Constructions of Complementary Sequences for Power-Controlled OFDM Transmission

Kai-Uwe Schmidt and Adolf Finger

Communications Laboratory
Technische Universität Dresden
01062 Dresden, Germany
schmidtk@ifn.et.tu-dresden.de

Abstract. We present constructions of polyphase sequences suitable for the use as codewords in orthogonal frequency-division multiplexing (OFDM) with strictly bounded peak-to-mean envelope power ratio (PMEPR). Our first construction establishes that each polyphase sequence of length 2^m lies in a complementary set, whose size depends on a special property of its associated generalized Boolean function. Thus we identify a large family of sequences with PMEPR at most 2^{k+1}, where k is a non-negative integer. Our second construction yields sequences that lie in so-called almost complementary pairs and have PMEPR at most 3. A number of coding schemes for OFDM with low PMEPR is then presented. These schemes extend and complement previously proposed coding options.

1 Introduction

Davis and Jedwab [1] established a link between Golay's complementary sequences [2] and certain second-order cosets of a generalized first-order Reed–Muller code. The union of these cosets yields a powerful code, which can be used to perform error correction and ensures a PMEPR not exceeding 2. Its main disadvantage is that the code rate rapidly decreases for larger block lengths. Therefore Davis and Jedwab proposed to include further cosets in order to increase the rate of the codes at the cost of a slightly higher PMEPR [1]. This raised the problem of finding explicit constructions for such cosets. Paterson [3] provided a construction for further second-order cosets comprising sequences lying in so-called complementary sets of size 2^{k+1} ($k > 0$), and thus, the resulting codes have PMEPR at most 2^{k+1}. Parker and Tellambura [4] proposed an elaborate method to construct higher-order cosets comprised of complementary sets. However their construction suffers from the lack of efficient encoding and decoding algorithms. We propose a construction for complementary sets of a given size lying in cosets of a given order. Our construction includes previous constructions in [1] and [3] as special cases. In a straightforward way we then obtain a wide range of coding options for OFDM with low PMEPR. In addition we address the problem of constructing sequences that have low PMEPR and do not necessarily lie in complementary sets. We present a construction for cosets comprised

of so-called almost complementary pairs and establish that their PMEPR is at most 3. In this way we prove a conjecture by Davis and Jedwab [1] and Nieswand and Wagner [5] and identify more cosets with PMEPR at most 3. These results further extend possible coding options for OFDM with low PMEPR.

2 Notation and Background

2.1 Problem Statement

We will study an OFDM system with n subcarriers. Let $A = (A_0\, A_1 \cdots A_{n-1})$ be a polyphase codeword of length n that is used to modulate the subcarriers. Its corresponding OFDM signal can be mathematically described by

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

where λ is a positive constant. An important characteristic of such a signal (or of the modulating codeword) is its PMEPR, which is defined to be

$$\mathrm{PMEPR}(A) := \frac{1}{n} \sup_{0\le\theta<1} |S(A)(\theta)|^2.$$

Due to engineering reasons there is a high motivation to keep the PMEPR of the transmitted OFDM signals low. A particular elegant solution to solve this power-control issue is to use a special OFDM block code across the subcarriers [6]. By defining the PMEPR for such a code $\mathcal{C}$ to be

$$\mathrm{PMEPR}(\mathcal{C}) := \max_{A\in\mathcal{C}} \mathrm{PMEPR}(A).$$

we can formulate the problem as follows. *Find codes with high code rates and high minimum distances for which the above-defined value is small.*

2.2 Complementary Sequences for OFDM

We begin with recalling and extending some well-known relations between the aperiodic auto-correlation and the PMEPR of a sequence (cf. e.g. [7], [1], [3]). Let $A = (A_0\, A_1 \cdots A_{n-1})$ and $B = (B_0\, B_1 \cdots B_{n-1})$ be two complex-valued sequences. Then the *aperiodic cross-correlation* of A and B at a displacement $\ell \in \mathbb{Z}$ is given by

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

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

$$A(A)(\ell) := C(A, A)(\ell).$$

In the sequel the following lemma will be essential for the construction of sequence sets for OFDM with low PMEPR.

Lemma 1. *Suppose a set of N polyphase sequences of length n is given by $\{A^0\, A^1 \cdots A^{N-1}\}$. Then the PMEPR of each individual sequence in the set is at most*

$$N + \frac{2}{n} \sum_{\ell=1}^{n-1} \left| \sum_{i=0}^{N-1} A(A^i)(\ell) \right|.$$

In particular, if the set is a complementary set, each sequence has PMEPR at most N [3].

Proof. It is well known (cf. e.g. [7], [1]) and straightforward to show that

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

Hence

$$\begin{aligned} \sum_{i=0}^{N-1} |S(A^i)(\theta)|^2 &= \sum_{i=0}^{N-1} \left(A(A^i)(0) + 2 \sum_{\ell=1}^{n-1} \Re\{A(A^i)(\ell)\, e^{\sqrt{-1}\, 2\pi\ell\theta}\} \right) \\ &= Nn + 2 \sum_{\ell=1}^{n-1} \Re \left\{ \sum_{i=0}^{N-1} A(A^i)(\ell)\, e^{\sqrt{-1}\, 2\pi\ell\theta} \right\} \\ &\le Nn + 2 \sum_{\ell=1}^{n-1} \left| \sum_{i=0}^{N-1} A(A^i)(\ell) \right|, \end{aligned}$$

where we used the fact that $A(A)(0) = n$ for polyphase sequences. The lemma follows then with the definition of the PMEPR. □

The above result motivates the construction of sequences lying in sets of sequences of small size, where the sum of the aperiodic auto-correlation sidelobes of all sequences in the set is small for all nonzero shifts. In this paper we shall particularly study two types of such sequences sets. The first one are the so-called complementary sets, which are defined as follows.

