SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Low-PMEPR Rotatable Pilot Sequences for MIMO-OFDM Systems

Yajing ZHOU[1], Zhengchun ZHOU[2*], Zhi GU[1] & Pingzhi FAN[3]

[1] *School of Information Science and Technology, Southwest Jiaotong University, Chengdu* 610031*, China;*
[2] *School of Mathematics, Southwest Jiaotong University, Chengdu* 610031*, China;*
[3] *Institute of Mobile Communications, Southwest Jiaotong University, Chengdu* 610031*, China*

Appendix A Preliminaries

The following are some notations which will be used throughout this paper.

- q is an even integer and $\mathbb{Z}_q = \{0, 1, ..., q-1\}$ denotes the ring of integers modulo q;
- ξ_q denotes the q-th primitive root of unity;
- $|a|$ and a^* denote the modulus and complex conjugation, respectively, of the complex number a;
- $\mathrm{Re}(a)$ denotes the real part of the complex number a;
- $|\mathbf{a}|$ denotes the magnitude of the vector $\mathbf{a}$;
- $|\mathcal{C}|$ denotes the set size of the sequence set $\mathcal{C}$;
- $dim(\cdot)$ denotes the dimension of the square matrix;
- $\mathbf{a}\Diamond\mathbf{b} = (a(0), b(0), a(1), b(1), ..., a(L-1), b(L-1))$ denotes the bit-interleaved sequence of L-length sequences $\mathbf{a}$ and $\mathbf{b}$;
- $\mathbf{a} \parallel \mathbf{b} = (a(0), a(1), ..., a(L-1), b(0), b(1), ..., b(L-1))$ denotes the concatenation of L-length sequences $\mathbf{a}$ and $\mathbf{b}$;
- $R(\mathcal{C})$ and $d(\mathcal{C})$ denote the code rate and minimum Hamming distance of the codebook $\mathcal{C}$, respectively.

Appendix A.1 Complementary Sequence Sets

Let $\mathbf{a} = (a(0), a(1), ..., a(L-1))$ and $\mathbf{b} = (b(0), b(1), ..., b(L-1))$ be two complex-valued sequences of length L. The aperiodic cross-correlation between $\mathbf{a}$ and $\mathbf{b}$ at a time shift τ is defined as

$$\rho_{\mathbf{a},\mathbf{b}}(\tau) = \begin{cases} \sum_{i=0}^{L-1-\tau} a(i)b^*(i+\tau), & 0 \leqslant \tau \leqslant L-1; \\ \sum_{i=0}^{L-1+\tau} a(i-\tau)b^*(i), & -(L-1) \leqslant \tau \leqslant -1; \\ 0, & |\tau| \geqslant L. \end{cases} \tag{A1}$$

When $\mathbf{a} = \mathbf{b}$, it is called the aperiodic autocorrelation of $\mathbf{a}$ and denoted by $\rho_{\mathbf{a}}(\tau)$.

Definition 1 ([1]). Let $\mathcal{A} = \{\mathbf{a}_i\}_{i=1}^N$ be a set of N sequences of length L. It is said to be a CSS of size N if $\sum_{i=1}^N \rho_{\mathbf{a}_i}(\tau) = 0$ for $0 < \tau < L$. In this case, every $\mathbf{a}_i$ in $\mathcal{A}$ is called a complementary sequence (CS). In particular, when $N = 2$, $\mathcal{A}$ is called a Golay complementary pair (GCP), and each of the constituent sequences in this pair is called a Golay complementary sequence (GCS).

Definition 2 (Golay Mates). A GCP $(\mathbf{c}, \mathbf{d})$ is called a mate of GCP $(\mathbf{a}, \mathbf{b})$ if

$$\rho_{\mathbf{a},\mathbf{c}}(\tau) + \rho_{\mathbf{b},\mathbf{d}}(\tau) = 0, \text{ for all } 0 \leqslant \tau < L. \tag{A2}$$

Appendix A.2 PMEPRs of OFDM Symbols

Let us consider an OFDM system with L subcarriers. For a complex-valued sequence $\mathbf{a} = (a(0), a(1), ..., a(L-1))$, the transmitted OFDM signal is the real part of the complex envelope, which can be written as

$$S_{\mathbf{a}}(t) = \sum_{i=0}^{L-1} a(i) e^{2\pi(f_c + i\Delta f)t\sqrt{-1}}, \ 0 \leqslant t < T \tag{A3}$$

, where f_c denotes the carrier frequency, and $\Delta f = \frac{1}{T}$ denotes the subcarrier spacing, with T being the OFDM symbol duration. The sequence $\mathbf{a}$ of length L is called the modulating codeword of the OFDM symbol.

