0, a_{11}, v_1), e_{10}' = (v_0', a_{10}', v_1), e_{11}' = (v_0', a_{11}', v_1)\}$,

as shown if Figure 1.

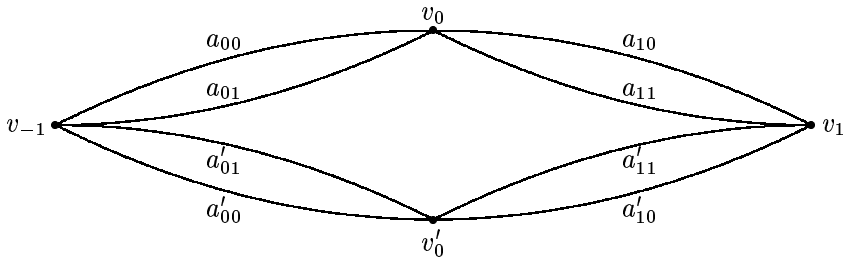

Figure 1: Trellis of Example 4.1

The number of parallel subtrellis in T is 2. We note that $\{\{(e_{00}, e_{10})\}, \{(e_{00}', e_{10}')\}\}$, $\{\{(e_{00}, e_{10}), (e_{01}, e_{11})\}, \{(e_{00}', e_{10}')\}\}$ and $\{\{(e_{00}, e_{10}), (e_{01}, e_{11})\}, \{(e_{00}', e_{10}'), (e_{01}', e_{11}')\}\}$ are all sets containing only parallel subtrellises (and that there are many more).

In [12] a minimal–span generator matrix for a code is used to determine the number of isomorphic parallel subtrellises in uniform sectionalisations of the minimal trellis of the code. Large numbers of such subtrellises are good for Viterbi decoding using parallel processing, [12]. We use our general form minimal–span generator matrix for $\mathcal{RM}$–codes and a result of [12] to calculate the number of isomorphic parallel subtrellises in uniform sectionalisations of the minimal trellises of $\mathcal{RM}$–codes. Uniform sectionalisations of trellises for $\mathcal{RM}$–codes are necessarily 2^u–way sectionalisations for some $0 \leq u \leq m$ and the sections are of length 2^{m-u}. All parallel subtrellises will be isomorphic so we just refer to parallel subtrellises.

DEFINITION 4.2 *For $0 \leq u \leq m$, we write $||r, m, 2^u||$ for the number of parallel subtrellises in the 2^u–way uniform sectionalisation of the minimal trellis of $\mathcal{RM}(r, m)$.*

For example, $||r, m, 1|| = 2^{\dim(r,m)}$, and each parallel subtrellis consists of a single path. From now on, we assume $u \geq 1$.

The *span* of a non–zero codeword c is defined to be $[\text{initial}(c), \text{final}(c)]$ (where as previously, $[i, j] = \{i, i+1, \ldots, j\}$). The following is an immediate consequence of [12, Remark 4, p. 55].

LEMMA 4.3 *The number of rows in a minimal–span generator matrix for $\mathcal{RM}(r, m)$ whose span contains $\{2^{m-u} - 1, 2^m - 2^{m-u}\}$ is $\log_2 ||r, m, 2^u||$.*

Thus we are interested in the spans of the rows of $G(r, m)$. By Theorem 3.1 and Lemma 3.3 a row of $G(r, m)$ has initial point with binary representation

$$(0\text{—}0, \underset{i_1}{1}, 0\text{—}0, \underset{i_2}{1}, 0\text{—}0, \ldots, 0\text{—}0, \underset{i_{r-k}}{1}, 0\text{—}0, \underset{m-k}{0}, \beta(m-k+1), \ldots, \beta(m)) \tag{9}$$

and final point with binary representation

$$(1\text{—}1, \underset{i_1}{0}, 1\text{—}1, \underset{i_2}{0}, 1\text{—}1, \ldots, 1\text{—}1, \underset{i_{r-k}}{0}, 1\text{—}1, \underset{m-k}{1}, \beta(m-k+1), \ldots, \beta(m)) \tag{10}$$

for some $0 \leq k \leq r$, $1 \leq i_1 < \cdots < i_{r-k} \leq m-k-1$ and $\beta(m-k+1), \ldots, \beta(m) \in \mathbb{F}_2$, and conversely for all such $k, i_1, \ldots, i_{r-k}$ and $\beta(m-k+1), \ldots, \beta(m)$ there exists a row of $G(r,m)$ with initial point with binary representation (9) and with final point with binary representation (10). We write $\rho(k, i_1, \ldots, i_{r-k}, \beta(m-k+1), \ldots, \beta(m))$ for this row. We note that $\rho = ev(p \cdot q)$ where $p = p(k, \beta(m-k+1), \ldots, \beta(m))$ and $q = q(k, i_1, \ldots, i_{r-k})$ are as given in the proof of Theorem 3.1.