Definition 2. *A set of N sequences is called a* complementary set of size N *if the aperiodic auto-correlations of its members sum up to zero except for the zero displacement. If $N = 2$, the two sequences are commonly termed a* Golay complementary pair *[2].*

By Lemma 1 the PMEPR of each polyphase sequence lying in a complementary set of size N is at most N. A construction for such sequence sets will be established in Section 3.

Definition 3. *A pair of sequences is called an* almost complementary pair *if the sum of the aperiodic auto-correlations of its members is zero except for the zero shift and for at most two more shifts (τ and $-\tau \mid 1 \le \tau < n$).*

It follows from Lemma 1 that the PMEPR of a polyphase sequence lying in an almost complementary pair is upper-bounded by 6. A tighter upper bound on the PMEPR can be obtained by taking the height of the out-of-phase peak in the auto-correlation sum into account. In Section 4 we shall construct almost complementary pairs comprising sequences with PMEPR at most 3.

2.3 Generalized Boolean Functions and Associated Sequences

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

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

where the weights $c_0, \ldots c_{2^m-1}$ are in $\mathbb{Z}_q$, and $(i_0\, i_1 \ldots i_{m-1})$ is the binary expansion of $0 \le i < 2^m$, such that $i = \sum_{j=0}^{m-1} i_j 2^j$. 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 $(f_0\, f_1 \cdots f_{2^m-1})$ as the *$\mathbb{Z}_q$-valued sequence associated with* f and the sequence $(\xi^{f_0}\, \xi^{f_1} \cdots \xi^{f_{2^m-1}})$ 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 expansion of the integer $0 \le i < 2^m$.

In the remainder of this subsection we recall the technique of the restriction of polyphase sequences of length 2^m and its application to the expansion of correlations of sequences. For details see [3] and [8].

Definition 4. *[3] Let $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ be a generalized Boolean function in the variables $x_0, x_1, \cdots, x_{m-1}$, and let F be its associated polyphase sequence. Suppose*

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

is a list of k indices and write $x = (x_{j_0}\, x_{j_1}\, \cdots\, x_{j_{k-1}})$. We shall call the entries of x the restricting variables. *Let $d = (d_0\, d_1\, \cdots\, d_{k-1}) \in \mathbb{Z}_2^k$, and let $(i_0\, i_1\, \cdots\, i_{m-1})$ be the binary expansion of the integer i. Then the* restricted sequence $F|_{x=d}$ *is a sequence of length 2^m with its elements $(F|_{x=d})_i$ being defined as*

$$(F|_{x=d})_i := \begin{cases} F_i & \text{if} \quad (i_{j_0}\, i_{j_1} \cdots i_{j_{k-1}}) = (d_0\, d_1 \cdots d_{k-1}) \\ 0 & \text{if} \quad (i_{j_0}\, i_{j_1} \cdots i_{j_{k-1}}) \ne (d_0\, d_1 \cdots d_{k-1}) \end{cases},$$

where $i = 0, 1, \cdots, 2^m - 1$. For the case $k = 0$ we fix $F|_{x=d} = F$.

Following [3] it is a consequence of the above definition that

$$F = \sum_{d \in \mathbb{Z}_2^k} F|_{x=d}. \tag{1}$$

A sequence that is restricted in k variables comprises $2^m - 2^{m-k}$ zero entries and 2^{m-k} nonzero entries. Those nonzero entries are determined by a function, which is denoted as $f|_{x=d}$ and called a *restricted function*. This function is a Boolean function in $m-k$ variables and is obtained by replacing the variables x_{j_α} by d_α for all $0 \le \alpha < k$ in the original function f. The restricted sequence $F|_{x=d}$ is then found by associating a polyphase sequence of length 2^{m-k} with $f|_{x=d}$ and inserting $2^m - 2^{m-k}$ zeros at the corresponding positions. Similarly to a disjunctive normal form of a Boolean function [9, Chapter 13], the original function f can be reconstructed from the functions $f|_{x=d}$ by

$$f = \sum_{d \in \mathbb{Z}_2^k} f|_{x=d} \prod_{\alpha=0}^{k-1} x_{j_\alpha}^{d_\alpha} (1 - x_{j_\alpha})^{(1-d_\alpha)}. \tag{2}$$

The following lemma will be useful to expand correlations of sequences of length 2^m.

Lemma 5. *[3],[8] Let $f, g : \mathbb{Z}_2^m \to \mathbb{Z}_q$ be two generalized Boolean functions in the variables $x_0, x_1, \cdots, x_{m-1}$, and let F and G be their associated polyphase sequences, respectively. Let $J = \{j_0, j_1, \cdots, j_{k-1}\}$ and $I = \{i_0, i_1, \cdots, i_{k'-1}\}$ be two sets of indices, such that $I \cap J = \varnothing$ and $I \cup J \subseteq \{0, 1, \cdots, m-1\}$. Write $x = (x_{j_0} x_{j_1} \cdots x_{j_{k-1}})$ and $x' = (x_{i_0} x_{i_1} \cdots x_{i_{k'-1}})$. Suppose d, d_1, d_2 are binary words of length k and c, c_1, c_2 are binary words of length k'. Then we have*

$$C(F|_{x=d_1}, G|_{x=d_2})(\ell) = \sum_{c_1, c_2} C(F|_{xx'=d_1c_1}, G|_{xx'=d_2c_2})(\ell)$$

and

$$A(F|_{x=d})(\ell) = \sum_{c} A(F|_{xx'=dc})(\ell) + \sum_{c_1 \neq c_2} C(F|_{xx'=dc_1}, F|_{xx'=dc_2})(\ell).$$

2.4 Generalized Reed–Muller Codes

We recall (slightly modified) definitions and some basic properties of the generalized Reed–Muller codes $\mathrm{RM}_q(r, m)$ and $\mathrm{ZRM}_q(r, m)$ (cf. [1] and [3]).

