Grassmannian Packings from Operator Reed-Muller Codes

Alexei Ashikhmin
Mathematics of Communications Research Department
Bell Labs, Alcatel-Lucent
Email: aea@research.bell-labs.com
A. Robert Calderbank
Department of Mathematics
Princeton University
Email: calderbank@math.princeton.edu

***Abstract*— This paper introduces multidimensional generalizations of binary Reed Muller codes where the codewords are projection operators, and the corresponding subspaces are widely separated with respect to the chordal distance on Grassmannian space. Parameters of these Grassmannian packings are derived and a low complexity decoding algorithm is developed by modifying standard decoding algorithms for binary Reed Muller codes. The subspaces are associated with projection operators determined by Pauli matrices appearing in the theory of quantum error correction and this connection with quantum stabilizer codes may be of independent interest.**

The Grassmannian packings constructed here find application in noncoherent wireless communication with multiple antennas, where separation with respect to the chordal distance on Grassmannian space guarantees closeness to the channel capacity. It is shown that the capacity of the noncoherent MIMO channel at both low and moderate SNR (under the constraint that only isotropically distributed unitary matrices are used for information transmission) is closely approximated by these packings.

Index Terms **Space-Time codes, Grassmannian packings, chordal distance, noncoherent MIMO channel, Reed-Muller codes.**

I. Introduction

One of the main problems of the theory of error correcting codes is the packing problem on the surface of a unit sphere - we would like to know how to arrange as many as possible points on the surface of a real or complex sphere so that they would be as far apart as possible. An extensive number of works was devoted to construction of good packings and derivation of bounds on the packing size and distance distribution. If now we draw lines from the origin of the sphere and the points we can reformulate the problem as the problem of packing one-dimensional subspaces in the embedding real or complex space. In other words we would like to arrange as many as possible one dimensional subspaces in the embedding real or complex space keeping angles between them as large as possible.

A natural generalization of the standard packing problem arises if we consider subspaces of dimension higher than one. In other words, we would like to arrange M-dimensional subspaces of T-dimensional real or complex space so that they are far apart from each other with respect to some metric. These packings are called packings in Grassmannian space, since the set of all M-dimensional subspaces in a real or complex T-dimensional space forms the Grassmannian space $G(T, M)$. They find applications in wireless communication and statistics and appear to be good candidates for use in quantum information theory.

In this article we suggest several constructions of Grassmannian packings in complex space and study their application for information transmission through noncoherent MIMO channel. We start with overviewing some well known notions from linear algebra.

Let Q be an M-dimensional subspace of $\mathbb{C}^T$. Basic linear algebra states that Q can be defined with the help of its orthogonal projection operator P. If, for instance, columns of a $T \times M$ matrix U form an orthonormal basis of Q then $P = UU^\dagger$ ($U^\dagger$ is the Hermitian adjoint of U). The advantage of using orthogonal projection operators is that any subspace has a unique orthogonal operator, while it has infinitely many bases. A linear operator P is an orthogonal projector iff it is Hermitian and idempotent, i.e.

$$P^\dagger = P \text{ and } PP = P.$$

The dimension of the subspace Q defined by P is

$$\dim(Q) = \mathrm{Tr}\,(P).$$

Any projection operator P has $\mathrm{Tr}\,(P)$ eigenvalues 1 and $T - \mathrm{Tr}\,(P)$ eigenvalues 0.

Let Q_1, Q_2 be two M-dimensional subspaces of $\mathbb{C}^T, T \geq M$. We can define the set of principal angles $\theta_1, \theta_2, \ldots, \theta_M$ ($\pi/2 \geq \theta_1 \geq \theta_2 \geq \ldots \geq \theta_M \geq 0$) between Q_1 and Q_2 as follows. First, we find one-dimensional subspaces, say R_1 and R_2, in Q_1 and Q_2 respectively, that are separated by the largest possible angle. This angle is the first principal angle θ_1. Next, we replace Q_1 and Q_2 with the orthogonal complement of R_1 in Q_1 and the orthogonal complement of R_2 in Q_2 respectively. Applying the fist step to the new Q_1 and Q_2 we obtain the second principal angle θ_2. Repeating this procedure M times we find all principal angles.

The following fact is often useful for determining the principal angles. Let U_1 and U_2 be orthonormal bases of Q_1 and Q_2 respectively. Let d_i be singular values of the matrix $U_1^\dagger U_2$ ordered in the increasing order, that is $d_1 \leq d_2 \leq \ldots \leq d_M$. Then $d_i = \cos\theta_i$.

Unlike the one-dimensional case, in which any meaningful distance between subspaces can be computed from the angle between the subspaces, in the case of $M > 1$ there are many nonequivalent notations of distance. For instance,

- $d(Q_1, Q_2) = (\sin^2(\theta_1) + \sin^2(\theta_2) + \ldots + \sin^2(\theta_m))^{1/2}$ is the chordal or Frobenius distance
- $d(Q_1, Q_2) = (\theta_1^2 + \theta_2^2 + \ldots + \theta_m^2)^{1/2}$ is the Geodesic distance,
- $d(Q_1, Q_2) = \arccos(\cos(\theta_1) \cdot \cos(\theta_2) \cdot \ldots \cdot \cos(\theta_m)$ is the Fubini-Study distance.

Many other distances can be defined in multidimensional case.

We argue in Section II that the chordal distance should be used as an optimization parameter of space-time codes in noncoherent MIMO channel. It is convenient to calculate the chordal distance between subspaces Q_1 and Q_2 through their orthogonal projection operators P_1 and P_2 [7]

$$\begin{aligned} & d(Q_1, Q_2)^2 \\ = \; & \sin^2(\theta_1) + \sin^2(\theta_2) + \ldots + \sin^2(\theta_m) \\ = \; & \mathrm{Tr}\,(P_1) - \mathrm{Tr}\,(P_1 P_2) = M - \mathrm{Tr}\,(P_1 P_2). \end{aligned}$$

In this article we will use a slightly different definition of distance.

Definition 1:

$$d(Q_1, Q_2)^2 = d(P_1, P_2) = \mathrm{Tr}\,(P_1) + \mathrm{Tr}\,(P_2) - 2\mathrm{Tr}\,(P_1 P_2).$$

The reasons for the change of distance definitions are the following. First, definition 1 is applicable for the case when $\mathrm{Tr}\,(P_1) \neq \mathrm{Tr}\,(P_2)$. Second, let us consider the vector space formed by linear operators acting in $\mathbb{C}^T$ with the inner product defined by $\mathrm{Tr}\,(A^\dagger B), A, B \in \mathbb{C}^T$. Then, remembering that for any projection operator $PP = P$ and $P^\dagger = P$, we see that definition 1 is the standard definition of the Euclidean distance in a vector space. Note also, that if $\mathrm{Tr}\,(P_1) = \mathrm{Tr}\,(P_2)$ then the distance computed according to definition 1 is just twice the chordal distance.

We will construct and analyze Grassmannian packings with respect to distance metric defined in Definition 1.

Definition 2: An $(T, \leq M, L, d)$ Grassmannian packing is a collection of orthogonal projectors $P_1, \ldots, P_L$ acting on $\mathbb{C}^T$ and such that

$$\begin{aligned} \mathrm{Tr}\,(P_i) &\leq M, \; 1 \leq i \leq L; \\ d(P_i, P_j) &\geq d, \; \text{for } 1 \leq i < j \leq L. \end{aligned}$$

The parameter

$$d = \min_{1 \leq i < j \leq L} \mathrm{Tr}\,(P_i) + \mathrm{Tr}\,(P_j) - 2\mathrm{Tr}\,(P_i P_j)$$

is the minimum distance of the packing. We will use notation (T, M, L, d) for the case when $\mathrm{Tr}\,(P_i) = M, \; 1 \leq i \leq L$.

We are motivated by two important applications. First, Grassmannian packings are highly structured examples of dictionaries, and the problem of finding sparse representations in non-orthogonal dictionaries has been considered by a number of authors (see for example [26] and references over there). The minimum distance of a packing determines the coherence of the dictionary, which is critical to the design of fast algorithms. The second application is to noncoherent wireless communication with multiple antennas. We consider this application in more details in the next section.

II. Information Transmission through the Non-Coherent MIMO Channel

The standard model of MIMO communication is the following. We assume that the transmitter and receiver are equipped with M and N antennas respectively. It is standard practice to transmit information through a space-time code, which is defined as a set $\{U_1, \ldots, U_L\}$ of $T \times M$ complex matrices. The matrices U_i are called *codewords* of the space-time code; the columns are indexed by transmit antennas, the rows are indexed by time slots in a data frame, and the entries are the symbols to be transmitted.

The received signal X can be written in the form

$$X = \sqrt{\frac{\rho}{M}} U H + W. \tag{1}$$

Here H is an $M \times N$ matrix whose $h_{i,j}$ entry is the fading coefficient between i-th transmitting and j-th receiving antennas. The $T \times N$ matrix W corresponds to the additive noise. Entries of matrices H and W are complex Gaussian random variables $\mathcal{CN}(0, 1)$. We assume that the matrix H does not change during time interval T, which is called the *coherence interval*. The average power of transmitted signal S is equal to one, that is

$$\frac{1}{M} \sum_{i=1}^{M} \mathbb{E}|u_{t,i}|^2 = 1, \; t = 1, \ldots, T.$$

The symbol ρ represents the expected signal-to-noise ratio at each receiver antenna.

In the case when the matrix H is known at the receiver, but not at the transmitter, and the coherence interval is small, so that the channel is changing rapidly, the capacity C was derived by Foschini [10] and Telatar [18] and is giving by

$$C = \mathbb{E}\left[\log_2 \det\left(\mathbf{I} + \frac{\rho}{M} H H^\dagger\right)\right].$$

The standard approach to coding for this channel (see ten Brink, Kramer, Ashikhmin [19],[20]) is to employ iterative decoding techniques such as low density parity check (LDPC) and repeat-accumulate (RA) codes. Here $T = 1$ in (1) and we combine a standard signal constellation, such as QPSK or M-QAM constellation, with a specially designed LDPC or RA code. Iterative decoding makes it possible to construct transmission schemes that are only 1dB away from the capacity.

Unlike the previous case the problem of constructing capacity achieving codes for noncoherent MIMO channels is far from being solved. In this section we suggest practical ways of attacking the problem.

The non-coherent MIMO channel represents the most extreme communication scenario and corresponds to the case when the matrix H is not known at the receiver and transmitter because of insufficient time to form estimates from a

pilot signal. This situation arises when the transmitter and/or receiver move rapidly. Non-coherent reception is also relevant for networks relying on frequency hopping for the purposes of secure transmission, time-division multiple access, or block-interleaved channels.

Estimation of the capacity of this channel is a difficult and open problem. Marzetta and Hochwald [16] have derived capacities for several particular cases, namely $N = M = 1$, $T \geq 1$, and $M, N \geq 1$, $T = 1$. They also derived several upper and lower bounds for some other cases. For the case when the signal to noise ratio tends to infinity the capacity was derived by Zheng and Tse [21]. [Some other related references can be found in [15].] It was proven in [21] that in the high SNR regime, channel capacity can be achieved if information is transmitted through isotropically distributed unitary matrices. A matrix U is said to be *isotropically distributed unitary* if its columns are orthonormal and the density is unchanged when it is left-multiplied by a $T \times T$ unitary matrix. In [12] the problem of computing the mutual information $I(U;X)$ between a random, according to the uniform distribution, isotropically distributed unitary matrix U and the output X of the noncoherent MIMO channel was considered. It is shown in [12] that the problem can be reduced to the numerical estimation of a definite integral and thus for any given M, N, T, and ρ the mutual information can be accurately estimated. Though there is no a formal proof that uniform distribution maximizes the mutual information for this type of signals, it is very unlikely that any skewness of the distribution would increase the mutual information. Numerical experiments also show that non uniform distributions used for these type of signals reduce $I(U;X)$. For these reasons we will call this capacity as the *unitary space-time code (USTC) capacity* and use it as a benchmark for qualifying space-time codes.

There are many papers devoted to construction and decoding of space-time codes for the noncoherent MIMO channels. Most of these papers assume the communication scenario shown in fig. 1 Space-time codes are constructed with the aim of achieving a small probability of error $\Pr(U_j \neq \widehat{U}_j)$ over a single coherence interval. At the same time as one of our next examples, and many other examples as well, shows this approach does not allow reliable information transmission close to the Shannon limit.

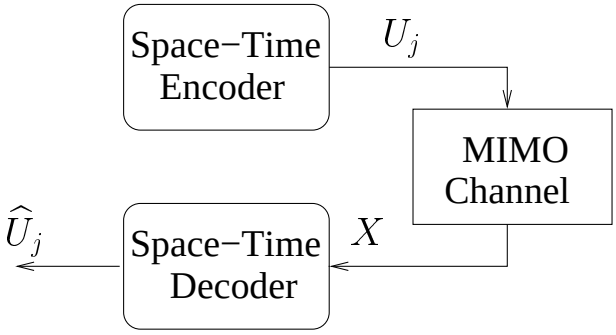


Fig. 1.

One of the ways to approach the capacity of a communication channel is concatenation of an outer code (for instance long LDPC or Repeat Accumulative) code with an appropriately chosen modulation constellation. In such concatenated constructions the outer code should be specifically designed for a given type of modulation detector. The modulation constellation and the demodulation detector should be chosen so that the mutual information between transmitted signals and the output of the detector is close to the channel capacity. One of available methods for construction of such kinds of concatenated schemes is the EXIT function method. For example in [19], [20], [14] concatenated schemes approaching capacity of the ergodic MIMO channel and fading channels were constructed.

We view space-time codes as a generalization of modulation constellations for the case of MIMO transmission. A concatenation scheme based on a space-time code and an outer code is shown in fig.2. We qualify space-time codes using the same criterion that are used for qualifying modulation constellations. Thus our aim is to construct space-time codes and low complexity decoding algorithms such that the mutual information $I(U_j, X)$ and $I(U_j, \widehat{U}_j)$ be close to the noncoherent MIMO capacity. The mutual information $I(U_j, \widehat{U}_j) \leq I(U_j, X)$ depends on the decoding algorithm. If we use a posteriori probability decoding we have $I(U_j, \widehat{U}_j) = I(U_j, X)$. One can design outer LDPC or RA codes, using techniques described in [19], [20], [14], [27], that would allow reliable information transmission with rates only slightly smaller than $I(U_j, \widehat{U}_j)$. Thus if $I(U_j, \widehat{U}_j)$, and therefore $I(U_j, X)$, is close to the noncoherent MIMO capacity one can construct a concatenation scheme that will allow reliable information transmission with rates close to the noncoherent MIMO capacity. For these reason our goal is construction of space-time codes with $I(U_j, X)$ close to the noncoherent MIMO capacity and their decoding error probability plays only a secondary role for us.

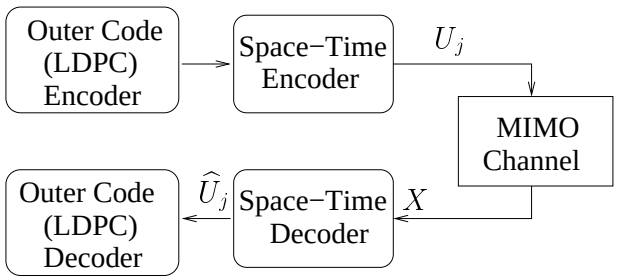


Fig. 2.

Hochwald and Marzetta [13] showed the probability of maximum likelihood (ML) detection error for two unitary signals, say U_1 and U_2, is given by

$$P_e = \sum_j \mathrm{Res}_{\omega=ia_j} \left\{ -\frac{1}{\omega + i/2} \prod_{m=1, d_m>0}^{M} \left[\frac{1+\rho TM}{(\rho T/M)^2(1-d_m^2)(\omega^2+a_m^2)} \right]^N \right\} \quad (2)$$

where

$$a_m = \sqrt{\frac{1}{4} + \frac{1+\rho TM}{(\rho T/M)^2(1-d_m^2)}},\ i=\sqrt{-1},$$

and $d_i = \cos(\theta_i), i = 1, M$, are the singular values of the matrix $U_1^\dagger U_2$. (It will be clear from the context whether i is an index or $\sqrt{-1}$.) They also formulated the rule for ML decoding of a unitary space-time code $\mathcal{U} = \{U_1, U_2, \ldots, U_L\}$ as follows

$$\arg\max_{U_j \in \mathcal{U}} \mathrm{Tr}\,(U_j U_j^\dagger \mathbf{X}\mathbf{X}^\dagger). \tag{3}$$

A simple analysis of expression (2) shows that to achieve a small value of P_e (that is in the high SNR regime) it is important that the product diversity $\prod_j (1-d_j^2)$ be as large as possible [17]. In the case of low SNR the chordal distance (the sum diversity) $\sum_j (1-d_j^2) = \sum_j \sin^2(\theta_i)$, can be significantly more important.

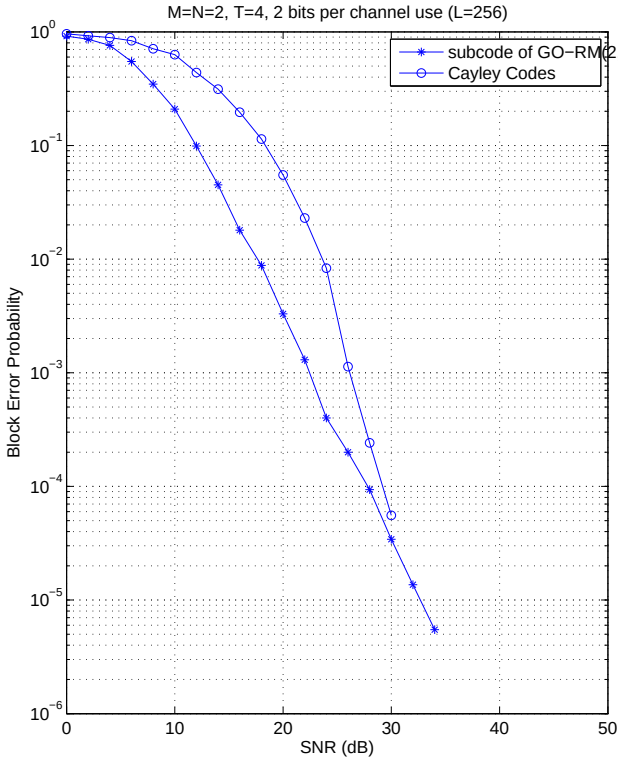


Fig. 3. The block error probability of the subcode of $(O\text{-}RM(1,2))$ and the Caley code

For example let us consider the case when $M = 2, N = 2$, and $T = 4$. In [11] the so-called Cayley space-time codes were suggested. These codes are optimized with respect to the product diversity. In fig.3 we plot the block error probability $\Pr(U_j \neq \widehat{U}_j)$ for the Cayley space-time code of size 256 and a subcode of $GO\text{-}RM(2,2)$ code constructed in IX, optimized with respect to the chordal distance (sum diversity). We take a subcode of $GO\text{-}RM(2,2)$ in order to obtain a space time code of the same size as the Cayley space-time code. We would like to point out two things. First the Cayley space time code becomes better than $GO\text{-}RM(2,2)$ only when the probability of error becomes smaller than, approximately, 10^{-5}. Before this point $GO\text{-}RM(2,2)$ gains several dBs over the Cayley space-time code, for instance it gains about 6dB at the probability of error 10^{-2}. The second thing is that both of these space time codes achieve a small probability of error, say 10^{-4}, about $20dB$ away from the USTC.

