On Cosets of the Generalized First-Order Reed–Muller Code with Low PMEPR

Kai-Uwe Schmidt

Abstract

Golay sequences are well suited for use as codewords in orthogonal frequency-division multiplexing (OFDM) since their peak-to-mean envelope power ratio (PMEPR) in q-ary phase-shift keying (PSK) modulation is at most 2. It is known that a family of polyphase Golay sequences of length 2^m organizes in $m!/2$ cosets of a q-ary generalization of the first-order Reed–Muller code, $\mathrm{RM}_q(1, m)$. In this paper a more general construction technique for cosets of $\mathrm{RM}_q(1, m)$ with low PMEPR is established. These cosets contain so-called near-complementary sequences. The application of this theory is then illustrated by providing some construction examples. First, it is shown that the $m!/2$ cosets of $\mathrm{RM}_q(1, m)$ comprised of Golay sequences just arise as a special case. Second, further families of cosets of $\mathrm{RM}_q(1, m)$ with maximum PMEPR between 2 and 4 are presented, showing that some previously unexplained phenomena can now be understood within a unified framework. A lower bound on the PMEPR of cosets of $\mathrm{RM}_q(1, m)$ is proved as well, and it is demonstrated that the upper bound on the PMEPR is tight in many cases. Finally it is shown that all upper bounds on the PMEPR of cosets of $\mathrm{RM}_q(1, m)$ also hold for the peak-to-average power ratio (PAPR) under the Walsh–Hadamard transform.

Index Terms

Aperiodic, code, complementary, correlation, Golay, OFDM, PAPR, PMEPR, Reed–Muller, Rudin–Shapiro, sequence, Walsh–Hadamard

I. INTRODUCTION

Despite many evident advantages of the orthogonal frequency-division multiplexing (OFDM) modulation technique, the widespread acceptance of OFDM mainly suffers from the usually high peak-to-mean envelope power ratio (PMEPR) of uncoded OFDM signals. A promising and elegant approach to solve this power control issue is to use a block code across the subcarriers

Kai-Uwe Schmidt is with Communications Laboratory, Dresden University of Technology, 01062 Dresden, Germany, e-mail: schmidtk@ifn.et.tu-dresden.de, web: http://www.ifn.et.tu-dresden.de/~schmidtk/

[13]. A well-designed code is able to provide a certain level of error protection and ensures a maximum PMEPR that is substantially reduced compared to uncoded transmission [12].

It has been proposed in a number of contributions [37], [36], [18] to use so-called *Golay sequences* [6] as codewords in OFDM since their PMEPR is at most 2 in q-ary phase-shift keying (PSK) modulation [27]. Major progress has been made by Davis and Jedwab [4]; slightly generalized by Paterson [23], it was proved that a large family of polyphase Golay sequences of length 2^m organizes in $m!/2$ cosets of a naturally generalized first-order Reed–Muller code $\mathrm{RM}_q(1, m)$ inside a generalized second-order Reed–Muller code $\mathrm{ZRM}_q(2, m)$ (see Definition 5). A so-called supercode was then constructed by taking the union of these $m!/2$ cosets. However, due to a vanishing rate, these codes are only suitable for, say, $m \leq 5$. For this reason, Davis and Jedwab [4] proposed to include further cosets of $\mathrm{RM}_q(1, m)$ with PMEPR greater than 2 (but still low), in order to increase the rate of the codes at the cost of a slightly larger PMEPR. An exhaustive search technique was employed to identify such cosets. However, due to high complexity, such a search becomes prohibitive when $m > 4$ and q increases.

Paterson's work [23] and the more general study by the author and Finger [32] provided some advanced theoretical background: it was shown that each coset of $\mathrm{RM}_q(1, m)$ is entirely comprised of sequences lying in so-called *complementary sets* of the same size. It follows that the PMEPR of the codewords in each coset of $\mathrm{RM}_q(1, m)$ is at most 2^{k+1}, where k depends on the algebraic structure of a coset representative. However this upper bound is not always tight, and at best it yields merely the lowest possible power of 2.

Further results were obtained by Parker and Tellambura. In [21] they proposed another construction technique for cosets of $\mathrm{RM}_q(1, m)$ comprised of sequences lying in complementary sets of small size. In [19], [20] the Rudin–Shapiro construction [33], [29] was used to obtain sequence sets with low PMEPR from a suitable pair of starting sequences (kernel). However such a construction generally does not yield cosets of $\mathrm{RM}_q(1, m)$. In order to obtain complete cosets, an exhaustive search (although less complex than that proposed in [4]) for suitable kernels was performed in [19] and [20].

Still, some phenomena cannot be fully explained. For example Davis and Jedwab observed [4] that there are cosets of $\mathrm{RM}_8(1, m)$ whose maximum PMEPR, taken over all words in the coset, is equal to 3. In addition numerical results suggest the existence of more coset classes whose PMEPR is bounded by low values not being a power of 2. This motivates further analyses of the PMEPR of cosets of $\mathrm{RM}_q(1, m)$. In particular it is of interest to identify cosets with PMEPR close to 2.

The remainder of this paper is structured as follows. In the next section we introduce the OFDM communication model and establish most of our notation. In Section III we use the

Rudin–Shapiro construction to obtain sets of so-called *near-complementary sequences* from a given kernel. This construction is then stated in Theorem 7 in the context of algebraic normal forms of generalized Boolean functions. These preliminary results serve as a stepping stone to establish our main result, summarized in Theorem 10 and Corollary 15, where we construct cosets of $\mathrm{RM}_q(1, m)$ from a kernel and prove an upper bound on the PMEPR of these cosets. In Section IV we prove a new lower bound on the PMEPR of arbitrary cosets of $\mathrm{RM}_q(1, m)$ and apply the results to those cosets constructed in Corollary 15. These observations lead to some relations between the PMEPR and the peak-to-average power ratio (PAPR) under the Walsh–Hadamard transform. In Section V we will comment on this issue and briefly discuss implications for the coding problem in multicode code-division multiple access (MC-CDMA) systems (cf. [24], [25]).

The application of our theory is then illustrated in Section VI by providing some construction examples. First, it is shown that the Davis–Jedwab construction of Golay sequences [4] just arises as a special case in our theory, since they originate from trivial kernels (of length 1). Second, we present further classes of cosets of $\mathrm{RM}_q(1, m)$ whose maximum PMEPR is between 2 and 4. In particular we identify a class with PMEPR at most 3, and therefore, we provide a proof for the conjecture by Davis and Jedwab on the PMEPR of a subset of this class [4]. The upper bound on the PMEPR is also compared with our lower bound, and it is shown that the upper bound is attained in many cases. Finally, in Section VII, we close with some concluding remarks and open problems.

II. NOTATION AND PRELIMINARIES

Throughout this paper $\xi = \exp(\sqrt{-1}\, 2\pi/q)$ denotes a primitive qth root of unity and q is an even positive integer.

A. Aperiodic Correlations and Complementarity

Let $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^n$ be two sequences with $\boldsymbol{A} = (A_0\, A_1\, \cdots\, A_{n-1})$ and $\boldsymbol{B} = (B_0\, B_1\, \cdots\, B_{n-1})$. If $A_i = \xi^{a_i}$ and $a_i \in \mathbb{Z}_q$ for all $i = 0, 1, \cdots, n-1$, we shall call $\boldsymbol{A}$ a *polyphase sequence*. The *aperiodic cross-correlation* of $\boldsymbol{A}$ and $\boldsymbol{B}$ at a displacement $\ell \in \mathbb{Z}$ is given by

$$
C(\boldsymbol{A}, \boldsymbol{B})(\ell) \triangleq \begin{cases} \sum_{i=0}^{n-\ell-1} A_{i+\ell} B_i^* & 0 \le \ell < n \\ \sum_{i=0}^{n+\ell-1} A_i B_{i-\ell}^* & -n < \ell < 0 \\ 0 & \text{otherwise,} \end{cases}
$$

where $()^*$ denotes complex conjugation. The *aperiodic auto-correlation* of $\boldsymbol{A}$ at a displacement $\ell \in \mathbb{Z}$ is then conveniently written as

$$A(\boldsymbol{A})(\ell) \triangleq C(\boldsymbol{A}, \boldsymbol{A})(\ell).$$

Definition 1: For two sequences $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^n$ we define the operator '$\star$' as follows

$$\boldsymbol{A} \star \boldsymbol{B} \triangleq \sum_{\ell=1-n}^{n-1} |A(\boldsymbol{A})(\ell) + A(\boldsymbol{B})(\ell)|.$$

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are polyphase sequences of length n, we obtain

$$\boldsymbol{A} \star \boldsymbol{B} = 2n + 2 \sum_{\ell=1}^{n-1} |A(\boldsymbol{A})(\ell) + A(\boldsymbol{B})(\ell)|.$$

A pair of polyphase sequences $(\boldsymbol{A}, \boldsymbol{B})$ is called a *complementary pair* if $\boldsymbol{A} \star \boldsymbol{B} = 2n$, which implies that the aperiodic auto-correlations of the two sequences sum up to zero for each nonzero shift. In tribute to Golay, who extensively studied binary complementary pairs in connection with multislit spectrometry [6], they are often called *Golay complementary pairs*, and each sequence lying in such a pair is termed a *Golay sequence*.

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are polyphase sequences, $\boldsymbol{A} \star \boldsymbol{B}$ is at most $2n^2$. We shall call a pair of polyphase sequences $(\boldsymbol{A}, \boldsymbol{B})$ a *near-complementary pair* and the sequences therein *near-complementary sequences* if $2n \leq (\boldsymbol{A} \star \boldsymbol{B}) \ll 2n^2$. In other words, the aperiodic auto-correlations of $\boldsymbol{A}$ and $\boldsymbol{B}$ sum up to small values at a few nonzero shifts.

B. OFDM Power Control

Let us consider an n-subcarrier OFDM system. The transmitted OFDM signal is the real part of the *complex envelope*, which can be written as

$$S(\boldsymbol{A})(\theta) \triangleq \sum_{i=0}^{n-1} A_i \, e^{\sqrt{-1}\, 2\pi(i+\zeta)\theta}, \quad 0 \leq \theta < 1,$$

where ζ is a positive constant. The sequence $\boldsymbol{A} = (A_0 \, A_1 \cdots A_{n-1})$ is called the modulating codeword of the OFDM symbol. Throughout this paper it is assumed that the elements of $\boldsymbol{A}$ are selected from a PSK constellation. Then $\boldsymbol{A}$ is a polyphase sequence.

The PMEPR of the OFDM signal (or of the modulating codeword $\boldsymbol{A}$) is then defined to be

$$\text{PMEPR}(\boldsymbol{A}) \triangleq \frac{1}{n} \sup_{0 \leq \theta < 1} |S(\boldsymbol{A})(\theta)|^2. \tag{1}$$

Notice that the PMEPR can be as large as n, which occurs, for example, if $\boldsymbol{A}$ is the all-one word. However it is desirable to use codewords with PMEPR that is substantially lower than n. For the construction of such codewords the following theorem will be essential in the sequel.

Theorem 2: Let $(\boldsymbol{A}, \boldsymbol{B})$ be a pair of polyphase sequences of length n. Then the PMEPR of $\boldsymbol{A}$ and $\boldsymbol{B}$ is at most $(\boldsymbol{A} \star \boldsymbol{B})/n$.