The instantaneous power of an OFDM sequence (codeword) $\mathbf{a}$ is given by

$$P_{\mathbf{a}}(t) \ = \ |S_{\mathbf{a}}(t)|^2 = \rho_{\mathbf{a}}(0) + 2\mathrm{Re}\left(\sum_{\tau=1}^{L-1} \rho_{\mathbf{a}}(\tau) e^{2\pi(\tau\Delta f)t\sqrt{-1}}\right). \tag{A4}$$

The PMEPR of the OFDM sequence $\mathbf{a}$ is then defined as

$$\mathrm{PMEPR}(\mathbf{a}) = \frac{\sup_{t\in[0,T)} P_{\mathbf{a}}(t)}{P_{av}(\mathbf{a})} \tag{A5}$$

* Corresponding author (email: zzc@swjtu.edu.cn)

, where $P_{av}(\mathbf{a}) = \frac{1}{T}\int_{[0,T]} P_{\mathbf{a}}(t)dt = \|\mathbf{a}\|^2 = \rho_{\mathbf{a}}(0)$ is the average power of $\mathbf{a}$, and $\sup_{t\in[0,T)} P_{\mathbf{a}}(t)$ is the maximum value of $\{P_{\mathbf{a}}(t) : t \in [0,T)\}$. Accordingly, the PMEPR of a sequence set $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N\}$ is defined as

$$\text{PMEPR}(\mathcal{A}) = \max_{\mathbf{a}_i\in\mathcal{A}} \text{PMEPR}(\mathbf{a}_i). \tag{A6}$$

Combining (A4) and (A5), it can be derived that for sequence set $\mathcal{A}$ of length L, we have

$$\begin{aligned}
PMEPR(\mathbf{a}_i) &= \frac{\sup_{t\in[0,T)} P_{\mathbf{a}_i}(t)}{P_{av}(\mathbf{a}_i)} = \frac{1}{\rho_{\mathbf{a}_i}(0)}\left(\rho_{\mathbf{a}_i}(0) + 2\sup_{t\in[0,T)}\left(\text{Re}\left(\sum_{\tau=1}^{L-1}\rho_{\mathbf{a}_i}(\tau)e^{2\pi(\tau\Delta f)t\sqrt{-1}}\right)\right)\right)\\
&\leqslant \frac{\sum_{k=1}^{N}\sup_{t\in[0,T)} P_{\mathbf{a}_k}(t)}{P_{av}(\mathbf{a}_i)} = \frac{1}{\rho_{\mathbf{a}_i}(0)}\left(\sum_{k=1}^{N}\rho_{\mathbf{a}_k}(0) + 2\sum_{k=1}^{N}\sup_{t\in[0,T)}\left(\text{Re}\left(\sum_{\tau=1}^{L-1}\rho_{\mathbf{a}_k}(\tau)e^{2\pi(\tau\Delta f)t\sqrt{-1}}\right)\right)\right)\\
&\leqslant \frac{1}{\rho_{\mathbf{a}_i}(0)}\left(\sum_{k=1}^{N}\rho_{\mathbf{a}_k}(0) + 2\sum_{\tau=1}^{L-1}\left|\sum_{k=1}^{N}\rho_{\mathbf{a}_k}(\tau)\right|\right) \quad (1\leqslant i\leqslant N).
\end{aligned}$$

Based on inequality above, Liu *et al.* generalized the bound [2] on the PMEPR of polyphase sequences and obtained the following result.

Lemma 1 ([3]). Let $\mathcal{A}$ be a CSS of size N in which all the sequences have the same energy. Then the PMEPR of $\mathcal{A}$ is upper bounded by N.

Lemma 1 is useful to evaluate the PMEPR of a sequence in the sequel.

Appendix A.3 Rotatable Sequence Set

Definition 3. For a complex-valued sequence $\mathbf{a} = (a(0), a(1), ..., a(L-1))$ and a q-th rotatable index set Γ, define a sequence set $\mathcal{A}_\Gamma$ as follows

$$\mathcal{A}_\Gamma = \begin{bmatrix}\mathbf{a}_{1-q}^\Gamma\\ \vdots\\ \mathbf{a}_0^\Gamma\\ \vdots\\ \mathbf{a}_{q-1}^\Gamma\end{bmatrix} = \begin{bmatrix} a_{1-q}^\Gamma(0),\ a_{1-q}^\Gamma(1),\ \ldots,\ a_{1-q}^\Gamma(L-1)\\ \vdots\\ a_0^\Gamma(0),\ a_0^\Gamma(1),\ \ldots,\ a_0^\Gamma(L-1)\\ \vdots\\ a_{q-1}^\Gamma(0),\ a_{q-1}^\Gamma(1),\ \ldots,\ a_{q-1}^\Gamma(L-1)\end{bmatrix}_{(2q-1)\times L}$$

, where

$$a_p^\Gamma(i) = \begin{cases} a(i), & i \notin \Gamma,\\ a(i)\cdot\xi_q^p, & i\in\Gamma.\end{cases} \tag{A7}$$

Then $\mathcal{A}_\Gamma$ is called the rotatable sequence set (RSS) of $\mathbf{a}$ corresponding to Γ. Note that $\mathbf{a}_0^\Gamma = \mathbf{a}$ for any Γ, hence, we denote $\mathbf{a}_0^\Gamma = \mathbf{a}$. Every sequence $\mathbf{a}_p^\Gamma$ $(1-q\leqslant p\leqslant q-1)$ in $\mathcal{A}_\Gamma$ is called a rotatable sequence (RS). In particular, $\mathcal{A}_\Gamma$ is called a (q,Γ,L,P)-RSS if $\text{PMEPR}(\mathbf{a}_p^\Gamma)\leqslant P$ for a given P. In line with the introduction, Γ can be regarded as the pilot index set and $q = dim(\mathbf{P})$, where $\mathbf{P}$ is the MIMO training mapping cover sequence matrix generally called the $\mathbf{P}$ matrix.