Definition 6. *(a) For $0 \le r \le m$ the code $\mathrm{RM}_q(r, m)$ is defined as the set of sequences associated with a generalized Boolean function $\mathbb{Z}_2^m \to \mathbb{Z}_q$ of order at most r. (b) For $q \ge 4$ and $1 \le r \le m$ the code $\mathrm{ZRM}_q(r, m)$ is defined as the set of sequences associated with 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.*

Clearly for $q \geq 4$ and $1 \leq r \leq m$ we have $\mathrm{ZRM}_q(r,m) \subset \mathrm{RM}_q(r,m)$. Now recall the classical definitions of the minimum Hamming distance d_H and Lee distance d_L of a code $\mathcal{C} \subseteq \mathbb{Z}_q^n$ (see e.g. [9]), and notice that in the binary case (i.e. $q = 2$) the minimum Hamming and Lee distances coincide. We have:

Result 7. *[1], [3] The minimum Lee distances of* $\mathrm{RM}_q(r,m)$ *and* $\mathrm{ZRM}_q(r,m)$ *are equal to* 2^{m-r} *and* 2^{m-r+1}*, respectively.*

2.5 A Known Construction for Complementary Pairs

We recall a construction for complementary pairs from [3]. With each quadratic form f over $\mathbb{Z}_q$ in the variables $x_{i_0}, x_{i_1}, \cdots, x_{i_{m-1}}$, generally given by

$$\sum_{0 \leq j < k < m} q_{jk} x_{i_j} x_{i_k} + L, \qquad q_{jk} \in \mathbb{Z}_q$$

with L being an affine form over $\mathbb{Z}_q$, one can associate a labeled graph $G(f)$. The vertices of this graph are labeled with $i_0, i_1, \cdots, i_{m-1}$, and the edge between vertex i_j and vertex i_k is labeled with q_{jk}. Such a graph is called a *path* (passing through m vertices) if $m = 1$ (then the graph consists of a single vertex) or if $m \geq 2$ and f is of the form

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

where π is a permutation of $\{0, 1, \cdots, m-1\}$. The indices $i_{\pi(0)}$ and $i_{\pi(m-1)}$ are called *end vertices* of the path. We are now in the position to quote:

Result 8. *[3] Let* $0 \leq j_0 < j_1 < \cdots < j_{k-1} < m$ *be a list of* k *indices, write* $x = (x_{j_0} x_{j_1} \cdots x_{j_{k-1}})$*, and let* $d \in \mathbb{Z}_2^k$*. Suppose* $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ *is a generalized Boolean function in the variables* $x_0, x_1, \cdots, x_{m-1}$*, such that* $f|_{x=d}$ *is quadratic and* $G(f|_{x=d})$ *is a path (in* $m - k$ *vertices). Let* F *and* F' *be the polyphase sequences associated with* f *and* $f + (q/2)x_a + c'$*, respectively. Then* $F|_{x=d}$ *and* $F'|_{x=d}$ *form a complementary pair. Here* a *is an end vertex of the path* $G(f|_{x=d})$ *and* $c' \in \mathbb{Z}_q$*.*

In particular, if $k = 0$, the above result identifies $(m!/2)q^{m+1}$ polyphase sequences lying in complementary pairs [3, Corollary 11], [1, Theorem 3], where in the latter reference $q = 2^h$.

3 A Construction for Complementary Sets

Theorem 9. *Let* $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ *be a generalized Boolean function in* m *variables* $x_0, x_1, \cdots, x_{m-1}$*. Define a list of* k *indices by*

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

and write $x = (x_{j_0}\, x_{j_1} \cdots x_{j_{k-1}})$. *Suppose that for each* $d \in \mathbb{Z}_2^k$ *the restricted function* $f|_{x=d}$ *is quadratic and the graph* $G(f|_{x=d})$ *is a path having an end vertex* a_d. *Then the polyphase sequences associated with the functions*

$$f \;+\; \frac{q}{2}\left(\sum_{\alpha=0}^{k-1} c_\alpha\, x_{j_\alpha} + c'\, e\right) \qquad c_0, \cdots, c_{k-1}, c' \in \mathbb{Z}_2$$

form a complementary set of size 2^{k+1}, *where*

$$e = \sum_{d \in \mathbb{Z}_2^k} x_{a_d} \prod_{\alpha=0}^{k-1} x_{j_\alpha}^{d_\alpha} (1 - x_{j_\alpha})^{(1-d_\alpha)}.$$

Proof. Write $c = (c_0\, c_1 \cdots c_{k-1})$ and denote the 2^{k+1} sequences in the set by $F_{cc'}$. We have to show that the sum of auto-correlations $\sum_{c,\,c'} A\,(F_{cc'})\,(\ell)$ is zero for $\ell \neq 0$. We employ Lemma 5 and write

$$\sum_{c,\,c'} A(F_{cc'})(\ell) = \underbrace{\sum_{c,\,c'} \sum_{d} A(F_{cc'}|_{x=d})(\ell)}_{S_1} + \underbrace{\sum_{c,\,c'} \sum_{d_1 \neq d_2} C(F_{cc'}|_{x=d_1}, F_{cc'}|_{x=d_2})(\ell)}_{S_2}.$$

We first focus on the term S_1, which becomes

$$S_1 = \sum_{c} \sum_{d} \left(A(F_{c0}|_{x=d})(\ell) + A(F_{c1}|_{x=d})(\ell)\right).$$

Recall that $e|_{x=d} = x_{a_d}$ is an end vertex of the graph $G(f|_{x=d})$. Thus the functions corresponding to $F_{c0}|_{x=d}$ and $F_{c1}|_{x=d}$ are

$$f|_{x=d} + \frac{q}{2}\sum_{\alpha=0}^{k-1} c_\alpha\, d_\alpha \qquad \text{and} \qquad f|_{x=d} + \frac{q}{2}\sum_{\alpha=0}^{k-1} c_\alpha\, d_\alpha + \frac{q}{2}\, x_{a_d},$$

respectively. Notice that the sum over α is just a constant occuring in both functions. Hence, by hypothesis and by Result 8, $F_{c0}|_{x=d}$ and $F_{c1}|_{x=d}$ form a complementary pair. It follows that the inner term of S_1 is zero for $\ell \neq 0$, and thus, also S_1 itself is zero for $\ell \neq 0$.