In fig.4 we plot the mutual information $I(U_j, X)$ between a uniformly randomly chosen signal U_j and the signal X received at the output of the noncoherent MIMO channel for these two codes. We also plot the USTC. One can see that at low SNR the subcode of $GO\text{-}RM(2,2)$ has the mutual information that closely follows the USTC. At higher values of SNRs the mutual information flattens out and approaches its maximum possible value, which is the code rate $R = 2$. At the same time, in the case of the Cayley space-time code, the mutual information is always far away from the USTC. As we noticed above at SNRs larger than 30dB the Cayley space-time code has smaller probability of error. Hence it is very probable that it will gain in terms of mutual information, compared to the subcode of $GO\text{-}RM(2,2)$, as well. This gain, however, is going to be very small, since the mutual information is already close to its maximum possible value. What is even more important the Cayley space-time code becomes negligibly better than the subcode of $GO\text{-}RM(2,2)$ only **at SNRs that are very far away from the USTC**, that is in the regime of very small importance.

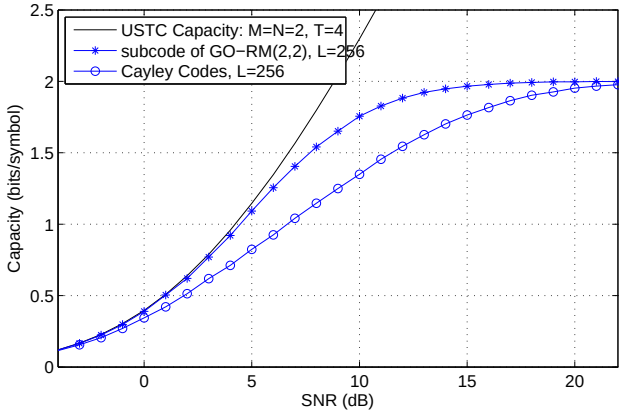


Fig. 4. Mutual information for a subcode of $GO\text{-}RM(2,2)$ and the Caley code

We believe that this example is very generic in the following sense. Let s^* be the SNR at which USTC is equal to the rate R of an unitary space-time code. The mutual information $I(U_j, X)$ of any space-time code with rate R can be close to the noncoherent MIMO channel capacity only at SNRs lower than s^*, because at higher SNR the mutual information is flattens out and approaches R. At these SNRs the decoding error probability of any space-time code of rate R is relatively large and therefore, as it follows from the analysis of (2), codes optimized with respect to the chordal distance (sum diversity) will significantly gain over space time codes optimized with respect to the product diversity.

A large number of papers are devoted to construction space time codes with large product diversity. Based on the above analysis we conjecture that this approach is not quite right if

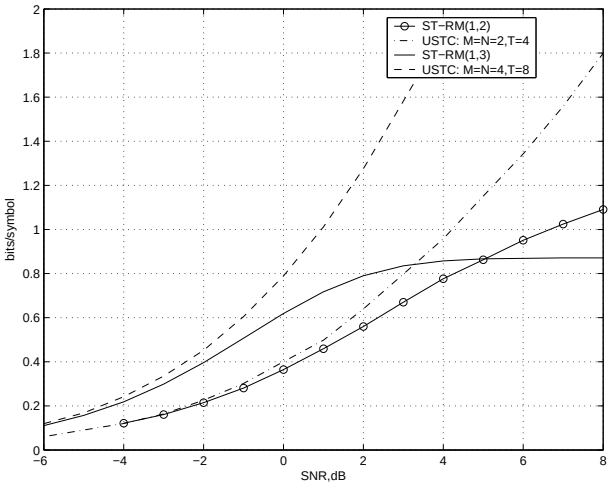


Fig. 5. Capacities of $\mathcal{U}_{unif}$, $I(O\text{-}RM(1,2))$, and $I(O\text{-}RM(1,3))$

we would like to build communication systems approaching the channel capacity. Space time codes with large chordal distance (sum diversity) should be constructed instead.

In fig.5 we plot the mutual information for two other space time codes constructed in section VI for parameters $M = 2, N = 2, T = 4$ and $M = 4, N = 4, T = 8$. One can see again that the mutual information for this codes follows closely the USTC before it starts to flatten out to the code rate.

III. Notation

Throughout this paper we use $\oplus$ to denote binary addition and we use $\otimes$ to denote the tensor or Kronecker product $A \otimes B = (a_{ij}B)$ of two matrices A and B.

The standard inner product of two vectors $\mathbf{v}$ and $\mathbf{w}$ is denoted $\mathbf{v}\cdot\mathbf{w}$, and the componentwise product is denoted $\mathbf{v}\bullet\mathbf{w}$. We will order the binary variables $x_1, x_2, \ldots, x_t$ appearing in a Boolean function $f(x_1, \ldots, x_m)$ lexicographically. For example, if $m = 3$ then the Boolean function $f(x_1, x_2, x_3) = x_1x_2 \oplus x_3 \oplus 1$ is specified by the table

$$\begin{array}{lcccccccc} x_3: & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ x_2: & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ x_1: & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ f: & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}$$

The last row specifies $f(x_1, \ldots, x_m)$ at all possible arguments, and is denoted by $\mathbf{f}$ (it will be clear from the context whether a vector is a row vector or column vector). The vector $\widehat{\mathbf{f}}$ is obtained from $\mathbf{f}$ by exponentiation, that is 0 is replaced by 1 and 1 is replaced by -1, so that $\widehat{f}(x_1, \ldots, x_m) = (-1)^{f(x_1,\ldots,x_m)}$. We will use $\mathbf{f}$ ($\widehat{\mathbf{f}}$) and $f(x_1, \ldots, x_m)$ ($\widehat{f}(x_1, \ldots, x_m)$) interchangeably and the meaning will be clear from the context. We denote the Boolean function $f(x_1, \ldots, x_m) = x_i$ by $\mathbf{x}_i$ (and $\widehat{\mathbf{f}}$ by $\widehat{\mathbf{x}}_i$)). The Boolean function such that $f(x_1, \ldots, x_m) = 0$ for any value of the arguments $x_1, \ldots, x_m$ will be denoted by $0(x_1, \ldots, x_m)$. In general, the value of f at the argument $\mathbf{v}$ is denoted either by $f(\mathbf{v})$ or $f_{\mathbf{v}}$.

Definition 3: The $2^m \times 2^m$ Hadamard matrix H_m is given by

$$H_m = [(-1)^{\mathbf{u}\cdot\mathbf{v}}]_{\mathbf{u},\mathbf{v}\in\mathbb{F}_2^m},$$

and satisfies the recursion

$$H_m = H_1 \otimes H_{m-1}. \tag{4}$$

Definition 4: The Hadamard transform $\widehat{F}(\mathbf{u})$ of a Boolean function $f(x_1, \ldots, x_{2t})$ is defined by

$$\widehat{F}(\mathbf{u}) = \sum_{\mathbf{x}\in\mathbb{F}_2^{2t}} (-1)^{\mathbf{u}\cdot\mathbf{x}\oplus f(\mathbf{x})}, \quad \mathbf{u} \in \mathbb{F}_2^{2t}.$$

or equivalently

$$\widehat{\mathbf{F}} = H_{2t}\widehat{\mathbf{f}}.$$

IV. The Heisenberg-Weyl or Pauli Group

Let σ_x, σ_y, and σ_z be the Pauli matrices, given by

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix},$$

and consider linear operators E of the following form:

$$E = e_1 \otimes \ldots \otimes e_m, \ e_i \in \{I_2, \sigma_x, \sigma_y, \sigma_z\}.$$

The Heisenberg-Weyl or Pauli group $\mathcal{E}_m$ is the group of 4^{m+1} linear operators $\alpha \cdot E, \alpha = \pm 1, \pm i$. (For a detailed description of the role this group plays in the construction of quantum codes see [1], [3], [4].) The center of $\mathcal{E}_m$ is $\mathcal{Z} = \{I_{2^m}, -I_{2^m}, iI_{2^m}, -iI_{2^m}\}$. The quotient group $\overline{\mathcal{E}}_m = \mathcal{E}_m/\mathcal{Z}$ is isomorphic to the binary vector space $\mathbb{F}_2^{2m}$. Binary vectors $(\mathbf{a}, \mathbf{b}) \in \mathbb{F}_2^{2m}$ index operators $E_{(\mathbf{a},\mathbf{b})}$ defined by

$$E_{(\mathbf{a},\mathbf{b})} = e_1 \otimes \ldots \otimes e_m, \tag{5}$$

$$\text{where } e_i = \begin{cases} I_2, & a_i = 0, b_i = 0, \\ \sigma_x, & a_i = 1, b_i = 0, \\ \sigma_z, & a_i = 0, b_i = 1, \\ \sigma_y, & a_i = 1, b_i = 1. \end{cases}$$

The operators $\alpha E_{(\mathbf{a},\mathbf{b})}, \alpha \in \{\pm 1, \pm i\}$, form a coset of $\mathcal{Z}$.

The operators $E_{\mathbf{v}}, \mathbf{v} \in \mathbb{F}_2^{2m}$, satisfy the following properties

- $E_{\mathbf{v}}$ is Hermitian, i.e.

$$E_{\mathbf{v}}^{\dagger} = E_{\mathbf{v}}, \tag{6}$$

- operators $E_{\mathbf{v}}, \mathbf{v} \neq \mathbf{0}$, are traceless, i.e.

$$\mathrm{Tr}\,(E_{\mathbf{v}}) = \begin{cases} 2^m, & \mathbf{v} = \mathbf{0}; \\ 0, & \mathbf{v} \neq \mathbf{0}. \end{cases} \tag{7}$$

- the set of operators $E_{\mathbf{v}}$ forms an orthonormal basis for the space of linear operators acting on $\mathbb{C}_{2^m}$, i.e.

$$\mathrm{Tr}\,(E_{\mathbf{v}}E_{\mathbf{w}}) = \begin{cases} 0, & \mathbf{v} \neq \mathbf{w}; \\ 2^m, & \mathbf{v} = \mathbf{w}. \end{cases} \tag{8}$$

It is easy to check that

$$E_{(\mathbf{a},\mathbf{b})}E_{(\mathbf{a}',\mathbf{b}')} = \alpha E_{(\mathbf{a}\oplus\mathbf{a}',\mathbf{b}\oplus\mathbf{b}')}, \tag{9}$$

where $\alpha \in \{1, -1, i, -1\}$ and $\alpha \in \{1, -1\}$ if $E_{(\mathbf{a},\mathbf{b})}$ and $E_{(\mathbf{a}',\mathbf{b}')}$ commute. From (9) it follows that

$$E_{\mathbf{a},\mathbf{b}}E_{\mathbf{a},\mathbf{b}} = I_{2^m}. \tag{10}$$

The symplectic inner product of vectors $(\mathbf{a}, \mathbf{b}), (\mathbf{a}', \mathbf{b}') \in \mathbb{F}_2^{2m}$ is defined by

$$(\mathbf{a}, \mathbf{b}) * (\mathbf{a}', \mathbf{b}') = \mathbf{a} \cdot \mathbf{b}' \oplus \mathbf{b} \cdot \mathbf{a}'. \tag{11}$$

We will say that vectors $(\mathbf{a},\mathbf{b}),(\mathbf{a}',\mathbf{b}')$ are orthogonal if the symplectic inner product (11) equals zero. It is easy to check that

$$E_{(\mathbf{a},\mathbf{b})}E_{(\mathbf{a}',\mathbf{b}')}=(-1)^{(\mathbf{a},\mathbf{b})*(\mathbf{a}',\mathbf{b}')}E_{(\mathbf{a}',\mathbf{b}')}E_{(\mathbf{a},\mathbf{b})}. \qquad (12)$$

Thus $E_{(\mathbf{a},\mathbf{b})}$ and $E_{(\mathbf{a}',\mathbf{b}')}$ commute if $(\mathbf{a},\mathbf{b})$ and $(\mathbf{a}',\mathbf{b}')$ are orthogonal and anticommute if $(\mathbf{a},\mathbf{b})*(\mathbf{a}',\mathbf{b}')=1$.

We will say that a binary linear code of length $2m$ is self-orthogonal if any two codewords are orthogonal. Any $[2m,k]$ self-orthogonal code C is isomorphic to a subgroup of $\bar{\mathcal{E}}_m$. We now show that the inverse image of this group splits as a direct product and that C is isomorphic to an Abelian subgroup S of $\mathcal{E}_m$.

If $G=\{\mathbf{c}_1,\ldots,\mathbf{c}_k\}$ is a basis of C, then $S=<E_{\mathbf{c}_1},\ldots,E_{\mathbf{c}_k}>$, and the subgroup S consists of operators of the form $\beta_{\mathbf{c}}E_{\mathbf{c}},\mathbf{c}\in C,\beta_{\mathbf{c}}=\pm 1$. If $\mathbf{c}=\mathbf{c}_{i_1}+\mathbf{c}_{i_2}+\ldots+\mathbf{c}_{i_t},\mathbf{c}\in C$, then $\beta_{\mathbf{c}}$ must satisfy $\beta_{\mathbf{c}}E_{\mathbf{c}}=E_{\mathbf{c}_{i_1}}\cdot\ldots\cdot E_{\mathbf{c}_{i_2}}$.

For example consider the self-orthogonal code C with codewords

$$\mathbf{c}_0=(0,0,0,0,0,0),\mathbf{c}_1=(0,0,0,0,1,1),$$
$$\mathbf{c}_2=(1,1,1,1,0,0),\mathbf{c}_3=(1,1,1,1,1,1).$$

If we choose $G=\{\mathbf{c}_1,\mathbf{c}_2\}$ as a basis of C then

$$S=\{E_{\mathbf{c}_0}=I_8,E_{\mathbf{c}_1},E_{\mathbf{c}_2},-E_{\mathbf{c}_3}\}.$$

If we choose $G=\{\mathbf{c}_1,\mathbf{c}_3\}$ then

$$S=\{E_{\mathbf{c}_0}=I_8,E_{\mathbf{c}_1},-E_{\mathbf{c}_2},E_{\mathbf{c}_3}\}.$$

In general there are several ways to construct S, and our results are independent of the choice of construction.

The normalizer $\mathcal{L}$ of $\mathcal{E}_m$ in the unitary group $U(2^m)$ (the subgroup of elements A from $U(2^m)$ such that $A^{-1}\mathcal{E}_mA=\mathcal{E}_m$) includes $\mathcal{E}_m$, all matrices of the form

$$I_2\otimes\ldots\otimes I_2\otimes H\otimes I_2\otimes\ldots\otimes I_2,$$

where

$$H=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\1&-1\end{bmatrix},$$

and all matrices $\mathrm{diag}\,(i^{\phi(\mathbf{v})}),\mathbf{v}\in\mathbb{F}_2^m$, where ϕ is any $\mathbb{Z}_4$-valued quadratic form on $\mathbb{F}_2^m$. The group $\mathcal{L}$ fixes the center $\mathcal{Z}(\mathcal{E}_m)$ and acts transitively on Abelian subgroups of $\mathcal{E}_m$ with a given dimension.

Finally, note that the matrices

$$V=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\-i&i\end{bmatrix},R=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\i&-i\end{bmatrix},$$
$$W=\frac{1}{2}\begin{bmatrix}1-i&1+i\\1+i&1-i\end{bmatrix},Q=\begin{bmatrix}1&0\\0&i\end{bmatrix}.$$

normalize $\mathcal{E}_1$ and satisfy

$$V\sigma_xV^\dagger=\sigma_z,W\sigma_yW^\dagger=\sigma_z, \qquad (13)$$

and

$$Q\sigma_zQ^\dagger=\sigma_z,Q\sigma_yQ^\dagger=\sigma_x,H\sigma_zH^\dagger=\sigma_x,R\sigma_yR^\dagger=-\sigma_x. \qquad (14)$$

V. Operator Reed-Muller Codes

A. Construction

The binary Reed-Muller codes are defined in terms of binary Boolean functions as follows.

Definition 5: The Reed-Muller code $RM(r,m)$ is the set of all vectors $\mathbf{f}$ associated with Boolean functions $f(x_1,x_2,\ldots,x_m)$ of degree at most r. The length of the code $RM(r,m)$ is 2^m, the dimension k is given by

$$k=\binom{m}{0}+\binom{m}{1}+\ldots+\binom{m}{r},$$

and the minimum distance is $d=2^{m-r}$.

Operators Reed-Muller are also defined in terms of binary Boolean functions, but the binary variables are replaced by certain projection operators.

Every binary vector $\mathbf{v}$ in $\mathbb{F}_2^{2m}$ determines projection operators

$$P_{\mathbf{v},\lambda}=\frac{1}{2}(I_{2^m}+\lambda E_{\mathbf{v}}),\lambda=\pm 1. \qquad (15)$$

Since $E_{\mathbf{v}}$ is Hermitian the operators $P_{\mathbf{v},\lambda}$ are also Hermitian. Further, $P_{\mathbf{v},\lambda}$ is idempotent since, using (10), we have

$$P_{\mathbf{v},\lambda}P_{\mathbf{v},\lambda}=\frac{1}{4}(I_{2^n}+\lambda E_{\mathbf{v}}+I_{2^n}+\lambda E_{\mathbf{v}})=P_{\mathbf{v},\lambda}.$$

In what follows we mostly use projections $P_{\mathbf{v},1}$, and we will abbreviate $P_{\mathbf{v},1}$ by $P_{\mathbf{v}}$.

For any two commuting operators P and P' we define the operation $\oplus$ by

$$P\oplus P'=P+P'-2PP' \qquad (16)$$

The reasons for this definition are the following. If operators P and P' commute, then we may choose a basis to make P and P' diagonal with all diagonal entries either 0 and 1. In this basis $\oplus$ is equivalent to modulo 2 summation of diagonal entries of P and P'. Let now $G=\{\mathbf{v}_1,\ldots,\mathbf{v}_m\},\mathbf{v}_i\in\mathbb{F}_2^{2m}$, be a set of m pairwise orthogonal and linearly independent binary vectors. Then it follows from (15) and (12) that any pair of operators $P_{\mathbf{v}_i}$ commute and therefore can be simultaneously diagonalized. It is not difficult to show that after diagonalization the diagonal entries of any $P_{\mathbf{v}_i}$ from a codeword of $RM(1,m)$. Hence the diagonal entries of the projection operator $P_{\mathbf{v}_i}\oplus P_{\mathbf{v}_j}$ also form a codeword of $RM(1,m)$. Thus the operation $\oplus$ allows us to operate with $P_{\mathbf{v}_i}$ as with codewords of $RM(1,m)$. Moreover, we will show later in this section that using standard matrix multiplication and the operation $\oplus$ we can construct projection operators whose diagonal entries, after diagonalization, form codewords of $RM(r,m)$ for any $1\le r\le m$.