Proposition 1. For given two q-th rotatable index sets Γ and Γ', let $\Gamma' = \{0,1,...,L-1\}\setminus\Gamma$, we have $\mathbf{a}_p^\Gamma = \mathbf{b}_{-p}^{\Gamma'}$, where $\mathbf{b} = \mathbf{a}\cdot\xi_q^p = \left(a(0)\cdot\xi_q^p, a(1)\cdot\xi_q^p, ..., a(L-1)\cdot\xi_q^p\right)$.

Note that $\text{PMEPR}(\mathcal{A}_\Gamma)$ may become much larger when the rotatable index set Γ is unsuitable for $\mathbf{a}$, which can be observed from the following example.

Example 1.

$$\mathbf{P} = (P(i,j))_{i,j=1}^4 = \begin{pmatrix}1 & -1 & 1 & 1\\ 1 & 1 & -1 & 1\\ 1 & 1 & 1 & -1\\ -1 & 1 & 1 & 1\end{pmatrix}$$

is the $\mathbf{P}$ matrix for MIMO-OFDM with 4 transmit antennas and 4 spatial streams, and

$$\mathbf{R} = (R(i,j))_{i,j=1}^4 = \begin{pmatrix}1 & -1 & 1 & 1\\ 1 & -1 & 1 & 1\\ 1 & -1 & 1 & 1\\ 1 & -1 & 1 & 1\end{pmatrix}$$

is the corresponding receive signal matrix $\mathbf{R}$ consisting of the first row of $\mathbf{P}$ matrix. Let

$$\mathbf{a} = (a(i))_{i=0}^{15} = (1,1,1,-1,1,1,-1,1,1,1,1,-1,-1,-1,1,-1)$$

be a preamble sequence with PMEPR($\mathbf{a}$)=1.7071 and $\Gamma = \{3, 6, 11, 12, 13, 15\}$ be the pilot tones. Then the preamble sequence for the second spatial stream on the second transmit antenna is

$$\mathbf{a}_1^{\Gamma} = \left(a_1^{\Gamma}(i)\right)_{i=0}^{15} = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$$

with PMEPR($\mathbf{a}_1^{\Gamma}$)=16 and it is terrible and intolerable for MIMO-OFDM systems, where

$$a_1^{\Gamma}(i) = \begin{cases} R(2,2) \cdot a(i) = -a(i), & i \in \Gamma \\ P(2,2) \cdot a(i) = a(i), & \text{otherwise} \end{cases}$$

and it is the corresponding sequence defined in (A7) when $q = 2$ and $p = 1$.

Example 1 motivates us to search some RSSs with low PMEPRs.

Appendix B Proposed Low-PMEPR Training Sequences for Rotatable Pilots

Appendix B.1 Generalized Boolean Functions

For $\mathbf{x} = (x_1, x_2, \cdots, x_m) \in \mathbb{Z}_2^m$, a generalized Boolean function $f(\mathbf{x})$ is defined as a mapping f from $\mathbb{Z}_2^m$ to $\mathbb{Z}_q$:

$$f(\mathbf{x}) = \sum_{i=0}^{2^m-1} a_i \prod_{k=1}^{m} x_k^{i_k}, \ a_i \in \mathbb{Z}_q$$

, where $(i_1, i_2, \cdots, i_m)$ is the binary representation of the integer $i = \sum_{k=1}^{m} 2^{k-1} i_k$. For any given $f(\mathbf{x})$, we can define a sequence

$$\mathbf{f} = (f(0), f(1), \cdots, f(2^m - 1)) = (f(0, 0, \cdots, 0), f(1, 0, \cdots, 0), \cdots, f(1, 1, \cdots, 1)).$$

One can naturally associate a complex-valued sequence $\psi(\mathbf{f})$ of length L with $\mathbf{f}$ as

$$\psi(\mathbf{f}) = (\xi_q^{f(0)}, \xi_q^{f(1)}, \cdots, \xi_q^{f(L-1)}). \tag{B1}$$

Lemma 2 ([2]). Let q be an even integer and m be a positive integer. Let

$$\begin{aligned} a(\mathbf{x}) &= \frac{q}{2} \sum_{k=1}^{m-1} x_{\pi(k)} x_{\pi(k+1)} + \sum_{k=1}^{m} c_k x_k + c \\ b(\mathbf{x}) &= a(\mathbf{x}) + \frac{q}{2} x_{\pi(1)} \\ c(\mathbf{x}) &= a(\mathbf{x}) + \frac{q}{2} x_{\pi(m)} \end{aligned}$$

where π is a permutation of $\{1, 2, \cdots, m\}$, and $\mathbf{x} \in \mathbb{Z}_2^m, c_k, c \in \mathbb{Z}_q$. Then $\{\psi(\mathbf{a}), \psi(\mathbf{b})\}$ and $\{\psi(\mathbf{a}), \psi(\mathbf{c})\}$ are GCPs of length 2^m with PMEPRs upper bounded by 2.