It remains to prove that the term S_2 is zero. This part of the proof follows more or less the same reasoning as the second part of the proof of [3, Theorem 12]. □

Theorem 9 generalizes [3, Theorem 12] from complementary sets that contain sequences corresponding to quadratic generalized Boolean functions to complementary sets comprised of sequences associated with arbitrary generalized Boolean functions. Hence Theorem 9 provides a general upper bound on the PMEPR of polyphase sequences of length 2^m. Moreover this bound remains the same for all words in a coset of $\mathrm{RM}_q(1, m)$.

Corollary 10. *Suppose that $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ is a generalized Boolean function in the variables $x_0, x_1, \cdots, x_{m-1}$. If there exists a set of k restricting variables $x = (x_{j_0}\, x_{j_1}\, \cdots\, x_{j_{k-1}})$ with*

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

such that for each $d \in \mathbb{Z}_2^k$, the restricted function $f|_{x=d}$ is quadratic and the graph $G(f|_{x=d})$ is a path, then the polyphase sequences in the coset $f + \mathrm{RM}_q(1,m)$ have PMEPR at most 2^{k+1}.

It should be noted that for large k the above corollary appears to be rather weak. Using Corollary 10 and (2), it is now straightforward to find an explicit construction for sequences having PMEPR at most 2^{k+1}. Therefore partition the m indices $\{0, 1, \cdots, m-1\}$ into sets $I = \{i_0, i_1, \cdots, i_{m-k-1}\}$ and $J = \{j_0, j_1, \cdots, j_{k-1}\}$. Let $\pi_0, \pi_1, \cdots, \pi_{2^k-1}$ be 2^k permutations of $\{0, 1, \cdots, m-k-1\}$, and let $g_0, \cdots, g_{m-k-1}, g' : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be $m-k+1$ generalized Boolean functions. Then each sequence associated with

$$\frac{q}{2} \sum_{d \in \mathbb{Z}_2^k} \sum_{\alpha=0}^{m-k-2} x_{i_{\pi_d(\alpha)}} x_{i_{\pi_d(\alpha+1)}} \prod_{\beta=0}^{k-1} x_{j_\beta}^{d_\beta} (1 - x_{j_\beta})^{(1-d_\beta)} + \sum_{\alpha=0}^{m-k-1} x_{i_\alpha} g_\alpha(x_{j_0}, \cdots, x_{j_{k-1}}) + g'(x_{j_0}, \cdots, x_{j_{k-1}}) \quad (3)$$

satisfies Corollary 10 for a particular k and, hence, has PMEPR at most 2^{k+1}. It is apparent that each such a sequence lies inside $\mathrm{RM}_q(k+2, m)$ and particularly inside $\mathrm{ZRM}_q(k+2, m)$ if $q \ge 4$.

In what follows we focus on a particular subset of the sequences associated with forms of type (3). In this way we can obtain a wide range of coding options for OFDM with low PMEPR that allow a trade-off between the size of the sequence set and minimum distance.

Corollary 11. *Let $g_0, \cdots, g_{m-k-1}, g' : \mathbb{Z}_2^k \to \mathbb{Z}_q$ be $m-k+1$ generalized Boolean functions. Suppose $2 \le r \le k+1$, $\deg(g_\alpha) \le r-1$ for $\alpha = 0, 1, \cdots m-k-1$, and $\deg(g') \le r$. Then the polyphase sequences associated with the forms*

$$\frac{q}{2} \sum_{\alpha=0}^{m-k-2} x_{\pi(\alpha)} x_{\pi(\alpha+1)} + \sum_{\alpha=0}^{m-k-1} x_\alpha g_\alpha(x_{m-k}, \cdots, x_{m-1}) + g'(x_{m-k}, \cdots, x_{m-1}),$$

where π is a permutation of $\{0, 1, \cdots, m-k-1\}$, have PMEPR at most 2^{k+1}. Moreover the sequences form cosets of $\mathrm{RM}_q(1,m)$, which are contained inside $\mathrm{RM}_q(r,m)$. In particular these cosets are contained in $\mathrm{ZRM}_q(r,m)$ if (i) $q \ge 4$ and $k = 0$ or if (ii) $q \ge 4$ and all coefficients of the monomials in the algebraic normal forms of g_α with degree equal to $r-1$ and in the algebraic normal form of g' with degree equal to r are even.

Now suppose that k is given. Then a simple counting argument shows that the above corollary identifies $2^{K_{\mathrm{RM}}}$ words inside $\mathrm{RM}_q(r,m)$, where

$$K_{\mathrm{RM}} = \log_2 \frac{(m-k)!}{2} + \left[\binom{k}{r} + (m-k+1) \sum_{i=0}^{r-1} \binom{k}{i} \right] \log_2 q$$

and $2^{K_{\mathrm{ZRM}}}$ words inside $\mathrm{ZRM}_q(r,m)$ with

$$K_{\mathrm{ZRM}} = K_{\mathrm{RM}} - \binom{k}{r} - (m-k)\binom{k}{r-1}.$$

4 A Construction for Almost Complementary Pairs

Theorem 12. *Let $k \le m-3$. Suppose $J = \{j_0, j_1, \cdots, j_{k-1}\}$ and $I = \{i_0, i_1, \cdots, i_{m-k-1}\}$ are two sets of indices, such that $I \cap J = \varnothing$ and $I \cup J = \{0, 1, \cdots, m-1\}$. Write $x = (x_{j_0} x_{j_1} \cdots x_{j_{k-1}})$ and let $d \in \mathbb{Z}_2^k$. Let $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ be a generalized Boolean function in the variables $x_0, x_1, \cdots, x_{m-1}$, such that $f|_{x=d}$ is of the form*