Example 1: Let $m=2$ and $\mathbf{v}_1=(0,0,1,0),\mathbf{v}_2=(0,0,0,1)$ and $\lambda_1=\lambda_2=1$. Corresponding $P_{\mathbf{v}_1}$ and $P_{\mathbf{v}_2}$ are

$$P_{\mathbf{v}_1}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix},P_{\mathbf{v}_2}=\begin{bmatrix}1&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix},\text{ and}$$

$$P_{\mathbf{v}_1} \oplus P_{\mathbf{v}_2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Lemma 1: In general, the operation $\oplus$ is associative, and it is commutative when applied to commuting operators.

Proof: Associativity is a straightforward computation for arbitrary linear operators, and if A_1 commutes with A_2 then

$$A_1 \oplus A_2 = A_1 + A_2 - 2A_1A_2 = A_1 + A_2 - 2A_2A_1 = A_2 \oplus A_1.$$

□

Example 2: Let $m = 3$ and $\mathcal{G} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where $\mathbf{v}_1 = (1,0,0,0,1,0), \mathbf{v}_2 = (0,1,1,1,1,0), \mathbf{v}_3 = (0,0,1,0,1,1)$, and $\lambda_1 = \lambda_2 = \lambda_3 = 1$. It is easy to check that $\mathbf{v}_1, \mathbf{v}_2$, and $\mathbf{v}_3$ are orthogonal with respect to the symplectic inner product (11). Then

$$P_{\mathbf{v}_1} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$P_{\mathbf{v}_2} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & i & 0 & 0 & 0 & 0 \\ 0 & 1 & i & 0 & 0 & 0 & 0 & 0 \\ 0 & -i & 1 & 0 & 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 & 0 & 1 & -i & 0 \\ 0 & 0 & 0 & 0 & 0 & i & 1 & 0 \\ 0 & 0 & 0 & 0 & i & 0 & 0 & 1 \end{bmatrix},$$

$$P_{\mathbf{v}_3} = \frac{1}{2}\begin{bmatrix} 1 & i & 0 & 0 & 0 & 0 & 0 & 0 \\ -i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -i & 0 & 0 & 0 & 0 \\ 0 & 0 & i & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & i & 0 & 0 \\ 0 & 0 & 0 & 0 & -i & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -i \\ 0 & 0 & 0 & 0 & 0 & 0 & i & 1 \end{bmatrix}.$$

Direct computations show that $P = (P_{\mathbf{v}_1}P_{\mathbf{v}_2} \oplus P_{\mathbf{v}_1}) \oplus P_{\mathbf{v}_3} = P_{\mathbf{v}_1}P_{\mathbf{v}_2} \oplus (P_{\mathbf{v}_1} \oplus P_{\mathbf{v_3}}) = P_{\mathbf{v}_1} \oplus (P_{\mathbf{v}_1}P_{\mathbf{v}_2} \oplus P_{\mathbf{v}_3})$ and the result is equal to

$$P = \frac{1}{4}\begin{bmatrix} 2 & i & -1 & 0 & 0 & -i & 1 & 0 \\ -i & 2 & 0 & 1 & i & 0 & 0 & -1 \\ -1 & 0 & 2 & -i & -1 & 0 & 0 & -i \\ 0 & 1 & i & 2 & 0 & 1 & i & 0 \\ 0 & -i & -1 & 0 & 2 & i & 1 & 0 \\ i & 0 & 0 & 1 & -i & 2 & 0 & -1 \\ 1 & 0 & 0 & -i & 1 & 0 & 2 & -i \\ 0 & -1 & i & 0 & 0 & -1 & i & 2 \end{bmatrix}. \quad (17)$$

Associativity and commutativity (Lemma 1) make it possibly unambiguously replace binary variables of a Boolean function by projection operators. Given an arbitrary Boolean function $f(x_1, \ldots, x_m)$ and any set $G = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ of pairwise orthogonal and linearly independent vectors, we define the function $f(P_{\mathbf{v}_1}, \ldots, P_{\mathbf{v}_m})$ where multiplication modulo 2 is replaced by multiplication of operators, addition modulo 2 is replaced by $\oplus$ defined in (16). Constant terms in $f(x_1, \ldots, x_m)$ are replaced by the identity operator I_{2^m}. For example if $f(x_1, \ldots, x_m) = (x_1 \oplus 1)x2 \oplus x_1x_m \oplus 1$, then $f(P_{\mathbf{v}_1}, \ldots, P_{\mathbf{v}_m}) = (P_{\mathbf{v}_1} \oplus I_{2^m})P_{\mathbf{v}_2} \oplus P_{\mathbf{v}_1}P_{\mathbf{v}_m} \oplus I_{2^m}$.

Theorem 1: The operator $P_{f,G} = f(P_{\mathbf{v}_1}, \ldots, P_{\mathbf{v}_m})$ is an orthogonal projection on a subspace of dimension $\mathrm{Tr}\,(P_{f,G}) = \mathrm{wt}\,(\mathbf{f})$.

Proof: It is easy to see that

$$\begin{aligned} &(1,1) \otimes \ldots \otimes (1,1) \otimes \underbrace{(1,-1)}_{\text{the } i\text{-th factor}} \otimes (1,1) \otimes \ldots \otimes (1,1) \\ = \; &(\underbrace{1,\ldots,1}_{2^{i-1}}) \otimes (1,-1) \otimes (\underbrace{1,\ldots,1}_{2^{m-i}}) = \widehat{\mathbf{x}}_i. \end{aligned} \quad (18)$$

In other words the tensor product (18) is the vector associated with the Boolean function $f(x_1, \ldots, x_m) = x_i$.

Let

$$\mathbf{a}_i = (\underbrace{0,0,\ldots,0}_{m}, \underbrace{0,\ldots,0}_{i-1}, 1, \underbrace{0,\ldots,0}_{m-i}) \in \mathbb{F}_2^{2m}.$$

The vectors $\mathbf{a}_i, i = 1, m$, are pairwise orthogonal and linearly independent. Hence we can assume $G = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\}$. Now taking into account (18), we obtain that

$$E_{(\mathbf{0},\mathbf{a}_i)} = I_2 \otimes \ldots \otimes I_2 \otimes \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}}_{\text{the } i\text{-th factor}} \otimes I_2 \otimes \ldots \otimes I_2$$

is a diagonal operator whose diagonal is the Boolean function $\widehat{\mathbf{x}}_i$. From this it follows that

$$P_{(\mathbf{0},\mathbf{a}_i)} = \frac{1}{2}(I_{2^{2m}} + E_{(\mathbf{0},\mathbf{a}_i)})$$

is an orthogonal projection in the diagonal form and

$$\mathrm{diag}\,(P_{(\mathbf{0},\mathbf{a}_i)}) = \mathbf{x}_i.$$

Hence $P_{(\mathbf{0},\mathbf{a}_i)}P_{(\mathbf{0},\mathbf{a}_j)}$ is a projection operator in the diagonal form and

$$\mathrm{diag}\,(P_{(\mathbf{0},\mathbf{a}_i)}P_{(\mathbf{0},\mathbf{a}_j)}) = \mathbf{x}_i \bullet \mathbf{x}_j. \quad (19)$$

Thus $\mathrm{diag}\,(P_{(\mathbf{0},\mathbf{a}_i)}P_{(\mathbf{0},\mathbf{a}_j)})$ is the vector associated with the Boolean function x_1x_2. Similarly one can see from the definition of the operation $\oplus$ that $P_{(\mathbf{0},\mathbf{a}_i)} \oplus P_{(\mathbf{0},\mathbf{a}_j)}$ is a projection operator in the diagonal form and

$$\mathrm{diag}\,(P_{(\mathbf{0},\mathbf{a}_i)} \oplus P_{(\mathbf{0},\mathbf{a}_j)}) = \mathbf{x}_i \oplus \mathbf{x}_j. \quad (20)$$

From (19) and (20) it follows that

$$P_{f,G} = f(P_{(\mathbf{0},\mathbf{a}_1)}, \ldots, P_{(\mathbf{0},\mathbf{a}_m)})$$

is an orthogonal projection in the diagonal form whose diagonal is the vector $\mathbf{f}$ associated with the Boolean function $f(x_1, \ldots, x_m)$. Hence, we have that

$$\mathrm{Tr}\,(P_{f,G}) = \mathrm{wt}\,(\mathbf{f}).$$

Let now $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{F}_2^{2m}$ be an arbitrary set of pairwise orthogonal and linearly independent vectors. The operators $E_{(\mathbf{0},\mathbf{a}_i)}, i = 1, m$, and $E_{\mathbf{v}_1}, \ldots, E_{\mathbf{v}_m}$ form generator sets of two Abelian subgroups, say $\mathcal{A}_1$ and $\mathcal{A}_2$, of $\mathcal{E}_m$. As we noticed in Section IV the group $\mathcal{L}$ acts transitively on such subgroups. Thus there exists a unitary operator $U \in \mathcal{L}$ that moves $\mathcal{A}_2$ into $\mathcal{A}_1$. Without loss of generality we can assume that $\mathbf{v}_i$ are enumerated so that

$$E_{\mathbf{v}_i} = U E_{(\mathbf{0},\mathbf{a}_i)} U^\dagger.$$

Hence

$$\begin{aligned} P_{\mathbf{v}_i} &= \frac{1}{2}(I_{2^m} + U E_{(\mathbf{0},\mathbf{a}_i)} U^\dagger) \\ &= \frac{1}{2} U (I_{2^m} + E_{(\mathbf{0},\mathbf{a}_i)}) U^\dagger \\ &= U P_{(\mathbf{0},\mathbf{a}_i)} U^\dagger. \end{aligned}$$

Similarly we can obtain

$$P_{\mathbf{v}_i} P_{\mathbf{v}_j} = U P_{(\mathbf{0},\mathbf{a}_i)} U^\dagger U P_{(\mathbf{0},\mathbf{a}_j)} U^\dagger = U P_{(\mathbf{0},\mathbf{a}_i)} P_{(\mathbf{0},\mathbf{a}_j)} U^\dagger. \tag{21}$$

Applying now (21) to each term of $f(P_{\mathbf{v}_1}, \ldots, P_{\mathbf{v}_m})$ we obtain

$$P_{f,\mathcal{G}} = f(P_{\mathbf{v}_1}, \ldots, P_{\mathbf{v}_m}) = U f(P_{(\mathbf{0},\mathbf{a}_1)}, \ldots, P_{(\mathbf{0},\mathbf{a}_m)}) U^\dagger.$$

Noting that if A is Hermitian then $UAU^\dagger$ is also Hermitian and that

$$\begin{aligned} & P_{f,G} P_{f,G} \\ = \; & U f(P_{(\mathbf{0},\mathbf{a}_1)}, \ldots, P_{(\mathbf{0},\mathbf{a}_m)}) U^\dagger U f(P_{(\mathbf{0},\mathbf{a}_1)}, \ldots, P_{(\mathbf{0},\mathbf{a}_m)}) U^\dagger \\ = \; & U f(P_{(\mathbf{0},\mathbf{a}_1)}, \ldots, P_{(\mathbf{0},\mathbf{a}_m)}) U^\dagger \\ = \; & P_{f,G} \end{aligned}$$

we conclude that $P_{f,G}$ is an orthogonal projector for any choice of orthogonal vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Finally using the identity $\mathrm{Tr}\,(ABC) = \mathrm{Tr}\,(CBA)$, we obtain

$$\begin{aligned} & \mathrm{Tr}\,(f(P_{\mathbf{v}_1}, \ldots, P_{\mathbf{v}_m})) \\ = \; & \mathrm{Tr}\,(U f(P_{(\mathbf{0},\mathbf{a}_1)}, \ldots, P_{(\mathbf{0},\mathbf{a}_m)}) U^\dagger) \\ = \; & \mathrm{Tr}\,(f(P_{(\mathbf{0},\mathbf{a}_1)}, \ldots, P_{(\mathbf{0},\mathbf{a}_m)})) \\ = \; & \mathrm{wt}\,(\mathbf{f}). \end{aligned}$$

□

Example 3: Let $m = 3$ and $f(x_1, x_2, x_3) = x_1 x_2 \oplus x_1 \oplus x_3$. Given vectors $\mathbf{v}_1, \mathbf{v}_2$, and $\mathbf{v}_3$ defined as in Example 2, then $f(P_{\mathbf{v}_1}, P_{\mathbf{v}_2}, P_{\mathbf{v}_3}) = P_{\mathbf{v}_1} P_{\mathbf{v}_2} \oplus P_{\mathbf{v}_1} \oplus P_{\mathbf{v}_3}$ and the result is given in (17). It is not difficult to check that $f(P_{\mathbf{v}_1}, P_{\mathbf{v}_2}, P_{\mathbf{v}_3})$ is a projection operator and that the dimension of the corresponding subspace is equal to $\mathrm{Tr}\,(P_{f,\mathcal{G}}) = \mathrm{wt}\,(\mathbf{f}) = 4$.

Now we are in a position to define operator Reed-Muller codes. Let $\mathcal{C}_m$ be the set of all $[2m, m]$ self-orthogonal codes, and for each $C \in \mathcal{C}_m$ choose a basis G. A canonical choice for G is the first m (in the lexicographical order) linearly independent vectors from C. Let $\mathcal{G}_m$ be the set of all bases G, and let $\mathcal{F}_{r,m}$ be the set of all Boolean functions $f(x_1, \ldots, x_m)$ of degree at most r.

Definition 6: The r-th order operator Reed-Muller code $O\text{-}RM(r, m)$ is defined as the set of all projection operators $P_{f,G}$ acting in $\mathbb{C}^{2^m}$, where $f \in \mathcal{F}_{r,m}$ and $G \in \mathcal{G}_m$.

Remark 1: Note that the same projection operator can result from different choices of basis and Boolean function (see Theorem 2 below). Multiple copies of projectors are excluded from $O\text{-}RM(r, m)$.

Theorem 2: Let G and G' be two different bases of a $[2m, m]$ self-orthogonal code. Then

$$\{P_{f,G} : f \in \mathcal{F}_{r,m}\} = \{P_{f,G'} : f \in \mathcal{F}_{r,m}\}.$$

Proof: It follows from (9) and the definition of $P_{\mathbf{v}}$ that for any $\mathbf{v}, \mathbf{w} \in C$ we have

$$P_{\mathbf{v}} \oplus P_{\mathbf{w}} = P_{\mathbf{v} \oplus \mathbf{w}} \text{ or } P_{\mathbf{v}} \oplus P_{\mathbf{w}} = P_{\mathbf{v} \oplus \mathbf{w}} \oplus I_{2^m} \tag{22}$$

according as $E_{\mathbf{v}} E_{\mathbf{w}} = -E_{\mathbf{v} \oplus \mathbf{w}}$ or $E_{\mathbf{v}} E_{\mathbf{w}} = E_{\mathbf{v} \oplus \mathbf{w}}$.

Let $\mathcal{G} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ and $\mathcal{G}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$. Any $\mathbf{w}_j$ can be represented as a linear combination

$$\mathbf{w}_j = \mathbf{v}_{i_1} \oplus \ldots \oplus \mathbf{v}_{i_t}$$

for appropriately chosen indices $i_1, \ldots, i_t$. Hence, from (22), we have

$$P_{\mathbf{w}_j} = P_{\mathbf{v}_{i_1}} \oplus \ldots \oplus P_{\mathbf{v}_{i_t}} \oplus a_j I_{2^m}, a_j = 0, 1.$$

This substitution yields

$$\begin{aligned} & P_{g,G'} = g(P_{\mathbf{w}_1}, \ldots, P_{\mathbf{w}_m}) \\ = \; & g(P_{\mathbf{v}_{r_1}} \oplus \ldots \oplus P_{\mathbf{v}_{r_h}} \oplus a_1 I_{2^m}, \ldots, P_{\mathbf{v}_{j_1}} \oplus \ldots \oplus P_{\mathbf{v}_{j_q}} \\ & \oplus a_m I_{2^m}) \\ = \; & f(P_{\mathbf{v}_1}, \ldots, P_{\mathbf{v}_m}) \end{aligned}$$

for some Boolean function $f(x_1, \ldots, x_m)$ and appropriately chosen indices $r_1, \ldots, r_h, \ldots, j_1, \ldots, j_q$ and variables $a_1, \ldots, a_m$. Since the degree of $f(x_1, \ldots, x_m)$ can not exceed the degree of $g(x_1, \ldots, x_m)$ we conclude that

$$P_{g,G'} \in \{P_{f,G} : f \in \mathcal{F}_{r,m}\}.$$

□

B. Encoding and Decoding of $O\text{-}RM(r, m)$

The following connection between $O\text{-}RM(r, m)$ and the binary Reed-Muller codes of length 2^{2m} is very important for efficient decoding of the associated space-time codes.

Let $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{2^{2m}-1}$ be vectors from $\mathbb{F}_2^{2m}$ ordered lexicographically. Let C be an $[2m, m]$ self-orthogonal code and G be a basis of C. Form $P_{f,\mathcal{G}} \in O\text{-}RM(r, m)$. Let operators $\beta_{\mathbf{c}} E_{\mathbf{c}}, \mathbf{c} \in C$, form the Abelian group S isomorphic to C, as it was described in Section IV. Note that the factors $\beta_{\mathbf{c}}$ depend on the choice of G. Define the set of operators $D_{\mathbf{v}}, \mathbf{v} \in \mathbb{F}_2^{2m}$, by

$$D_{\mathbf{v}} = \begin{cases} E_{\mathbf{v}}, & \mathbf{v} \notin C, \\ \beta_{\mathbf{v}} E_{\mathbf{v}}, & \mathbf{v} \in C. \end{cases} \tag{23}$$

Note that the operators $D_{\mathbf{v}}$ hold the properties (6), (7), (8) and therefore they form an orthonormal basis of the vector space of linear operators acting on $\mathbb{C}^{2^m}$ with the inner product $\mathrm{Tr}\,(A^\dagger B)$. Let us form the vector $\mathbf{w}$ with entries enumerated by $\mathbf{v}_i$ and defined by

$$w_{\mathbf{v}_i} = \mathrm{Tr}\,(D_{\mathbf{v}_i} P), i = 0, \ldots, 2^{2m} - 1. \tag{24}$$

Since operators $D_{\mathbf{v}}$ form an orthonormal basis we have

$$P = \sum_{i=0}^{2^{2m}} w_{\mathbf{v}_i} D_{\mathbf{v}_i}. \tag{25}$$

Let H_{2m} be the Hadamard matrix of order 2^{2m} formed according to Definition 3.

Theorem 3: The rotation of the vector $\mathbf{w}$ by the Hadamard matrix H_{2m} is a code word of the r-th order Reed-Muller code of length 2^{2m}, i.e.

$$\mathbf{z} = \frac{1}{2^m} H_{2m}\mathbf{w} \in RM(r, 2m). \tag{26}$$

Example 4: Using the same settings as in Example 2 we obtain

$$\begin{aligned}\mathbf{w}^\dagger = (&4,0,0,0,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,-2,\\ &0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-2,0,0,\\ &0,0,0,0,0,0,0,0,0,0,0,0,-2,0,0,0,0,0,0,0,0,0)\end{aligned}$$

and

$$\begin{aligned}&\mathbf{z}^\dagger = H_6\mathbf{w}\\ = \;&(0,1,0,1,1,0,0,1,0,1,1,0,1,0,1,0,1,0,0,1,0,1,0,\\ &1,1,0,1,0,0,1,1,0,1,0,0,1,1,0,1,0,0,1,0,1,0,1,\\ &1,0,1,0,1,0,1,0,0,1,0,1,1,0,0,1,0,1) \in RM(2,6).\end{aligned}$$

Before proving Theorem 3 let us make the following note. Denote by $H_{2m}^{(0)}$ the matrix obtained from H_{2m} be replacing 1 with 0 and -1 with 1. We will enumerate rows and columns of H_{2m} and $H_{2m}^{(0)}$ by vectors $\mathbf{v}_i$ and say that is the i-th column (row) corresponds to $\mathbf{v}_i$.