Lemma 3 ([4]). Let $m \geqslant 2$ be a positive integer and q be an even integer. Let

$$\begin{aligned} a(\mathbf{x}) &= \frac{q}{2} \sum_{k=1}^{m-1} x_{\pi(k)} x_{\pi(k+1)} + \lambda x_{\pi(t)} x_{\pi(t+1)} + \sum_{k=1}^{m} c_k x_k + c \\ b(\mathbf{x}) &= a(\mathbf{x}) + \frac{q}{2} x_{\pi(t)} \\ c(\mathbf{x}) &= a(\mathbf{x}) + \frac{q}{2} x_{\pi(t+1)} \\ d(\mathbf{x}) &= a(\mathbf{x}) + \frac{q}{2} x_{\pi(t)} + \frac{q}{2} x_{\pi(t+1)} \end{aligned}$$

where π is a permutation of $\{1, 2, ..., m\}$, $\mathbf{x} \in \mathbb{Z}_2^m, c_k, c, \lambda \in \mathbb{Z}_q$, and $1 \leqslant t \leqslant m-1$. Then $\{\psi(\mathbf{a}), \psi(\mathbf{b}), \psi(\mathbf{c}), \psi(\mathbf{d})\}$ is a CSS of length $L = 2^m$.

Lemma 4 ([4]). Let $m \geqslant 2$ and $2 \leqslant r \leqslant m$ be positive integers and q be an even integer. Let

$$f(\mathbf{x}) = \frac{q}{2} \sum_{k=1}^{m-1} x_{\pi(k)} x_{\pi(k+1)} + \lambda x_{\pi(m)} \prod_{k \in \mathcal{S}} x_{\pi(k)} + \sum_{k=1}^{m} c_k x_k + c \tag{B2}$$

, where π is a permutation of $\{1, 2, ..., m\}$, $\mathcal{S} = \{s_1, s_2, ..., s_{r-1}\}$ is any subset of $\{1, 2, ..., m-1\}$ with $|\mathcal{S}| = r-1$, and $\mathbf{x} \in \mathbb{Z}_2^m, \lambda, c_k, c \in \mathbb{Z}_q$. Then $\{\psi(\mathbf{S}_d) : 0 \leqslant d \leqslant 2^r - 1\}$ is a CSS of length $L = 2^m$ and size 2^r where $\mathbf{S}_d = \mathbf{f} + \frac{q}{2} \sum_{k=1}^{r-1} d_k x_{\pi(s_k)} + \frac{q}{2} d_r x_{\pi(m)}$, and $d = \sum_{k=1}^{r} d_k 2^{k-1}$ is the binary representation of integer d.

Appendix B.2 Proof of *Construction 1*

Proof.

1. Since for any $1-q \leqslant p \leqslant q-1$, $1 \leqslant k \leqslant m$, $\mathbf{a}_p^{\Gamma_k}$ is given by

$$a_p^{\Gamma_k}(i) = \begin{cases} \xi_q^{a(i)}, & i_k = 0 \\ \xi_q^{a(i)+p}, & i_k = 1 \end{cases} \tag{B3}$$

then it is easy to verify that $\mathbf{a}_p^{\Gamma_k}$ is the corresponding complex-valued sequence of $g(\mathbf{x}) = a(\mathbf{x}) + px_k$. Let $g_2(\mathbf{x}) = g(\mathbf{x}) + \frac{q}{2}x_{\pi(m)}$, then according to **Lemma 2**, we have $(\psi(\mathbf{g}), \psi(\mathbf{g}_2))$ is a GCP. Hence, by **Lemma 1**, we have that $\mathcal{A}_{\Gamma_k}$ is a $(q, \Gamma_k, 2^m, 2)$-RSS.

2. Since for any $1-q \leqslant p \leqslant q-1$, $1 \leqslant t \leqslant m-1$, $\mathbf{a}_p^{\Gamma_t}$ is given by

$$a_p^{\Gamma_t}(i) = \begin{cases} \xi_q^{a(i)}, & i_{\pi(t)}i_{\pi(t+1)} = 0 \\ \xi_q^{a(i)+p}, & i_{\pi(t)}i_{\pi(t+1)} = 1 \end{cases} \tag{B4}$$

then it is easy to verify that $\mathbf{a}_p^{\Gamma_t}$ is the corresponding complex-valued sequence of $g(\mathbf{x}) = a(\mathbf{x}) + (\lambda+p)x_{\pi(t)}x_{\pi(t+1)}$. Let $g_2(\mathbf{x}) = g(\mathbf{x}) + \frac{q}{2}x_{\pi(t+1)}$, $g_3(\mathbf{x}) = g(\mathbf{x}) + \frac{q}{2}x_{\pi(t)}$, $g_4(\mathbf{x}) = g(\mathbf{x}) + \frac{q}{2}x_{\pi(t+1)} + \frac{q}{2}x_{\pi(t)}$, then according to **Lemma 3**, we have $\{\psi(\mathbf{g}), \psi(\mathbf{g}_2), \psi(\mathbf{g}_3), \psi(\mathbf{g}_4)\}$ is a CSS. Hence, by **Lemma 1**, we have that $\mathcal{A}_{\Gamma_t}$ is a $(q, \Gamma_t, 2^m, 4)$-RSS.