Proof: It is well known (cf., e.g., [35], [4]) that

$$|S(\boldsymbol{A})(\theta)|^2 = A(\boldsymbol{A})(0) + 2 \sum_{\ell=1}^{n-1} \Re \left\{ A(\boldsymbol{A})(\ell) \, e^{\sqrt{-1} \, 2\pi\ell\theta} \right\},$$

where $\Re\{.\}$ is the real part of a complex number. Hence

$$|S(\boldsymbol{A})(\theta)|^2 + |S(\boldsymbol{B})(\theta)|^2 = 2n + 2 \sum_{\ell=1}^{n-1} \Re \left\{ [A(\boldsymbol{A})(\ell) + A(\boldsymbol{B})(\ell)] \, e^{\sqrt{-1} \, 2\pi\ell\theta} \right\}$$

$$\leq 2n + 2 \sum_{\ell=1}^{n-1} |A(\boldsymbol{A})(\ell) + A(\boldsymbol{B})(\ell)| = \boldsymbol{A} \star \boldsymbol{B},$$

by Definition 1. Since $|S(.)(\theta)|^2$ of each individual sequence is non-negative, the PMEPR of $\boldsymbol{A}$ and $\boldsymbol{B}$ is at most $(\boldsymbol{A} \star \boldsymbol{B})/n$. $\qquad\square$

The above theorem is consistent with the results in [27] in the special case where $(\boldsymbol{A}, \boldsymbol{B})$ is a Golay complementary pair. Then the PMEPR of $\boldsymbol{A}$ and $\boldsymbol{B}$ is at most 2. In [37], [36], [18], [4] it has been proposed to exclusively use Golay sequences as codewords in OFDM. Consequently tight power control for OFDM is ensured, however, the code rate rapidly decreases for larger lengths. Theorem 2 motivates the use of larger sequence families with slightly higher PMEPR as codewords in OFDM.

C. Generalized Boolean Functions and Associated Sequences

A *generalized Boolean function* f is defined as a mapping $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$. Such a function can be written uniquely in its *algebraic normal form*, i.e., f is the sum of 2^m weighted *monomials*

$$f(\boldsymbol{x}) = f(x_0, x_1, \cdots, x_{m-1}) = \sum_{i=0}^{2^m-1} c_i \prod_{\alpha=0}^{m-1} x_\alpha^{i_\alpha},$$

where the weights $c_0, \cdots, c_{2^m-1}$ are in $\mathbb{Z}_q$, and $(i_0 \, i_1 \cdots i_{m-1})$ is the binary representation of $0 \leq i < 2^m$, such that $i = \sum_{j=0}^{m-1} i_j 2^j$ is its binary expansion. The *order of the ith monomial* is defined to be $\sum_{j=0}^{m-1} i_j$, and the *order, or algebraic degree, of a generalized Boolean function f*, denoted by $\deg(f)$, is equal to the highest order of the monomials with a nonzero coefficient in the algebraic normal form of f.

A generalized Boolean function may be equally represented by sequences of length 2^m. We shall define the sequence

$$\psi(f) \triangleq (f_0 \, f_1 \cdots f_{2^m-1})$$

as the $\mathbb{Z}_q$-*valued sequence associated with* f and the sequence

$$\Psi(f) \triangleq (\xi^{f_0} \xi^{f_1} \cdots \xi^{f_{2^m-1}})$$

as the *polyphase sequence associated with* f. Here we denote $f_i = f(i_0, i_1, \cdots, i_{m-1})$, where $(i_0\, i_1 \cdots i_{m-1})$ is the binary representation of $0 \le i < 2^m$.

In the remainder of this subsection we introduce the technique of extending polyphase sequences of length 2^m and their corresponding generalized Boolean functions. It will be used in the next section to prove our results on near-complementary sequences.

Definition 3: Suppose $m > k$, let $f : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be a generalized Boolean function in the variables $x_0, x_1, \cdots, x_{k-1}$, and write $\boldsymbol{F} = \Psi(f)$. Let

$$0 \le i_0 < i_1 < \cdots < i_{k-1} < m$$

and write

$$0 \le j_0 < j_1 < \cdots < j_{m-k-1} < m$$

for the remaining indices. Also denote $\boldsymbol{x} = (x_{j_0} \cdots x_{j_{m-k-1}})$ and let $\boldsymbol{d} = (d_0\, d_1 \cdots d_{m-k-1})$ be a binary word of length $m - k$. We define the *extended sequence* $\boldsymbol{F}_{[\boldsymbol{x}=\boldsymbol{d}]}$ of length 2^m as follows. As $(u_0\, u_1 \cdots u_{k-1})$ ranges over $\mathbb{Z}_2^k$, at position

$$\sum_{\alpha=0}^{k-1} u_\alpha 2^{i_\alpha} + \sum_{\alpha=0}^{m-k-1} d_\alpha 2^{j_\alpha}$$

the sequence $\boldsymbol{F}_{[\boldsymbol{x}=\boldsymbol{d}]}$ is equal to $\xi^{f(u_0, u_1, \cdots, u_{k-1})}$ and equal to zero otherwise. We also define the *extended generalized Boolean function* $f_{[\boldsymbol{x}]}$ that is formally regarded as a generalized Boolean function in m variables, i.e., it is of type $\mathbb{Z}_2^m \to \mathbb{Z}_q$. This function is obtained from f by replacing the variables x_α by x_{i_α} in the algebraic normal form of f for $\alpha = 0, 1, \cdots, k - 1$.

Notice that $\boldsymbol{F}_{[\boldsymbol{x}=\boldsymbol{d}]}$ comprises 2^k nonzero elements and $2^m - 2^k$ zeros. It is a consequence of the above definition that at the positions where $\boldsymbol{F}_{[\boldsymbol{x}=\boldsymbol{d}]}$ is nonzero the sequence $\boldsymbol{F}_{[\boldsymbol{x}=\boldsymbol{d}]}$ is equal to the polyphase sequence associated with $f_{[\boldsymbol{x}]}$.

Example 4: Let $f : \mathbb{Z}_2^2 \to \mathbb{Z}_2$ be given by

$$f = x_0 x_1 + x_1.$$

Writing '+' for '+1' and '−' for '−1', we have $\boldsymbol{F} = \Psi(f) = (++-+)$. Now take $\boldsymbol{x} = (x_0\, x_2)$ and $\boldsymbol{d} = (10)$. Then $\boldsymbol{F}_{[\boldsymbol{x}=\boldsymbol{d}]} = (0+0+00000-0+0000)$. By relabeling the variable indices in f according to $0 \mapsto 1, 1 \mapsto 3$, we obtain the extended function $f_{[\boldsymbol{x}]} = x_1 x_3 + x_3$. Regarding this function as a generalized Boolean function in m variables, we obtain its associated polyphase sequence $(++++++++--++--++)$. This sequence is equal to $\boldsymbol{F}_{[\boldsymbol{x}=\boldsymbol{d}]}$ at the positions where $\boldsymbol{F}_{[\boldsymbol{x}=\boldsymbol{d}]}$ is nonzero.

D. Generalized Reed–Muller Codes

A code $\mathcal{C}$ of length n over the ring $\mathbb{Z}_q$ is defined as a subset $\mathcal{C} \subseteq \mathbb{Z}_q^n$. Such a code is called $\mathbb{Z}_q$-*linear* if each $\mathbb{Z}_q$-linear combination of the codewords of $\mathcal{C}$ yields again a codeword of $\mathcal{C}$. If $\mathcal{C}$ is $\mathbb{Z}_q$-linear, a *coset* of $\mathcal{C}$ is defined to be

$$\boldsymbol{f} + \mathcal{C} \triangleq \{\boldsymbol{f} + \boldsymbol{c} \,|\, \boldsymbol{c} \in \mathcal{C}\},$$

where $\boldsymbol{f} \in \mathbb{Z}_q^n$ is called its *coset representative*, and of course, the additions are taken modulo q. We say that a coset of a code $\mathcal{C}_1$ lies inside a code $\mathcal{C}_2$ if $\boldsymbol{f} + \mathcal{C}_1 \subseteq \mathcal{C}_2$.

We are interested in codes defined by generalized Boolean functions. In what follows we recall the definitions and some basic properties of the generalized Reed–Muller codes $\mathrm{RM}_q(r, m)$ and $\mathrm{ZRM}_q(r, m)$ (cf. [4] and [23]).

Definition 5: (a) For $0 \leq r \leq m$ the code $\mathrm{RM}_q(r, m)$ is defined as the set of sequences $\psi(f)$, where f is a generalized Boolean function $\mathbb{Z}_2^m \to \mathbb{Z}_q$ of order at most r.
(b) For $q \geq 4$ and $1 < r \leq m$ the code $\mathrm{ZRM}_q(r, m)$ is defined as the set of sequences $\psi(f)$, where f is a generalized Boolean function $\mathbb{Z}_2^m \to \mathbb{Z}_q$ with algebraic normal form containing monomials of order at most $r - 1$ and two times the monomials of order r.

The codes $\mathrm{RM}_q(r, m)$ and $\mathrm{ZRM}_q(r, m)$ are $\mathbb{Z}_q$-linear, and their minimum Lee distances are equal to 2^{m-r} and 2^{m-r+1}, respectively [4], [23]. In this paper we will particularly study cosets of the code $\mathrm{RM}_q(1, m)$, which is comprised of the codewords corresponding to all affine forms over $\mathbb{Z}_q$ in m two-state variables. Hence the number of words in such cosets is equal to q^{m+1}.

III. Constructions of Sequence Families with Low PMEPR

A. Rudin–Shapiro Constructions

In what follows we present a slightly generalized version of the Rudin–Shapiro construction [33], [29], and exhibit its application to the construction of near-complementary pairs. Our main argument is the following lemma.

Lemma 6: Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two sequences of the same length and let $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ and $\boldsymbol{D} = \boldsymbol{A} - \boldsymbol{B}$. Then $\boldsymbol{C} \star \boldsymbol{D} = 2(\boldsymbol{A} \star \boldsymbol{B})$.

Proof: It is straightforward to show

$$A(\boldsymbol{C})(\ell) = A(\boldsymbol{A})(\ell) + A(\boldsymbol{B})(\ell) + C(\boldsymbol{A}, \boldsymbol{B})(\ell) + C(\boldsymbol{B}, \boldsymbol{A})(\ell)$$

$$A(\boldsymbol{D})(\ell) = A(\boldsymbol{A})(\ell) + A(\boldsymbol{B})(\ell) - C(\boldsymbol{A}, \boldsymbol{B})(\ell) - C(\boldsymbol{B}, \boldsymbol{A})(\ell).$$

Combining the relations above and Definition 1, the lemma follows. $\qquad\square$

It is well known that the Rudin–Shapiro construction can be employed to recursively construct longer complementary pairs starting from a known complementary pair [6], [2]. Using the notation in the present paper, this will be illustrated in the context of a more general framework.

Let A and B be two polyphase sequences of length 2^k. The extended sequence $A_{[x_k=0]}$ is a sequence of length 2^{k+1} and contains the sequence A in the left half, while its right half contains only zeros. Likewise the sequence $B_{[x_k=1]}$ contains zeros in the left half and the sequence B in the right half. Observe that $A(A_{[x_k=0]})(\ell) = A(A)(\ell)$ and $A(B_{[x_k=1]})(\ell) = A(B)(\ell)$ for each $\ell \in \mathbb{Z}$. Let us construct

$$C = A_{[x_k=0]} + B_{[x_k=1]}$$
$$D = A_{[x_k=0]} - B_{[x_k=1]},$$

and notice that the same pair could be constructed by

$$C = (A|B), \quad D = (A|-B),$$

where $(.|.)$ denotes concatenation. The reader may recognize the classical Rudin–Shapiro construction. By Lemma 6 we know that $C \star D = 2(A \star B)$. Hence, if (A, B) is a complementary pair, so will be (C, D). This construction is also known as Golay's concatenation technique for synthesizing complementary pairs [6].

We may also find a new sequence pair by

$$C = A_{[x_0=0]} + B_{[x_0=1]}$$
$$D = A_{[x_0=0]} - B_{[x_0=1]}.$$

Notice that $A_{[x_0=0]}$ is obtained from A by inserting zeros in A at alternating positions starting at the second position. Likewise $B_{[x_0=1]}$ is obtained from B by inserting zeros in B at alternating positions starting at the first position. It is then easy to verify that $A_{[x_0=0]} \star B_{[x_0=1]} = A \star B$, and by Lemma 6, we have $C \star D = 2(A \star B)$. Therefore, if (A, B) is a complementary pair, so will be (C, D). This construction is essentially the same as Golay's interleaving technique for synthesizing complementary pairs [6].