Lemma 2: Let C be an $[2m, r]$ linear code. The columns $\mathbf{h}^{(0)}$ of $H_{2m}^{(0)}$ corresponding to vectors $\mathbf{v} \in C$ form an $[2^{2m}, r]$ linear code.

Proof: The entry of H_{2m} at the intersection of the row and column corresponding to $\mathbf{u}, \mathbf{v} \in \mathbb{F}_2^{2m}$ respectively is $(-1)^{\mathbf{v}\cdot\mathbf{u}}$. The corresponding entry of $H_{2m}^{(0)}$ is $\mathbf{v}\cdot\mathbf{u}$. Let $n = 2^{2m}$ and vectors $\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1} \in \mathbb{F}_2^{2m}$ be ordered lexicographically. If $\mathbf{v}, \mathbf{v}'$ are codewords of C then corresponding columns of $H_{2m}^{(0)}$ are

$$\begin{aligned}\mathbf{h}_{\mathbf{v}}^{(0)} &= (\mathbf{v}\cdot\mathbf{u}_0, \mathbf{v}\cdot\mathbf{u}_1, \ldots, \mathbf{v}\cdot\mathbf{u}_{n-1})^\dagger,\\ \mathbf{h}_{\mathbf{v}'}^{(0)} &= (\mathbf{v}'\cdot\mathbf{u}_0, \mathbf{v}'\cdot\mathbf{u}_1, \ldots, \mathbf{v}'\cdot\mathbf{u}_{n-1})^\dagger,\end{aligned}$$

and their sum

$$\mathbf{h}_{\mathbf{v}}^{(0)} \oplus \mathbf{h}_{\mathbf{v}'}^{(0)} = ((\mathbf{v}\oplus\mathbf{v}')\cdot\mathbf{u}_0, (\mathbf{v}\oplus\mathbf{v}')\cdot\mathbf{u}_1, \ldots, (\mathbf{v}\oplus\mathbf{v}')\cdot\mathbf{u}_{n-1})^\dagger$$

corresponds to the codeword $\mathbf{v}_i \oplus \mathbf{v}_j \in C$. □

Proof: (Theorem 3) First let us consider Boolean functions that are monomials, i.e. functions of the form

$$f(x_1, \ldots, x_m) = x_{i_1} x_{i_2} \ldots x_{i_r}.$$

Let $G = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m\}$ be a set of m orthogonal and linearly independent vectors from $\mathbb{F}_2^{2m}$. Then

$$\begin{aligned}P_{f,G} &= f(P_{\mathbf{c}_{i_1}}, \ldots, P_{\mathbf{c}_{i_r}}) = P_{\mathbf{c}_{i_1}} \ldots P_{\mathbf{c}_{i_r}}\\ &= \frac{1}{2}(I_{2^m} + E_{\mathbf{c}_{i_1}}) \ldots \frac{1}{2}(I_{2^m} + E_{\mathbf{c}_{i_r}})\\ &= \frac{1}{2^r}\sum_{\mathbf{c}\in C} D_{\mathbf{c}},\end{aligned}$$

where C is the self-orthogonal code generated by the vectors $\mathbf{c}_{i_1}, \ldots, \mathbf{c}_{i_r}$.

As we mentioned before the operators $D_{\mathbf{v}}$ have the property (7) and therefore

$$w_{\mathbf{v}} = \mathrm{Tr}\,(D_{\mathbf{v}} P_{f,G}) = \begin{cases} 2^{m-r}, & \mathbf{v} \in C,\\ 0, & \mathbf{v} \notin C.\end{cases}$$

Hence the entries of $\mathbf{z}$ are

$$z_{\mathbf{u}} = \frac{2^{m-r}}{2^m} \sum_{\mathbf{c}\in C} h_{\mathbf{u},\mathbf{c}}, \mathbf{u} \in \mathbb{F}_2^{2m}.$$

According to Lemma 2 the columns of $H_{2m}^{(0)}$ corresponding to nonzero elements of $\mathbf{w}$ form an $[2^{2m}, r]$ linear code, say A. Obviously A can be generated by $\mathbf{h}_{\mathbf{c}_{i_1}}^{(0)}, \ldots, \mathbf{h}_{\mathbf{c}_{i_r}}^{(0)}$. Since A is a linear code we know that either $a_{\mathbf{u}} = 0$ for all codewords $\mathbf{a} \in A$ or half codewords $\mathbf{a} \in A$ have $a_{\mathbf{u}} = 0$ and another half have $a_{\mathbf{u}} = 1$. Now, keeping in mind, that 0 and 1 of $H_{2m}^{(0)}$ correspond to 1 and -1 in H_{2m} we obtain

$$z_{\mathbf{u}} = \begin{cases} 1, & h_{\mathbf{u},\mathbf{c}_{i_1}}^{(0)} = 0, h_{\mathbf{u},\mathbf{c}_{i_2}}^{(0)} = 0, \ldots, h_{\mathbf{u},\mathbf{c}_{i_r}}^{(0)} = 0;\\ 0, & \text{else.}\end{cases}$$

It is straightforward to see that the same vector can be defined as follows

$$z_{\mathbf{u}}^\dagger = (1 \oplus h_{\mathbf{u},\mathbf{c}_{i_1}}^{(0)}) \cdot \ldots \cdot (1 \oplus h_{\mathbf{u},\mathbf{c}_{i_r}}^{(0)}), \mathbf{u} \in \mathbb{F}_2^{2m}.$$

Columns of $H_{2m}^{(0)}$ are Boolean functions of degree 1. Hence the vector $\mathbf{z}$ is a Boolean function of degree r and therefore $\mathbf{z} \in RM(2m, r)$.

Let now $\mathbf{z}'$ and $\mathbf{z}''$ be vectors obtained according to (24) and (26) from Boolean functions

$$\begin{aligned}f'(x_1, \ldots, x_m) &= x_{i_1} \ldots x_{i_r}\\ &\text{and}\\ f''(x_1, \ldots, x_m) &= x_{l_1} \ldots x_{l_t},\end{aligned}$$

respectively. Note that we allow the functions f' and f'' to have common factors. Let $\mathbf{z}$ be the vector obtained from $P = P_{f',G} P_{f'',G}$ according to (24) and (26). Using arguments similar to the ones used above, we obtain that

$$\mathbf{z} = \mathbf{z}' \bullet \mathbf{z}''. \tag{27}$$

Let finally

$$P = P_{f',G} \oplus P_{f'',G} = P_{f',G} + P_{f'',G} - 2P_{f',G}P_{f'',G}$$

and $\mathbf{z}$ be the vector obtained from P according to (24) and (26). With the help of (27) we obtain

$$\mathbf{z} = \mathbf{z}' + \mathbf{z}'' - 2\mathbf{z}' \bullet \mathbf{z}'' = \mathbf{z}' \oplus \mathbf{z}''.$$

This finishes the proof. □

Now decoding of O-$RM(r, m)$ can be accomplished using the following algorithm.

Decoding of O-$RM(r, m)$

1) A matrix X is received from the channel;
2) Compute $P_X = XX^\dagger$;
3) Compute $\mathbf{w} = (w_0, w_1, \ldots, w_{2^{2m}-1})$, where $w_i = |\mathrm{Tr}\,(E_{\mathbf{v}_i} P_X)|$;

4) Compute $\mathbf{z} = \frac{1}{2^m} H_{2^{2m}} \mathbf{w}$;
5) Decode $\mathbf{z}$ with the help of any decoder of $RM(r, 2m)$;
6) Compute $\widehat{\mathbf{w}} = H\mathbf{z}$;
7) Form a linear code C from vectors $\mathbf{v} \in \mathbb{F}_2^{2m}$ such that $\widehat{\mathbf{w}}_{\mathbf{v}} \neq 0$ and compute $k = \log_2 |C|$;
8) In C find the first independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in the lexicographical ordering;
9) With the help of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ find $\beta_{\mathbf{v}}$ such that $\beta_{\mathbf{v}} E_{\mathbf{v}}$ form the subgroup S of $\mathcal{E}_m$ isomorphic to C (see section IV);
10) Form the vector $\mathbf{y}$ with entries $y_{\mathbf{v}} = \beta_{\mathbf{v}} \widehat{\mathbf{w}}_{\mathbf{v}}, \mathbf{v} \in C$;
11) Decode $\mathbf{y}$ with of any decoder of $RM(t, k), t = \min\{r, k\}$.

The result of the decoding is a self-orthogonal code C (found in step 7) and a code word (equivalently a Boolean function) of the code $RM(t, k)$ found in step 10.

There has been intense interests in efficient decoding of binary Reed-Muller codes for decades. The best results were achieved by Dumer in [9], who has demonstrated decoding algorithms for codes of moderate lengths with performance that is very close to ML decoding and complexity proportional to $n \log_2 n$. If we use one of Dumer's algorithms the complexity of the described decoding of $O\text{-}RM(1, 2)$ codes will be approximately $T^2 M + 2M \log_2 M$.

C. Parameters of $O\text{-}RM(r, m)$

The binary $RM(r, m)$ code has minimum distance 2^{m-r}. We will prove a similar result for an $O\text{-}RM(r, m)$ code.

Theorem 4: Let $P, P' \in O\text{-}RM(r, m)$ then

$$d(P, P') = \mathrm{Tr}\,(P) + \mathrm{Tr}\,(P') - 2PP' \geq 2^{m-r}.$$

Proof: Let C and C' be $[2m, m]$ self-orthogonal codes (not necessary distinct). Define the linear code $B = C \cap C'$ and let G_B be a basis of B. According to Theorem 2, w.l.o.g., we can choose a basis G of C and basis G' of C' such that

$$G_B \subseteq G \text{ and } G_B \subseteq G'. \tag{28}$$

Let f and f' be two arbitrary Boolean functions of degree less than or equal to r. Consider projectors $P_{f,\mathcal{G}}$ and $P_{f',\mathcal{G}'}$.

Define the sets of operators $D_{\mathbf{v}}$ and $D'_{\mathbf{v}}$ according to (23) for G and G' respectively. From (28) it follows that

$$D_{\mathbf{v}} = D'_{\mathbf{v}}, \mathbf{v} \in B. \tag{29}$$

Define vectors $\mathbf{w}$ and $\mathbf{w}'$ by $w_{\mathbf{v}} = \mathrm{Tr}\,(D_{\mathbf{v}} P)$ and $w'_{\mathbf{v}} = \mathrm{Tr}\,(D'_{\mathbf{v}})$ respectively. Note that from (22) it follows that

$$w_{\mathbf{v}} = 0 \text{ if } \mathbf{v} \notin C, \text{ and similarly } w'_{\mathbf{v}} = 0 \text{ if } \mathbf{v} \notin C'. \tag{30}$$

According to Theorem 3 we

$$\mathbf{z} = \frac{1}{2^m} H_{2m} \mathbf{w}, \ \mathbf{z}' = \frac{1}{2^m} H_{2m} \mathbf{w}' \in RM(2m, r).$$

Taking into account that $\frac{1}{2^m} H_{2m}$ is unitary and that the minimum distance of $RM(2m, r)$ is 2^{2m-r} we obtain

$$\begin{aligned} d^2(\mathbf{w}, \mathbf{w}') &= \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} (w_{\mathbf{v}} - w'_{\mathbf{v}})^2 = d^2(\mathbf{z}, \mathbf{z}') \\ &= \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} (z_{\mathbf{v}} - z'_{\mathbf{v}})^2 = d_{\mathrm{Ham}}(\mathbf{z}, \mathbf{z}') \geq 2^{2m-r}, \end{aligned} \tag{31}$$

where $d_{\mathrm{Ham}}(\cdot, \cdot)$ is the Hamming distance.

Using (25) and taking into account that (7) holds for $D_{\mathbf{v}}$, we obtain

$$\begin{aligned} \mathrm{Tr}\,(P) &= \mathrm{Tr}\,(PP) = \frac{1}{2^m} \frac{1}{2^m} \sum_{\mathbf{v}, \mathbf{u} \in \mathbb{F}_2^{2m}} w_{\mathbf{v}} w_{\mathbf{u}} \mathrm{Tr}\,(D_{\mathbf{v}} D_{\mathbf{u}}) \\ &= \frac{1}{2^m} \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} w_{\mathbf{v}}^2. \end{aligned}$$

Similarly

$$\mathrm{Tr}\,(P') = \frac{1}{2^m} \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} w'^2_{\mathbf{v}}.$$

Further,

$$\begin{aligned} \mathrm{Tr}\,(PP') &= \frac{1}{2^m} \frac{1}{2^m} \sum_{\mathbf{v}, \mathbf{u} \in \mathbb{F}_2^{2m}} w_{\mathbf{v}} w'_{\mathbf{u}} \mathrm{Tr}\,(D_{\mathbf{v}} D'_{\mathbf{u}}) \\ &= \frac{1}{2^m} \frac{1}{2^m} \sum_{\mathbf{v}, \mathbf{u} \in B} w_{\mathbf{v}} w'_{\mathbf{u}} \mathrm{Tr}\,(D_{\mathbf{v}} D_{\mathbf{u}}) \\ &= \frac{1}{2^m} \sum_{\mathbf{v} \in B} w_{\mathbf{v}} w'_{\mathbf{v}} = \frac{1}{2^m} \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} w_{\mathbf{v}} w'_{\mathbf{v}}, \end{aligned}$$

where the second equality follows from (29) and (30), and the last equality follows from (8) and the definition of $D_{\mathbf{v}}$. Finally, taking into account (31), we obtain

$$\begin{aligned} & d(P, P') = \mathrm{Tr}\,(P) + \mathrm{Tr}\,(P') - 2\mathrm{Tr}\,(PP') \\ = \ & \frac{1}{2^m} \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} w_{\mathbf{v}}^2 + \frac{1}{2^m} \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} w'^2_{\mathbf{v}} - 2 \frac{1}{2^m} \sum_{\mathbf{v} \in \mathbb{F}_2^{2m}} w_{\mathbf{v}} w'_{\mathbf{v}} \\ = \ & \frac{1}{2^m} d^2(\mathbf{w}, \mathbf{w}') \geq 2^{m-r}. \end{aligned}$$

□

According to Theorem 1 the dimension of the subspace corresponding to $P_{f,G}$ equals the Hamming weight of the binary vector f. Thus choosing Boolean function of weight M we obtain a Grassmannian packings of M-dimensional subspaces in $\mathbb{C}^{2^m}$. It is well known that the binary Reed-Muller code $RM(r, m)$ consists of even vectors with weights in the interval $[2^{m-r}, 2^m - 2^{m-r}]$, one codeword of weight 0, and one codeword of weight 2^m. The codewords of weight 0 and 2^m correspond to Boolean functions $f(x_1, \ldots, x_m) = 0$ and $f(x_1, \ldots, x_m) = 1$, which are not used in Definition 6. Thus $O\text{-}RM(r, m)$ contains only s-dimension subspaces, where s is even $2^{m-r} \leq s \leq 2^m - 2^{m-r}$.

In order to define the number A_s of s-dimensional subspaces in $O\text{-}RM(r, m)$ we need the following lemmas.

Lemma 3: [24, Chapter 15.2] The number of $[n, k]$ binary linear codes is equal to the Gaussian binomial coefficient

$$\begin{bmatrix} n \\ k \end{bmatrix} = \frac{(2^n - 1)(2^{n-1} - 1) \cdot \ldots \cdot (2^{n-k+1} - 1)}{(2^k - 1)(2^{k-1} - 1) \cdot \ldots \cdot (2^2 - 1)}.$$

Lemma 4: (see [25, Section 2.2] The number of $[n, k]$ self-orthogonal codes is

$$\mathcal{N}(n, k) = \prod_{j=0}^{k-1} \frac{2^{n-2j} - 1}{2^{j+1} - 1}.$$

Proof: Let C_t be an $[n, t]$ self-orthogonal code and $C_t^{\perp}$ be its dual code. If $\mathbf{v} \in C_t^{\perp} \setminus C_t$ then $C_{t+1} = C_t \cup (\mathbf{v} \oplus C_t)$

is an $[n, t+1]$ self-orthogonal code. Since any vector from the coset $\mathbf{v} \oplus C_t$ results in one and the same code C_{t+1} and since $|C_t^{\perp} \setminus C_t| = 2^{n-t} - 2^t$ one can obtain $(2^{n-t} - 2^t)/2^t$ different codes C_{t+1} from a given C_t. In any C_{t+1} there are $2^{t+1} - 1$ different codes C_t. Thus

$$\mathcal{N}(n, t+1) = \frac{2^{n-t} - 2^t}{2^t(2^{t+1} - 1)} \mathcal{N}(n, t) = \frac{2^{n-2t} - 1}{2^{t+1} - 1} \mathcal{N}(n, t).$$

For any vector $\mathbf{v}$ we have $\mathbf{v} * \mathbf{v} = 0$. Hence $\mathcal{N}(n, 1) = 2^n - 1$.

Now the assertion follows by induction. □

It is easy to see that some Boolean functions of t variables can be mapped into Boolean function that involve only $j, j < t$, variables by a linear transformation of their arguments. For example, the function $f(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3$ can be mapped into the function $g(y_1, y_2, y_3) = y_1$ by the linear transformation

$$(y_1, y_2, y_3) = (x_1, x_2, x_3) \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We will call a Boolean function of t variables irreducible if it can not be mapped into a function involving smaller number of variables by any linear transform of its variables.

Let $n_{t,j}^w$ be the number of functions of weight w and t variables that can be mapped into a function involving only j variables, but not $j-1$ or smaller number of variables. Let $S_{w,t}$ be the set of irreducible functions of weight w of t variables and $n_{t,t}^w = |S_{w,t}|$.

If we consider a function $f(x_1, \ldots, x_j)$ of weight w as a function of t variables then $f(x_1, \ldots, x_t)$ will have the weight

$$w \cdot 2^{t-j}. \tag{32}$$

Let A and B be binary codes of length t generated by vectors $\mathbf{e}_1, \ldots, \mathbf{e}_j$ and $\mathbf{e}_{j+1}, \ldots, \mathbf{e}_t$ respectively, where $\mathbf{e}_j$ is the vector whose the j-th entry is 1 and all other entries are 0s. Let $\mathbf{b}_0 = \mathbf{0}, \mathbf{b}_1, \ldots, \mathbf{b}_{2^{t-j}-1}$ be codewords of B. It is not difficult to see that the space $\mathbb{F}_2^t$ can be represented as the union of the cosets of A, i.e.

$$\{\mathbf{v} : \mathbb{F}_2^t\} = \{A \cup A \oplus \mathbf{b}_1 \cup \ldots \cup A \oplus \mathbf{b}_{2^{t-j}-1}\}$$

and that $f()$ repeats itself on the cosets of A, that is

$$f(\mathbf{b}_j \oplus \mathbf{a}) = f(\mathbf{b}_r \oplus \mathbf{a}), \text{ for any } \mathbf{a} \in A.$$

It is also easy to see that if $f(x_1, \ldots, x_j)$ is irreducible then the code A is the smallest code with this property. The contrary is also true - if a function $f(x_1, \ldots, x_t)$ repeated itself on the cosets of an $[t, j]$ code then it can be mapped into a function of j variables.