3. Since for any $1-q \leqslant p \leqslant q-1$, $\mathbf{a}_p^{\Gamma}$ is given by

$$a_p^{\Gamma_k}(i) = \begin{cases} \xi_q^{a(i)}, & i_{\pi(m)}\prod_{k\in\mathcal{S}} i_{\pi(k)} = 0 \\ \xi_q^{a(i)+p}, & i_{\pi(m)}\prod_{k\in\mathcal{S}} i_{\pi(k)} = 1 \end{cases} \tag{B5}$$

then it is easy to verify that $\mathbf{a}_p^{\Gamma}$ is the corresponding complex-valued sequence of $f(\mathbf{x}) = g(\mathbf{x}) + (\lambda+p)x_{\pi(m)}\prod_{k\in\mathcal{S}} x_{\pi(k)}$, define $\mathbf{S}_d = \mathbf{f} + \frac{q}{2}\sum_{k=1}^{r-1} d_k\mathbf{x}_{\pi(s_k)} + \frac{q}{2}d_r\mathbf{x}_{\pi(m)}$ and $d = \sum_{k=1}^{r} d_k 2^{k-1}$ is the binary representation of integer d, then according to **Lemma 4**, we have $\{\psi(\mathbf{S}_d) : 0 \leqslant d \leqslant 2^r - 1\}$ is a CSS of length $L = 2^m$ and size 2^r. Hence, by **Lemma 1**, we have for any $1-q \leqslant p \leqslant q-1$,

$$\mathrm{PMEPR}(\mathbf{a}_p^{\Gamma}) \leqslant 2^r \tag{B6}$$

, and $\mathcal{A}_{\Gamma}$ is a $(q, \Gamma, 2^m, 2^r)$-RSS.

Example 2. For $q = 4$ and $m = 4$, let π be the identical permutation of $\{1,2,3,4\}$, and $a(\mathbf{x}) = \sum_{k=1}^{3} x_k x_{k+1}$, then

$$\mathbf{a} = (1,1,1,-1,1,1,-1,1,1,1,1,-1,-1,-1,1,-1).$$

(Note that for any given Γ, we have $\mathbf{a}_{-1}^{\Gamma} = \mathbf{a}_3^{\Gamma}$ and $\mathbf{a}_{-2}^{\Gamma} = \mathbf{a}_2^{\Gamma}$.)

- For $\Gamma_1 = \{i : i_1 = 1\}$, $\mathbf{a}_p^{\Gamma_1} = (1, \xi_4^p, 1, -\xi_4^p, 1, \xi_4^p, -1, \xi_4^p, 1, \xi_4^p, 1, -\xi_4^p, -1, -\xi_4^p, 1, -\xi_4^p)$.
- For $\Gamma_1 = \{i : i_1 i_2 = 1\}$, $\mathbf{a}_p^{\Gamma_1} = (1,1,1,-\xi_4^p,1,1,-1,\xi_4^p,1,1,1,-\xi_4^p,-1,-1,1,-\xi_4^p)$.
- For $\Gamma_1 = \{i : i_1 i_2 i_4 = 1\}$, $\mathbf{a}_p^{\Gamma_1} = (1,1,1,-1,1,1,-1,1,1,1,1,-\xi_4^p,-1,-1,1,-\xi_4^p)$. where the elements marked in red are different from those in $\mathbf{a}$ where $1 \leqslant p \leqslant 3$.

Table B1 shows the PMEPRs of $\mathbf{a}_p^{\Gamma_k}$ for different rotatable index sets Γ_k, which are consistent with *Construction 1*.

Table B1 PMEPR($\mathbf{a}_p^{\Gamma}$) in **Example 2** for $p = 1,2,3$ and different Γ

	$p=0$	$p=1$	$p=2$	$p=3$
$\Gamma = \{i : i_1 = 1\}$	1.707	1.707	1.707	1.707
$\Gamma = \{i : i_1 i_2 = 1\}$	1.707	2.914	4.000	2.914
$\Gamma = \{i : i_1 i_2 i_4 = 1\}$	1.707	2.910	4.000	2.910

Remark 1. The third construction in *Construction 1* presents a construction of CSS with power-of-two lengths and size 2^r where $r \geqslant 2$. Note that when $r = 2$ and $S = \{m-1\}$, the CSS generated by the third construction is the same as that of the second construction by setting $t = m-1$.

Appendix B.3 The Code Rate and Minimum Hamming Distance of Proposed Sequences

In this subsection, first, we compute the set sizes of the codebooks constructed in *Construction 1*. These results can then be immediately applied to calculate the code rates of our proposed training sequences. Then the minimum Hamming distances of these codebooks are discussed.