If A and B are extended in more than one variable and Lemma 6 is applied multiple times, then we can obtain even more general methods to construct longer sequence pairs from short ones. Again, if we restrict our attention to complementary pairs, we obtain Golay's general interleaving method to construct complementary pairs of length 2^m [7], [6], [2].

We have now established the following. Starting from an initial polyphase sequence pair (A, B) of length 2^k we can use the notion of extended sequences and a generalized Rudin–Shapiro construction to compose polyphase sequence pairs (C, D) of length 2^m with $m > k$. Since by Theorem 2 we have $(C \star D)/2^m = (A \star B)/2^k$, the sequences A, B, C, and D have the same PMEPR upper bound $(A \star B)/2^k$. If the sequence pair (A, B) cannot be obtained from a shorter sequence pair in this way, then it is referred to as a *kernel*.

B. Explicit Constructions of Near-Complementary Sequences

Using the language of generalized Boolean functions, an explicit construction for near-complementary sequences is presented in the following.

Theorem 7: Let $m > k$ and write $m - k = s + t$ for non-negative integers s, t. Define the sets

$$J = \{0, \cdots, s - 1, m - t, \cdots, m - 1\}$$

and $I = \mathbb{Z}_m \backslash J$. Denote the elements of J and I by

$$0 \leq j_0 < j_1 < \cdots < j_{m-k-1} < m$$

and

$$0 \leq i_0 < i_1 < \cdots < i_{k-1} < m,$$

respectively. Let $a, b : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be two generalized Boolean functions and define $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ by

$$f(x_0, \cdots, x_{m-1}) = a(x_{i_0}, \cdots, x_{i_{k-1}})(1 - x_{j_{\pi(0)}}) + b(x_{i_0}, \cdots, x_{i_{k-1}})x_{j_{\pi(0)}}$$
$$+ \frac{q}{2} \sum_{\alpha=0}^{m-k-2} x_{j_{\pi(\alpha)}} x_{j_{\pi(\alpha+1)}} + \sum_{\alpha=0}^{m-k-1} w_\alpha x_{j_{\pi(\alpha)}} + w,$$

where $w_0, \cdots, w_{m-k-1}, w \in \mathbb{Z}_q$ and π is a permutation of $\{0, 1, \cdots, m - k - 1\}$. Then

$$\Psi(f) \star \Psi\left(f + \frac{q}{2} x_{j_{\pi(m-k-1)}}\right) = 2^{m-k}\left[\Psi(a) \star \Psi(b)\right],$$

and

$$\text{PMEPR}(\Psi(f)) \leq \frac{\Psi(a) \star \Psi(b)}{2^k}.$$

Proof: Write $\boldsymbol{x} = (x_{j_{\pi(0)}} \cdots x_{j_{\pi(m-k-1)}})$ and let $\boldsymbol{d} = (d_0 \cdots d_{m-k-1})$ be a binary word of length $m - k$. Moreover let $\boldsymbol{A} = \Psi(a)$ and $\boldsymbol{B} = \Psi(b)$. Consider the extended sequences $\boldsymbol{A}_{[\boldsymbol{x}=\boldsymbol{d}]}$ and $\boldsymbol{B}_{[\boldsymbol{x}=\boldsymbol{d}]}$. These sequences can be constructed by successively inserting zeros at alternating positions, at the beginning, or at the end of the sequences $\boldsymbol{A}$ and $\boldsymbol{B}$. It is then straightforward to establish that

$$\boldsymbol{A}_{[\boldsymbol{x}=\boldsymbol{d}]} \star \boldsymbol{B}_{[\boldsymbol{x}=\boldsymbol{d}]} = \boldsymbol{A} \star \boldsymbol{B}. \tag{2}$$

We now use a method similar to Golay's general interleaving construction, which is here applied to pairs that are not necessarily complementary. Let us define the recurrence formulae

$$\boldsymbol{C}^{(\mu+1)}(d_{\mu+1}, \cdots, d_{m-k-1}) = \boldsymbol{C}^{(\mu)}(0, d_{\mu+1}, \cdots, d_{m-k-1}) + \xi^{w_\mu} \boldsymbol{D}^{(\mu)}(1, d_{\mu+1}, \cdots, d_{m-k-1}) \tag{3}$$

$$\boldsymbol{D}^{(\mu+1)}(d_{\mu+1}, \cdots, d_{m-k-1}) = \boldsymbol{C}^{(\mu)}(0, d_{\mu+1}, \cdots, d_{m-k-1}) - \xi^{w_\mu} \boldsymbol{D}^{(\mu)}(1, d_{\mu+1}, \cdots, d_{m-k-1}), \tag{4}$$

where $\mu = 0, 1, \cdots, m - k - 1$ and

$$\boldsymbol{C}^{(0)}(d_0, d_1, \cdots, d_{m-k-1}) = \boldsymbol{A}_{[\boldsymbol{x}=\boldsymbol{d}]}$$
$$\boldsymbol{D}^{(0)}(d_0, d_1, \cdots, d_{m-k-1}) = \boldsymbol{B}_{[\boldsymbol{x}=\boldsymbol{d}]}.$$

Since the positions of the nonzero components in $\boldsymbol{C}^{(\mu)}(0, d_{\mu+1}, \cdots, d_{m-k-1})$ and $\boldsymbol{D}^{(\mu)}(1, d_{\mu+1}, \cdots, d_{m-k-1})$ are distinct for each $\mu = 0, \cdots, m - k - 1$, $\boldsymbol{C}^{(m-k)}$ and $\boldsymbol{D}^{(m-k)}$ are polyphase sequences of length 2^m. Observe that $A(\xi^{w_\mu} \boldsymbol{D}^{(\mu)})(\ell) = A(\boldsymbol{D}^{(\mu)})(\ell)$ for each $\ell \in \mathbb{Z}$. Therefore, by repeated application of Lemma 6 and using (2), we have

$$\boldsymbol{C}^{(m-k)} \star \boldsymbol{D}^{(m-k)} = 2^{m-k}(\boldsymbol{A} \star \boldsymbol{B}). \tag{5}$$

We know that $\boldsymbol{A}_{[\boldsymbol{x}=\boldsymbol{d}]}$ and the polyphase sequence corresponding to $a_{[\boldsymbol{x}]}$ are equal in those positions where $\boldsymbol{A}_{[\boldsymbol{x}=\boldsymbol{d}]}$ is nonzero. An analogous statement holds for $\boldsymbol{B}_{[\boldsymbol{x}=\boldsymbol{d}]}$ and $b_{[\boldsymbol{x}]}$. Similarly we can find functions, say $c^{(\mu)}$ and $d^{(\mu)}$, whose associated polyphase sequences are equal to $\boldsymbol{C}^{(\mu)}$ and $\boldsymbol{D}^{(\mu)}$ at their respective nonzero components. Using (3) and (4) it can be verified that the functions $c^{(\mu)}$ and $d^{(\mu)}$ can be recursively constructed as follows

$$c^{(\mu+1)} = c^{(\mu)}(1 - x_{j_{\pi(\mu)}}) + (d^{(\mu)} + w_\mu)x_{j_{\pi(\mu)}}$$
$$d^{(\mu+1)} = c^{(\mu+1)} + \frac{q}{2}x_{j_{\pi(\mu)}},$$

where $c^{(0)} = a_{[\boldsymbol{x}]} = a(x_{i_0}, \cdots, x_{i_{k-1}})$ and $d^{(0)} = b_{[\boldsymbol{x}]} = b(x_{i_0}, \cdots, x_{i_{k-1}})$. Explicitly we obtain

$$c^{(\mu)} = a(x_{i_0}, \cdots, x_{i_{k-1}})(1 - x_{j_{\pi(0)}}) + b(x_{i_0}, \cdots, x_{i_{k-1}})x_{j_{\pi(0)}}$$
$$+ \frac{q}{2}\sum_{\alpha=0}^{\mu-2} x_{j_{\pi(\alpha)}} x_{j_{\pi(\alpha+1)}} + \sum_{\alpha=0}^{\mu-1} w_\alpha x_{j_{\pi(\alpha)}}$$
$$d^{(\mu)} = c^{(\mu)} + \frac{q}{2}x_{j_{\pi(\mu-1)}},$$

where $\mu = 1, 2, \cdots, m - k$. We have $f = c^{(m-k)} + w$ and $f + \frac{q}{2}x_{j_{\pi(m-k-1)}} = d^{(m-k)} + w$. The theorem follows then from (5) and Theorem 2. $\qquad \square$

We refer to Section VII for a discussion on the relation of Theorem 7 to previous results in [19] and [20].

C. Cosets of $\mathrm{RM}_q(1, m)$ with Low PMEPR

Once suitable kernels are known, Theorem 7 identifies a large family of sequences with low PMEPR. However it would be desirable to construct sequence families that naturally form unions of cosets of $\mathrm{RM}_q(1, m)$ inside a higher-order generalized Reed–Muller code. In this way we could quickly obtain error-correcting codes, for which well-designed encoding and decoding

algorithms exist (see, e.g., [4], [26], [8], [30], [31]). The remainder of this section is dedicated to finding such sequence sets. First we require some preliminaries.

Definition 8: Let $f : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be a generalized Boolean function. We define the sequence $\Phi(f)$ of length $(4^k + 2)/3$ as follows. As $(u_0\, u_1 \cdots u_{k-1})$ ranges over $\mathbb{Z}_2^k$, at position

$$\sum_{\alpha=0}^{k-1} u_\alpha 2^{2\alpha}$$

the sequence $\Phi(f)$ is equal to $\xi^{f(u_0, u_1, \cdots, u_{k-1})}$ and equal to zero otherwise.

We remark that the sequence $\Phi(f)$ may also be obtained from the extended sequence $\Psi(f)_{[\boldsymbol{x}=\boldsymbol{d}]}$ by setting $\boldsymbol{x} = (x_1\, x_3 \cdots x_{2k-3})$, letting $\boldsymbol{d}$ be the all-zero word of length $k - 1$, and deleting the trailing zeros.

Definition 8 implies that for any generalized Boolean function $f : \mathbb{Z}_2^k \to \mathbb{Z}_q$ at its 2^k nonzero elements the sequence $\Phi(f)$ coincides with the polyphase sequence associated with a function that is obtained by replacing x_α by $x_{2\alpha}$ for each $\alpha = 0, 1, \cdots, k - 1$ in the algebraic normal form of f, where this new function is regarded as the algebraic normal form of a generalized Boolean function in $2k - 1$ variables.

Example 9: Let $f : \mathbb{Z}_2^2 \to \mathbb{Z}_4$ be given by

$$f(x_0, x_1) = 2x_0x_1 + 3x_0 + x_1.$$

We obtain:

$$\Phi(f) = (1\ -j\ 0\ 0\ j\ -1), \quad j = \sqrt{-1}.$$

At the nonzero positions the above sequence coincides with the polyphase sequence associated with the function $g : \mathbb{Z}_2^3 \to \mathbb{Z}_4$ whose algebraic normal form is given by $g(x_0, x_1, x_2) = 2x_0x_2 + 3x_0 + x_2$.

We are now in the position to state the main theorem of this paper.