Lemma 5: Let us consider a function $f()$ from $S_{w,j}$ as functions of t variables. By appropriate linear transformations $f(x_1, \ldots, x_t)$ can be mapped into $\begin{bmatrix} t \\ j \end{bmatrix}$ *different* functions of t variables. These linear transformations are those that map the code A into $\begin{bmatrix} t \\ j \end{bmatrix}$ different $[t, j]$ linear codes.

Proof: Let T_C be a nondegenerate linear transformation, that is a $t \times t$ binary matrix of rank t that maps codewords of A into codewords of a $[t, j]$ code C by

$$\mathbf{c} = \mathbf{a} \cdot T_C, \ \mathbf{a} \in A.$$

Let T_D be another linear transformation that maps A into a $[t, j]$ code D. Define

$$\begin{aligned} g_C(x_1, \ldots, x_t) &= f((x_1, \ldots, x_t)T_C) \text{ and} \\ g_D(x_1, \ldots, x_t) &= f((x_1, \ldots, x_t)T_D). \end{aligned}$$

We will show that $g_C(x_1, \ldots, x_t) \neq g_D(x_1, \ldots, x_t)$.

It is easy to see that

$$\{\mathbf{v} : \mathbb{F}_2^t\} = \{C \cup C \oplus \mathbf{u}_1 \cup \ldots \cup C \oplus \mathbf{u}_{2^{t-j}-1}\}$$

and

$$\{\mathbf{v} : \mathbb{F}_2^t\} = \{D \cup D \oplus \mathbf{v}_1 \cup \ldots \cup D \oplus \mathbf{v}_{2^{t-j}-1}\},$$

where $\mathbf{u}_j = \mathbf{b}_j T_C$ and $\mathbf{v}_j = \mathbf{b}_j T_D$.

It is easy to check that $g_C()$ and $g_D()$ repeat themselves on cosets of C and D respectively, that is

$$g_C(\mathbf{u}_j \oplus \mathbf{c}) = g_C(\mathbf{u}_r \oplus \mathbf{c}) \text{ and } g_D(\mathbf{u}_j \oplus \mathbf{d}) = g_D(\mathbf{u}_r \oplus \mathbf{d})$$

for any $\mathbf{c} \in C$ and $\mathbf{d} \in D$. Assume that $g_C(x_1, \ldots, x_t) = g_D(x_1, \ldots, x_t)$. Since $C \neq D$ this can be the case only if $g_C(x_1, \ldots, x_t)$ repeats itself on cosets of the linear code $R = C \cap D$ of size $2^r, r < j$. Hence, there exists a linear transformation T that maps R into the code generated by $\mathbf{e}_1, \ldots, \mathbf{e}_r$. Thus $f((x_1, \ldots, x_t) \cdot T_C \cdot T)$ is a Boolean function of r variables, which contradicts with our assumption that $f()$ is an irreducible function of j variables. □

Let $g(x_1, \ldots, x_r)$ and $f(x_1, \ldots, x_t)$ be irreducible functions of r and j variables respectively and let $r < j$.

Lemma 6: There are no linear transformations T and R such that

$$f((x_1, \ldots, x_t)T) = g((x_1, \ldots, x_t)R).$$

Proof: Let $\mathbf{x}_0, \ldots, \mathbf{x}_{2^t-1}$ are binary t-tuples ordered lexicographically. We will use notation $\mathbf{f}(\mathbf{x}T)$ for the vector $(f(\mathbf{x}_0 T), f(\mathbf{x}_1 T), \ldots, f(\mathbf{x}_{2^t-1} T))$.

Let P_T and P_R be the $2^t \times 2^t$ permutation matrices corresponding to T and R, that is

$$P_T \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}T) \text{ and } P_R \mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{x}R).$$

If $f(\mathbf{x}T) = g(\mathbf{x}R)$ then

$$\mathbf{f}(\mathbf{x}TR^{-1}) = P_R^{-1} \mathbf{f}(\mathbf{x}T) = P_R^{-1} \mathbf{g}(\mathbf{x}R) = RR^{-1} \mathbf{g}(\mathbf{x}) = \mathbf{g}(\mathbf{x}).$$

Thus we obtain a contradiction with our assumption that $f()$ is an irreducible function of j variables. □

From these Lemmas and (32) it follows that

$$n_{t,j}^{w \cdot 2^{t-j}} = n_{j,j}^w \cdot \begin{bmatrix} t \\ j \end{bmatrix}, 0 \leq j \leq t-1. \tag{33}$$

Let us denote by $A_{w,t}$ the number of codewords of weight w in the code $RM(r, t)$. $A_{w,t}$ can be viewed as the total number of Boolean function of t variables, degree not exceeding r, and weight w . Hence, we can compute $n_{t,t}^w$ as

$$n_{t,t}^w = A_{w,t} - \sum_{j=0}^{t-1} n_{t,j}^w. \tag{34}$$

Finally we obtain the following theorem.

Theorem 5: The number A_w of subspaces of dimension w in $O\text{-}RM(r,m)$, within the embedding space $\mathbb{C}^{2^m}$, is equal to

$$A_w = \sum_{t=0}^{m} \mathcal{N}(2m,t) \cdot n_{t,t}^{w/2^{m-t}},$$

where for an integer $w/2^{m-t}$ $n_{t,t}^{w/2^{m-t}}$ is defined by (34) and (33) and is zero otherwise.

Some values of A_w for codes $O\text{-}RM(r,m)$ with small m are presented in the following tables.

$m=2$

w/r	1	2
1	0	30
2	30	60
min d	2	1

$m=3$

w/r	1	2	3
1	0	0	1080
2	0	1260	1260
3	0	0	7560
4	126	7686	7686
min d	4	2	1

$m=4$

w/r	1	2	3	4
1	0	0	0	36720
2	0	0	91800	91800
3	0	0	0	1285200
4	0	21420	3877020	3877020
5	0	0	0	10024560
6	0	1028160	17093160	17093160
7	0	0	0	26254800
8	510	643110	28183110	28183110
min d	8	4	2	1

By taking from $O\text{-}RM(r,m)$ subspaces of certain dimensions we obtain different Grassmannian packings. For example, by taking from $O\text{-}RM(2,4)$ all subspaces of dimension 6 we obtain a Grassmannian packing $(16,6,1028160,4)$ and by taking subspaces whose dimension does not exceed 8 we obtain a Grassmannian packing $(16,\leq 8,1692690,4)$.

VI. The first order operator Reed-Muller codes

A. Parameters

In the case of $O\text{-}RM(1,m)$ codes the number A_w of subspaces of dimension w can be found explicitly.

Theorem 6: The first order operator Reed-Muller codes $O\text{-}RM(1,m)$ consists of $2^{2m+1}-2$ subspaces of dimension 2^{m-1} of the space $\mathbb{C}^{2^m}$, one subspace of dimension zero and the space $\mathbb{C}^{2^m}$ itself.

Proof: Projection operators of $O\text{-}RM(1,m)$ are defined by Boolean functions of degree one. Let $\mathbf{v}_1,\ldots,\mathbf{v}_t$ be some orthogonal vectors. According to (22) we have

$$P_{\mathbf{v}} = P_{\mathbf{v}_1} \oplus \ldots \oplus P_{\mathbf{v}_t} \oplus aI_{2^m},\ a=0,1,$$

where $\mathbf{v} = \mathbf{v}_1 \oplus \ldots \oplus \mathbf{v}_t$. Since $P_{\mathbf{v}} \in O\text{-}RM(1,m)$ for any nonzero $\mathbf{v} \in \mathbb{F}_2^{2m}$ we conclude that the total number of projection operators $P \in O\text{-}RM(1,m)$ such that $\mathrm{Tr}\,(P) = 2^{m-1}$ is equal to the doubled number of nonzero vectors of $\mathbb{F}_2^{2m}$.

The functions $f(x_1,\ldots,x_m)=0$ and $f(x_1,\ldots,x_m)=1$ give us projectors $P=0$ and $P=I_{2^m}$ respectively. □

We also can find the distance distribution $O\text{-}RM(1,m)$. Let $B_t(Q') = |\{Q \in O\text{-}RM(1,m) : \mathrm{dist}\,(Q,Q')^2 = t\}|$ for $Q' \in O\text{-}RM(1,m)$.

Theorem 7: For any $Q' \in O\text{-}RM(1,m)$ we have

$$B_0(Q')=1, B_{2^m}(Q')=1, B_{2^{m-1}}(Q') = 2^{2m+1}-2.$$

Proof: Let $P' = 1/2(I_{2^m} + E_{\mathbf{v}_1})$ be the projection operator on Q'. Let $P = 1/2(I_{2^m} + E_{\mathbf{v}_2})$ for some $\mathbf{v}_2 \neq \mathbf{v}_1$. Then, according to (9) and (7), we have

$$\begin{aligned} & d(P',P) \\ = \ & \mathrm{Tr}\,(P') - \mathrm{Tr}\,(P) \\ & -2\mathrm{Tr}\,(\frac{1}{4}(I_{2^m} + E_{\mathbf{v}_1} + E_{\mathbf{v}_2} + \alpha E_{\mathbf{v}_1 \oplus E_{\mathbf{v}_2}})) \\ = \ & 2^{m-1}. \end{aligned}$$

The distance between Q' and the subspace with $P = 1/2(I_{2^m} - E_{\mathbf{v}_1})$ is 2^m. Finally $\mathrm{dist}\,(Q',Q')^2 = 0$. □

In the case of $O\text{-}RM(1,m)$ codes we can completely characterize all possible distributions of the principal angles between the code subspaces. Let

$$\begin{aligned} P_{\mathbf{v}_1} &= \frac{1}{2}(I_{2^m} + E_{\mathbf{v}_1}) \oplus a_1 I_{2^m} \text{ and} \\ P_{\mathbf{v}_2} &= \frac{1}{2}(I_{2^m} + E_{\mathbf{v}_2}) \oplus a_2 I_{2^m},\ a_1,a_2 = 0,1. \end{aligned}$$

Let $\theta_1,\ldots,\theta_m$ be the principal angles between subspaces defined by P_1 and P_2 respectively. Denote by U_1 and U_2 arbitrary chosen orthonormal bases of these subspaces and let d_i be the singular values of the matrix $U_1^\dagger U_2$ in the increasing order. Remind that $d_i = \cos\theta_i$.

Theorem 8: The principal angles θ_i are defined by the following possible values of d_i

- if $\mathbf{v}_1 = \mathbf{v}_2$ and $P_{\mathbf{v}_1} \neq P_{\mathbf{v}_2}$ then $d_i = 0, i = 1,\ldots,2^m$;
- if $\mathbf{v}_1 \neq \mathbf{v}_2$ and $\mathbf{v}_1 * \mathbf{v}_2 = 0$ then half of d_i are equal to 1 and another half are equal to 0;
- if $\mathbf{v}_1 \neq \mathbf{v}_2$ and $\mathbf{v}_1 * \mathbf{v}_2 = 1$ then $d_i = \frac{1}{\sqrt{2}}, i = 1,\ldots,2^m$;

Proof: The first claim of the theorem corresponds to the case of subspaces with the orthogonal projections

$$P_{\mathbf{v}_1} = \frac{1}{2}(I_{2^m} + E_{\mathbf{v}_1}) \text{ and } P_{\mathbf{v}_2} = \frac{1}{2}(I_{2^m} + E_{\mathbf{v}_1}) \oplus I_{2^m}.$$

Hence the subspaces are orthogonal that is $\mathbf{u}_1^\dagger \mathbf{u}_2 = 0$ for any $\mathbf{u}_1 \in U_1$ and $\mathbf{u}_2 \in U_2$ and the first claim of the theorem follows.

To prove the two other claims of the theorem we first note that since the group $\mathcal{L}$, defined in Section IV acts transitively on totally singular subspaces of $\overline{\mathcal{E}}_m$ it is enough to consider any fixed subspace from $O\text{-}RM(1,m)$. Hence, without loss of generality we take

$$\mathbf{v}_1 = (\underbrace{0,\ldots,0}_{m},\underbrace{1,0,\ldots,0}_{m})$$

and an arbitrary nonzero vector

$$\mathbf{v}_2 = (a_1, \ldots, a_m, b_1, \ldots, b_m) \in \mathbb{F}_2^{2m}.$$

The projection operator $P_{\mathbf{v}_1}$ is diagonal with

$$\text{diag}\,(P_{\mathbf{v}_1}) = (\underbrace{1, \ldots, 1}_{2^{m-1}}, \underbrace{0, \ldots, 0}_{2^{m-1}}). \quad (35)$$

Let vectors $\mathbf{e}_i \in \mathbb{C}^{2^m}$ be vectors whose i-th entry is 1 and all other entries are 0s. From (35) it follows that U_1 can be formed as

$$U_1 = \left[\begin{array}{c} \mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_{2^{m-1}-1} \end{array} \right].$$

Let $\mathbf{v}_1 * \mathbf{v}_2 = 0$. In this case operators $E_{\mathbf{v}_1}$ and $E_{\mathbf{v}_2}$ commute. Therefore we can simultaneously diagonalized $E_{\mathbf{v}_1}$ and $E_{\mathbf{v}_2}$. Indeed, let

$$U = I_2 \otimes u_2 \otimes \ldots \otimes u_m,$$

where

$$u_i = \left\{ \begin{array}{ll} V, & a_i = 1, b_i = 0; \\ W, & a_i = 1, b_i = 1; \\ I_2, & a_i = 0, b_i = 1, \text{ or } a_i = 0, b_i = 0. \end{array} \right.$$

Then from (13) it follows that

$$\begin{aligned} U E_{\mathbf{v}_1} U^\dagger &= E_{\mathbf{v}_1} && (36) \\ U E_{\mathbf{v}_2} U^\dagger &= I_2 \otimes t_2 \otimes \ldots \otimes t_m, \ t_i = I_2 \text{ or } \sigma_z. && (37) \end{aligned}$$

Let us rotate subspaces defined by $P_{\mathbf{v}_1}$ and $P_{\mathbf{v}_2}$ by U. This is equivalent to the change of bases: $U_1' = UU_1$ and $U_2' = UU_2$. The rotation does not change the set of the principal angles. From (36) it follows that

$$U P_{\mathbf{v}_1} U^\dagger = \frac{1}{2} U (I_{2^m} + E_{\mathbf{v}_1}) U^\dagger = P_{\mathbf{v}_1}.$$

From (37) it follows that

$$P'_{\mathbf{v}_2} = U_2' U_2'^\dagger = U U_2 U_2^\dagger U^\dagger = U P_{\mathbf{v}_2} U^\dagger$$

and using (37) we conclude that $P'_{\mathbf{v}_2}$ is diagonal. Let

$$\text{diag}\,(P'_{\mathbf{v}_2}) = (x_1, \ldots, x_{2^{m-1}}, y_1, \ldots, y_{2^{m-1}}), \ x_j, y_j = 0, 1.$$

Since, according to Theorem 7, $d(P_{\mathbf{v}_1}, P'_{\mathbf{v}_2}) = 2^{m-1}$ and $\text{Tr}\,(P'_{\mathbf{v}_2}) = 2^{m-1}$ we conclude that half of x_i are equal to 1 and another half are equal to 0. The same is valid for y_i. Let $l_1, \ldots, l_{2^{m-2}}$ be such that $x_{l_j} = 1$ and let $r_1, \ldots, r_{2^{m-2}}$ be such that $y_{r_j} = 1$. Then vectors $\mathbf{e}_{l_1}, \ldots, \mathbf{e}_{l_{2^{m-2}}}$ and $\mathbf{e}_{r_1}, \ldots, \mathbf{e}_{r_{2^{m-2}}}$ form a basis of the subspace defined by $P'_{\mathbf{v}_2}$. Next, we note that vectors $\mathbf{e}_{l_1}, \ldots, \mathbf{e}_{l_{2^{m-2}}}$ belong to the subspace defined by $P_{\mathbf{v}_1}$ and vectors $\mathbf{e}_{r_1}, \ldots, \mathbf{e}_{r_{2^{m-2}}}$ are orthogonal to this subspace. Hence, according the definition of the principle angles, we conclude that the half of the principle angles are equal to $\pi/2$ and half of them are 0s and the second claim of the theorem follows.

Let now $\mathbf{v}_1 * \mathbf{v}_2 = 1$. Define

$$U = u_1 \otimes u_2 \otimes \ldots \otimes u_m,$$

where $u_1 = Q$ or $u_1 = I_2$ depending on whether $\mathbf{v}_2$ has $a_1 = 1, b_1 = 1$ or $a_1 = 1, b_1 = 0$ respectively.

$$u_i = \left\{ \begin{array}{ll} Q, & a_i = 1, b_i = 1; \\ H, & a_i = 0, b_i = 1; \\ I_2, & else. \end{array} \right.$$

It is easy to see that the subspace defined by $P_{\mathbf{v}_1}$ is invariant to the rotation by this U and

$$\begin{aligned} P' &= \frac{1}{2} U P_{\mathbf{v}_2} U^\dagger = \frac{1}{2} \left(I_{2^m} + U E_{\mathbf{v}_2} U^\dagger \right) \\ &= \frac{1}{2} (I_{2^m} + \sigma_x \otimes B') \\ &= \frac{1}{2} \left(I_{2^m} + \left[\begin{array}{cc} \mathbf{0} & B' \\ B' & \mathbf{0} \end{array} \right] \right), \end{aligned}$$

where B' is a symmetric permutation matrix. Let $\mathbf{b}'$ be a row of B'. If the j-th entry of $\mathbf{b}'$ is nonzero, i.e. $b'_j = 1$, then the vector

$$\frac{1}{\sqrt{2}} (\mathbf{b}', \mathbf{b}_j),$$

where $\mathbf{b}_j$ is the j-th row of B', is an eigen vector of P'. The norm of the vector is 1. Let us form the matrix U' from these eigenvectors. Then

$$U' U_1^\dagger = \frac{1}{\sqrt{2}} A,$$

were A is a permutation matrix and therefore all singular values of $U'U_1^\dagger$ are $\frac{1}{\sqrt{2}}$. □

Note that for any given $\mathbf{v}_1 \in \mathbb{F}_2^{2m}$ the number of $\mathbf{v}_2 \in \mathbb{F}_2^{2m}$ such that $\mathbf{v}_1 * \mathbf{v}_2 = 1$ equals 2^{2m-1}. Indeed w.l.o.g. we can choose $\mathbf{v}_1 = (1, 0, \ldots, 0)$ for which the claim is obvious.