PROPOSITION 4.4 *For $0 \leq r \leq m-1$ and $1 \leq u \leq m$, $\log_2 ||r, m, 2^u|| = \binom{m-u}{r}$.*

PROOF. Take a row $\rho = \rho(k, i_1, \ldots, i_{r-k}, \beta(m-k+1), \ldots, \beta(m))$ of $G(r,m)$. Then $\text{initial}(\rho) \leq 2^{m-u} - 1$ if and only if

$$\alpha_{\text{initial}(\rho)} \leq (\underbrace{1, \ldots, 1}_{m-u}, \underbrace{0, \ldots, 0}_{u}) \tag{11}$$

and $\text{final}(\rho) \geq 2^m - 2^{m-u}$ if and only if

$$\alpha_{\text{final}(\rho)} \geq (\underbrace{0, \ldots, 0}_{m-u}, \underbrace{1, \ldots, 1}_{u}). \tag{12}$$

In this case $u \geq 1$ implies that $k = 0$ (for otherwise $0 = \beta(m) = 1$) and from (11) (or (12)), $i_r = i_{r-k} \leq m-u$. Conversely if $k = 0$ and $i_r \leq m-u$, then (11) and (12) hold.

Thus the span of ρ contains $\{2^{m-u} - 1, 2^m - 2^{m-u}\}$ if and only if $\rho = \rho(0, i_1, \ldots, i_r)$ for some $1 \leq i_1 < \ldots < i_r \leq m-u$. The number of such rows of $G(r,m)$ is $\binom{m-u}{r}$ and so the result follows from Lemma 4.3. □

COROLLARY 4.5 *For $0 \leq r \leq m-1$, $||r, m, 2^u|| \geq 2$ if and only if $u \leq m-r$.*

EXAMPLE 4.6 *Propositions 2.7 and 4.4 imply that $s_{2^{m-1}-1}(r,m) = \binom{m-1}{r} = \log_2 ||r, m, 2||$. Thus the 2–way uniform sectionalisation of $\mathcal{RM}(r,m)$ consists of $||r, m, 2||$ parallel subtrellises each with a single vertex at each depth.*

Similarly it follows from Proposition 4.4 that for $m \geq 2$,

$$s_{2^{m-2}-1}(r,m) = s_{2^{m-1}-1}(r,m) = s_{2^{m-1}+2^{m-2}-1}(r,m) = \log_2 ||r, m, 2||$$

so that the 4–way uniform sectionalisation of $\mathcal{RM}(r,m)$ consists of $||r, m, 4||$ parallel subtrellises with $||r-1, m, 4||$ vertices at each depth excepting the first and last.

As Corollary 4.5 suggests the minimal trellis of a low–rate $\mathcal{RM}$–code has a higher degree of parallelism than its dual. More formally,

COROLLARY 4.7 *Let $r < m-r-1$. Then $||r, m, 2|| = ||m-r-1, m, 2||$ and for $u \geq 2$,*

$$||r, m, 2^u|| \begin{cases} \geq 2||m-r-1, m, 2^u|| & \text{if } u \leq m-r \\ = ||m-r-1, m, 2^u|| = 1 & \text{otherwise.} \end{cases}$$

PROOF. For $u = 1$, we have

$$\log_2 ||r, m, 2|| = \binom{m-1}{r} = \binom{m-1}{m-1-r} = \log_2 ||m-r-1, m, 2||$$

(as we would expect from Example 4.6 and the fact that $s_i(r,m) = s_i(m-r-1,m)$ for each $-1 \leq i \leq 2^m - 1$).

Also if $u \geq m-r+1$ then $u > r+2 = m-(m-r-1)+1$ so that from Corollary 4.5 $||r, m, 2^u|| = ||m-r-1, m, 2^u|| = 1$.

For $2 \leq u \leq m-r$ we have

$$\log_2 ||r, m, 2^u|| - \log_2 ||m-r-1, m, 2^u|| = \frac{(m-u)!}{(m-u-r)!r!} - \frac{(m-u)!}{(r-u+1)!(m-r-1)!}$$

which equals

$$\frac{(m-u)!}{(m-r-1)!r!}[(m-r-1)(m-r-2)\cdots(m-u-r+1) - r(r-1)\cdots(r-u+2)]$$

which is positive since $m-r-1-j > r-j$ and $m-r-u+1 > 0$. □

For $1 \leq j \leq 2^u$, the vertices at depths $(j-1)2^{m-u} - 1$ and $j2^{m-u} - 1$ in the 2^u–way uniform sectionalisation of the minimal trellis for $\mathcal{RM}(r,m)$ are *adjacent*. Adjacent vertices v_{j-1} (at depth $(j-1)2^{m-u} - 1$) and v_j (at depth $j2^{m-u} - 1$) are *connected* if there exists a branch from v_{j-1} to v_j. We note that there may be more than one branch between adjacent connected vertices (the branches having different labels) but that for fixed j the numbers of branches between pairs of connected vertices at depths $(j-1)2^{m-u} - 1$ and $j2^{m-u} - 1$ are all equal. In [12], it is noted that sectionalisations with more than two branches between adjacent connected vertices are disadvantageous for decoding purposes.

DEFINITION 4.8 *Let $1 \leq j \leq 2^u$. We write $\langle r, m, 2^u, j\rangle$ for the number of branches between connected vertices at depths $(j-1)2^{m-u} - 1$ and $j2^{m-u} - 1$ in the 2^u–way uniform sectionalisation of the minimal trellis of $\mathcal{RM}(r,m)$.*

The following is a corollary of well–known facts about sectionalisations of minimal trellises.