Theorem 10: Suppose that $m > k$. Let $a, b : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be two generalized Boolean functions, and let $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$, a generalized Boolean function in the variables $x_0, x_1, \cdots, x_{m-1}$, be given by

$$f(x_0, \cdots, x_{m-1}) = a(x_{\pi(0)}, \cdots, x_{\pi(k-1)})(1 - x_{\pi(k)}) + b(x_{\pi(0)}, \cdots, x_{\pi(k-1)})x_{\pi(k)}$$

$$+ \frac{q}{2}\sum_{\alpha=k}^{m-2} x_{\pi(\alpha)}x_{\pi(\alpha+1)} + \sum_{\alpha=0}^{m-1} w_\alpha x_{\pi(\alpha)} + w,$$

where π is a permutation of $\{0, 1, \cdots, m - 1\}$ and $w_0, \cdots, w_{m-1}, w \in \mathbb{Z}_q$. Then

$$\Psi(f) \star \Psi\left(f + \frac{q}{2}x_{\pi(m-1)}\right) \leq 2^{m-k}\left[\Phi(a) \star \Phi(b)\right]$$

and

$$\text{PMEPR}(\Psi(f)) \leq \frac{\Phi(a) \star \Phi(b)}{2^k}.$$

We need a series of lemmas in order to prove the theorem. It is well known that each integer $0 \leq i < 2^m$ has a unique binary representation $(i_0 \cdots i_{m-1})$, such that $i = \sum_{\alpha=0}^{m-1} i_\alpha 2^\alpha$ and $i_\alpha \in \{0, 1\}$. If we now allow $i_\alpha \in \{-1, 0, +1\}$, we obtain a *signed-digit representation* (SDR) of i [11]. Such a representation is in general not unique. An SDR is called *sparse* if it contains no adjacent nonzero entries. We have the following lemma, which is a well-known result in number-representation theory (cf. [9], [11]). A proof is included, since the lemma will play a crucial role in the sequel.

Lemma 11: Every nonzero integer i has a unique sparse SDR $(i_0 \cdots i_{m-1})$ with $i_{m-1} \neq 0$.

Proof: Without loss of generality we assume that $i > 0$. We first show that for each nonzero integer there exists a sparse SDR. For $i = 1$ this is obvious. The remaining cases are proved by induction. There are three cases: (i) if i is even, take a sparse SDR for $i/2$ and prepend '0', (ii) if $i \equiv 1 \pmod 4$, take a sparse SDR for $(i-1)/4$ and prepend '1, 0', (iii) if $i \equiv -1 \pmod 4$, take a sparse SDR for $(i+1)/4$ and prepend '−1, 0'.

Now let us prove the uniqueness of a sparse SDR for a given positive integer. We first show that there is a unique sparse SDR for $i = 1$ and then proceed by induction. Suppose $i_{m-1} = 1$. Then

$$i \geq \begin{cases} 2^{m-1} - \displaystyle\sum_{\alpha=0}^{(m-3)/2} 2^{2\alpha} \quad = \dfrac{2^m + 1}{3} & \text{if } m \text{ is odd} \\ 2^{m-1} - \displaystyle\sum_{\alpha=0}^{(m-4)/2} 2^{2\alpha+1} = \dfrac{2^m + 2}{3} & \text{if } m \text{ is even.} \end{cases}$$

If $i_{m-1} = -1$, a similar argument yields $i < 0$. We conclude that, if $i = 1$, then $m = 1$ and $i_0 = 1$. Hence a sparse SDR for $i = 1$ is unique.

Now suppose that $i > 1$ is the smallest positive integer with two sparse SDRs, namely $(u_0, \ldots, u_{m-1})$ and $(v_0, \ldots, v_{n-1})$. If i is even, then $u_0 = v_0 = 0$. By shifting the two SDRs one position to the left, we obtain two sparse SDRs for $i/2 < i$, which is a contradiction. If $i \equiv 1 \pmod 4$ or $i \equiv -1 \pmod 4$, then $u_0 = v_0 = 1$ or $u_0 = v_0 = -1$, respectively. Consequently, we would have two sparse SDRs for $(i - u_0)/4 < i$. Again we arrive at a contradiction, which completes the proof. $\square$

We are now able to prove the following three lemmas.

Lemma 12: Suppose $a, b : \mathbb{Z}_2^k \to \mathbb{Z}_q$ are two generalized Boolean functions. Then

$$\Psi(a)_{[\boldsymbol{x}=\boldsymbol{d}]} \star \Psi(b)_{[\boldsymbol{x}=\boldsymbol{d}]} \leq \Phi(a) \star \Phi(b)$$

holds for any appropriate list of variables $\boldsymbol{x}$ and for any $\boldsymbol{d}$ of suitable length.

Proof: Let

$$0 \le i_0 < i_1 < \cdots < i_{k-1} < m$$

and write

$$0 \le j_0 < j_1 < \cdots < j_{m-k-1} < m$$

for the remaining indices. Denote $\boldsymbol{x} = (x_{j_0} \cdots x_{j_{m-k-1}})$ and let $\boldsymbol{d} = (d_0 \cdots d_{m-k-1})$ be a fixed binary word of length $m - k$. Write $\boldsymbol{A} = \Psi(a)_{[\boldsymbol{x}=\boldsymbol{d}]}$, $\widetilde{\boldsymbol{A}} = \Phi(a)$, $\boldsymbol{B} = \Psi(b)_{[\boldsymbol{x}=\boldsymbol{d}]}$, and $\widetilde{\boldsymbol{B}} = \Phi(b)$. We claim that

$$A(\boldsymbol{A})(\ell) = \sum_{\ell' \in \mathcal{L}(\ell)} A(\widetilde{\boldsymbol{A}})(\ell') \tag{6}$$

for some disjoint sets $\mathcal{L}(\ell)$, which we shall now prove. Let u and u' have binary expansion

$$u = \sum_{\alpha=0}^{k-1} u_\alpha 2^{i_\alpha} + \sum_{\alpha=0}^{m-k-1} d_\alpha 2^{j_\alpha}$$

$$u' = \sum_{\alpha=0}^{k-1} u_\alpha 2^{2\alpha},$$

respectively, and use an analogous notation for v and v'. Then u and u' define the positions of the nonzero components of $\boldsymbol{A}$ and $\widetilde{\boldsymbol{A}}$, respectively. By Definition 3 and Definition 8 we have $A_u = \tilde{A}_{u'}$ for any u and corresponding u'. Now consider the nonzero product $\tilde{A}_{u'}\tilde{A}_{v'}^* = A_u A_v^*$ with $u' \ne v'$ (and therefore $u \ne v$) occuring in the expression $A(\widetilde{\boldsymbol{A}})(u' - v')$ and also in $A(\boldsymbol{A})(u - v)$. An SDR of $\ell' = u' - v'$ is given by

$$(u_0 - v_0 \ \ 0 \ \ u_1 - v_1 \ \cdots \ 0 \ \ u_{k-1} - v_{k-1}).$$

This SDR is sparse and, by Lemma 11, unique. In other words, for each $\alpha = 0, \cdots, k - 1$ the differences $u_\alpha - v_\alpha$ are uniquely determined by ℓ'. Hence $\ell = u - v$ (which is independent of $\boldsymbol{d}$) is also uniquely determined by ℓ'. This means that all nonzero products contributing to the sum in the expression $A(\widetilde{\boldsymbol{A}})(\ell')$ also contribute to the sum in the expression $A(\boldsymbol{A})(\ell)$ for exactly one ℓ, which proves (6).

Now we use (6) to establish

$$\boldsymbol{A} \star \boldsymbol{B} = 2^{k+1} + 2 \sum_{\ell=1}^{2^m - 1} |A(\boldsymbol{A})(\ell) + A(\boldsymbol{B})(\ell)|$$

$$= 2^{k+1} + 2 \sum_{\ell=1}^{2^m - 1} \left| \sum_{\ell' \in \mathcal{L}(\ell)} A(\widetilde{\boldsymbol{A}})(\ell') + \sum_{\ell' \in \mathcal{L}(\ell)} A(\widetilde{\boldsymbol{B}})(\ell') \right|$$

$$\le 2^{k+1} + 2 \sum_{\ell=1}^{2^m - 1} \sum_{\ell' \in \mathcal{L}(\ell)} \left| A(\widetilde{\boldsymbol{A}})(\ell') + A(\widetilde{\boldsymbol{B}})(\ell') \right|.$$

Since the sets $\mathcal{L}(\ell)$ are disjoint and each product contributing to $A(\widetilde{\boldsymbol{A}})(\ell')$ for some ℓ' also contributes to $A(\boldsymbol{A})(\ell)$ for some ℓ, the latter expression is equal to $\widetilde{\boldsymbol{A}} \star \widetilde{\boldsymbol{B}}$, which completes the proof. $\qquad\square$

Lemma 13: Let $a, b : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be two generalized Boolean functions, and define $\tilde{a} = a + L$ and $\tilde{b} = b + L$, where

$$L = \sum_{\alpha=0}^{k-1} w_\alpha x_\alpha, \qquad w_0, \cdots, w_{k-1} \in \mathbb{Z}_q.$$

Then

$$\Phi(\tilde{a}) \star \Phi(\tilde{b}) = \Phi(a) \star \Phi(b).$$

Proof: Write $\boldsymbol{A} = \Phi(a)$, $\widetilde{\boldsymbol{A}} = \Phi(\tilde{a})$, $\boldsymbol{B} = \Phi(b)$, and $\widetilde{\boldsymbol{B}} = \Phi(\tilde{b})$. Let u and v have binary representation $(u_0 \cdots u_{2k-2})$ and $(v_0 \cdots v_{2k-2})$, respectively. For $u \neq v$ we consider the products appearing in $A(\widetilde{\boldsymbol{A}})(u - v)$

$$
\begin{aligned}
\tilde{A}_u \tilde{A}_v^* &= A_u \xi^{\sum_{\alpha=0}^{k-1} w_\alpha u_{2\alpha}} A_v^* \xi^{-\sum_{\alpha=0}^{k-1} w_\alpha v_{2\alpha}} \\
&= A_u A_v^* \xi^{\sum_{\alpha=0}^{k-1} w_\alpha (u_{2\alpha} - v_{2\alpha})}.
\end{aligned}
\tag{7}
$$

Let us focus our attention on the word $(u_0 - v_0 \cdots u_{2k-2} - v_{2k-2})$, which is an SDR of $u - v$. By Definition 8, A_u is equal to zero if and only if there exists an $\alpha \in \{0, \cdots, k - 2\}$ such that $u_{2\alpha+1} \neq 0$. Therefore, if the product $A_u A_v^*$ is nonzero, an SDR of $u - v$ is sparse and, by Lemma 11, unique. In this case the differences $u_{2\alpha} - v_{2\alpha}$ are uniquely determined by the difference $u - v$. Then we have $\tilde{A}_u \tilde{A}_v^* = A_u A_v^* \xi^{K(u-v)}$, where $K(u - v)$ only depends on $u - v$ and not explicitly on u and v themselves. Thus

$$A(\widetilde{\boldsymbol{A}})(\ell) = \xi^{K(\ell)} A(\boldsymbol{A})(\ell),$$

and similarly,

$$A(\widetilde{\boldsymbol{B}})(\ell) = \xi^{K(\ell)} A(\boldsymbol{B})(\ell).$$

We conclude that

$$|A(\widetilde{\boldsymbol{A}})(\ell) + A(\widetilde{\boldsymbol{B}})(\ell)| = |A(\boldsymbol{A})(\ell) + A(\boldsymbol{B})(\ell)|$$

for all $\ell \in \mathbb{Z}$, and the lemma follows. $\qquad\square$

Lemma 14: Let $a, b : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be two generalized Boolean functions and define

$$
\begin{aligned}
\tilde{a}(x_0, \cdots, x_{k-1}) &= a(x_{\sigma(0)}, \cdots, x_{\sigma(k-1)}) \\
\tilde{b}(x_0, \cdots, x_{k-1}) &= b(x_{\sigma(0)}, \cdots, x_{\sigma(k-1)}),
\end{aligned}
$$

where σ is a permutation of $\{0, 1, \cdots, k-1\}$. Then

$$\Phi(\tilde{a}) \star \Phi(\tilde{b}) = \Phi(a) \star \Phi(b).$$

Proof: Write $\boldsymbol{A} = \Phi(a)$, $\widetilde{\boldsymbol{A}} = \Phi(\tilde{a})$, $\boldsymbol{B} = \Phi(b)$, and $\widetilde{\boldsymbol{B}} = \Phi(\tilde{b})$. Let u and u' have binary representation

$$(u_0 \, u_1 \, u_2 \cdots u_{2k-3} \, u_{2k-2})$$

and

$$(u_{2\sigma(0)} \, u_1 \, u_{2\sigma(1)} \cdots u_{2k-3} \, u_{2\sigma(k-1)}),$$

respectively. In an analogous manner we denote the binary representations of v and v'. Let us consider the product $\tilde{A}_u \tilde{A}_v^* = A_{u'} A_{v'}^*$. By the same reasoning as in the proof of Lemma 13 we know that, if $A_{u'} A_{v'}^*$ is nonzero, an SDR of $u' - v'$ is sparse and uniquely determined by $u' - v'$. Similarly an SDR of $u - v$ is sparse if $\tilde{A}_u \tilde{A}_v^*$ is nonzero. It is then clear that, if $A_{u'} A_{v'}^*$ is nonzero, there is a one-to-one correspondence between the sparse SDR of $u - v$ and the sparse SDR of $u' - v'$, and hence, $u - v$ is uniquely determined by the difference $u' - v'$ and does not explicitly depend on u' and v' themselves. This means that each nonzero product appearing in $A(\boldsymbol{A})(u' - v')$ also appears in $A(\widetilde{\boldsymbol{A}})(u - v)$ and, by setting $\sigma := \sigma^{-1}$ and arguing analogously, the converse holds too. Hence $A(\widetilde{\boldsymbol{A}})(u - v) = A(\boldsymbol{A})(u' - v')$ and, similarly, $A(\widetilde{\boldsymbol{B}})(u - v) = A(\boldsymbol{B})(u' - v')$. The lemma follows then immediately. $\qquad\square$