B. Encoding of O-$RM(1, m)$

Let U_i be orthogonal bases of subspaces from O-$RM(1, m)$, that is U_i are codewords of the corresponding space-time code. The following theorem, shows that we can use QPSK modulation for transmitting entries of codewords of O-$RM(r, m)$ code. See also example 5.

Theorem 9: $U_i \in O$-$RM(1, m)$ can be chosen in such an way that all their entries will be proportional to elements of the set $\{0, 1, -1, i, -i\}$.

Proof: The following vectors

$$\begin{aligned} \mathbf{v}_0^{(x)} &= (1, 1)^\top, & \mathbf{v}_1^{(x)} &= (1, -1)^\top, \\ \mathbf{v}_0^{(z)} &= (1, 0)^\top, & \mathbf{v}_1^{(z)} &= (0, 1)^\top, \\ \mathbf{v}_0^{(y)} &= (1, -i)^\top, & \mathbf{v}_1^{(y)} &= (1, i)^\top, \end{aligned}$$

are eigen vectors of σ_2, σ_z, and σ_y respectively. The subindices 0 and 1 indicate that the corresponding eigenvalue is 1 or -1 respectively. Let us define also vectors

$$\mathbf{v}_0^{(I)} = (1, 0)^\top, \qquad \mathbf{v}_1^{(I)} = (0, 1)^\top.$$

Let $\mathbf{v} \in \mathbb{F}_2^{2m}$ and

$$E_{\mathbf{v}} = e_1 \otimes e_2 \otimes \ldots \otimes e_m,$$

where e_i are chosen according to (5). Let

$$\mathbf{u} = \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \ldots \otimes \mathbf{u}_m,$$

where $\mathbf{u}_i$ is equal to $\mathbf{v}_{j_i}^{(I)}, \mathbf{v}_{j_i}^{(x)} \mathbf{v}_{j_i}^{(z)}$, or $\mathbf{v}_{j_i}^{(y)}$ according to whether e_i is I_2, σ_x, σ_z, or σ_y respectively, and $j_i = 0$ or 1.

Using the identity

$$(A \otimes B)(C \otimes D) = AC \otimes BD,$$

it is easy to see that $\mathbf{u}$ is an eigenvector of $E_{\mathbf{v}}$ with the eigenvalue 1 if

$$j_1 \oplus j_2 \oplus \ldots \oplus j_m = 0$$

and the eigenvalue -1 otherwise. From this it follows that there are 2^{m-1} ways to choose values j_i so that the vector $\mathbf{u}$ is an eigenvector with the eigenvalue 1 and it is easy to see that all such vectors are linearly independent and their entries belong to the set $\{0, 1, -1, i, -i\}$. In the case of the eigenvector -1 one can use the same arguments. □

C. ML Decoding of O-$RM(1, m)$

In the case of the first order Operator Reed-Muller codes there exists a simple maximum likelihood decoding algorithm.

We assume that all the subspaces of dimension 2^{m-1} of O-$RM(1, m)$ are used in the MIMO transmission described in Section II. These subspaces have the following projection operators

$$P_{\mathbf{v},\lambda} = \frac{1}{2}(I_{2m} + \lambda E_{\mathbf{v}}), \mathbf{v} \in \mathbb{F}_2^{2^m} \setminus \{\mathbf{0}\}, \lambda = \pm 1.$$

Let us consider the case of $\mathbf{v} = (\mathbf{0}, \mathbf{b})$. In this case the matrices $E_{(\mathbf{0},\mathbf{b})}$ are diagonal, their diagonal elements are 1 and -1 and they form codewords of the first order Reed-Muller code of the length 2^m (except for the all-one and all-minus one code words, since we can not choose $\mathbf{b} = \mathbf{0}$). For example for $m = 2$

$$E_{(0,0,0,1)} = I_2 \otimes \sigma_z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Let $\mathbf{X}$ be the matrix received from the MIMO channel. In order to find a maximum in the expression

$$\begin{aligned} & \arg \max_{\lambda=\pm 1, \mathbf{b}\in\mathbb{F}_2^m, \mathbf{b}\neq\mathbf{0}} \mathrm{Tr}\left(P_{(\mathbf{0},\mathbf{b}),\lambda} XX^\dagger\right) \\ = & \arg \max_{\lambda=\pm 1, \mathbf{b}\in\mathbb{F}_2^m, \mathbf{b}\neq\mathbf{0}} \frac{1}{2}\left(\mathrm{Tr}\left(I_{2^m} XX^\dagger\right) + \mathrm{Tr}\left(\lambda E_{(\mathbf{0},\mathbf{b})} XX^\dagger\right)\right) \\ = & \arg \max_{\lambda=\pm 1, \mathbf{b}\in\mathbb{F}_2^m, \mathbf{b}\neq\mathbf{0}} \mathrm{Tr}\left(\lambda E_{(\mathbf{0},\mathbf{b})} XX^\dagger\right) \end{aligned}$$

we can form the vector $\mathbf{x}$ of the diagonal elements of $XX^\dagger$ and to decode it with the help of an ML decoder of the binary $RM(1, m)$ code. (In this decoding we have to assume that the all-one and the all-minus one codewords have zero probability to be transmitted.)

The matrices $E_{(\mathbf{a},\mathbf{0})}$ are permutation matrices, they are unitary and Hermitian. Hence

$$\begin{aligned} \mathrm{Tr}\left(\lambda E_{(\mathbf{a},\mathbf{b})} XX^\dagger\right) &= \mathrm{Tr}\left(E_{(\mathbf{a},\mathbf{0})} \lambda E_{(\mathbf{a},\mathbf{b})} XX^\dagger E_{(\mathbf{a},\mathbf{0})}\right) \\ &= \mathrm{Tr}\left(\alpha E_{(\mathbf{0},\mathbf{b})} XX^\dagger E_{(\mathbf{a},\mathbf{0})}\right), \alpha = \{1, -1\}. \end{aligned}$$

$XX^\dagger E_{(\mathbf{a},\mathbf{0})}$ is a permutation of $XX^\dagger$. Thus in order to find the maximum in the expression

$$\arg \max_{\lambda=\pm 1, b\in\mathbb{F}_2^m} \mathrm{Tr}\left(\lambda E_{(\mathbf{a},\mathbf{b})} XX^\dagger\right)$$

for any given $\mathbf{a}$, it is enough to form the vector $\mathbf{x}$ of diagonal elements of the matrix $XX^\dagger E_{(\mathbf{a},\mathbf{0})}$ and to decode $\mathbf{x}$ with the help of an ML decoder of the binary $RM(1, m)$ code. (Here we assume that the all-one code word and the all-minus one code word have the same probability to be transmitted as all other code words.)

In order to estimate the complexity of this decoding we note that we have to make 2^m ML decodings of the $RM(1, m)$ code of length 2^m and choose the most reliable λ and $(\mathbf{a}, \mathbf{b})$ among all decodings. The complexity of the ML decodings based on the Fast Hadamard Transform (FHT) is $2^m \times m2^m = m2^{2m}$. The complexity of computation $XX^\dagger$ is $O(2^{3m})$. Hence the overall complexity is $O(2^{3m})$. For comparison the complexity of the brute force decoding according to (3) is $O(2^{5m})$.

The formal description of the algorithm is following.

Maximum Likelihood Decoding of O-$RM(1, m)$

1) Form $\mathbf{c} = \mathrm{diag}(XX^\dagger)$;
2) Use FHT to compute $\mathbf{v} = (v_0, \ldots, v_{2^m-1})^\top = H_{2^m}\mathbf{c}$;
3) Find
$$\begin{aligned} t_{\mathbf{0}} &= \max\{|v_1|, \ldots, |v_{2^m-1}|\}, \\ j_{\mathbf{0}} &= \arg\max\{|v_1|, \ldots, |v_{2^m-1}|\}, \end{aligned}$$
4) For all $\mathbf{a} \in \mathbb{F}_2^m, \mathbf{a} \neq \mathbf{0}$, do steps 5,6,7
5) Form $\mathbf{c} = \mathrm{diag}(XX^\dagger E_{(\mathbf{a},\mathbf{0})})$;
6) Use FHT to compute $\mathbf{v} = (v_0, \ldots, v_{2^m-1})^\top = H_{2^m}\mathbf{c}$;
7) Find
$$\begin{aligned} t_{\mathbf{a}} &= \max\{|v_0|, \ldots, |v_{2^m-1}|\}, \\ j_{\mathbf{a}} &= \arg\max\{|v_0|, \ldots, |v_{2^m-1}|\}; \end{aligned}$$
8) Find $\widehat{\mathbf{a}} = \arg\max_{\mathbf{a}\in\mathbb{F}_2^m}\{t_{\mathbf{a}}\}$;
9) Assign the binary representation of the integer $j_{\mathbf{a}}$ to $\widehat{\mathbf{b}}$;
10) Assign
$$\widehat{\lambda} = \begin{cases} 1, & \mathrm{Tr}\left(XX^\dagger E_{(\widehat{\mathbf{a}},\widehat{\mathbf{b}})}\right) > 0 \\ -1, & \mathrm{Tr}\left(XX^\dagger E_{(\widehat{\mathbf{a}},\widehat{\mathbf{b}})}\right) < 0 \end{cases}$$
11) The result of decoding is $\widehat{\lambda}, (\widehat{\mathbf{a}}, \widehat{\mathbf{b}})$.

Example 5: Let us consider codes O-$RM(1, 2)$. According to Theorem 6 this code consists of 30 subspaces of dimension 2 of the space $\mathbb{C}^4$. Typical projection operators, corresponding to vectors $\mathbf{v}_1 = (0, 0|1, 0)$ and $\mathbf{v}_2 = (0, 1|1, 1)$ and $\lambda = 1$, are

$$P_{\mathbf{v}_1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, P_{\mathbf{v}_2} = \frac{1}{2}\begin{bmatrix} 1 & i & 0 & 0 \\ -i & 1 & 0 & 0 \\ 0 & 0 & 1 & -i \\ 0 & 0 & i & 1 \end{bmatrix}.$$

Let us take the operator $P_{\mathbf{v}_2}$. The following two vectors can be chosen to form an orthonormal basis of the corresponding subspace

$$\mathbf{u}_1^\top = \frac{1}{\sqrt{2}}(1, -i, 0, 0), \ \mathbf{u}_2^\top = \frac{1}{\sqrt{2}}(0, 0, 1, i).$$

In the idealized case of noiseless MIMO channel with the channel matrix $H = I_2$ and no additive noise the intermediate results in the decoding algorithm described above will be

- Step 2: $\mathbf{v} = (4, 0, 0, 0), \ t_0 = 0$ (since we do not use v_0)
- Steps 6 and 7:

- $\mathbf{a}=(0,1)$: $\mathbf{v}=(0,0,0,2i),\ t_{\mathbf{a}}=2, j_{\mathbf{a}}=3$
- $\mathbf{a}=(1,0)$: $\mathbf{v}=(0,0,0,0),\ t_{\mathbf{a}}=0$
- $\mathbf{a}=(1,1)$: $\mathbf{v}=(0,0,0,0),\ t_{\mathbf{a}}=0$

Hence $\widehat{\mathbf{a}}=(0,1), \widehat{\mathbf{b}}=(1,1)$. Since $\mathrm{Tr}\,(XX^{\dagger}E_{(\widehat{\mathbf{a}},\widehat{\mathbf{b}})})=4$ we have $\widehat{\lambda}=1$.

VII. Bent Functions

Bent functions can be defined as Boolean functions that are maximally far away from any linear Boolean function. Below we remind the definition and main properties of bent functions, which will be useful for us in the following sections.

Definition 7: A Boolean function $g(x_1,\ldots,x_{2t})$ of $2t$ variables is called bent if its Hadamard transform has the following property

$$\widehat{G}(\mathbf{u})=\sum_{\mathbf{x}\in\mathbb{F}_2^{2t}}(-1)^{\mathbf{u}\cdot\mathbf{x}\oplus g(\mathbf{x})}=H_{2t}\widehat{\mathbf{g}}=\pm 2^t,\quad \mathbf{u}\in\mathbb{F}_2^{2t}. \tag{38}$$

Lemma 7: [24, Chapter 14.5]

$$\mathrm{wt}\,(\mathbf{g})=2^{2t-1}\pm 2^{t-1}.$$

Lemma 8: If $g(\mathbf{x})$ is a bent function then for any $\mathbf{v}\in\mathbb{F}_2^{2t}$ the function $g'(\mathbf{x})=g(\mathbf{x}\oplus\mathbf{v})$ is also bent.

Proof:

$$\begin{aligned}\widehat{G'}(\mathbf{u}) &= \sum_{\mathbf{x}\in\mathbb{F}_2^{2t}}(-1)^{\mathbf{u}\cdot\mathbf{x}\oplus g(\mathbf{x}\oplus\mathbf{v})}\\ &= \sum_{\mathbf{y}\in\mathbb{F}_2^{2t}}(-1)^{\mathbf{u}\cdot(\mathbf{y}\oplus\mathbf{v})\oplus g(\mathbf{y})}\\ &= (-1)^{\mathbf{u}\cdot\mathbf{v}}\widehat{G}(\mathbf{u})=\pm 2^t.\end{aligned}$$

□

Let $g_1(y_1,\ldots,y_r)$ and $g_2(z_1,\ldots,z_s)$ be Boolean functions of r and s arguments respectively. We say that the function $g(y_1,\ldots,y_r,z_1,\ldots,z_s)=g_1(y_1,\ldots,y_r)\oplus g_2(z_1,\ldots,z_s)$ of $r+s$ variables is the direct sum of $g_1(y_1,\ldots,y_r)$ and $g_2(z_1,\ldots,z_s)$. It is easy to see that

$$\widehat{g}(y_1,\ldots,y_r,z_1,\ldots,z_s)=\widehat{g_1}(y_1,\ldots,y_r)\widehat{g_2}(z_1,\ldots,z_s).$$

Lemma 9: We have a similar correspondence for the Hadamard transforms of these functions

$$\begin{aligned}\widehat{G}(\mathbf{u}) &= \widehat{G_1}(\mathbf{v})\widehat{G_2}(\mathbf{w})\\ &\quad \mathbf{u}=(\mathbf{v},\mathbf{w})\in\mathbb{F}_2^{r+s},\ \mathbf{v}\in\mathbb{F}_2^r,\ \mathbf{w}\in\mathbb{F}_2^s.\end{aligned}$$

Proof:

$$\begin{aligned}\widehat{G}(\mathbf{u}) &= \sum_{\mathbf{x}\in\mathbb{F}_2^{r+s}}(-1)^{\mathbf{u}\cdot\mathbf{x}\oplus g(\mathbf{x})}\\ &= \sum_{\mathbf{y}\in\mathbb{F}_2^r}\sum_{\mathbf{z}\in\mathbb{F}_2^s}(-1)^{\mathbf{v}\cdot\mathbf{y}\oplus\mathbf{w}\cdot\mathbf{z}\oplus g_1(\mathbf{y})\oplus g_2(\mathbf{z})}\\ &= \sum_{\mathbf{y}\in\mathbb{F}_2^r}(-1)^{\mathbf{v}\cdot\mathbf{y}\oplus g_1(\mathbf{y})}\sum_{\mathbf{z}\in\mathbb{F}_2^s}(-1)^{\mathbf{w}\cdot\mathbf{z}\oplus g_2(\mathbf{z})}\\ &= \widehat{G_1}(\mathbf{v})\widehat{G_2}(\mathbf{w})\end{aligned}$$

□

Lemma 10: Building the Hadamard transform two times yields almost the same function

$$\widehat{\widehat{G}}=2^{2t}\widehat{g}.$$

Proof: The claim follows from the fact that $H_mH_m=2^mI_{2^m}$. □

From this lemma it follows that $\widehat{g}^*=1/2^t\widehat{G}$ is again a bent function. We will say that g^* is the dual function of g.

We distinguish a canonical bent function

$$g(x_1,\ldots,x_{2t})=x_1x_2\oplus x_3x_4\oplus\cdots\oplus x_{2t-1}x_{2t}. \tag{39}$$

Lemma 11: [8],[6] Canonical bent function Bent function is dual to itself, i.e. $g^*=g$, and therefore

$$\widehat{G}(\mathbf{x})<0 \quad \text{iff} \quad g(\mathbf{x})=1,\ \mathbf{x}\in\mathbb{F}_2^{2t}.$$

Proof: We prove the claim by induction. Let $g(x_1,x_2)=x_1x_2$. It is easy to check that $\widehat{\mathbf{g}}=2\widehat{\mathbf{G}}$. Using Lemma 9 for $g(x_1,\ldots,x_{2t})=x_1x_2\oplus x_3x_4\oplus\cdots\oplus x_{2t-1}x_{2t}$ we obtain by induction that

$$\widehat{\mathbf{g}}=2^t\widehat{\mathbf{G}}. \tag{40}$$

□

Let $s(\mathbf{x})=\mathbf{x}A+\mathbf{p}$ be an affine transform. Here A is an invertible binary matrix and $\mathbf{p}$ is a binary vector. The following lemma is very important for us. It establishes a connection between the Hadamard transforms of a function and an affine transform of the function.

Lemma 12: Let f be a Boolean function and $g=f\circ s=f(s(x))$ its affine transform. Then

$$\widehat{G}(\mathbf{u})=(-1)^{(\mathbf{u}(A^{-1})^T)\cdot\mathbf{p}}\widehat{F}(\mathbf{u}\left(A^{-1}\right)^T)$$

Proof: Taking into account that $s^{-1}(\mathbf{y})=\mathbf{y}A^{-1}\oplus\mathbf{p}A^{-1}$ we obtain

$$\begin{aligned}\widehat{G}(\mathbf{u}) &= \sum_{\mathbf{x}\in\mathbb{F}_2^{2m}}(-1)^{\mathbf{u}\cdot\mathbf{x}\oplus f(s(\mathbf{x}))}\\ &= \sum_{\mathbf{y}\in\mathbb{F}_2^{2m}}(-1)^{\mathbf{u}\cdot s^{-1}(\mathbf{y})\oplus f(\mathbf{y})}\\ &= \sum_{\mathbf{y}\in\mathbb{F}_2^{2m}}(-1)^{\mathbf{u}\cdot(\mathbf{y}A^{-1}\oplus\mathbf{p}A^{-1})\oplus f(\mathbf{y})}\\ &= (-1)^{\mathbf{u}\cdot(\mathbf{p}A^{-1})}\sum_{\mathbf{y}\in\mathbb{F}_2^{2m}}(-1)^{(\mathbf{u}(A^{-1})^T)\cdot\mathbf{y}\oplus f(\mathbf{y})}\\ &= (-1)^{(\mathbf{u}(A^{-1})^T)\cdot\mathbf{p}}\widehat{F}(\mathbf{u}(A^{-1})^T).\end{aligned}$$

□

Sometimes it is convenient to formulate the result of the previous lemma in a slightly different form:

$$(-1)^{\mathbf{v}\cdot\mathbf{p}}\widehat{G}(\mathbf{v}A^T)=\widehat{F}(\mathbf{v}) \tag{41}$$

Note also that if A is the identity matrix, that is if $g(\mathbf{x})=f(\mathbf{x}\oplus\mathbf{p})$ we have

$$\widehat{G}(\mathbf{u})=(-1)^{\mathbf{u}\cdot\mathbf{p}}\widehat{F}(\mathbf{u}). \tag{42}$$

From

$$|\widehat{G}(\mathbf{u})|=|\widehat{F}(\mathbf{u})|$$

it follows that g is bent iff f is bent.