$$\frac{q}{2} \sum_{\gamma=0}^{m-k-2} x_{i_{\pi(\gamma)}} x_{i_{\pi(\gamma+1)}} + \alpha x_b x_d + \beta x_c x_d + \sum_{\gamma=0}^{m-k-1} c_\gamma x_{i_{\pi(\gamma)}} + c,$$
$$c_0, \cdots, c_{m-k-1}, c, \alpha, \beta \in \mathbb{Z}_q,$$

where π is a permutation of $\{0, 1, \cdots, m-k-1\}$ and $(a\,b\,c\,d)$ is either $(i_{\pi(m-k-1)}\, i_{\pi(0)}\, i_{\pi(1)}\, i_{\pi(2)})$ or $(i_{\pi(0)}\, i_{\pi(m-k-1)}\, i_{\pi(m-k-2)}\, i_{\pi(m-k-3)})$. Let F and F' be the polyphase sequences associated with the functions f and $f' = f + (q/2)x_a + c'$, respectively, where $c' \in \mathbb{Z}_q$. Then, provided that $\alpha - \beta = 0$ or $\alpha + \beta = 0$, $F|_{x=d}$ and $F'|_{x=d}$ form an almost complementary pair. In particular we have

$$|A(F|_{x=d})(\ell) + A(F'|_{x=d})(\ell)| = \begin{cases} 2^{m-k} & \ell = 0 \\ 2^{m-k-1} & \ell = \pm\tau, \quad 1 \le \tau < n \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. We take $(a\,b\,c\,d) = (i_{\pi(m-k-1)}\, i_{\pi(0)}\, i_{\pi(1)}\, i_{\pi(2)})$. With a similar reasoning we can perform the proof for $(a\,b\,c\,d) = (i_{\pi(0)}\, i_{\pi(m-k-1)}\, i_{\pi(m-k-2)}\, i_{\pi(m-k-3)})$.

Substitute: $G = F|_{x=d}$, $G' = F'|_{x=d}$, $g = f|_{x=d}$, and $g' = f'|_{x=d}$, and keep in mind that G and G' are in general restricted sequences. We are interested in the expression $A(G)(\ell) + A(G')(\ell)$, which is expanded using Lemma 5

$$\begin{aligned} &A(G)(\ell) + A(G')(\ell) \\ &= A(G|_{x_a=0})(\ell) + A(G|_{x_a=1})(\ell) + A(G'|_{x_a=0})(\ell) + A(G'|_{x_a=1})(\ell) \\ &+ C(G|_{x_a=0}, G|_{x_a=1})(\ell) + C(G|_{x_a=1}, G|_{x_a=0})(\ell) \\ &+ C(G'|_{x_a=0}, G'|_{x_a=1})(\ell) + C(G'|_{x_a=1}, G'|_{x_a=0})(\ell). \end{aligned} \tag{4}$$

Let us first consider the case $k = m - 3$, and take this as the base case for the proof. Then we proceed to prove the theorem for $k < m - 3$ by induction on k. For $k = m - 3$ we have $a = d$, and the list of restricting indices writes $J = \{0, 1, \cdots, m-1\} \backslash \{a, b, c\}$. The functions corresponding to the sequences $G|_{x_a=0}$, $G|_{x_a=1}$, $G'|_{x_a=0}$, and $G'|_{x_a=1}$ are then

$$\begin{aligned} g|_{x_a=0} &= r & g'|_{x_a=0} &= g|_{x_a=0} + c' \\ g|_{x_a=1} &= r + \left(\alpha + \frac{q}{2}\right) x_b + \beta x_c + c_a & g'|_{x_a=1} &= g|_{x_a=1} + c' + \frac{q}{2} \end{aligned}$$

with

$$r = \frac{q}{2} x_b x_c + c_b x_b + c_c x_c + c.$$

Since we have $g'|_{x_a=0} = g|_{x_a=0} + c'$ and $g'|_{x_a=1} = g|_{x_a=1} + c' + q/2$, it follows that $G'|_{x_a=0} = \xi^{c'} G|_{x_a=0}$ and $G'|_{x_a=1} = -\xi^{c'} G|_{x_a=1}$. This implies that the cross-correlations in (4) are related as follows

$$\begin{aligned} C(G|_{x_a=0}, G|_{x_a=1})(\ell) &= -C(G'|_{x_a=0}, G'|_{x_a=1})(\ell) \\ C(G|_{x_a=1}, G|_{x_a=0})(\ell) &= -C(G'|_{x_a=1}, G'|_{x_a=0})(\ell). \end{aligned} \tag{5}$$

Hence the cross-correlations in (4) sum up to zero for all ℓ. We also conclude

$$A(G|_{x_a=0})(\ell) + A(G|_{x_a=1})(\ell) = A(G'|_{x_a=0})(\ell) + A(G'|_{x_a=1})(\ell). \tag{6}$$

Next we consider the left-hand side of (6). Write $u = u_a\, u_b\, u_c$, $v = v_a\, v_b\, v_c$, $G_u = G|_{x_a x_b x_c = u_a u_b u_c}$, and $g_u = g|_{x_a x_b x_c = u_a u_b u_c}$. Then, by Lemma 5, we have

$$A(G|_{x_a=0})(\ell) + A(G|_{x_a=1})(\ell) = \sum_u A(G_u)(\ell) + \sum_{\substack{u,v \\ u_a = v_a \\ (u_b u_c) \neq (v_b v_c)}} C(G_u, G_v)(\ell). \tag{7}$$