Proof: [Proof of Theorem 10.] Let a set of $m - k$ indices be given by $\{j_0, j_1, \cdots, j_{m-k-1}\}$, where $0 \le j_\alpha < m$ for each $0 \le \alpha < m - k$. Write $\boldsymbol{x} = (x_{j_0} \cdots x_{j_{m-k-1}})$ and let $\boldsymbol{d}$ be a binary word of length $m - k$. First we construct a number of initial function pairs $(\tilde{a}, \tilde{b})$, where

$$\tilde{a}(x_0, \cdots, x_{k-1}) = a(x_{\sigma(0)}, \cdots, x_{\sigma(k-1)}) + L(x_0, \cdots, x_{k-1})$$
$$\tilde{b}(x_0, \cdots, x_{k-1}) = b(x_{\sigma(0)}, \cdots, x_{\sigma(k-1)}) + L(x_0, \cdots, x_{k-1}),$$

σ is a permutation of $\{0, 1, \cdots, k-1\}$, and

$$L(x_0, \cdots, x_{k-1}) = \sum_{\alpha=0}^{k-1} w_\alpha x_\alpha, \qquad w_0, \cdots, w_{k-1} \in \mathbb{Z}_q.$$

From Lemma 12, Lemma 13, and Lemma 14 it follows

$$\Psi(\tilde{a})_{[\boldsymbol{x}=\boldsymbol{d}]} \star \Psi(\tilde{b})_{[\boldsymbol{x}=\boldsymbol{d}]} \le \Phi(a) \star \Phi(b). \tag{8}$$

We can then proceed with the same constructive reasoning as in the proof of Theorem 7 to obtain the function pair $(c^{(m-k)}, d^{(m-k)})$ from $(\tilde{a}, \tilde{b})$. Using (8) and the arguments from the proof of Theorem 7 we conclude

$$\Psi(c^{(m-k)}) \star \Psi(d^{(m-k)}) \le 2^{m-k}[\Phi(a) \star \Phi(b)]. \tag{9}$$

Then $c^{(m-k)}$ is a function where the addition of all linear terms in m variables is possible. Moreover any permutation can be applied to the indices of the m variables in $c^{(m-k)}$, because σ and the indices $j_0, j_1, \cdots, j_{m-k-1}$ can be chosen arbitrarily, so that all $0 \leq j_\alpha < m$ are distinct. We can, therefore, replace j_α by $\pi(k+\alpha)$ for each $\alpha = 0, 1, \cdots, m-k-1$ and set $f = c^{(m-k)} + w$ and $f + \frac{q}{2} x_{\pi(m-1)} = d^{(m-k)} + w$. The theorem follows then from (9) and Theorem 2. $\qquad\square$

We have a useful corollary of Theorem 10.

Corollary 15: Let $m > k$ and suppose $a, b : \mathbb{Z}_2^k \to \mathbb{Z}_q$ are two generalized Boolean functions in k variables. Then each polyphase codeword in the cosets $\psi(f) + \mathrm{RM}_q(1, m)$ with $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ given by

$$f = f(x_0, \cdots, x_{m-1}) = \frac{q}{2} \sum_{\alpha=k}^{m-2} x_{\pi(\alpha)} x_{\pi(\alpha+1)}$$
$$+ a(x_{\pi(0)}, \cdots, x_{\pi(k-1)})(1 - x_{\pi(k)}) + b(x_{\pi(0)}, \cdots, x_{\pi(k-1)}) x_{\pi(k)},$$

where π is a permutation of $\{0, 1, \cdots, m-1\}$, has PMEPR at most $[\Phi(a) \star \Phi(b)]/2^k$ and lies inside $\mathrm{RM}_q(r, m)$ with

$$r = \begin{cases} \max\{\deg(b-a)+1, \deg(a)\} & \text{if} \quad m = k+1 \\ \max\{\deg(b-a)+1, \deg(a), 2\} & \text{if} \quad m > k+1. \end{cases}$$

In particular, if $q \geq 4$ and all coefficients of the monomials in the algebraic normal form of f with degree equal to r are even, then the cosets are contained in $\mathrm{ZRM}_q(r, m)$. The number of distinct functions f, and therefore the number of distinct cosets, is at least $(m-k)!/2$ and at most $m!$ (the true value depends on how many permutations can be applied to the variable indices in the algebraic normal forms of a and b such that distinct pairs (a, b) are generated).

In order to construct families of cosets of $\mathrm{RM}_q(1, m)$ with low PMEPR, we just have to find two generalized Boolean functions $a, b : \mathbb{Z}_2^k \to \mathbb{Z}_q$ with $[\Phi(a) \star \Phi(b)]/2^k$ being low. This can be accomplished by an exhaustive search and, if k is small enough, such functions could even be found using a 'by hand' construction technique. Several examples will be given in Section VI.

IV. Lower Bounds on the PMEPR

Several lower bounds on the maximum PMEPR taken over all the words in a coset of $\mathrm{RM}_q(1, m)$ have been proposed in [23], [34], [16]. These approaches rely on the examination of the OFDM signal at time $\theta = 0$ or at some other θ. Initially this method was proposed in [3]. However existing results apply to second-order cosets of $\mathrm{RM}_q(1, m)$, where the coset representative is binary (in the q-ary context this means that it has values '0' and '$q/2$').

In the following we provide a general lower bound on the PMEPR of the cosets $\psi(f) +$ $\mathrm{RM}_q(1, m)$, where $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ is an arbitrary generalized Boolean function. A q-ary generalization of the Walsh–Hadamard transform plays an important role in our study.

Definition 16: We define the q-ary Walsh–Hadamard transform (WHT) of a generalized Boolean function $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ (or of its polyphase sequence $\Psi(f)$) to be

$$F(\boldsymbol{w}) \triangleq \sum_{\boldsymbol{x} \in \mathbb{Z}_2^m} \xi^{f(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x}},$$

where $\boldsymbol{w} \in \mathbb{Z}_q^m$ and '$\cdot$' denotes the scalar product of vectors. By convention, in the special case where $m = 0$ and $f = a$ ($a \in \mathbb{Z}_q$), we define F to be equal to ξ^a.

If $q = 2$, the above definition coincides with that of the classical WHT (see, e.g., [15, Chapter 14]). The following lemma will be useful in the sequel.

Lemma 17: Let $f, g : \mathbb{Z}_2^m \to \mathbb{Z}_q$ be two generalized Boolean functions that are related by

$$g(\boldsymbol{x}) = f(A\boldsymbol{x}) + \boldsymbol{v} \cdot \boldsymbol{x} + v$$

with $\boldsymbol{v} \in \mathbb{Z}_q^m$, $v \in \mathbb{Z}_q$, and A being an $m \times m$ permutation matrix. Then the sets $\{G(\boldsymbol{w}) \,|\, \boldsymbol{w} \in \mathbb{Z}_q^m\}$ and $\{\xi^v F(\boldsymbol{w}) \,|\, \boldsymbol{w} \in \mathbb{Z}_q^m\}$ are equal.

Proof: Write

$$\begin{aligned}
G(\boldsymbol{w}) &= \sum_{\boldsymbol{x} \in \mathbb{Z}_2^m} \xi^{g(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x}} \\
&= \sum_{\boldsymbol{x} \in \mathbb{Z}_2^m} \xi^{f(A\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x} + \boldsymbol{v} \cdot \boldsymbol{x} + v}.
\end{aligned}$$

Setting $\boldsymbol{x} = A^{-1}\boldsymbol{y}$, we obtain

$$\begin{aligned}
G(\boldsymbol{w}) &= \xi^v \sum_{\boldsymbol{y} \in \mathbb{Z}_2^m} \xi^{f(\boldsymbol{y}) + (\boldsymbol{w} + \boldsymbol{v}) \cdot (A^{-1}\boldsymbol{y})} \\
&= \xi^v \sum_{\boldsymbol{y} \in \mathbb{Z}_2^m} \xi^{f(\boldsymbol{y}) + (A(\boldsymbol{w} + \boldsymbol{v})) \cdot \boldsymbol{y}} \\
&= \xi^v F(A(\boldsymbol{w} + \boldsymbol{v})),
\end{aligned}$$

and the lemma follows. $\qquad\square$

Remark: In the particular case where $q = 2$, Lemma 17 still holds for A being an arbitrary invertible binary matrix (cf. [15, Chapter 14]).

Now we are ready to formulate the following theorem.

Theorem 18: Let $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ be a generalized Boolean function. Then there exists a polyphase codeword in the coset $\psi(f) + \mathrm{RM}_q(1, m)$ having PMEPR at least

$$\frac{1}{2^m} \max_{\boldsymbol{w} \in \mathbb{Z}_q^m} |F(\boldsymbol{w})|^2. \tag{10}$$

Proof: Consider the coset $\psi(f) + \mathrm{RM}_q(1, m)$, which contains the polyphase codewords $\boldsymbol{F}_{\boldsymbol{w}w} = \Psi(f(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x} + w)$, where $\boldsymbol{w} \in \mathbb{Z}_q^m$ and $w \in \mathbb{Z}_q$. The complex envelope corresponding to $\boldsymbol{F}_{\boldsymbol{w}w}$ reads for $\theta = 0$

$$
\begin{aligned}
S(\boldsymbol{F}_{\boldsymbol{w}w})(0) &= \sum_{i=0}^{2^m-1} F_{\boldsymbol{w}w,i} \\
&= \sum_{\boldsymbol{x} \in \mathbb{Z}_2^m} \xi^{f(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x} + w} \\
&= F(\boldsymbol{w})\, \xi^w.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\max_{\boldsymbol{w},w} \sup_{0 \le \theta < 1} |S(\boldsymbol{F}_{\boldsymbol{w}w})(\theta)|^2 &\ge \max_{\boldsymbol{w},w} |S(\boldsymbol{F}_{\boldsymbol{w}w})(0)|^2 \\
&= \max_{\boldsymbol{w},w} |F(\boldsymbol{w})\, \xi^w|^2 \\
&= \max_{\boldsymbol{w}} |F(\boldsymbol{w})|^2,
\end{aligned}
$$

and the theorem follows from (1). $\qquad\square$

Remark: In the case where f is a quadratic Boolean function and $q = 2$, it can be shown that [16, Theorem 1] coincides with the above theorem.