Lemma 13: [24, Chapter 14.5] The directional derivative of g in the direction $\mathbf{b}\neq\mathbf{0}$ is zero, i.e.

$$\sum_{\mathbf{x}\in\mathbb{F}_2^{2t}}g(\mathbf{x})\oplus g(\mathbf{x}\oplus\mathbf{b})=0.$$

VIII. Hadamard Transform of Binary Boolean Functions

A typical codeword of $RM(2,2m)$ is defined by a Boolean function of the form [24, Chapter 15.2]

$$f(x_1,\dots,x_{2m}) = \mathbf{x}Q\mathbf{x}^T + \mathbf{r}\mathbf{x}^T + c,$$

where Q is a $2m\times 2m$ upper triangular binary matrix, $\mathbf{r}$ is a binary vector of dimension $2m$, and c is 0 or 1. The quadratic part of f can be characterized by the symmetric matrix $B = Q+Q^T$. The following lemma is a modification of Dickson's theorem (see [24, Chapter 15.2] for details).

Lemma 14: Let $f : \mathbb{F}_2^{2m} \to \mathbb{F}_2$ be a quadratic function with the symmetric matrix B of rank $2t$. There exists an affine transformation s, so that $f\circ s$ has one of the following algebraic forms:

$$x_1x_2 \oplus x_3x_4 \oplus \dots \oplus x_{2t-1}x_{2t}, \qquad 1\le t\le m \tag{43}$$

$$x_1x_2 \oplus \dots \oplus x_{2t-1}x_{2t} \oplus 1, \qquad 1\le t\le m \tag{44}$$

$$x_1x_2 \oplus \dots \oplus x_{2t-1}x_{2t} \oplus x_{2t+1}, \qquad 1\le t\le m-1. \tag{45}$$

Functions of these forms have the Hamming weights $2^{2m-1} - 2^{2m-t-1}, 2^{2m-1}+2^{2m-t-1}$, and 2^{2m-1} respectively.

Next we consider how the Hadamard transform of a quadratic function $f(x_1,\dots,x_{2m})$ depends on the rank of the matrix B. This is closely related to the more common practice of taking the Hadamard transform of the exponentiated vector $\widehat{\mathbf{f}}$.

Let $f(x_1,\dots,x_{2m})$ be a Boolean function that can be transformed to the form (43) or (44). In other words f, after the affine transformation, does not have a linear part. Recall that $\mathrm{wt}(\mathbf{f}) = 2^{2m-1}\pm 2^{2m-t-1}$. Set $\mathbf{F} = H_{2m}\mathbf{f}$.

Theorem 10: [1] The entries $F_{\mathbf{b}}, \mathbf{b}\in\mathbb{F}_2^{2m}$, have the following properties.

1) The nonzero entries of $\mathbf{F}$ have the values
$$F_{\mathbf{b}} = \begin{cases} 2^{2m-1} + \widehat{a}_{\mathbf{b}}2^{2m-t-1}, & \mathbf{b}=\mathbf{0},\\ \widehat{a}_{\mathbf{b}}2^{2m-t-1}, & \mathbf{b}\neq\mathbf{0},\end{cases}$$
where $\widehat{a}_{\mathbf{b}} = \pm 1$.
2) The vectors $\mathbf{b}$ such that $F_{\mathbf{b}}\neq\mathbf{0}$ form a linear subspace C of dimension $2t$ within $\mathbb{F}_2^{2m}$.
3) Let ψ be a linear isomorphism from $\mathbb{F}_2^{2t}$ into C. Set $\widehat{\mathbf{z}} = [\widehat{a}_{\psi(\mathbf{u})}]$, where the vectors $\mathbf{u}$ are ordered lexicographically in $\mathbb{F}_2^{2t}$. (In other words $\widehat{\mathbf{z}}$ is the vector of signs of the nonzero entries of $\mathbf{F}$ ordered lexicographically.) Then $\widehat{\mathbf{z}}$, or equivalently the vector $\mathbf{z}$ obtained from $\widehat{\mathbf{z}}$ by reverse exponentiation, is a bent function of degree 2.

It is easy to understand the claims of the theorem with the help of the following example.

Example 6: Let $m=3, t=2$ and

$$f(x_1,\dots,x_6) = x_1x_2\oplus x_3x_4.$$

Then

$$\begin{aligned}\mathbf{F} &= H_6\mathbf{f}\\ &= (24,-8,-8,8,-8,-8,-8,8,-8,-8,-8,8,8,8,\\ &\quad 8,-8,0,\dots,0)^{\dagger}.\end{aligned}$$

[1] Theorems 10 and 11 are proven in more general form for partially bent functions in [6]

One can see that vectors $\mathbf{b}$ such that $F_{\mathbf{b}}\neq 0$ are

$$(0,0,0,0,0,0),(0,0,0,0,0,1),(0,0,0,0,1,0),\dots,$$
$$(0,0,1,1,1,1)$$

and that they form a $[6,4]$ linear code. It is easy to check that

$$\mathbf{z}^{\top} = (-1,-1,-1,1,-1,-1,-1,1,-1,-1,-1,1,1,1,1,-1)$$

and that it is a bent function.

Proof: Denote $\widehat{\mathbf{F}} = H_{2m}\widehat{\mathbf{f}}$. Using the simple observation that $\mathbf{f} = -(\widehat{\mathbf{f}}-1)/2$, we obtain

$$\mathbf{F} = H_{2m}\mathbf{f} = H_{2m}(-\frac{\widehat{\mathbf{f}}-\mathbf{1}}{2}) = -\frac{\widehat{\mathbf{F}}}{2} + \frac{H_{2m}\mathbf{1}}{2}. \tag{46}$$

The second term in the above expression is defined by

$$(H_{2m}\mathbf{1})_{\mathbf{b}} = \begin{cases} 2^{2m}, & \mathbf{b}=\mathbf{0},\\ 0, & \mathbf{b}\neq\mathbf{0}.\end{cases}$$

Let us analyze the first term. According to Lemma 14 there exists an affine transform $s(\mathbf{x}) = \mathbf{x}A+\mathbf{p}$ so that $f\circ s = g$ or $f\circ s = g\oplus 1$, where $g(x_1,\dots,x_{2m})$ is the canonical bent function defined in (39). We will consider only the case $f\circ s = g$; the case $f\circ s = g\oplus 1$ can be proven similarly.

The function g can be written in the form

$$g(x_1,\dots,x_{2m}) = g_1(x_1,\dots,x_{2t})\oplus 0(x_{2t+1},\dots,x_{2m}),$$

where $g_1(x_1,\dots,x_{2t}) = x_1x_2\oplus\dots\oplus x_{2t-1}x_{2t}$.

Taking into account that

$$\widehat{0}(\mathbf{w}) = \begin{cases} 2^{2m-2t}, & \mathbf{w}=\mathbf{0},\\ 0, & \mathbf{w}\neq\mathbf{0},\end{cases}$$

and using Lemma 9 we obtain

$$\widehat{G}(\mathbf{u}) = \begin{cases} 2^{2m-2t}\widehat{G_1}(\mathbf{v}), & \mathbf{u}=(\mathbf{v},\mathbf{0})\\ 0, & \mathbf{u}=(\mathbf{v},\mathbf{w}),\mathbf{w}\neq 0.\end{cases} \tag{47}$$

Note that vectors $(\mathbf{v},\mathbf{0})$ form a linear subspace of dimension $2t$.

Let $\mathbf{z} = (z_1,\dots,z_{2m}) = \mathbf{p}A^{-1}$ and let

$$\begin{aligned} g'(\mathbf{x}) &= g(\mathbf{x}\oplus\mathbf{z})\\ &= g_1(x_1\oplus z_1,\dots,x_{2t}\oplus z_{2t})\\ &\oplus\ 0(x_{2t+1}\oplus z_{2t+1},\dots,x_{2m}\oplus z_{2m}).\end{aligned}$$

According to Lemma 8 the first term in this expression is a bent function. Hence we have that

$$|\widehat{G'}(\mathbf{u})| = \begin{cases} 2^{2m-2t}|\widehat{G_1}(\mathbf{v})| = 2^{2m-t}, & \mathbf{u}=(\mathbf{v},\mathbf{0})\\ 0, & \mathbf{u}=(\mathbf{v},\mathbf{w}),\mathbf{w}\neq 0,\end{cases} \tag{48}$$

and the signs of $\widehat{G'}(\mathbf{v}), \mathbf{v}=(\mathbf{u},\mathbf{0})$, form a bent function.

Finally, according to (42) and (41) we have

$$(-1)^{\mathbf{u}\cdot\mathbf{p}}\widehat{G}(\mathbf{u}A^T) = \widehat{G'}(\mathbf{u}A^T) = \widehat{F}(\mathbf{u}).$$

From this expression, from (48), from (46), and from the fact that nonzero elements of $\widehat{G'}(\mathbf{v})$ form a bent function we obtain a proof of all three claim of the theorem. □

Let $f(x_1,\dots,x_{2m})$ be a Boolean function for which the matrix B has rank $2t$, and assume the linear component is nonzero when $\mathbf{f}$ is reduced to the form (45). Recall that in this case $\mathrm{wt}(\mathbf{f}) = 2^{2m-1}$. Again, set $\mathbf{F} = H_{2m}\mathbf{f}$.

Theorem 11: The entries $F_{\mathbf{b}}, \mathbf{b} \in \mathbb{F}_2^{2m}$, have the following properties.

- The nonzero entries of $\mathbf{F}$ have values
$$F_{\mathbf{b}} = \begin{cases} 2^{2m-1}, & \mathbf{b} = \mathbf{0}, \\ a_{\mathbf{b}} 2^{2m-t-1}, \widehat{a}_{\mathbf{b}} = \pm 1, & \mathbf{b} \neq \mathbf{0}. \end{cases}$$
- The vectors $\mathbf{b}, \mathbf{b} \neq \mathbf{0}$, such that $F_{\mathbf{b}} \neq \mathbf{0}$ form a coset $C \oplus \mathbf{h}$ of a linear subspace C of dimension $2t$ within $\mathbb{F}_2^{2m}$
- Let ψ be a linear isomorphism from $\mathbb{F}_2^{2t}$ into C. Set $\widehat{\mathbf{z}} = [\widehat{a}_{\psi(\mathbf{u})\oplus\mathbf{h}}]$ where the vectors $\mathbf{u} \in \mathbb{F}_2^{2t}$ are ordered lexicographically in $\mathbb{F}_2^{2t}$. (In other words $\widehat{\mathbf{z}}$ is the vector of signs of nonzero entries of $\mathbf{F}$, excluding the zero entry, ordered lexicographically.) Then $\widehat{\mathbf{z}}$, or equivalently the vector $\mathbf{z}$ obtained from $\widehat{\mathbf{z}}$ by reverse exponentiation, is a bent function.

Example 7: Let $m = 3, t = 2$ and

$$f(x_1, \ldots, x_6) = x_1 x_2 \oplus x_3 x_4 \oplus x_5.$$

Then

$$\begin{aligned} \mathbf{F} &= \mathbf{f} H_6 \\ &= (32, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \\ &\quad -8, -8, -8, 8, -8, -8, -8, 8, -8, -8, -8, 8, 8, 8, 8, -8, \\ &\quad 0, \ldots, 0)^{\dagger}. \end{aligned}$$

It is easy to see that nonzero entries of $\mathbf{F}$, excluding the zero entry, form a coset of a linear subspace of dimension 4 and their signs form a bent function of length 16.

Proof: This proof follows closely the proof of the previous theorem. For this reason we only point out the differences between the proofs. According to Lemma 14 there exists an affine transform s so that

$$f \circ s = g(x_1, \ldots, x_{2m}) = g_1(x_1, \ldots, x_{2t}) \oplus l(x_{2t+1}, \ldots, x_{2m}),$$

where

$$g_1(x1, \ldots, x_{2t}) = x_1 x_2 \oplus \ldots \oplus x_{2t-1} x_{2t}$$

and

$$l(x_{2t+1}, \ldots, x_{2m}) = x_{2t+1}.$$

The Hadamard transform of $l(x_{2t+1}, \ldots, x_{2m})$ is

$$\widehat{L}(\mathbf{w}) = \begin{cases} 2^{2m-2t}, & \mathbf{w} = (\underbrace{0, \ldots, 0}_{2t}, 1, \underbrace{0, \ldots, 0}_{2m-2t-1}) \\ 0, & \text{else.} \end{cases}$$

Hence, from Lemma 9 we obtain

$$\widehat{G}(\mathbf{u}) = \begin{cases} 0, & \mathbf{u} = (\mathbf{v}, 0, \mathbf{w}), \mathbf{w} = 0 \\ 2^{2m-2t} \widehat{G_1}(\mathbf{v}) \neq 0, & \mathbf{u} = (\mathbf{v}, 1, \mathbf{w}), \mathbf{w} = 0 \\ 0, & \mathbf{u} = (\mathbf{v}, 0, \mathbf{w}), \mathbf{w} \neq 0 \\ 0, & \mathbf{u} = (\mathbf{v}, 1, \mathbf{w}), \mathbf{w} \neq 0. \end{cases}$$
$$\mathbf{v} \in \mathbb{F}_2^{2t}, \mathbf{w} \in \mathbb{F}_2^{2m-2t-1}.$$

Thus we have that the vectors $\mathbf{u}$ such that $\widehat{G}(\mathbf{u}) \neq 0$ form a coset of linear subspace of dimension $2t$, while in the proof of Theorem 10, the equation (47), they formed a linear subspace. This is the only difference between the proofs. Now using the same steps as in the proof of Theorem 10 we finish the proof of Theorem 11. □

The converses of these two theorems are given below.

Theorem 12: If $\mathbf{F} \in \mathbb{R}^{2^{2m}}$ has all the properties listed in Theorem 10, then

$$\mathbf{c} = \frac{1}{2^{2m}} H_{2m} \mathbf{F} \in RM(2, 2m)$$

and $\mathrm{wt}(\mathbf{c}) = 2^{2m-1} \pm 2^{2m-t-1}$.

Before proving this theorem we remind the following facts.

Lemma 15: [24, Chapter 15.2] The number of symplectic $2m \times 2m$ matrices of rank $2t$ over $\mathbb{F}_2$ is

$$\begin{aligned} &N(2m, 2t) \\ &= \frac{(2^{2m}-1)(2^{2m-1}-1) \cdot \ldots \cdot (2^{2m-2t+1}-1)}{(2^{2t}-1)(2^{2t-2}-1) \cdot \ldots \cdot (2^2-1)} \cdot 2^{t(t-1)}. \end{aligned}$$

Lemma 16: [24, Chapter 15.2] The number of quadratic functions $f(x_1, \ldots, x_{2m})$ that can be transformed to the form (43) or (44) is equal to

$$N(2m, 2t) 2^{2t+1}.$$

Lemma 17: [24, Chapter 15.2] The number of bent functions of $2t$ variables of degree 2 is equal to

$$N(2t, 2t) \cdot 2^{2t+1}.$$

Proof: (Theorem 12.) From Lemmas 15 and 16 we have that the number of boolean functions $\mathbf{f}$ used in Theorem 10 is

$$\begin{aligned} &\frac{(2^{2m}-1)(2^{2m-1}-1) \cdot \ldots \cdot (2^{2m-2t+1}-1)}{(2^{2t}-1)(2^{2t-2}-1) \cdot \ldots \cdot (2^2-1)} \\ &\cdot \ 2^{t(t-1)} \cdot 2^{2t+1}. \end{aligned} \tag{49}$$

The number of vectors $\mathbf{F}$ that have all the properties listed in Theorem 10 is equal to the number of $[2m, 2t]$ linear codes times the number of bent functions of $2t$ variables of degree 2. Hence, using the Lemma 17 and 3, we obtain

$$\begin{aligned} &\frac{(2^{2m}-1)(2^{2m-1}-1) \cdot \ldots \cdot (2^{2m-2t+1}-1)}{(2^{2t}-1)(2^{2t-1}-1) \cdot \ldots \cdot (2-1)} \\ &\cdot \ \frac{(2^{2t}-1)(2^{2t-1}-1) \cdot \ldots \cdot (2-1)}{(2^{2t}-1)(2^{2t-2}-1) \cdot \ldots \cdot (2^2-1)} \\ &\cdot \ 2^{t(t-1)} \cdot 2^{2t+1}. \end{aligned} \tag{50}$$

Simplifying (50), we conclude that it is equal to (49). □

Next theorem can be proven in a similar way.

Theorem 13: If $\mathbf{F} \in \mathbb{R}^{2^{2m}}$ has all the properties listed in Theorem 11, then

$$\mathbf{c} = \mathbf{F} \frac{1}{2^{2m}} H_{2m} \in RM(2, 2m)$$

and $\mathrm{wt}(\mathbf{c}) = 2^{2m-1}$.

IX. Generalization of O-RM codes

We will use some facts from group representation theory to construct (T, M, L, d) Grassmannian packings with $T = 2^m, M = 2^{m-1}$ and $d = 2^{m-2}$ for which the orthogonal projectors can be mapped into codewords of the binary second order Reed-Muller code $RM(2, 2m)$. Hence the notation $GO\text{-}RM(2, m)$. This construction extends naturally to the following values of $M = 2^{m-1} \pm 2^{m-t-1}, 1 \leq t \leq \lfloor m/2 \rfloor$.

A. Projectors from Bent Functions

Let G be an Abelian group of linear operators and λ be a linear character of G. The following fact is well known in group representation theory, (see for example [23]).

Theorem 14: The operator

$$P = \frac{1}{|G|}\sum_{g\in G}\lambda^{\dagger}(g)g$$

is an orthogonal projector.

We provide an extension of this theorem, where instead of an Abelian group we use a coset of the group and instead of linear characters we use bent functions.

Let C be a $[2m, 2t]$ linear self-orthogonal code, and let S be the corresponding Abelian subgroup. If $\mathbf{b} \notin C$ then the operator $E_{\mathbf{b}} \notin S$. If $\mathbf{b} \in C^{\perp}$, where $C^{\perp}$ is the dual code of C with respect to the symplectic inner product (11), then $E_{\mathbf{b}}$ commutes with all operators from S. Otherwise $E_{\mathbf{b}}$ commutes with half the operators in S and anticommute with the other half. For brevity we set $E = E_{\mathbf{b}}$. Define

$$S_0 = \{A \in S : AE = EA\}, |S_0| = 2^{2t-1}.$$

Note that if $\mathbf{b} \in C^{\perp}$ then S_0 can be defined in multiple ways; any of them works for our purposes. Let

$$S_1 = \{BA : A \in S_0\}, \text{ for some } B \in S \setminus S_0.$$

If $\mathbf{b} \notin C^{\perp}$ then

$$A' \in S_1 \Rightarrow A'E = -EA'. \tag{51}$$

and it is easy to see that $S = S_0 \cup S_1$.