Appendix B.3.1 *The Code Rates of Proposed Codebooks*

Lemma 5 ([2]). With the same notation as *Construction 1*, let

$$\mathcal{F}_1[\mathbf{x}] \triangleq \left\{ \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + \sum_{k=1}^{m} c_k x_k + c : c_k, c \in \mathbb{Z}_q \right\} \tag{B7}$$

, and $\mathcal{C}_1 = \{\psi(\mathbf{f}) : f(\mathbf{x}) \in \mathcal{F}_1[\mathbf{x}]\}$. Then $|\mathcal{C}_1| = \frac{m!}{2} q^{m+1}$.

Theorem 1. With the same notation as *Construction 1*, let

$$\mathcal{F}_2[\mathbf{x}] \triangleq \left\{ \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + \lambda x_{\pi(t)}x_{\pi(t+1)} + \sum_{k=1}^{m} c_k x_k + c : c_k, c, \lambda \in \mathbb{Z}_q, 1 \leqslant t \leqslant m-1 \right\} \tag{B8}$$

, and $\mathcal{C}_2 = \{\psi(\mathbf{f}) : f(\mathbf{x}) \in \mathcal{F}_2[\mathbf{x}]\}$. Then $|\mathcal{C}_2| = \frac{m!}{2}(m-1)q^{m+2}$.

Proof. Suppose that there exist $f_1(\mathbf{x}) \neq f_2(\mathbf{x}) \in \mathcal{F}_2[\mathbf{x}]$, where

$$f_1(\mathbf{x}) = \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + \lambda x_{\pi(t)}x_{\pi(t+1)} + \sum_{k=1}^{m} c_k x_k + c \tag{B9}$$

$$f_2(\mathbf{x}) = \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi'(k)}x_{\pi'(k+1)} + \lambda' x_{\pi'(t')}x_{\pi'(t'+1)} + \sum_{k=1}^{m} c'_k x_k + c' \tag{B10}$$

such that $\psi(\mathbf{f}_1) = \psi(\mathbf{f}_2)$, i.e., $f_1(i) \equiv f_2(i) \pmod q$ $(0 \leqslant i \leqslant 2^m - 1)$.

When $i = 0$, we have $f_1(i) = c$ and $f_1(i) = c'$, which implies that $c = c'$.

When $i = 2^{k-1}$ $(1 \leqslant k \leqslant m)$, we have $f_1(i) = c_k + c$ and $f_2(i) = c'_k + c'$. Since $c = c'$, then we have $c_k = c'_k$ where $1 \leqslant k \leqslant m$.

According to the facts that $c = c'$ and $c_k = c'_k$ $(1 \leqslant k \leqslant m)$, it is straightforward that

$$f_1(\mathbf{x}) - f_2(\mathbf{x}) = \frac{q}{2}\sum_{k=1}^{m-1}\left(x_{\pi(k)}x_{\pi(k+1)} - x_{\pi'(k)}x_{\pi'(k+1)}\right) + \lambda x_{\pi(t)}x_{\pi(t+1)} - \lambda' x_{\pi'(t')}x_{\pi'(t'+1)}. \tag{B11}$$

If $\pi \neq \pi'$, then there exists $a \neq t$ and $a \in \{1, 2, ..., m-1\}$, such that $(\pi(a), \pi(a+1)) \neq (\pi'(k), \pi'(k+1))$ for $1 \leqslant k \leqslant m-1$. Let $i = 2^{\pi(a)-1} + 2^{\pi(a+1)-1}$, then $f_1(i) - f_2(i) = \frac{q}{2} \neq 0 \pmod q$, which contradicts the assumption that $f_1(i) \equiv f_2(i) \pmod q$. So, we have $\pi = \pi'$, and $f_1(\mathbf{x}) - f_2(\mathbf{x}) = \lambda x_{\pi(t)}x_{\pi(t+1)} - \lambda' x_{\pi(t')}x_{\pi(t'+1)}$. It can be easily get that $f_1(i) - f_2(i) \equiv 0 \pmod q$ for all $0 \leqslant i \leqslant 2^m - 1$ if and only if $\lambda = \lambda' = 0$ or $(\lambda, t) = (\lambda', t')$, which implies $f_1(\mathbf{x}) = f_2(\mathbf{x})$ and contradicts the the assumption that $f_1(\mathbf{x}) \neq f_2(\mathbf{x})$. Hence, $|\mathcal{C}_2| = |\mathcal{F}_2[\mathbf{x}]| = \frac{m!}{2}(m-1)q^{m+2}$.

Theorem 2. With the same notation as *Construction 1*, let

$$\mathcal{F}_3[\mathbf{x}] \triangleq \left\{ \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + \lambda x_{\pi(m)}\prod_{k\in\mathcal{S}} x_{\pi(k)} + \sum_{k=1}^{m} c_k x_k + c : c_k, c, \lambda \in \mathbb{Z}_q \right\} \tag{B12}$$

, and $\mathcal{C}_3 = \{\psi(\mathbf{f}) : f(\mathbf{x}) \in \mathcal{F}_3[\mathbf{x}]\}$. Then

$$|\mathcal{C}_3| = \begin{cases} \frac{m!}{2}q^{m+1}\left[\frac{m}{2} + \frac{1}{m} + (q-2)(m-1)\right], & r = 2 \\ C_{m-1}^{r-1}\frac{m!}{2}q^{m+2} & r > 2 \end{cases} \tag{B13}$$

where C_{m-1}^{r-1} is the combinatorial number.