General lower bounds on the expression in (10) directly give lower bounds on the achievable PMEPR of cosets of $\mathrm{RM}_q(1, m)$. Let us first discuss the binary case, i.e., $q = 2$. It is well known (cf., e.g., [15, Chapter 14]) that

$$
\max_{\boldsymbol{w} \in \mathbb{Z}_2^m} |F(\boldsymbol{w})| = 2^m - 2 \min_{\boldsymbol{c} \in \mathrm{RM}_2(1,m)} \mathrm{d}_H\left(\psi(f), \boldsymbol{c}\right),
$$

where $\mathrm{d}_H(\cdot, \cdot)$ is the Hamming distance between two binary sequences. The expression

$$
\rho(m) \triangleq \max_{\boldsymbol{f} \in \mathbb{Z}_2^{2^m}} \min_{\boldsymbol{c} \in \mathrm{RM}_2(1,m)} \mathrm{d}_H\left(\boldsymbol{f}, \boldsymbol{c}\right)
$$

is the *covering radius* of $\mathrm{RM}_2(1, m)$. Therefore, provided that $\rho(m) \le 2^{m-1}$, we have

$$
\min_{f} \max_{\boldsymbol{w} \in \mathbb{Z}_2^m} |F(\boldsymbol{w})| = 2^m - 2\rho(m),
$$

where the minimum is taken over all Boolean functions $f : \mathbb{Z}_2^m \to \mathbb{Z}_2$. Results on the covering radius of $\mathrm{RM}_2(1, m)$ can now be used to lower-bound (10) for $q = 2$ (cf., e.g., [1], [25]), and therefore, to lower bound the PMEPR of cosets of $\mathrm{RM}_2(1, m)$. If m is even, we have $\rho(m) = 2^{m-1} - 2^{m/2-1}$, which leads to the trivial lower bound of 1 for (10). For odd m it is known that $\rho(m) \ge 2^{m-1} - 2^{(m-1)/2}$, where equality holds when $m \le 7$. We conclude that the PMEPR of cosets of $\mathrm{RM}_2(1, m)$ is at least 2 if m is odd and $m \le 7$. For odd $m \ge 9$, the exact

value of $\rho(m)$ is unknown. However, if $m \geq 15$, it is known that $\rho(m) \geq 2^{m-1} - \frac{27}{32}2^{(m-1)/2}$, which implies that (10) can be strictly smaller than 2 in this case.

Much less can be said when $q > 2$. However at least we can assert that the PMEPR of cosets of $\mathrm{RM}_q(1, m)$ must be strictly greater than 1 if q is a multiple of 4. To this end observe

$$\max_{\boldsymbol{w} \in \mathbb{Z}_q^m} |F(\boldsymbol{w})| \geq \max_{\boldsymbol{w} \in (q/4)\mathbb{Z}_4^m} |F(\boldsymbol{w})|. \tag{11}$$

It can be shown that the set $\{F(\boldsymbol{w}) \,|\, \boldsymbol{w} \in \frac{q}{4}\mathbb{Z}_4\}$ is equal to the set of values of the 2^m possible $\{H, N\}^m$ transforms [28] of a function $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$. In [28] it was proved that there is no function $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ for which all 2^m $\{H, N\}^m$ transforms contain only absolute values $2^{m/2}$. Moreover we conjecture that at least for small values of m and for q being a multiple of 4 the right-hand side in (11) is lower-bounded by $2^{(m+1)/2}$, and therefore the PMEPR of cosets of $\mathrm{RM}_q(1, m)$ is at least 2.

In what follows we apply Theorem 18 to obtain lower bounds on the PMEPR of the cosets of $\mathrm{RM}_q(1, m)$ constructed in the previous section.

Theorem 19: Let $a, b : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be two generalized Boolean functions, and let $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ be given by

$$f(x_0, \cdots, x_{m-1}) = \frac{q}{2} \sum_{\alpha=k}^{m-2} x_{\pi(\alpha)} x_{\pi(\alpha+1)}$$
$$+ a(x_{\pi(0)}, \cdots, x_{\pi(k-1)})(1 - x_{\pi(k)}) + b(x_{\pi(0)}, \cdots, x_{\pi(k-1)}) x_{\pi(k)},$$

where π is a permutation of $\{0, 1, \cdots, m-1\}$. Then there exists a polyphase codeword in the coset $\psi(f) + \mathrm{RM}_q(1, m)$ having PMEPR at least

$$\frac{1}{2^{k+2}} \max_{\boldsymbol{w} \in \mathbb{Z}_q^{k+2}} \left| A(\boldsymbol{w}')(1 + \xi^{w_{k+1}}) + B(\boldsymbol{w}')\xi^{w_k}(1 - \xi^{w_{k+1}}) \right|^2 \tag{12}$$

if $m - k$ is even and at least

$$\frac{1}{2^{k+1}} \max_{\boldsymbol{w} \in \mathbb{Z}_q^{k+1}} \left| A(\boldsymbol{w}') + B(\boldsymbol{w}')\, \xi^{w_k} \right|^2 \tag{13}$$

if $m - k$ is odd. Here $A(\boldsymbol{w}')$ and $B(\boldsymbol{w}')$ are the q-ary WHTs of a and b, respectively, and $\boldsymbol{w}' = (w_0 \cdots w_{k-1})$.

Proof: We intend to find a lower bound for

$$\frac{1}{2^m} \max_{\boldsymbol{w} \in \mathbb{Z}_q^m} |F(\boldsymbol{w})|^2,$$

where $F(\boldsymbol{w})$ is the q-ary WHT of f. Using Lemma 17 we conclude that, in order to prove the theorem, it is sufficient to assume that π is the identity permutation. So we are interested in the

coset representative corresponding to

$$f_m(\boldsymbol{x}) = f_m(x_0, \cdots, x_{m-1}) = \frac{q}{2} \sum_{\alpha=k}^{m-2} x_\alpha x_{\alpha+1} + a(x_0, \cdots, x_{k-1})(1 - x_k) + b(x_0, \cdots, x_{k-1})x_k.$$

Denote the q-ary WHT of f_m by $F_m(\boldsymbol{w})$. We require the following expansion

$$F_m(\boldsymbol{w}) = \sum_{\boldsymbol{x} \in \mathbb{Z}_2^m} \xi^{f_m(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x}}$$

$$= \sum_{\boldsymbol{u} \in \mathbb{Z}_2^{m-1}} \xi^{f_m(\boldsymbol{u},0) + \boldsymbol{v} \cdot \boldsymbol{u}} + \xi^{w_{m-1}} \sum_{\boldsymbol{u} \in \mathbb{Z}_2^{m-1}} \xi^{f_m(\boldsymbol{u},1) + \boldsymbol{v} \cdot \boldsymbol{u}}, \qquad (14)$$

where $\boldsymbol{v} = (w_0 \cdots w_{m-2})$. Let us first consider the case $m = k + 1$, so $m - k$ is odd. We have

$$f_{k+1} = a(x_0, \cdots, x_{k-1})(1 - x_k) + b(x_0, \cdots, x_{k-1})x_k,$$

and thus,

$$f_{k+1}(x_0, \cdots, x_{k-1}, 0) = a(x_0, \cdots, x_{k-1})$$

$$f_{k+1}(x_0, \cdots, x_{k-1}, 1) = b(x_0, \cdots, x_{k-1}).$$

With (14) we obtain

$$F_{k+1}(\boldsymbol{w}', w_k) = A(\boldsymbol{w}') + B(\boldsymbol{w}')\xi^{w_k}. \qquad (15)$$

Now let $m > k + 1$. Then we have

$$f_m(x_0, \cdots, x_{m-2}, 0) = f_{m-1}(x_0, \cdots, x_{m-2})$$

$$f_m(x_0, \cdots, x_{m-2}, 1) = f_{m-1}(x_0, \cdots, x_{m-2}) + \frac{q}{2}x_{m-2},$$

and with (14) it follows that

$$F_m(w_0, \cdots, w_{m-1}) = F_{m-1}(w_0, \cdots, w_{m-2}) + F_{m-1}(w_0, \cdots, w_{m-2} + \frac{q}{2})\xi^{w_{m-1}}. \qquad (16)$$

For $m - k$ odd suppose

$$F_m(\boldsymbol{w}', w_k, 0, \cdots, 0) = 2^{(m-k-1)/2}(A(\boldsymbol{w}') + B(\boldsymbol{w}')\xi^{w_k}), \qquad (17)$$

which is true for $m = k + 1$ (see (15)). We will use this expression as a hypothesis for the following induction. We employ (16) to obtain

$$F_{m+1}(\boldsymbol{w}', w_k, 0, \cdots, 0, w_m) = 2^{(m-k-1)/2}\left[A(\boldsymbol{w}')(1 + \xi^{w_m}) + B(\boldsymbol{w}')\xi^{w_k}(1 - \xi^{w_m})\right] \qquad (18)$$

and

$$F_{m+2}(\boldsymbol{w}', w_k, 0, \cdots, 0, w_m, w_{m+1}) = 2^{(m-k-1)/2}\left[A(\boldsymbol{w}')(1 + \xi^{w_m}) + B(\boldsymbol{w}')\xi^{w_k}(1 - \xi^{w_m})\right.$$

$$\left. + \xi^{w_{m+1}}\left(A(\boldsymbol{w}')(1 - \xi^{w_m}) + B(\boldsymbol{w}')\xi^{w_k}(1 + \xi^{w_m})\right)\right].$$

Consequently we have

$$F_{m+2}(\boldsymbol{w}', w_k, 0, \cdots, 0) = 2^{(m-k+1)/2}\left(A(\boldsymbol{w}') + B(\boldsymbol{w}')\xi^{w_k}\right),$$

which proves by induction that (17) and (18) hold in general. Now we can write for $m - k$ even

$$\frac{1}{2^m}\max_{\boldsymbol{w}\in\mathbb{Z}_q^m}\left|F_m(\boldsymbol{w})\right|^2 \geq \frac{1}{2^m}\max_{\boldsymbol{w}\in\mathbb{Z}_q^{k+2}}\left|F_m(\boldsymbol{w}', w_k, 0, \cdots, 0, w_{k+1})\right|^2$$

$$= \frac{1}{2^{k+2}}\max_{\boldsymbol{w}\in\mathbb{Z}_q^{k+2}}\left|A(\boldsymbol{w}')(1 + \xi^{w_{k+1}}) + B(\boldsymbol{w}')\xi^{w_k}(1 - \xi^{w_{k+1}})\right|^2$$

and for $m - k$ odd

$$\frac{1}{2^m}\max_{\boldsymbol{w}\in\mathbb{Z}_q^m}\left|F_m(\boldsymbol{w})\right|^2 \geq \frac{1}{2^m}\max_{\boldsymbol{w}\in\mathbb{Z}_q^{k+1}}\left|F_m(\boldsymbol{w}', w_k, 0, \cdots, 0)\right|^2$$

$$= \frac{1}{2^{k+1}}\max_{\boldsymbol{w}\in\mathbb{Z}_q^{k+1}}\left|A(\boldsymbol{w}') + B(\boldsymbol{w}')\xi^{w_k}\right|^2.$$

Then the statements in the theorem follow from Theorem 18. $\qquad\square$

V. THE PAPR UNDER THE WALSH–HADAMARD TRANSFORM

Suppose that $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ is a generalized Boolean function and p is a divisor of q. Then we define the *peak-to-average power ratio* (PAPR) of $\Psi(f)$ under the p-ary WHT to be

$$\mathrm{PAPR}_p(\Psi(f)) \triangleq \frac{1}{2^m}\max_{\boldsymbol{w}\in(q/p)\mathbb{Z}_p^m}|F(\boldsymbol{w})|^2.$$

If p is omitted, we shall refer to the PAPR of $\Psi(f)$ under the classical (2-ary) WHT. It is apparent that

$$\mathrm{PAPR}_p(\Psi(f)) \leq \mathrm{PAPR}_q(\Psi(f)).$$

We remark that this measure in fact arises from a more general definition of the PAPR under unitary transforms (cf. [21], [28]). Also the PMEPR can be restated in this context [21]. The PAPR under the p-ary WHT is an important measure in cryptographic applications [22], and it is of interest in MC-CDMA communications systems [25], particularly when $p = 2$.