LEMMA 4.9 *For $1 \leq j \leq 2^u$, $\log_2\langle r, m, 2^u, j\rangle$ is the number of rows in a minimal–span generator matrix for $\mathcal{RM}(r,m)$ whose spans are contained in $[(j-1)2^{m-u}, j2^{m-u} - 1]$.*

PROOF. Recall that our depths are labelled from -1 to $2^m - 1$. Set $\epsilon = (j-1)2^{m-u} - 1$ and $\eta = j2^{m-u} - 1$. From [5, p. 1751], $\log_2\langle r, m, 2^u, j\rangle = \dim \mathcal{RM}(r,m)_{\epsilon+1,\eta}$ where

$$\mathcal{RM}(r,m)_{\epsilon+1,\eta} = \{c \in RM(r,m) \ : \ c_k = 0 \text{ for } k \notin [\epsilon+1, \eta]\}.$$

From [9, Property 4, p. 1930], $\log_2\langle r, m, 2^u, j\rangle$ is therefore equal to the number of 'atomic classes' whose span is contained in $[\epsilon+1, \eta]$ i. e. the number of rows in a minimal–span generator matrix whose span is contained in $[\epsilon+1, \eta]$. □

We use a combinatorial lemma to determine $\langle r, m, 2^u, j\rangle$:

LEMMA 4.10 $\sum_{k=u}^{r} 2^{k-u}\binom{m-k-1}{r-k} = \sum_{k=0}^{r-u}\binom{m-u}{k}$.

PROOF. We can assume that $u \le r$. We start with the left–hand side of the required equation. Writing $2^{k-u} = \sum_{l=0}^{k-u} \binom{k-u}{l}$ and reversing the order of summation, we get

$$\sum_{l=0}^{r-u} \sum_{k=l+u}^{r} \binom{k-u}{l} \binom{m-k-1}{r-k} = \sum_{l=0}^{r-u} \sum_{k=l}^{r-u} \binom{k}{l} \binom{m-k-u-1}{m-r-1} =$$
$$\sum_{l=0}^{r-u} \sum_{k=0}^{m-u-1} \binom{k}{l} \binom{m-k-u-1}{m-r-1} = \sum_{l=0}^{r-u} \binom{m-u}{m-r+l} = \sum_{l=0}^{r-u} \binom{m-u}{r-u-l},$$

which is the right–hand side on putting $k = r-u-l$. □

PROPOSITION 4.11 *For $0 \le r \le m-1$ and $1 \le u \le m$,*

$$\log_2 \langle r, m, 2^u, j \rangle \ = \sum_{k=0}^{r-u} \binom{m-u}{k}.$$

PROOF. We count the number of rows $\rho = \rho(k, i_1, \ldots, i_{r-k}, \beta(m-k+1), \ldots, \beta(m))$ of $G(r,m)$ whose spans are contained in $[(j-1)2^{m-u}, j2^{m-u}-1]$.

Let $1 \le j \le 2^u$, with $j-1 = \sum_{l=1}^{u} \alpha_{j-1}(l) 2^{l-1}$ for some $\alpha_{j-1}(1), \ldots, \alpha_{j-1}(u) \in \mathbb{F}_2$. Set $\epsilon = (j-1)2^{m-u} - 1$ and $\eta = j2^{m-u} - 1$. Since $\epsilon + 1 = \sum_{l=m-u+1}^{m} \alpha_{j-1}(l-m+u)2^{l-1}$ and $\eta = (2^{m-u}-1) + \epsilon + 1 = \sum_{l=1}^{m-u} 2^{l-1} + \epsilon + 1$, we have

$$\alpha_{\epsilon+1} = (\underbrace{0, \ldots, 0}_{m-u}, \alpha_{j-1}(1), \ldots, \alpha_{j-1}(u)) \tag{13}$$

and

$$\alpha_{\eta} = (\underbrace{1, \ldots, 1}_{m-u}, \alpha_{j-1}(1), \ldots, \alpha_{j-1}(u)). \tag{14}$$

Thus if $[\text{initial}(\rho), \text{final}(\rho)] \subseteq [\epsilon+1, \eta]$, (13) and (14) imply that $m-k \le m-u$ (since from (9) and (10), $\alpha_{\text{initial}(\rho)} \ne \alpha_{\text{final}(\rho)}$) and that $\beta(m) = \alpha_{j-1}(u), \ldots, \beta(m-u+1) = \alpha_{j-1}(1)$. Conversely if $k \ge u$ and $\beta(m) = \alpha_{j-1}(u), \ldots, \beta(m-u+1) = \alpha_{j-1}(1)$ then $\alpha_{\text{initial}(\rho)}$ is at least (13) and $\alpha_{\text{final}(\rho)}$ is no more than (14).