Proof. With the similar argument as **Lemma 1**, we have $|\mathcal{C}_3| = |\mathcal{F}_3[\mathbf{x}]|$. When $r > 2$, with the same argument as **Theorem 1**, it can be easily obtained that $|\mathcal{F}_3[\mathbf{x}]| = C_{m-1}^{r-1}\frac{m!}{2}q^{m+2}$.

When $r = 2$, given that π be a permutation of $\{1, 2, ..., m\}$ and $\mathcal{S} = \{s\}$ where $s \in \{1, 2, ..., m-1\}$, $\mathcal{F}_3[\mathbf{x}]$ can be divided into three disjoint subsets:$\mathcal{F}_{31}[\mathbf{x}]$, $\mathcal{F}_{32}[\mathbf{x}]$ and $\mathcal{F}_{33}[\mathbf{x}]$, where

$$\mathcal{F}_{31}[\mathbf{x}] \triangleq \left\{ \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + \sum_{k=1}^{m} c_k x_k + c : c_k, c \in \mathbb{Z}_q \right\} \tag{B14}$$

$$\mathcal{F}_{32}[\mathbf{x}] \triangleq \left\{ \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + \frac{q}{2}x_{\pi(m)}x_{\pi(s)} + \sum_{k=1}^{m} c_k x_k + c : c_k, c \in \mathbb{Z}_q \right\} \tag{B15}$$

$$\mathcal{F}_{33}[\mathbf{x}] \triangleq \left\{ \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + \lambda x_{\pi(m)}x_{\pi(s)} + \sum_{k=1}^{m} c_k x_k + c : c_k, c, \lambda \in \mathbb{Z}_q, \lambda \neq \frac{q}{2}, \lambda \neq 0 \right\} \tag{B16}$$

It is straightforward that $|\mathcal{F}_{31}[\mathbf{x}]| = \frac{m!}{2}q^{m+1}$ and $|\mathcal{F}_{33}[\mathbf{x}]| = \frac{m!}{2}(q-2)(m-1)q^{m+1}$.

Let $f_1(\mathbf{x}), f_2(\mathbf{x}) \in \mathcal{F}_{32}[\mathbf{x}]$, and

$$f_1(\mathbf{x}) = \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + \frac{q}{2}x_{\pi(m)}x_{\pi(s)} + \sum_{k=1}^{m} c_k x_k + c \tag{B17}$$

$$f_2(\mathbf{x}) = \frac{q}{2}\sum_{k=1}^{m-1} x_{\pi'(k)}x_{\pi'(k+1)} + \frac{q}{2}x_{\pi'(m)}x_{\pi'(s')} + \sum_{k=1}^{m} c'_k x_k + c' \tag{B18}$$

Then, we have $f_1(\mathbf{x}) = f_2(\mathbf{x})$ if and only if $c = c', c_k = c'_k$ $(1 \leqslant k \leqslant m)$ and

$$\sum_{k=1}^{m-1} x_{\pi(k)}x_{\pi(k+1)} + x_{\pi(m)}x_{\pi(s)} = \sum_{k=1}^{m-1} x_{\pi'(k)}x_{\pi'(k+1)} + x_{\pi'(m)}x_{\pi'(s')}. \tag{B19}$$

Eq. (B19) holds if and only if $\pi = \pi', s = s'$ or

$$\begin{cases} s = 1,\ s' \in \{2, 3, ..., m\}; & \\ \pi'(k) = \pi(k + s' - 1), & 1 \leqslant k \leqslant m + 1 - s'; \\ \pi'(k) = \pi\left(k + s' - m - 1\right), & m + 2 - s' \leqslant k \leqslant m; \end{cases} \tag{B20}$$

or

$$\begin{cases} s = s', & 2 \leqslant s \leqslant m - 1; \\ \pi'(k) = \pi(k), & 1 \leqslant k \leqslant s; \\ \pi'(k) = \pi\left(m - (k - (s + 1))\right), & s + 1 \leqslant k \leqslant m; \end{cases} \tag{B21}$$

note that when $s = m - 1$, (B21) results in $\pi = \pi'$. Hence, $|\mathcal{F}_{32}[\mathbf{x}]| = \frac{m!}{2} q^{m+1} \left(\frac{m-2}{2} + \frac{1}{m}\right)$, and

$$\begin{aligned} |\mathcal{F}_3[\mathbf{x}]| &= |\mathcal{F}_{31}[\mathbf{x}]| + |\mathcal{F}_{32}[\mathbf{x}]| + |\mathcal{F}_{33}[\mathbf{x}]| \\ &= \frac{m!}{2} q^{m+1} + \frac{m!}{2} q^{m+1} \left(\frac{m-2}{2} + \frac{1}{m}\right) + \frac{m!}{2}(q-2)(m-1)q^{m+1} \\ &= \frac{m!}{2} q^{m+1} \left[\frac{m}{2} + \frac{1}{m} + (q-2)(m-1)\right]. \end{aligned} \tag{B22}$$

This completes the proof.