OFDM and MC-CDMA enjoy several similarities: in both cases codewords are used to modulate simultaneously a number of orthogonal signals, which are continuous in OFDM and discrete in MC-CDMA. In MC-CDMA systems the PAPR (under the classical WHT) turns out to be an analogous measure of the PMEPR in OFDM systems [24]. In [24] and [25] Paterson studied binary codes whose codewords have low PAPR, whereas constructing nonbinary codes is left as an open problem.

Theorem 18 relates the PMEPR of cosets of $\mathrm{RM}_q(1, m)$ with the q-ary WHT of its coset representative. It is a consequence of Lemma 17 that every word in a coset of $\mathrm{RM}_q(1, m)$ has

the same PAPR under the q-ary WHT (but generally not under the p-ary WHT). Using this fact, Theorem 18 states that cosets with low PMEPR must have also low PAPR under the q-ary WHT, and therefore, also under the p-ary WHT. Thus we have established the following corollary.

Corollary 20: Let $f : \mathbb{Z}_2^m \rightarrow \mathbb{Z}_q$ be a generalized Boolean function, and suppose that p divides q. Then the PAPR under the p-ary WHT of the polyphase codewords in the coset $\psi(f) + \mathrm{RM}_q(1, m)$ is at most the PMEPR of the polyphase codewords in the coset.

The above corollary shows that all codes with low PMEPR that are unions of cosets of $\mathrm{RM}_q(1, m)$ (e.g., codes that arise from the results in the present paper and the codes constructed in [4] and [23]) also enjoy low PAPR. Hence with Corollary 20 a number of (generally nonbinary) coding options for MC-CDMA with low PAPR can be derived. However, as opposed to many of the codes in [24] and [25] having PAPR equal to 1, Corollary 20 in combination with known results (e.g., Corollary 15) yields an upper bound on the PAPR that is at least 2.

We refer to the final section of [25] for further discussions on the connection between codes for OFDM and MC-CDMA.

VI. CONSTRUCTION EXAMPLES

In the following we will apply Corollary 15 in order to find families of cosets of $\mathrm{RM}_q(1, m)$ with low PMEPR. By Corollary 20 such cosets also have low PAPR under the p-ary WHT, where p is a divisor of q. Trivially we may choose our kernel functions a and b from the complete set of generalized Boolean functions in k variables. In such a way, we can construct large sequence families with their PMEPR bounded by 2^{k+1}, since the trivial bound states

$$\frac{\Phi(a) \star \Phi(b)}{2^k} \leq 2^{k+1}.$$

However it can be shown that these sequence sets are just subsets of a larger family of complementary sets of size 2^{k+1} as constructed in [32] and [23]. Hence this approach does not yield any new sequence sets.

A. Sequence Sets with PMEPR at most 2

An immediate consequence of Corollary 15 (and in this case also of Theorem 7) is the well-known construction of Golay sequences over $\mathbb{Z}_q$, which has been stated by Davis and Jedwab in [4] for q being a power of 2.

Corollary 21: Each of the $m!/2$ cosets of $\mathrm{RM}_q(1, m)$ with coset representatives corresponding to

$$\frac{q}{2} \sum_{i=0}^{m-2} x_{\pi(i)} x_{\pi(i+1)}, \tag{19}$$

where π is a permutation of $\{0, 1, \cdots, m-1\}$, is entirely comprised of Golay sequences. In particular the maximum PMEPR of the polyphase words in these cosets is (i) exactly 2 if m is odd and q arbitrary (even) and if m is even and $q \equiv 0 \pmod 4$, (ii) at least $1 + \cos^2 \frac{\pi}{q}$ and at most 2 if m is even and $q \equiv 2 \pmod 4$.

Proof: The corollary follows by setting $k = 0$, $a = 0$, and $b = 0$ in Corollary 15 and Theorem 19. Since $\Phi(a) \star \Phi(b) = 2$, the upper bound on the PMEPR is an immediate consequence of Corollary 15. Note that a and b have q-ary WHT equal to 1. For odd m the lower bound on the PMEPR follows then immediately from (13). For even m and $q \equiv 0 \pmod 4$ put

$$\boldsymbol{w} = \left(\frac{q}{4}, \frac{q}{4} \right)$$

in the maximum calculation in (12), which leads to a lower bound of 2. For even m and $q \equiv 2$ $\pmod 4$ it is not hard to see that the maximum in (12) is attained for

$$\boldsymbol{w} = \left(\frac{q-2}{4}, \frac{q+2}{4} \right).$$

Hence, with $j = \sqrt{-1}$, (12) becomes

$$
\frac{1}{4} \max_{\boldsymbol{w} \in \mathbb{Z}_q^2} \left| (1 + \xi^{w_1}) + \xi^{w_0} (1 - \xi^{w_1}) \right|^2 = \frac{1}{4} \left| (1 + j e^{j\pi/q}) + j e^{-j\pi/q} (1 - j e^{j\pi/q}) \right|^2
$$
$$
= \frac{1}{4} \left| 2 + 2j \Re \left\{ e^{j\pi/q} \right\} \right|^2
$$
$$
= 1 + \cos^2 \frac{\pi}{q},
$$

which completes the proof. □

We remark that the upper bound is well known: it was first proved by Davis and Jedwab for q being a power of 2 [4, Corollary 6, Corollary 9] and generalized to arbitrary q by Paterson [23, Corollary 11]. The lower bound in Corollary 21 (i) has been first proved Cammarano and Walker [3], and the case when m is odd also by Paterson [23, Theorem 21]. Except for the trivial bound 1 when $q = 2$, the lower bound in Corollary 21 (ii) is new. Note that these upper and lower bounds on the PMEPR now arise in a uniform way from a general framework. From the discussion following Theorem 18 we also conclude that the considered cosets are optimal at least for $q = 2$ and odd $m \le 7$ in the sense that for these parameters there cannot exist cosets with PMEPR lower than 2.

Li and Chu [14] reported 1024 additional quaternary (with elements in $\{1, -1, \sqrt{-1}, -\sqrt{-1}\}$) Golay sequences of length 16. While Fiedler and Jedwab [5] recently provided an explanation for these sequences, it has been observed earlier by Holzmann and Kharaghani [10] that there exists essentially one quaternary kernel of length 8

$$
((+ + + - + + - +), (+ j \, j - + i \, i-)), \qquad j = \sqrt{-1}, i = -j.
$$

These two sequences are associated with generalized Boolean functions $\mathbb{Z}_2^3 \to \mathbb{Z}_4$ corresponding to the forms

$$2x_0x_1 + 2x_1x_2 \qquad \text{and} \qquad 2x_0x_2 + 2x_1x_2 + x_0 + x_1,$$

respectively. Following [10] from this kernel 512 equivalent kernels can be generated by applying several operations that preserve the complementary property. These kernels can be used in conjunction with Theorem 7 to obtain an explicit construction for Golay sequences of length 2^m, where $m > 3$. Notice that all of them correspond to cubic forms. Indeed, if $m = 4$, we obtain exactly the 1024 Golay sequences of length 16 reported by Li and Chu [14]. Unfortunately, in this way, we cannot construct complete cosets of $\mathrm{RM}_q(1, m)$ whose PMEPR is bounded by 2. Instead Corollary 15 reveals that the PMEPR is at most 5 if the 512 kernels are used in conjunction with Corollary 15 to build cosets of $\mathrm{RM}_q(1, m)$.

B. Sequence Sets with Maximum PMEPR between 2 and 4

Consider the functions $a, b : \mathbb{Z}_2^2 \to \mathbb{Z}_q$ with

$$a = a(x_0, x_1) = \frac{q}{2}x_0x_1$$

$$b = b(x_0, x_1) = \frac{q}{2}x_0x_1 + \left(\alpha + \frac{q}{2}\right)x_0 + \beta x_1,$$

where $\alpha, \beta \in (q/p)\mathbb{Z}_p$ and p divides q. We have

$$\Phi(a) = (1 \;\; 1 \;\; 0 \;\; 0 \;\; 1 \;\; -1)$$

$$\Phi(b) = (1 \;\; -\xi^\alpha \;\; 0 \;\; 0 \;\; \xi^\beta \;\; \xi^{\alpha+\beta})$$

and

$$(A(\Phi(a))(\ell) \,|\, 0 \le \ell < 6) = (4 \;\; 0 \;\; 0 \;\; 1 \;\; 0 \;\; -1)$$

$$(A(\Phi(b))(\ell) \,|\, 0 \le \ell < 6) = (4 \;\; 0 \;\; 0 \;\; -\xi^{\beta-\alpha} \;\; 0 \;\; \xi^{\beta+\alpha}).$$

Hence

$$\Phi(a) \star \Phi(b) = 2(4 + |1 - \xi^{\beta-\alpha}| + |1 - \xi^{\beta+\alpha}|) \le 16.$$

So, by Corollary 15, the PMEPR of the cosets of $\mathrm{RM}_q(1, m)$ $(m > 2)$ with coset representatives corresponding to

$$\frac{q}{2}\sum_{i=0}^{m-2} x_{\pi(i)}x_{\pi(i+1)} + \alpha x_{\pi(0)}x_{\pi(2)} + \beta x_{\pi(1)}x_{\pi(2)}, \tag{20}$$

is at most 4. These cosets lie inside $\mathrm{RM}_q(2, m)$ and in particular inside $\mathrm{ZRM}_q(2, m)$ if $q > p$.

There are p^2 choices for the pair α, β and for each pair we can construct $m!/2$ distinct cosets by applying to the variable indices $m!/2$ permutations that are invariant under reversal. However

setting $\alpha = \beta = 0$ or $\alpha = \beta = q/2$ in (20) will generate the same set of coset representatives. Hence the total number of distinct coset representatives amounts to $(p^2 - 1)(m!/2)$.

For particular values for α and β the upper bound on the PMEPR of the cosets can be tightened. If $\alpha = \beta = 0$ or $\alpha = \beta = q/2$, we obtain sequences with PMEPR at most 2. However in this case the coset representatives are of the form given in (19). More generally, the $(p^2 - 1)(m!/2)$ cosets of $\mathrm{RM}_q(1, m)$ organize in classes with the same PMEPR upper bound. For $p = 2$ we obtain just two classes: $m!/2$ cosets with maximum PMEPR 2 and $m!$ cosets with PMEPR at most 4. Using Theorem 19 it can be shown that the latter bound is tight if $q \equiv 0 \pmod{4}$.

For $q \geq 4$ and $p = 4$ we have $\alpha, \beta \in \left\{ 0, \frac{q}{4}, \frac{q}{2}, \frac{3q}{4} \right\}$ and the classes are:

$\beta + \alpha$	$\beta - \alpha$	max. PMEPR	# cosets
0	0	2	$m!/2$
arbitrary (0)	0 (arbitrary)	3	$2m!$
$\frac{q}{4} \pmod{\frac{q}{2}}$	$\frac{q}{4} \pmod{\frac{q}{2}}$	$2 + \sqrt{2}$	$4m!$
arbitrary	arbitrary	4	$m!$

Notice that the number of cosets are counted by considering only those cosets that are not contained in a class with a lower PMEPR bound. For example the total number of cosets with PMEPR at most 3 is equal to $5m!/2$, while there are $13m!/2$ cosets with PMEPR at most $(2 + \sqrt{2})$. Employing Theorem 19 it can be shown that the PMEPR bounds given above are tight if $q \equiv 0 \pmod{8}$. If $q = 4$, Theorem 19 yields a lower bound of 2 for the first class, 2.5 for the second and third class, and 4 for the fourth class.