Thus the total number of rows of $G(r,m)$ whose span is contained in $[\epsilon+1, \eta]$ is $\sum_{k=u}^{r} 2^{k-u} \binom{m-k-1}{r-k}$ and the result follows from Lemmas 4.9 and 4.10. □

We note that Proposition 4.11 implies that $\langle r, m, 2^u, j \rangle = \langle r, m, 2^u, j' \rangle$ for all $1 \le j \le j' \le 2^u$, as in [7, Corollary 1].

COROLLARY 4.12 *For $0 \le r \le m-1$ and $1 \le u \le m$, there are no more than two branches between adjacent connected vertices in the 2^u–way uniform sectionalisation of the minimal trellis of $\mathcal{RM}(r,m)$ if and only if $r \le u$.*

Corollaries 4.5 and 4.12 suggest that for parallel decoding purposes, the most interesting uniform 2^u–way sectionalisations of the minimal trellis for $\mathcal{RM}(r,m)$ are those for which $r \le u \le m-r$. In particular for $r > m-r$, all sectionalisations of $\mathcal{RM}(r,m)$ having parallel subtrellises have at least four branches between adjacent connected vertices.

5 Trellises for $\mathcal{RM}(r,m)$ with parallel subtrellises

As noted in Section 4 parallel subtrellises in trellises can be utilised to speed up decoding using parallel processing. In [12] knowledge of a minimal–span generator matrix for a code C is used to design trellises for C with more parallel subtrellises than the minimal trellis, but with the same state complexity. An analysis of the advantages of such non–minimal trellises is given in [12, Section IV]. We apply a 'coset trellis construction' to the minimal–span generator matrix, $G(r,m)$, obtaining a trellis $T(r,m)$ for $\mathcal{RM}(r,m)$ with complexity $s(r,m)$ (Theorem 5.5) and with $2^{t(r,m)}$ parallel subtrellises, where $t(r,m)$ is determined in Lemma 5.2.

5.1 Coset trellises

We describe the trellis construction of [12]. For all trellises in this section v_{-1} is the vertex at depth -1. We recall that our definition of a trellis allows for more than one final vertex.

Given a trellis T' with vertex set $V' = \bigcup_{i=-1}^{n-1} V'_i$ and edge set $E' = \bigcup_{i=0}^{n-1} E'_i$ over an additive alphabet $(\mathbb{A},+)$ and $a = (a_1,\ldots,a_n) \in \mathbb{A}$ we wish to define the coset trellis $T'+a$. For $a' \neq a$ we would like $T'+a$ and $T'+a'$ to have the same number of vertices at each depth and the same vertex at depth -1, but disjoint sets of vertices at depth i for $0 \leq i \leq n-1$. Thus we put the vertex set of $T'+a$ equal to $V'+a = \bigcup_{i=-1}^{n-1}(V'+a)_i$ where

$$(V'+a)_i = \begin{cases} \{v_{-1}\} & \text{if } i=-1 \\ \{v'_i + a : v'_i \in V'_i\} & \text{if } 0 \leq i \leq n-1 \end{cases}$$

and $v'_i + a$ is merely the formal adjunction of a to v'_i. The edge set of $T'+a$ is $E'+a = \bigcup_{i=0}^{n-1}(E'+a)_i$ where

$$(E'+a)_i = \begin{cases} \{(v_{-1}, a'_0+a_0, v'_0+a) : (v_{-1},a'_0,v'_0) \in E'_0\} & \text{if } i=0 \\ \{(v'_{i-1}+a, a'_i+a_i, v'_i+a) : (v'_{i-1},a'_i,v'_i) \in E'_i\} & \text{if } 1 \leq i \leq n-1. \end{cases}$$

Let C be a length n code over a field $\mathbb{F}$. If C is the union of cosets $C = \bigcup_{k=1}^{N}(C'+c_k)$, where $C' \subseteq C$ and $c_1,\ldots,c_N \in C$ are distinct, and T' is a trellis for C' then we can form a trellis T for C by taking the trellises $T'+c_1,\ldots,T'+c_N$ in parallel. Thus T has vertex set $V = \bigcup_{i=-1}^{n-1} V_i$ where $V_i = \bigcup_{k=1}^{N}(V'+c_k)_i$ and edge set $E = \bigcup_{i=0}^{n-1} E_i$ where $E_i = \bigcup_{k=1}^{N}(E'+c_k)_i$. We note that T has a single vertex at depth -1 and $N|V'_i|$ vertices at depth i for $0 \leq i \leq n-1$. Thus if C is a binary code then $s(T) = \log_2 N + s(T')$. Also $||T|| = N||T'||$. We wish to construct trellises with state complexity no more than C but with parallel trellises.

Now let G be a minimal–span generator matrix for C. If ρ is a row of G we write $\rho \in G$. The *active span* of $\rho \in G$ is defined as $AS(\rho) = [\text{initial}(\rho), \text{final}(\rho)-1]$. It is straightforward to see that for $0 \leq i \leq n-1$,

$$\begin{aligned} |\{\rho \in G : i \in AS(\rho)\}| &= \dim(C) - |\{\rho \in G : \text{final}(\rho) \leq i\}| - |\{\rho \in G : \text{initial}(\rho) \geq i+1\}| \\ &= \dim(C) - \delta_i(C) - (\dim(C) - \gamma_i(C)) = s_i(C). \end{aligned} \tag{15}$$

(Actually this is well–known, e. g. [11].)