Let C_0 be the set of vectors corresponding to operators from S_0 and let C_1 be the set of vectors corresponding to operators from S_1, so that $C = C_0 \cup C_1$. Given $\mathbf{c} \in C$ let $\alpha_{\mathbf{c}}$ be such that $\alpha_{\mathbf{c}}E_{\mathbf{c}} \in S$ ($\alpha_{\mathbf{c}}$ is either 1 or -1). Set $A_{\mathbf{c}} = \alpha_{\mathbf{c}}E_{\mathbf{c}}$, and observe that from the construction of S and (9) it follows that $A_{\mathbf{d}}A_{\mathbf{c}} = A_{\mathbf{c}}A_{\mathbf{d}} = A_{\mathbf{d}\oplus\mathbf{c}}$.

Theorem 15: Let ψ be an isomorphism from $\mathbb{F}_2^{2t-1}$ into C_0 and let $f(x_1 \ldots, x_{2t})$ be a bent function. Let $\mathbf{w} = \psi^{-1}(\mathbf{b})$. If $\mathbf{b} \in C^{\perp}$ define

$$P = \frac{1}{2^{t+1}}\left(2^t I_{2^m} + \sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u})EA_{\psi(\mathbf{u})} + \sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u}\oplus\mathbf{w})EBA_{\psi(\mathbf{u})}\right), \tag{52}$$

and if $\mathbf{b} \notin C$ then define

$$P = \frac{1}{2^{t+1}}\left(2^t I_{2^m} + \sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u})EA_{\psi(\mathbf{u})} + i\sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u}\oplus\mathbf{w})EBA_{\psi(\mathbf{u})}\right). \tag{53}$$

Then P is an orthogonal projector and $\mathrm{Tr}\,(P) = 2^{m-1}$.

Proof: We will consider only the case $\mathbf{b} \notin C^{\perp}$. To prove that P is an orthogonal projector we have to show that $P = P^{\dagger}$ and $PP = P$. Since any operator in the extraspecial group is Hermitian, we have $P = P^{\dagger}$. After expansion of PP and cancellation of similar terms, with the help of (8), we obtain

$$\begin{aligned} PP &= \frac{1}{2^{t+1}}\Big(2^t I_{2^m} + \sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u})EA_{\psi(\mathbf{u})} \\ &+ i\sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u}\oplus\mathbf{w})EBA_{\psi(\mathbf{u})} \\ &+ \frac{1}{2^{t+1}}\sum_{\substack{\mathbf{h}\in\mathbb{F}_2^{2t-1}\\ \mathbf{h}\neq 0}} A_{\psi(\mathbf{h})}\sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u})\widehat{f}(\mathbf{u}\oplus\mathbf{h}) \\ &+ \frac{1}{2^{t+1}}\sum_{\substack{\mathbf{h}\in\mathbb{F}_2^{2t-1}\\ \mathbf{h}\neq 0}} A_{\psi(\mathbf{h})}\sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u}\oplus\mathbf{w})\widehat{f}(\mathbf{u}\oplus\mathbf{w}\oplus\mathbf{h})\Big). \end{aligned}$$

We combine the last two terms by taking the summation over C

$$\begin{aligned} PP = \frac{1}{2^{t+1}}\Big(&2^t I_{2^{2m}} + \sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u})EA_{\psi(\mathbf{u})} \\ &+ i\sum_{\mathbf{u}\in\mathbb{F}_2^{2t-1}} \widehat{f}(\mathbf{u}\oplus\mathbf{w})EBA_{\psi(\mathbf{u})} \\ &+ \frac{1}{2^{t+1}}\sum_{\substack{\mathbf{h}\in\mathbb{F}_2^{2t-1}\\ \mathbf{h}\neq 0}} A_{\psi(\mathbf{h})}\sum_{\mathbf{u}\in\mathbb{F}_2^{2t}} \widehat{f}(\mathbf{u})\widehat{f}(\mathbf{u}\oplus\mathbf{h})\Big). \end{aligned}$$

Since $\widehat{f}(\cdot)$ is a bent function, it follows from Lemma 13 that $\sum_{\mathbf{u}\in\mathbb{F}_2^{2t}} \widehat{f}(\mathbf{u})\widehat{f}(\mathbf{u}\oplus\mathbf{h}) = 0$, and we have $PP = P$. The claim $\mathrm{Tr}\,(P) = 2^{m-1}$ follows from (7). □

It is important to note that projection operators defined in Theorem 15 can be associated with codewords of $RM(2, 2m)$. Let P is defined according either to (52) or (53). Then

$$P = \frac{1}{2}I_{2^m} + \frac{1}{2^{t+1}}\sum_{\mathbf{u}\in C\oplus\mathbf{b}} \beta_{\mathbf{u}}\gamma_{\mathbf{u}}E_{\mathbf{u}},$$

where $\beta_j = \pm 1$ and $\gamma_{\mathbf{u}} = 1$ or i. Define the set of operators $\mathcal{P} = \{D_{\mathbf{v}} : \mathbf{v} \in \mathbb{F}_2^{2m}\}$ by

$$D_{\mathbf{v}} = \begin{cases} E_{\mathbf{v}}, & \mathbf{v} \notin C \oplus \mathbf{b}, \\ \gamma_{\mathbf{v}}E_{\mathbf{v}}, & \mathbf{v} \in C \oplus \mathbf{b}. \end{cases}$$

From (6) and (8) it follows that these operators form an orthonormal basis of the space of linear operators acting on $\mathbb{C}^{2^{2m}}$ with respect to the inner product $\mathrm{Tr}\,(AB^{\dagger})$. From this observation the following theorem follows.

Theorem 16: Let $P \in GO\text{-}RM(2, m)$ and let the vector $\mathbf{F}$ be defined by

$$F_{\mathbf{v}} = 2^m \mathrm{Tr}\,(PD_{\mathbf{v}}).$$

Then $\mathbf{F}$ has all the properties listed in Theorem 11 and therefore (according to Theorem 13) $\mathbf{c} = \mathbf{F}H_{2m}$ is a codeword of $RM(2, 2m)$.

This theorem shows that, as in the case of $O\text{-}RM(r, m)$, the decoding of $O\text{-}RM(r, m)$ can be conducted with the help of any decoding of the binary $RM(2, 2m)$ code.

Example 8: Let $m=3, t=1$, and $f=x_1x_2$. Let

$$\begin{aligned} C &= \{(0,0,0,0,0,0),(0,0,0,0,0,1),\\ &\quad (0,0,0,0,1,0),(0,0,0,0,1,1)\}. \end{aligned}$$

Let us choose $\mathbf{b}=(0,0,0,1,0,0)$. It is easy to see that $\mathbf{b}\in C^{\perp}$ and therefore there are several possible choices for C_0. Let

$$C_0=\{(0,0,0,0,0,0),(0,0,0,0,0,1)\} \text{ and}$$

$$C_1=\{(0,0,0,0,1,0),(0,0,0,0,1,1)\}.$$

Then according to (52) we obtain

$$\begin{aligned} P &= \frac{1}{4}(2I_8\\ &+ E_{(0,0,0,1,0,0)}E_{(0,0,0,0,0,0)}+E_{(0,0,0,1,0,0)}E_{(0,0,0,0,0,1)}\\ &+ E_{(0,0,0,1,0,0)}E_{(0,0,0,0,1,0)}-E_{(0,0,0,1,0,0)}E_{(0,0,0,0,1,1)})\\ &= \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 0 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & \\ & & & & & & & 1 \end{bmatrix} \end{aligned}$$

and

$$\mathbf{F}^{\dagger}=(32,0,0,0,16,16,16,-16,0,0,\ldots,0).$$

Let now $\mathbf{b}=(0,0,1,0,0,0)$. Then

$$C_0=\{(0,0,0,0,0,0),(0,0,0,0,1,0)\} \text{ and}$$

$$C_1=\{(0,0,0,0,0,1),(0,0,0,0,1,1)\}.$$

According to (53) we have

$$\begin{aligned} P &= \frac{1}{4}(2I_8\\ &+ E_{(0,0,1,0,0,0)}E_{(0,0,0,0,0,0)}+E_{(0,0,1,0,0,0)}E_{(0,0,0,0,1,0)}\\ &+ iE_{(0,0,1,0,0,0)}E_{(0,0,0,0,0,1)}+iE_{(0,0,1,0,0,0)}E_{(0,0,0,0,1,1)})\\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 & & & & & & \\ 1 & 1 & & & & & & \\ & & 1 & -i & & & & \\ & & i & 1 & & & & \\ & & & & 1 & 1 & & \\ & & & & 1 & 1 & & \\ & & & & & & 1 & -i \\ & & & & & & i & 1 \end{bmatrix}, \end{aligned}$$

and

$$\mathbf{F}^{\dagger}=(32,0,0,0,0,0,0,0,16,-16,16,16,0,0,\ldots,0).$$

Note that these projection operators do not belong to the code $O\text{-}RM(2,3)$. In general it is possible to show that $O\text{-}RM(2,m)\subset GO\text{-}RM(2,m)$.

B. Parameters

We have $T=2^m$, $M=2^{m-1}$, and we now estimate L using Lemmas 3 and 4.

We will set $\mathcal{N}(n,0)=1$. We will set $N(0,0)=1$.

Theorem 17: The size of $GO\text{-}RM(2,m)$ is

$$L=\sum_{t=0}^{\lfloor m/2\rfloor}\mathcal{N}(2m,2t)(2^{2m-2t}-1)N(2t,2t)2^{2t+1}, \quad (54)$$

where the functions $\mathcal{N}(\cdot,\cdot)$ and $N(\cdot,\cdot)$ are defined in Lemma 4 and Lemma 15 respectively.

Proof: According to Theorem 15, for construction of a projector P, we can use only self-orthogonal $[2m,2t]$ codes. This defines the lower and upper limits of summation in (54). Let $t\geq 1$ (the 0-th term of (54) requires a separate consideration, which we omit). According to Lemma 4 the number of self-orthogonal $[2m,2t]$ codes is $\mathcal{N}(2m,2t)$. The number of different cosets of an $[2m,2t]$ code is $2^{2m-2t}-1$. It follows from Theorem 5 of [24, Chapter 15.2] that the number of bent functions (of degree 2) of $2t$ variables is $N(2t,2t)\cdot 2^{2t+1}$. This completes the proof. □

In the next theorem we determine the minimum distance d of the packing $GO\text{-}RM(2,m)$.

Theorem 18: $d(GO\text{-}RM(2,m))=2^{m-2}$.

Proof: We will use notation introduced in Section V-C. Let $P,P'\in GO\text{-}RM(2,m)$. Define vectors $\mathbf{w}$ and $\mathbf{w}'$ by

$$w_{\mathbf{v}}=\frac{1}{2^m}\mathrm{Tr}\,(PD_v^{\dagger}) \text{ and } w'_{\mathbf{v}}=\frac{1}{2^m}\mathrm{Tr}\,(P'D_v^{\dagger})$$

respectively. Let $\mathbf{x}=2^m\mathbf{w}$ and $\mathbf{x}'=2^m\mathbf{w}'$ and

$$\mathbf{c}=\frac{1}{2^m}H_{2m}\mathbf{x}, \mathbf{c}'=\frac{1}{2^m}H_{2m}\mathbf{x}'.$$

According to Theorem 13 we have that

$$\mathbf{c},\mathbf{c}'\in RM(2,2m), \text{ and } \mathrm{wt}\,(\mathbf{c})=\mathrm{wt}\,(\mathbf{c}')=2^{2m-1}.$$

Taking into account that $1/2^mH_{2m}$ is unitary and that $c_j=0,1$ we obtain

$$||\mathbf{c}||_2=\sum_{j=0}^{2^{2m}-1}c_j^2=||\mathbf{x}||_2=2^{2m}||\mathbf{w}||_2.$$

Similarly

$$||\mathbf{c}'||_2=2^{2m}||\mathbf{w}'||_2, \text{ and } \mathbf{c}\cdot\mathbf{c}'=2^{2m}\mathbf{w}\cdot\mathbf{w}'.$$

Since the minimum distance of $RM(2,2m)$ is 2^{2m-2} we have

$$\mathrm{dist}\,(\mathbf{c},\mathbf{c}')^2=\sum_{j=0}^{2^{2m}-1}(c_j-c'_j)^2=||\mathbf{c}||_2+||\mathbf{c}'||_2-2\mathbf{c}\cdot\mathbf{c}'\geq 2^{2m-2}.$$

Hence

$$\mathbf{c}\cdot\mathbf{c}'\leq 2^{2m-1}-2^{2m-3} \text{ and } \mathbf{w}\cdot\mathbf{w}'\leq 2^{-1}-2^{-3}.$$

Next

$$\begin{aligned} \mathrm{Tr}\,(P'P)=\mathrm{Tr}\,(P'P^{\dagger}) &= \sum_{\mathbf{v},\mathbf{u}\in\mathbb{F}_2^{2m}}w'_{\mathbf{v}}w_{\mathbf{u}}\mathrm{Tr}\,(D_{\mathbf{v}}D_{\mathbf{u}})\\ &= 2^m\sum_{\mathbf{v}\in\mathbb{F}_2^{2m}}w'_{\mathbf{v}}w_{\mathbf{v}}=2^m\mathbf{w}'\cdot\mathbf{w}\leq 2^{m-1}-2^{m-3}. \end{aligned}$$

Finally, we have

$$\begin{aligned} & d(P',P) = \mathrm{Tr}\,(P') + \mathrm{Tr}\,(P) - 2\mathrm{Tr}\,(P'P) \\ \geq \quad & 2^{m-1} + 2^{m-1} - 2(2^{m-1} - 2^{m-3}) = 2^{m-2}. \end{aligned}$$

□

Example 9: The codes $GO\text{-}RM(2,m), m = 2,3,4$ define Grassmannian packings with the following parameters:

$$(4,2,390,1), (8,4,37926,2), \text{ and } (16,8,33544230,4).$$

Note that according to this example the sizes of the Grassmannian packings defined by $GO\text{-}RM(2,m), m = 2,3,4$ are significantly larger than the sizes of the corresponding Grassmannian packings obtained from $O\text{-}RM(2,m)$. At the same time it follows from Theorems 4 and 18 that $GO\text{-}RM(2,m)$ and $O\text{-}RM(2,m)$ have the same minimum distance. According to the following theorem, which we formulate without a proof, this is true for all values of m.

Theorem 19: $O\text{-}RM(2,m)$ and $GO\text{-}RM(2,m)$ have the same minimum distance and

$$O\text{-}RM(2,m) \subset GO\text{-}RM(2,m).$$

REFERENCES

[1] A. Ashikhmin and E. Knill, " Nonbinary quantum stabilizer codes," *IEEE Trans. Information Theory,* **47**, (2001), pp. 3065–3072.

[2] A. Ashikhmin and S. Litsyn, "Fast MAP Decoding of First Order Reed-Muller and Hamming Codes", *IEEE Trans. on Inform.Theory,* **50**, (2004), pp. 1812–1818.

[3] A. R. Calderbank, E. M. Rains, P. W. Shor and N. J. A. Sloane, "Quantum error correction and orthogonal geometry," *Phys. Rev. Lett.,* **78**, (1997), pp. 405–409.

[4] A. R. Calderbank, E. M. Rains, P. W. Shor and N. J. A. Sloane, "Quantum error correction via codes over $GF(4)$," *IEEE Trans. on Inform.Theory,* **44**, (1998), pp. 1369–1387.

[5] A. R. Calderbank, R. H. Hardin, E. M. Rains, P. W. Shor and N. J. A. Sloane, "A Group-Theoretic Framework for the Construction of Packings in Grassmannian Spaces," *J. Algebraic Combinatorics*, **9**, (1999), pp. 129-140.

[6] C. Carlet, "Boolean Functions for Cryptography and Error Correcting Codes," to appear as a chapter of the volume *Boolean Methods and Models*, Cambridge University Press.

[7] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packings lines, planes, etc.: packings in Grassmannian Spaces," *Experimental Math.,* **5**, pp.139–159, (1996).

[8] J. F. Dillon. Elementary Hadamard Difference Sets. Ph. D. Thesis, Univ. of Maryland, 1974.

[9] I. Dumer, "Recursive decoding and its performance for low-rate Reed-Muller codes", *IEEE Trans. on Inform. Theory*, **50**, (2004), pp. 811–823.

[10] G. J. Foschini, "Layered space–time architecture for wireless communication in a fading environment when using multi–element antennas," *Bell Labs. Tech. J.*, **1**, no. 2, pp. 41–59, 1996.

[11] Y. Jing, B. Hassibi, "Unitary space-time modulation via Cayley transform," *IEEE Transactions on Signal Processing,* **51**, no. 11, pp. 2891–2904, 2003.

[12] B. Hassibi and T. L. Marzetta, "Multiple-Antennas and isotropically random unitary inputs: the received signal density in closed form," *IEEE Trans. on Inform. Theory*, **48**, (2002), pp. 1473–1484.

[13] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communication in Rayleigh flat fading," *IEEE Trans. on Inform. Theory*, **46**, (2000), pp. 543–564.

[14] Y. Jiang, A. Ashikhmin, N. Sharma, "LDPC codes design for flat rayleigh fading channels with channel side information," *IEEE Trans. on Communication*, submitted. (Available at http://cm.bell-labs.com/who/aea/)

[15] A. Lozano, A. M. Tulino, S. Verdu, "Multiple-Antenna capacity in the low-power regime," *IEEE Trans. on Inform. Theory*, **49**, (2003), pp. 2527–2544.

[16] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *IEEE Trans. on Inform. Theory*, **45**, (1999), pp. 139–157.

[17] V. Tarokh, N. Seshadri, A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. on Inform. Theory*, **44**, (1998), pp.744–765.

[18] I. E. Telatar, "Capacity of multi–antenna Gaussian channels," *Eur. Trans. Telecom.*, **10**, pp. 585–595, Nov. 1999.

[19] S. ten Brink, G. Kramer, A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. on Commun.*, **52**, (2004), pp. 670 - 678.

[20] S. ten Brink and G. Kramer, "Design of repeat-accumulate codes for iterative detection and decoding," *IEEE Trans. on Signal Proc.*, **51**, (2003), pp. 2764-2772.

[21] L. Zheng and D. Tse, "Communication on the Grassman manifold: a geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. on Inform. Theory*, **48**, (2002), pp. 359–383.

[22] A. Ashikhmin and A. R. Calderbank, "Space-Time Reed-Muller Codes for Noncoherent MIMO Transmission," *International Symposium on Information Theory*, Adelaide, Australia, 2005, pp. 1952–1956.

[23] G. James and M. Liebeck, "Representations and Characters of Groups," Cambridge University Press, 1993.
Inc, 2004.

[24] F.J. MacWilliams and N.J.A Sloane. *The Theory of Error-Correcting Codes*. North-Holland Publishing Company, 1977.

[25] E. M. Rains and N. J. A. Sloane, "Self-Dual Codes,?? in *Handbook of Coding Theory*, V. S. Pless and W. C. Huffman (editors), Elsevier, Amsterdam, 1998, pp.177?294.

[26] J. A. Tropp, "Greed is Good: Algorithmic Results for Sparse Approximation,"*IEEE Trans. on Inform.Theory,* **50**, (2004), pp. 2231–2242.

[27] L. Yu, A. Ashikhmin, A. R. Calderbank, "LDPC codes for noncoherent MIMO transmission," in preparation.