We remark that the second class includes those 48 cosets of $\mathrm{RM}_8(1, 4)$ inside $\mathrm{ZRM}_8(2, 4)$ for which Davis and Jedwab observed that their PMEPR is exactly 3 [4, Tables V and VI]. While Nieswand and Wagner [17] could explain this behaviour partially, a complete proof arises as a special case of Corollary 15 and Theorem 19[1]. Moreover we identified further cosets of $\mathrm{RM}_q(1, m)$ with PMEPR at most 3.

For $q \geq 8$ and $p = 8$ there are already 9 classes with different PMEPR upper bounds. These are:

---

[1]The upper bound for this particular case has been proved alternatively by the author and Finger in [32].

max. PMEPR			# cosets
		2	$m!/2$
$2 + 1/\sqrt{2}$	$\approx$	2.707	$4m!$
$2 + \sqrt{2 - \sqrt{2}}$	$\approx$	2.765	$4m!$
		3	$2m!$
$2 + \sqrt{1 + 1/\sqrt{2}}$	$\approx$	3.307	$8m!$
$2 + \sqrt{2}$	$\approx$	3.414	$4m!$
$3 + 1/\sqrt{2}$	$\approx$	3.707	$4m!$
$2 + \sqrt{2 + \sqrt{2}}$	$\approx$	3.848	$4m!$
		4	$m!$

Applying Theorem 19 we are able to show that the PMEPR bounds are tight if $q \equiv 0 \pmod{16}$.

The number of suitable kernels of length 4 producing cosets of $\mathrm{RM}_q(1, m)$ with maximum PMEPR between 2 and 4 can be increased further. To this end, consider the pairs (a, b), where $a, b : \mathbb{Z}_2^2 \rightarrow \mathbb{Z}_q$ are given by

$$a = a(x_0, x_1) = \gamma x_0 x_1$$
$$b = b(x_0, x_1) = \delta x_0 x_1 + \left(\alpha + \frac{q}{2} \right) x_0 + \beta x_1,$$

$\alpha, \beta, \gamma, \delta \in \mathbb{Z}_q$, and γ and δ are selected such that $\Phi(a) \star \Phi(b) \leq 16$. However notice that the coset representatives of the constructed cosets of $\mathrm{RM}_q(1, m)$ generally correspond to cubic forms.

It is noteworthy that a computer-based search failed to find kernels that generate more cosets of $\mathrm{RM}_q(1, m)$ with PMEPR less than 4. Our search included all kernel functions in k variables for $k = 4$ and $q = 2$, $k = 3$ and $q = 4$, and quadratic functions with $k = 5$ and $q = 2$. We leave the identification of more cosets of $\mathrm{RM}_q(1, m)$ with low PMEPR (preferably close to 2) to further work.

VII. CONCLUDING REMARKS

In this paper we have shown how cosets of $\mathrm{RM}_q(1, m)$ with low PMEPR can be obtained in a systematic way once suitable kernels are known. The PMEPR of these cosets is low because such cosets are comprised of sequences lying in so-called near-complementary pairs. It was demonstrated that those cosets of $\mathrm{RM}_q(1, m)$ comprised entirely of Golay sequences simply arise as a special case in a more general theory. We presented some suitable kernels and, in this way, several previously unexplained phenomena arising in earlier works by Davis and Jedwab [4], by Paterson [23], and by Parker and Tellambura [19], [20] can now be understood in a general framework. Moreover we provided at least a partial answer to the question stated in

[23]: *What other regularities appear in the PMEPRs of cosets as we move to higher alphabets, and how can they be explained in general?* We have also established a connection between the OFDM and MC-CDMA coding problems, and suggested (generally nonbinary) coding solutions for MC-CDMA with low PAPR.

We wish to point out some relations to the work by Parker and Tellambura [19], [20]. Indeed Theorem 7 in the present paper is similar to [20, Theorem 5] and [19, Theorem 6], which also exploit the Rudin–Shapiro construction. However these references do not identify the crucial connection between the PMEPR and the aperiodic auto-correlation, which allows us to extend the PMEPR bound to a slightly weaker bound in Theorem 10 and Corollary 15 that holds for complete cosets of $\mathrm{RM}_q(1, m)$. In contrast, in order to obtain cosets of $\mathrm{RM}_2(1, m)$ with low PMEPR, the approach in [19] and [20] involves a computational search for the maximum PMEPR over a number of kernels. Indeed in [20, Table 6] it is implicitly shown that the cosets with coset representatives given in (20) have PMEPR at most 4 when $q = 2$. Notice that their semi-computationally obtained upper bound is exactly as predicted by Corollary 15.

We presented some kernels of short length and we also sought good kernels of medium length. From the complexity point of view it is also realistic to perform an exhaustive search for good kernels of larger lengths, say 32 or 64, and defined over larger alphabets. We have not attempted such a search and leave it to further work. It would be desirable to find kernels producing cosets of $\mathrm{RM}_q(1, m)$ with PMEPR close to 2, or even with PMEPR of exactly 2 (which implies the discovery of new complementary pairs). Moreover our theory would benefit from having an efficient way to construct good kernels.

References

[1] R. Brualdi, S. Litsyn, and V. Pless, "Covering radius," in *Handbook of Coding Theory*, V. S. Pless and W. C. Huffman, Eds. Elsevier, 1998.

[2] S. Z. Budisin, "New complementary pairs of sequences," *IEE Electron. Lett.*, vol. 26, no. 13, pp. 881–883, Jun. 1990.

[3] M. W. Cammarano and M. L. Walker, "Integer maxima in power envelopes of Golay codewords," 1999. [Online]. Available: http://www.mathcs.richmond.edu/~jad/summer.html

[4] J. A. Davis and J. Jedwab, "Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed–Muller codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 7, pp. 2397–2417, Nov. 1999.

[5] F. Fiedler and J. Jedwab, "How do more Golay sequences arise?" *submitted to IEEE Trans. Inform. Theory*, 2005.

[6] M. J. E. Golay, "Complementary series," *IRE Trans. Inform. Theory*, vol. 7, no. 2, pp. 82–87, Apr. 1961.

[7] ——, "Sieves for low autocorrelation binary sequences," *IEEE Trans. Inform. Theory*, vol. 23, no. 1, pp. 43–51, Jan. 1977.

[8] A. J. Grant and R. D. van Nee, "Efficient maximum-likelihood decoding of Q-ary modulated Reed–Muller codes," *IEEE Commun. Lett.*, vol. 2, no. 5, pp. 134–136, May 1998.

[9] U. Güntzer and M. Paul, "Jump interpolation search trees and symmetric binary numbers," *Inf. Process. Lett.*, vol. 26, no. 4, pp. 193–204, Dec. 1987.

[10] W. H. Holzmann and H. Kharaghani, "A computer search for complex Golay sequences," *Australasian Journal of Combinatorics*, vol. 10, pp. 251–258, 1994.

[11] J. Jedwab and C. J. Mitchell, "Minimum weight modified signed-digit representations and fast exponentation," *IEE Electron. Lett.*, vol. 25, no. 17, pp. 1171–1172, Aug. 1989.

[12] A. E. Jones and T. A. Wilkinson, "Combined coding for error control and increased robustness to system nonlinearities in OFDM," *Proc. of IEEE 46th Vehicular Technology Conf. (VTC)*, Apr. 1996.

[13] A. E. Jones, T. A. Wilkinson, and S. K. Barton, "Block coding scheme for reduction of peak to mean envelope power ratio of multicarrier transmission schemes," *IEE Electron. Lett.*, vol. 30, no. 25, pp. 2098–2099, Dec. 1994.

[14] Y. Li and W. B. Chu, "More Golay sequences," *IEEE Trans. Inform. Theory*, vol. 51, no. 3, pp. 1141–1145, Mar. 2005.

[15] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. North Holland Mathematical Library, 1977.

[16] K. Manji and B. S. Rajan, "On the PAPR of binary Reed–Muller OFDM codes," *Proc. of IEEE Int. Symp. Inform. Theory*, p. 423, Jul. 2004.

[17] K. M. Nieswand and K. N. Wagner, "Octary codewords with power envelopes of $3 * 2^m$," 1998. [Online]. Available: http://www.mathcs.richmond.edu/~jad/summer.html

[18] H. Ochiai and H. Imai, "Block coding scheme based on complementary sequences for multicarrier signals," *IEICE Transactions on Fundamentals*, vol. E80-A, pp. 2136–2146, 1997.

[19] M. G. Parker and C. Tellambura, "Generalised Rudin–Shapiro constructions," *Proc. of Workshop on Coding and Cryptography (WCC), Paris, France*, 2001.

[20] ——, "Golay–Davis–Jedwab complementary sequences and Rudin–Shapiro constructions," 2001. [Online]. Available: http://www.ii.uib.no/~matthew/ConstaBent2.pdf

[21] ——, "A construction for binary sequence sets with low peak-to-average power ratio," *Report No. 242, Department of Informatics, University of Bergen, Norway*, Feb. 2003. [Online]. Available: http://www.ii.uib.no/~matthew/ConstructReport.pdf

[22] M. G. Parker, "Generalised S-box nonlinearity," *NESSIE Public Document – NES/DOC/UIB/WP5/020/A*, 2003. [Online]. Available: https://www.cosic.esat.kuleuven.ac.be/nessie/reports/phase2/SBoxLin.pdf

[23] K. G. Paterson, "Generalized Reed–Muller codes and power control in OFDM modulation," *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 104–120, Jan. 2000.

[24] ——, "Sequences for OFDM and multi-code CDMA: Two problems in algebraic coding theory," *Proc. of SETA 2001, Discrete Mathematics and Theoretical Computer Science Series, Springer*, pp. 46–71, 2002.

[25] ——, "On codes with low peak-to-average power ratio for multicode CDMA," *IEEE Trans. Inform. Theory*, vol. 50, no. 3, pp. 550–559, Mar. 2004.

[26] K. G. Paterson and A. E. Jones, "Efficient decoding algorithms for generalized Reed–Muller codes," *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1272–1285, Aug. 2000.

[27] B. M. Popović, "Synthesis of power efficient multitone signals with flat amplitude spectrum," *IEEE Trans. Commun.*, vol. 39, no. 7, pp. 1031–1033, Jul. 1991.

[28] C. Riera and M. G. Parker, "Generalized bent criteria for Boolean functions," *to appear in IEEE Trans. Inform. Theory*, 2006.

[29] W. Rudin, "Some theorems on Fourier coefficients," *Proc. Amer. Math. Soc.*, vol. 10, pp. 855–859, 1959.

[30] K.-U. Schmidt, "Comment on 'Efficient maximum-likelihood decoding of Q-ary modulated Reed–Muller codes'," *IEEE Commun. Lett.*, vol. 10, no. 1, p. 52, Jan. 2006.

[31] K.-U. Schmidt and A. Finger, "Simple maximum-likelihood decoding of generalized first-order Reed–Muller codes," *IEEE Commun. Lett.*, vol. 9, no. 10, pp. 912–914, Oct. 2005.

[32] ——, "Constructions of complementary sequences for power-controlled OFDM transmission," *Proc. of Workshop on Coding and Cryptography (WCC) 2005, Bergen, Norway, to appear in Lecture Notes in Computer Science*, 2006.

[33] H. S. Shapiro, "Extremal problems for polynomials and power series," *Master Thesis, MIT*, 1951.

[34] T. E. Stinchcombe, "Aperiodic autocorrelations of length 2^m sequences, complementarity, and power

control for OFDM," Ph.D. dissertation, University of London, Apr. 2000. [Online]. Available: http://www.isg.rhul.ac.uk/alumni/thesis/stinchcombe_t.pdf

[35] C. Tellambura, "Upper bound on the peak factor of n-multiple carriers," *IEE Electron. Lett.*, vol. 33, no. 19, pp. 1608–1609, Sep. 1997.

[36] R. D. J. van Nee, "OFDM codes for peak-to-average power reduction and error correction," *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, Nov. 1996.

[37] T. A. Wilkinson and A. E. Jones, "Minimisation of the peak to mean envelope power ratio of multicarrier transmission schemes by block coding," *Proc. of IEEE Vehicular Technology Conference (VTC)*, Jul. 1995.