For $\rho_1,\ldots,\rho_t \in G$ we put $G \setminus [\rho_1,\ldots,\rho_t]$ equal to the $(\dim(C)-t) \times n$ matrix whose rows are the rows of G excepting $\rho_1,\ldots,\rho_t$ and we put $C \setminus [\rho_1,\ldots,\rho_t]$ equal to the code generated by $G \setminus [\rho_1,\ldots,\rho_t]$. With

$$M_l(C) = \{i : s_i(C) \geq s(C) - l + 1\} \subseteq \{-1,\ldots,n-1\}$$

we have

PROPOSITION 5.1 *If $\rho_1, \ldots, \rho_t$ are rows of a minimal–span generator matrix G for C such that $M_l(C) \subseteq AS(\rho_l)$ for each $1 \leq l \leq t$ then $s(C \setminus [\rho_1, \ldots, \rho_t]) = s(C) - t$.*

PROOF. We write C_t for $C \setminus [\rho_1, \ldots, \rho_t]$. Let $0 \leq i \leq n-1$. From (15) clearly $s_i(C_t) \geq s_i(C) - t$ so that $s(C_t) \geq s(C) - t$.

Now if $s_i(C) \leq s(C) - t$ then $s_i(C_t) \leq s_i(C) \leq s(C) - t$. Otherwise, $s_i(C) = s(C) - l + 1$ for some $1 \leq l \leq t$ and $i \in \bigcap_{k=l}^{t} AS(\rho_k)$ so that from (15), $s_i(C_t) \leq s_i(C) - (t - l + 1) = s(C) - t$. □

Let $\rho_1, \ldots, \rho_t$ be as in the statement of Proposition 5.1 and put $C_t = C \setminus [\rho_1, \ldots, \rho_t]$. Since G is a generator matrix for C, C is the union of the $N = 2^t$ cosets of C_t of the form $C_t + a_1\rho_1 + \ldots + a_t\rho_t$, where $a_1, \ldots, a_t \in \mathbb{F}_2$. Thus if T_t is the minimal trellis for C_t we can form a trellis T for C by taking the 2^t coset trellises $T_t + a_1\rho_1 + \ldots a_t\rho_t$ in parallel. Then $s(T) = t + s(C_t) = s(C)$ by Proposition 5.1 and T has $2^t||T_t||$ parallel subtrellises.

We note that from [12, Remark 4, p. 55] the minimal trellis of a code has two or more parallel subtrellises only if a minimal–span generator matrix for the code contains the all one vector. Thus $||T_t|| \neq 1$ only if C is the code containing only the all zero and all one vectors. In this case the minimal trellis for C has two parallel subtrellises and no trellis for C can have more.

5.2 A maximal submatrix of $G(r, m)$

We specialize to $C = \mathcal{RM}(r, m)$ and write $M_l(r, m)$ for $M_l(\mathcal{RM}(r, m))$. The case $r = 0$ is described in the last paragraph of Section 5.1, *so we take* $r \geq 1$. Thus no subcode of $\mathcal{RM}(r, m)$ generated by rows of a minimal–span generator matrix for $\mathcal{RM}(r, m)$ has a minimal trellis with two or more parallel subtrellises.

Now recall from Section 2 that

(i) $s_i(r, m) = s_{2^m - i - 2}(r, m)$ (Corollary 2.6)

(ii) $i(r, m)$ was defined by $\alpha_{i(r,m)} = (0, \ldots, 0, \underbrace{1, 0, 1, 0, \ldots, 1, 0}_{2\min\{r, m-r-1\}})$ and

(iii) $s_{i(r,m)}(r, m) = s(r, m)$ (Theorem 2.11), so that $i(r, m)$ and $2^m - i(r, m) - 2$ are in $M_1(r, m)$.

Recall also from Section 3 that the rows of $G(r, m)$ are those length 2^m vectors of the form

$$\rho = \rho(k, i_1, \ldots, i_{r-k}, \beta(m - k + 1), \ldots, \beta(m))$$

for some $0 \leq k \leq r$, $1 \leq i_1 < \cdots < i_{r-k} \leq m - k - 1$ and $\beta(m - k + 1), \ldots, \beta(m) \in \mathbb{F}_2$ and that the initial and final points of ρ are given by (9) and (10).

If $M_1(r, m) \subseteq AS(\rho)$ then $\text{initial}(\rho) \leq i(r, m)$ and $\text{final}(\rho) \geq 2^m - i(r, m) - 1$. Now $\alpha_{i(r,m)}(m) = 0$ and $\alpha_{2^m - i(r,m) - 1}(m) = 1$ and for $k > 0$, $\alpha_{\text{initial}(\rho)}(m) = \alpha_{\text{final}(\rho)}(m) = \beta(m)$, (where as usual for $\alpha \in \mathbb{F}_2^m$, $\alpha(m)$ is the m^{th} entry of α). Thus if $M_1(r, m) \subseteq AS(\rho)$, we must have $k = 0$ and $\rho = \rho(0, i_1, \ldots, i_r)$ for some $1 \leq i_1 < \cdots < i_r \leq m - 1$. We put

$$t(r, m) = |\{\rho(0, i_1, \ldots, i_r) \in G(r, m) : \text{initial}(\rho(0, i_1, \ldots, i_r)) \leq i(r, m)\}|$$

and

$$\{\rho_1, \ldots, \rho_{t(r,m)}\} = \{\rho(0, i_1, \ldots, i_r) \in G(r, m) : \text{initial}(\rho(0, i_1, \ldots, i_r)) \leq i(r, m)\}.$$

We assume that the ρ_l are ordered such that $\text{initial}(\rho_1) > \cdots > \text{initial}(\rho_{t(r,m)})$ (and hence $\text{final}(\rho_1) < \cdots < \text{final}(\rho_{t(r,m)})$).