The functions corresponding to the sequences G_u are

$$\begin{aligned} g_{000} &= c & g_{100} &= c \\ g_{010} &= c + c_b & g_{110} &= c + c_b + \alpha + \frac{q}{2} \\ g_{001} &= c + c_c & g_{101} &= c + c_c + \beta \\ g_{011} &= c + c_b + c_c + \frac{q}{2} & g_{111} &= c + c_b + c_c + \alpha + \beta. \end{aligned}$$

Each of the eight sequences G_u contains exactly one nonzero element occuring at position $\sum_{\gamma=0}^{k-1} d_\gamma 2^{j_\gamma} + u_a 2^a + u_b 2^b + u_c 2^c$. Thus we have

$$\sum_u A(G_u)(\ell) = \begin{cases} 8 & \ell = 0 \\ 0 & \ell \neq 0 \end{cases}. \tag{8}$$

It follows also that the twelve cross-correlations of type $C(G_u, G_v)(\ell)$ in the second sum of (7) have exactly one nonzero element, which is located at

$(u_b - v_b)2^b + (u_c - v_c)2^c$. We next collect the cross-correlations having nonzero contributions at the same shift.
Shift 2^b ($u_b = 1$, $v_b = 0$, $u_c = v_c$):

$$C(G_{000}, G_{010})(2^b)+C(G_{100}, G_{110})(2^b)+C(G_{001}, G_{011})(2^b)+C(G_{101}, G_{111})(2^b) \\ = \xi^{-c_b} + \xi^{\frac{q}{2}-c_b-\alpha} + \xi^{\frac{q}{2}-c_b} + \xi^{-c_b-\alpha} = 0$$

Shift 2^c ($u_b = v_b$, $u_c = 1$, $v_c = 0$):

$$C(G_{000}, G_{001})(2^c)+C(G_{100}, G_{101})(2^c)+C(G_{010}, G_{011})(2^c)+C(G_{110}, G_{111})(2^c) \\ = \xi^{-c_c} + \xi^{-c_c-\beta} + \xi^{\frac{q}{2}-c_c} + \xi^{\frac{q}{2}-c_c-\beta} = 0$$

Shift $2^b + 2^c$ ($u_b = u_c = 1$, $v_b = v_c = 0$):

$$C(G_{000}, G_{011})(2^b + 2^c) + C(G_{100}, G_{111})(2^b + 2^c) = \xi^{\frac{q}{2}-c_c-c_b} + \xi^{-c_c-c_b-(\alpha+\beta)}$$

Shift $2^b - 2^c$ ($u_b = v_c = 1$, $u_c = v_b = 0$):

$$C(G_{001}, G_{010})(2^b - 2^c) + C(G_{101}, G_{110})(2^b - 2^c) = \xi^{c_c-c_b} + \xi^{\frac{q}{2}+c_c-c_b-(\alpha-\beta)}$$

Moreover there are contributions at the shifts -2^b, -2^c, $-2^b - 2^c$, and $-2^b + 2^c$, which are just the complex conjugated values of those at shifts 2^b, 2^c, $2^b + 2^c$, and $2^b - 2^c$, respectively. Now the additional condition $\alpha + \beta = 0$ or $\alpha - \beta = 0$ comes into play. Then the auto-correlations of G and G' cancel out at either $2^b + 2^c$ or $2^b - 2^c$. Considering (6), we have to count the auto-correlation in (8) and all the contributions from the cross-correlations twice in order to calculate the left-hand side of (4). By carefully counting the contributions, we observe that the left-hand side of (4) is nonzero for at most two nonzero shifts and has absolute values of at most 4 at those positions.

Now consider the case where x contains $k < m - 3$ restricting variables, and suppose that the theorem is true for x containing $k + 1$ variables. We focus on the expanded auto-correlations in (4). For $k < m-3$ the functions corresponding to the sequences $G|_{x_a=0}$, $G|_{x_a=1}$, $G'|_{x_a=0}$, and $G'|_{x_a=1}$ are

$$g|_{x_a=0} = p \qquad\qquad g'|_{x_a=0} = g|_{x_a=0} + c'$$
$$g|_{x_a=1} = p + \frac{q}{2}x_{i_{\pi(m-k-2)}} + c_a \qquad\qquad g'|_{x_a=1} = g|_{x_a=1} + c' + \frac{q}{2}$$

with

$$p = \frac{q}{2}\sum_{\gamma=0}^{m-k-3} x_{i_{\pi(\gamma)}}x_{i_{\pi(\gamma+1)}} + \alpha x_b x_d + \beta x_c x_d + \sum_{\gamma=0}^{m-k-2} c_\gamma x_{i_{\pi(\gamma)}} + c.$$

By the same reasoning leading to (5), it turns out that the cross-correlations in (4) sum up to zero for all ℓ. Using the above theorem as a hypothesis, we know that the sequences $G|_{x_a=0}$ and $G|_{x_a=1}$ form an almost complementary pair. Hence the sum of their auto-correlations has at most two nonzero values at

nonzero shifts. The same accounts for the pair of sequences $G'|_{x_a=0}$ and $G'|_{x_a=1}$. Note that the position and the phase of the nonzero components are independent of the restricting variables, while their magnitudes depend on the number of nonzero entries in the sequences $G|_{x_a=0}$, $G|_{x_a=1}$, $G'|_{x_a=0}$, and $G'|_{x_a=1}$. Hence the auto-correlations in (4) always superimpose at the considered positions and sum up to absolute values of at most 2^{m-k-1}. □

Setting $k = 0$ in the above theorem and applying Lemma 1, we obtain the following simple corollary.