Now $\{\rho_1, \ldots, \rho_{t(r,m)}\}$ is the largest possible set of rows of $G(r, m)$ each containing $M_1(r, m)$. Also if $\rho'_1, \ldots, \rho'_t \in G(r, m)$ are such that $M_l(r, m) \subseteq AS(\rho'_l)$, for $1 \leq l \leq t$, then $M_1(r, m) \subseteq AS(\rho'_l)$ for $1 \leq l \leq t$ so that $\{\rho'_1, \ldots, \rho'_t\} \subseteq \{\rho_1, \ldots, \rho_{t(r,m)}\}$. Thus $\{\rho_1, \ldots, \rho_{t(r,m)}\}$ is the largest possible

set of rows of $G(r,m)$ (or any other minimal–span generator matrix for $\mathcal{RM}(r,m)$) satisfying the conditions of Proposition 5.1. Hence the construction of [12] described above could not be used to produce a trellis T for $\mathcal{RM}(r,m)$ with $s(T) = s(r,m)$ and more than $2^{t(r,m)}$ parallel subtrellises. Although we do not explicitly show that $M_l(r,m) \subseteq AS(\rho_l)$ for each $1 \le l \le t(r,m)$, we will use a construction similar to that of [12] to produce a trellis $T(r,m)$ for $\mathcal{RM}(r,m)$ with $s(T(r,m)) = s(r,m)$ and $2^{t(r,m)}$ parallel subtrellises. First we evaluate $t(r,m)$.

We note that the initial points of $\rho_1, \ldots, \rho_{t(r,m)}$ are those i, $0 \le i \le i(r,m)$, with binary representation

$$(0\text{—}0, \underset{i_1}{1}, 0\text{—}0, \ldots, 0\text{—}0, \underset{i_r}{1}, 0\text{—}0, 0) \tag{16}$$

for some $1 \le i_1 < \cdots < i_r \le m-1$.

LEMMA 5.2 *With $\{\rho_1, \ldots, \rho_{t(r,m)}\}$ defined as above,*

$$t(r,m) = \sum_{k=0}^{\min\{r, m-r-1\}} \binom{m-2k-2}{r-k}.$$

PROOF. Obviously $t(r,m) = |\{\text{initial}(\rho_1), \ldots, \text{initial}(\rho_{t(r,m)})\}|$. Now $\{i : 0 \le i < i(r,m)\}$ is equal to the disjoint union, $\bigcup_{k=0}^{\min\{r-1, m-r-2\}} I(k)$ where

$$I(k) = \{i : (\underbrace{0, \ldots, 0}_{m-2k}, \underbrace{1,0,1,0,\ldots,1,0}_{2k}) \le \alpha_i < (\underbrace{0, \ldots, 0}_{m-2(k+1)}, \underbrace{1,0,1,0,\ldots 1,0}_{2(k+1)})\}$$

(e. g. by induction). Also $i \in I(k)$ if and only if

$$\alpha_i = (\alpha_i(1), \ldots, \alpha_i(m-2k-2), \underbrace{0,0,1,0,1,0,\ldots,1,0}_{2(k+1)})$$

for some $\alpha_i(1), \ldots, \alpha_i(m-2k-2) \in \mathbb{F}_2$, which is of the form (16) if and only if $|\alpha_i^{(m-2k-2)}|_1 = r-k$. Thus the number of $i < i(r,m)$ with α_i of the form (16) is

$$\sum_{k=0}^{\min\{r-1, m-r-2\}} \binom{m-2k-2}{r-k}.$$

Finally $\alpha_{i(r,m)}$ is of the form (16) if and only if $\min\{r, m-r-1\} = r$ and it is easy to check that

$$\binom{m-2\min\{r, m-r-1\}-2}{r-\min\{r, m-r-1\}} = \begin{cases} 1 & \text{if} \quad r \le m-r-1 \\ 0 & \text{if} \quad m-r-1 \le r. \end{cases}$$