Corollary 13. *For $m > 2$ let $f : \mathbb{Z}_2^m \to \mathbb{Z}_q$ be given by*

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

where $\alpha, \beta \in \mathbb{Z}_q$, π is a permutation of $\{0, 1, \cdots, m-1\}$, and $\alpha+\beta = 0$ or $\alpha-\beta = 0$. Then the PMEPR of the polyphase sequence in the coset $f + \mathrm{RM}_q(1, m)$ is at most 3. These cosets lie inside $\mathrm{RM}_q(2, m)$ and particularly inside $\mathrm{ZRM}_q(2, m)$ if $q \geq 4$ and α and β are even.

A simple counting argument shows that Corollary 13 identifies $(2q - 3)(m!/2)q^{m+1}$ sequences with PMEPR at most 3. Notice that $(m!/2)q^{m+1}$ of them have PMEPR bounded by 2, since they are also identified by setting $k = 0$ in Corollary 11. In particular, Corollary 13 applies to the $2m!\, q^{m+1}$ ($q = 2^h$, $h \geq 3$) sequences for which it was conjectured in [1] and [5] that their PMEPR is bounded by 3. Corollary 13 provides a proof for this conjecture and identifies further sequences with PMEPR at most 3. We remark that for $q = 2$ Corollary 13 merely restates the construction of complementary pairs in a pure sense [1, Theorem 3] (where $q = 2^h$), [3, Corollary 11].

5 OFDM Coding Schemes from Reed–Muller Codes

In what follows we apply Corollary 11 and Corollary 13 to construct a number of coding options for OFDM with low PMEPR. Since these codes are unions of cosets of $\mathrm{RM}_q(1, m)$, well known algorithms are readily applicable to encode and decode the codes (see [1], [10] and references therein). Alternatively in [11] the structure of the codes obtained from Corollary 11 is further exploited in order to simplify encoding and decoding. This is particularly efficient if the codes contain a large number of cosets of $\mathrm{RM}_q(1, m)$. We define the code rate and the information rate of a code $\mathcal{C}$ of length $n = 2^m$ over a q-ary alphabet to be $\lfloor \log_q |\mathcal{C}| \rfloor / n$ and $\lfloor \log_2 |\mathcal{C}| \rfloor / n$, respectively.

We remark that all our coding options directly arise just from two simple corollaries. This should be compared with the constructions in [1] and [3], which rely on a variety of techniques, including a computational search in [1]. Nevertheless there is some overlap between our codes and those in [1] and [3]. In the subsequent tables the superscript [1] means that the corresponding coding option

Table 1. Binary Coding Options

n	Option	max. PMEPR	# info bits	code rate	info rate	d_L
16	B1[1]	2	8	0.50	0.50	4
	B2[2]	4	9	0.56	0.56	4
	B3	8	12	0.75	0.75	2
	B4[2]	8	10	0.63	0.63	4
32	B1[1]	2	11	0.34	0.34	8
	B2	4	13	0.41	0.41	8
	B3	8	17	0.53	0.53	4
	B4	8	14	0.44	0.44	8
64	B1[1]	2	15	0.23	0.23	16
	B2	4	17	0.27	0.27	16
	B3	8	23	0.36	0.36	8
	B4	8	19	0.30	0.30	16

is identical to the Davis–Jedwab construction [1], which can be restated by setting $k = 0$ in Corollary 11. A [2] indicates that a better code has been reported in [1], while a [3] means that the same code appears in [1]. These latter results are based on a computational search, which becomes infeasible for large n. Therefore this situation occurs only for $n = 16$. Coding options marked by a [4] are weaker than codes constructed in [3]. A [5] indicates that a better code can be obtained using the techniques in [3], though the code itself has not been explicitely mentioned in [3]. All other coding options seem to be new or outperform previously reported results at least for the lengths considered here.

Table 5 shows possible coding options for binary signaling. Option B1 is the Davis–Jedwab construction. Option B2 is obtained from Corollary 11 by setting $k = 1$, and Options B3 and B4 arise from Corollary 11 by setting $k = 2$ and taking those codewords lying inside $\mathrm{RM}_2(3, m)$ and $\mathrm{RM}_2(2, m)$, respectively. The choice between the Options B1, B2, and B4 allows a trade-off between information rate and maximum PMEPR of the code, while the minimum Lee distance is the same. Options B3 and B4 provide a trade-off between minimum Lee distance and informations rate, while the PMEPR is constant.

Table 5 contains a list of coding options for quaternary signaling. Option Q1 is the Davis–Jedwab construction. Option Q2 is obtained from Corollary 13. Option Q3 uses Corollary 11 with $k = 1$ to construct a code lying in $\mathrm{RM}_4(2, m)$, while Option Q4 takes its subcode in $\mathrm{ZRM}_4(2, m)$. Option Q5 is a code inside $\mathrm{RM}_4(3, m)$ and obtained by setting $k = 2$ in Corollary 11. Option Q6 and Option Q7 are the subcodes lying in $\mathrm{ZRM}_4(3, m)$ and $\mathrm{ZRM}_4(2, m)$, respectively. Notice that we may also construct a subcode of the code in Option Q5 inside $\mathrm{RM}_4(2, m)$, however, this code contains less codewords than that in Option Q6, while the minimum distances and the maximum PMEPRs are the same. Moving to a quaternary constellation widely extends the possible coding options, which results in an increased number of possible trade-offs between PMEPR,

Table 2. Quaternary Coding Options

n	Option	max. PMEPR	# info bits	code rate	info rate	d_L
16	Q1[1]	2	13	0.41	0.81	8
	Q2	3	15	0.47	0.94	4
	Q3[4]	4	17	0.53	1.06	4
	Q4[2]	4	14	0.44	0.88	8
	Q5	8	24	0.75	1.50	2
	Q6	8	22	0.69	1.37	4
	Q7[2]	8	15	0.47	0.94	8
32	Q1[1]	2	17	0.27	0.53	16
	Q2	3	20	0.31	0.63	8
	Q3[4]	4	23	0.36	0.72	8
	Q4	4	19	0.30	0.59	16
	Q5	8	33	0.52	1.03	4
	Q6	8	30	0.47	0.94	8
	Q7	8	20	0.31	0.63	16
64	Q1[1]	2	22	0.17	0.34	32
	Q2	3	24	0.19	0.38	16
	Q3[4]	4	29	0.23	0.45	16
	Q4	4	24	0.19	0.38	32
	Q5	8	43	0.34	0.67	8
	Q6	8	39	0.30	0.61	16
	Q7	8	26	0.20	0.41	32

information rate, and minimum distance. Moreover the information rate can be increased up to twice that of the binary coding schemes. However this goes generally at the cost of a smaller minimum Euclidean distance of the codes, which results in an increased transmission error probability.