□

5.3 The coset trellis $T(r,m)$

We recall that $G(r,m) \setminus [\rho_1, \ldots, \rho_{t(r,m)}]$ is the $(\dim(r,m) - t(r,m)) \times 2^m$ matrix whose rows are the rows of $G(r,m)$ excepting $\rho_1, \ldots, \rho_{t(r,m)}$ and that $\mathcal{RM}(r,m) \setminus [\rho_1, \ldots, \rho_{t(r,m)}]$ is the code generated by $G(r,m) \setminus [\rho_1, \ldots, \rho_{t(r,m)}]$. We write $G_t(r,m)$ for $G \setminus [\rho_1, \ldots, \rho_{t(r,m)}]$ and $C_t(r,m)$ for $\mathcal{RM}(r,m) \setminus [\rho_1, \ldots, \rho_{t(r,m)}]$. Also we write $T_t(r,m)$ for the minimal trellis of $C_t(r,m)$. Next we form $T(r,m)$ by taking the $2^{t(r,m)}$ coset trellises $T_t(r,m) + a_1\rho_1 + \ldots + a_{t(r,m)}\rho_{t(r,m)}$, where $a_1, \ldots, a_{t(r,m)} \in \mathbb{F}_2$, in parallel. Clearly $T(r,m)$ consists of $2^{t(r,m)}$ parallel subtrellises. It remains to show that $s(T(r,m)) = s(r,m)$, i. e. that $s(C_t(r,m)) = s(r,m) - t$. We begin with

LEMMA 5.3 *For* $-1 \le i \le 2^m - 1$, $s_i(C_t(r,m)) = s_{2^m-2-i}(C_t(r,m))$. *In particular there is an* i, $0 \le i \le 2^{m-1} - 1$, *with* $s_i(C_t(r,m)) = s(C_t(r,m))$.

PROOF. From Proposition 2.2, i is a point of gain of $\mathcal{RM}(r,m)$ if and only if $2^m - i - 1$ is a point of fall of $\mathcal{RM}(r,m)$. Also $\text{initial}(\rho_k) = 2^m - \text{final}(\rho_k) - 1$ for $1 \le k \le t(r,m)$ (since this is true for all $\rho(0, i_1, \ldots, i_r) \in G(r,m)$ by Equations (9) and (10)). Since the points of gain of $C_t(r,m)$ are the points of gain of $\mathcal{RM}(r,m)$ excepting $\text{initial}(\rho_1), \ldots, \text{initial}(\rho_{t(r,m)})$ and the points of fall of $C_t(r,m)$ are the points of fall of $\mathcal{RM}(r,m)$ excepting $\text{final}(\rho_1), \ldots, \text{final}(\rho_{t(r,m)})$, we deduce that i is a point of gain of $C_t(r,m)$ if and only if $2^m - i - 1$ is a point of fall of $C_t(r,m)$. The proof of the lemma is then similar to that of Proposition 2.5 with $i = j$. □

For $I \subseteq \{0, 2^m - 1\}$ we write $s_I(C_t(r,m))$ for $\max\{s_i(C_t(r,m)) : i \in I\}$.

LEMMA 5.4 *For* $1 \le r \le m-1$, $s(C_t(r,m)) = s(r-1, m-1)$.

PROOF. In view of Lemma 5.3 we need only find $s_{[0,2^{m-1}-1]}(C_t(r,m))$. We use Proposition 2.2 without reference.

We first find $s_{[0,i(r,m)]}(C_t(r,m))$. So take $0 \le j \le i(r,m)$. Now the points of gain of $C_t(r,m)$ are the points of gain of $\mathcal{RM}(r,m)$ excepting $\text{initial}(\rho_1), \ldots, \text{initial}(\rho_{t(r,m)})$. Hence j is a point of gain of $C_t(r,m)$ if and only if j is a point of gain of $\mathcal{RM}(r,m)$ for which α_j is not of the form (16) if and only if $|\alpha_j^{(m-1)}|_1 = |\alpha_j^{(m)}|_1 \le r-1$ if and only if j is a point of gain of $\mathcal{RM}(r-1, m-1)$. Also, since $\text{final}(\rho_{t(r,m)}) > \ldots > \text{final}(\rho_1) > 2^{m-1} - 1$, j is a point of fall of $C_t(r,m)$ if and only if j is a point of fall of $\mathcal{RM}(r,m)$ if and only if $|\alpha_j^{(m)}|_0 \le r$ if and only if $|\alpha_j^{(m-1)}|_0 \le r-1$ if and only if j is a point of fall of $\mathcal{RM}(r-1, m-1)$. Thus for $0 \le i \le i(r,m)$, $s_i(C_t(r,m)) = s_i(r-1, m-1)$, and in particular $s_{[0,i(r,m)]}(C_t(r,m)) = s_{[0,i(r,m)]}(r-1, m-1)$. Since $i(r-1, m-1) \le i(r,m)$, we have that $s_{[0,i(r,m)]}(C_t(r,m)) = s(r-1, m-1)$.

Finally, if $i(r,m) < j \le 2^{m-1} - 1$ then j is a point of gain (respectively point of fall) of $C_t(r,m)$ if and only if j is a point of gain (respectively point of fall) of $\mathcal{RM}(r,m)$. Since $s_i(r,m) \le s_{i(r,m)}(r,m)$ for $i(r,m) < i \le 2^{m-1} - 1$ it follows that $s_i(C_t(r,m)) \le s_{i(r,m)}(C_t(r,m))$ for $i(r,m) < i \le 2^{m-1} - 1$ and the lemma is proved. □

THEOREM 5.5 *For* $1 \le r \le m-1$, $s(T(r,m)) = s(r,m)$.

PROOF. From Lemma 5.4, Theorem 2.11 and Lemma 5.2 we have

$$\begin{aligned} s(C_t(r,m)) + t(r,m) &= s(r-1, m-1) + t(r,m) \\ &= \sum_{j=0}^{\min\{r-1, m-r-1\}} \binom{m-2j-2}{r-j-1} + \sum_{j=0}^{\min\{r, m-r-1\}} \binom{m-2j-2}{r-j} \\ &= \sum_{j=0}^{\min\{r, m-r-1\}} \binom{m-2j-1}{r-j} = s(r,m). \end{aligned}$$

Thus $s(C_t(r,m)) = s(r,m) - t(r,m)$ and so $s(T(r,m)) = s(r,m)$. □

We note that for $r = m-1$ we get $t(r,m) = 0$.

EXAMPLE 5.6 *For the length* 32 $\mathcal{RM}$*-codes we get* $t(1,5) = 4$, $t(2,5) = 5$ *and* $t(3,5) = 2$ *and for the length* 64 $\mathcal{RM}$*-codes we get* $t(1,6) = 5$, $t(2,6) = 9$, $t(3,6) = 9$ *and* $t(4,6) = 1$. *These values agree with the left most values of* $P_{\max,L}$ *in [12, Table IV].*

An alternative approach to that of this section is to form a generator matrix $G'(r,m)$ from $G(r,m)$ by replacing the rows $\rho_1, \ldots, \rho_{t(r,m)}$ with $1+\rho_1, \ldots, 1+\rho_{t(r,m)}$. That $G'(r,m)$ is a generator matrix for $\mathcal{RM}(r,m)$ is not hard to prove. Also, since $r \geq 1$, $\text{initial}(1+\rho_1) = \cdots = \text{initial}(1+\rho_{t(r,m)}) = 0$ and $\text{final}(1+\rho_1) = \cdots = \text{final}(\rho_{t(r,m)}) = 2^m - 1$. Thus the trellis construction of [9] produces a trellis with $2^{t(r,m)}$ divergences at $i = 0$ and $2^{t(r,m)}$ convergences at $i = 2^m - 1$. The proof that the trellis has state complexity $s(r,m)$ is similar to the proof for $T(r,m)$. This approach has the advantage that the trellis has only one vertex at depth $2^m - 1$ whereas $T(r,m)$ has $2^{t(r,m)}$ vertices at depth $2^m - 1$. Thus decoding using $T(r,m)$ requires an extra comparison at the end. However the approach via coset–trellises gives immediately that $T(r,m)$ consists of $2^{t(r,m)}$ parallel subtrellises.

Acknowledgements The authors gratefully acknowledge financial support from the U.K. Engineering and Physical Sciences Research Council under grant K27728. The first author was supported by the EPSRC. We would also like to thank the referees for useful comments which improved the paper.

References

[1] Ed F. Assmus and Jenny Key (1992) Designs and their codes. Cambridge University Press.

[2] Yuval Berger and Yair Be'ery (1993) *Bounds on the trellis size of linear block codes.* IEEE Trans. Information Theory **39**, 203–209.

[3] Tim Blackmore and Graham Norton (1999) *On the state complexity of some long codes,* in Finite Fields: Theory, Applications and Algorithms, Eds. R. C. Mullin and G. L. Mullen, American Math. Soc. Series in Contemporary Maths. **225**, 203–214.

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

[5] G. David Forney, Jr. (1994) *Dimension/Length profiles and trellis complexity of linear block codes.* IEEE Trans. Information Theory **40**, 1741–1752.

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

[7] Tadao Kasami, Toyoo Takata, Toru Fujiwara and Shu Lin (1993) *On complexity of trellis structure of linear block codes.* IEEE Trans. Information Theory **39**, 1057–1064.

[8] Aaron B. Kiely, Samuel J. Dolinar, Robert J. McEliece, Laura L. Ekroot and Wei Lin (1996) *Trellis decoding complexity of linear block codes.* IEEE Trans. Information Theory **42**, 1687–1697.

[9] Frank R. Kschischang and Vladislav Sorokine (1995) *On the trellis structure of block codes.* IEEE Trans. Information Theory **41**, 1924–1937.

[10] Chung–Chin Lu and Sy–Hann Huang (1995) *On bit–level trellis complexity of Reed–Muller codes.* IEEE Trans. Information Theory **41**, 2061–2064.

[11] Robert J. McEliece (1996) *On the BCJR trellis for linear block codes.* IEEE Trans. Information Theory **42**, 1072–1092.

[12] Hari T. Moorthy, Shu Lin and Gregory T. Uehara (1997) *Good trellises for IC implementation of Viterbi decoders for linear block codes.* IEEE Trans. Communications **45**, 52–63.

[13] Douglas J. Muder (1988) *Minimal trellises for block codes.* IEEE Trans. Information Theory **34**, 1049–1053.

[14] Jack. K. Wolf (1978) *Efficient maximum likelihood decoding of linear block codes using a trellis.* IEEE Trans. Information Theory **24**, 76–80.