Table 5 shows a number of coding options for octary phase-shift keying. These options are obtained in a similar fashion as the quaternary coding options. Option O1 is the Davis–Jedwab construction, Option O2 uses Corollary 13 to construct a code inside $\mathrm{RM}_8(2, m)$, while Option O3 takes its subcode in $\mathrm{ZRM}_8(2, m)$. Option O4 uses Corollary 11 with $k = 1$ to construct a code inside $\mathrm{RM}_8(2, m)$. Option O5 is obtained by taking its subcode contained in $\mathrm{ZRM}_8(2, m)$. Option O6 arises by setting $k = 2$ in Corollary 11 to construct a code inside $\mathrm{RM}_8(3, m)$, while Options O7 and O8 use its subcodes in $\mathrm{ZRM}_8(3, m)$ and $\mathrm{ZRM}_8(2, m)$, respectively. Notice that although, based on numerical evaluation of the PMEPR, Option O3 was already mentioned in [1] and [12] for $n = 16$, Corollary 13 settles the theory behind this coding option and validates it for any $m > 2$. By moving to an octary constellation, the number of possible coding options is further extended. Also the information rate can be increased further, which in turn leads to a smaller minimum Euclidean distance of the codes compared to their binary or quaternary counterparts.

Table 3. Octary Coding Options

n	Option	max. PMEPR	# info bits	code rate	info rate	d_L
16	O1[1]	2	18	0.38	1.13	8
	O2	3	22	0.46	1.38	4
	O3[3]	3	20	0.42	1.25	8
	O4[4]	4	25	0.52	1.56	4
	O5[5]	4	22	0.46	1.38	8
	O6	8	36	0.75	2.25	2
	O7	8	34	0.71	2.13	4
	O8	8	25	0.52	1.56	8
32	O1[1]	2	23	0.24	0.72	16
	O2	3	27	0.28	0.84	8
	O3	3	26	0.27	0.81	16
	O4[4]	4	33	0.34	1.03	8
	O5[5]	4	29	0.30	0.91	16
	O6	8	49	0.51	1.53	4
	O7	8	46	0.48	1.44	8
	O8	8	33	0.34	1.03	16
64	O1[1]	2	29	0.15	0.45	32
	O2	3	33	0.17	0.52	16
	O3	3	31	0.16	0.48	32
	O4[4]	4	41	0.21	0.64	16
	O5[5]	4	36	0.19	0.56	32
	O6	8	63	0.33	0.98	8
	O7	8	59	0.31	0.92	16
	O8	8	42	0.22	0.66	32

6 Conclusion

We have established constructions of sequences lying in complementary sets of a given size (Theorem 9) and in almost complementary pairs (Theorem 12). An upper bound for the PMEPR of these sequences follows then immediately from Lemma 1. These results led to a number of coding options for OFDM with low PMEPR, which extend and complement existing schemes previously reported in [1], [3], and [12]. Corollary 13 also provides an answer to an open problem stated by Davis and Jedwab in [1]. Recent results have shown that this corollary in fact arises in a more general context, and we refer to [13] for details.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments. Especially one referee provided very detailed suggestions, which led to several improvements of the paper.

References

1. Davis, J.A., Jedwab, J.: Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed–Muller codes. IEEE Trans. Inform. Theory **45** (1999) 2397–2417
2. Golay, M.J.E.: Complementary series. IRE Trans. Inform. Theory **7** (1961) 82–87
3. Paterson, K.G.: Generalized Reed–Muller codes and power control in OFDM modulation. IEEE Trans. Inform. Theory **46** (2000) 104–120
4. Parker, M.G., Tellambura, C.: A construction for binary sequence sets with low peak-to-average power ratio. Report No. 242, Department of Informatics, University of Bergen, Norway, `http://www.ii.uib.no/~matthew/` (2003)
5. Nieswand, K.M., Wagner, K.N.: Octary codewords with power envelopes of $3 * 2^m$. `http://www.mathcs.richmond.edu/~jad/summer.html` (1998)
6. Jones, A.E., Wilkinson, T.A.: Combined coding for error control and increased robustness to system nonlinearities in OFDM. Proc. of IEEE 46th Vehicular Technology Conf. (VTC) (1996)
7. Tellambura, C.: Upper bound on the peak factor of n-multiple carriers. IEE Electron. Lett. **33** (1997) 1608–1609
8. Stinchcombe, T.E.: Aperiodic Autocorrelations of Length 2^m Sequences, Complementarity, and Power Control for OFDM. PhD thesis, University of London (2000)
9. MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes. North Holland Mathematical Library (1977)
10. Paterson, K.G., Jones, A.E.: Efficient decoding algorithms for generalized Reed–Muller codes. IEEE Trans. Commun. **48** (2000) 1272–1285
11. Schmidt, K.U.: Complementary sets, generalized Reed–Muller codes, and power control for OFDM. submitted to IEEE Trans. Inform. Theory (2005)
12. Paterson, K.G.: Coding techniques for power-controlled OFDM. Proc. of IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC) (1998) 801–805
13. Schmidt, K.U.: On cosets of the generalized first-order Reed–Muller code with low PMEPR. submitted to IEEE Trans. Inform. Theory (2005)