The code rates of our proposed codebooks are then given by the following corollaries which are direct results obtained from **Lemma 5**, **Theorem 1**, and **Theorem 2**.

Corollary 1. With the same notation as **Lemma 5**, the code rate of $\mathcal{C}_1$ equals $R(\mathcal{C}_1) = \frac{m+1+\lfloor \log_q (m!/2) \rfloor}{2^m}$.

Corollary 2. With the same notation as **Lemma 1**, the code rate of $\mathcal{C}_2$ equals

$$R(\mathcal{C}_2) = \frac{m + 2 + \lfloor \log_q \left[(m!/2)(m-1)\right] \rfloor}{2^m}. \tag{B23}$$

Corollary 3. With the same notation as **Lemma 2**, the code rate of $\mathcal{C}_3$ equals

$$R(\mathcal{C}_3) = \begin{cases} \frac{m+1+\left\lfloor \log_q \left\{ (m!/2)\left[\frac{m}{2} + \frac{1}{m} + (q-2)(m-1)\right] \right\} \right\rfloor}{2^m}, & r = 2, \\ \frac{m+2+\left\lfloor \log_q \left[(m!/2) C_{m-1}^{r-1}\right] \right\rfloor}{2^m}, & r > 2. \end{cases} \tag{B24}$$

Appendix B.3.2 *The Minimum Hamming Distances of Proposed Codebooks*

The minimum Hamming distance of $\mathcal{C}$ is identified by

$$d_{min}(\mathcal{C}) = min\{d\left(F_1, F_2\right) : F_1 \neq F_2, F_1, F_2 \in \mathcal{C}\}, \tag{B25}$$

where $d\left(F_1, F_2\right) = \omega\left(F_1 - F_2\right)$, and $\omega\left(F_1 - F_2\right)$ denotes the Hamming weight of the vector $F_1 - F_2$.

The generalized rth-order Reed-Muller code [2] of length 2^m, denoted by $\mathrm{RM}_q(r; m)$, is defined as the linear code over $\mathbb{Z}_q$ generated by the monomials in the Boolean functions x_i of degree at most r.

The minimum Hamming distance of $\mathrm{RM}_q(r; m)$ is given in [2].

Lemma 6 ([2])**.** For $q \geqslant 2$, the generalized rth-order Reed-Muller code, $\mathrm{RM}_q(r; m)$, has minimum Hamming distance 2^{m-r}.

Theorem 3. With the same notation as **Lemma 5**, we have that $d_{min}(\mathcal{C}_1) = 2^{m-2}$.

Proof. Since $\mathcal{C}_1 \subseteq \mathrm{RM}_q(2; m)$ and $d_{min}(\mathrm{RM}_q(2; m)) = 2^{m-2}$, which can be obtained from **Lemma 6**, then, we have $d_{min}(\mathcal{C}_1) \geqslant d_{min}(\mathrm{RM}_q(2; m)) = 2^{m-2}$. To demonstrate that $d_{min}(\mathcal{C}_1) = 2^{m-2}$, it is sufficient to prove that there exist $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}_1$, such that $d(\mathbf{c}_1, \mathbf{c}_2) = 2^{m-2}$. Let

$$f_1(\mathbf{x}) = \frac{q}{2}\left(x_1x_2 + x_2x_3 + \cdots + x_{m-2}x_{m-1} + x_{m-1}x_m\right) \tag{B26}$$

$$f_2(\mathbf{x}) = \frac{q}{2}\left(x_1x_2 + x_2x_3 + \cdots + x_{m-2}x_m + x_mx_{m-1}\right) \tag{B27}$$

and $\mathbf{c}_1 = \psi(\mathbf{f}_1), \mathbf{c}_2 = \psi(\mathbf{f}_2)$. It can be easily observed that

$$d(\mathbf{c}_1, \mathbf{c}_2) = \omega\left(\frac{q}{2}\left(\mathbf{x}_{m-2}\mathbf{x}_{m-1} - \mathbf{x}_{m-2}\mathbf{x}_m\right)\right) = 2^{m-2}. \tag{B28}$$

This completes the proof.

Corollary 4. Let $\mathcal{C}_2$ and $\mathcal{C}_3$ be as defined in **Theorem 1** and **Theorem 2**, respectively. Then $d_{min}(\mathcal{C}_2) = 2^{m-2}$ and $d_{min}(\mathcal{C}_3) = 2^{m-r}$.

Appendix C General RSS Construction Methods for Small 4-th Rotatable Index Sets with Low PMEPRs

Definition 4 (Insertion Function). Consider a sequence $\mathbf{a} = (a(0), a(1), \ldots, a(L-1))$ of length N. Then $\mathfrak{I}(\mathbf{a}, r, x)$, given by

$$\mathfrak{I}(\mathbf{a}, r, x) = \begin{cases} (x, a(0), a(1), \ldots, a(L-1)), & r = 0 \\ (a(0), a(1), \ldots, a(L-1), x), & r = L \\ (a(0), \ldots, a(r-1), x, a(r), \ldots, a(L-1)), & 0 < r < L \end{cases} \tag{C1}$$

is an insertion function which generates a sequence of length $(L+1)$ with an additional element x at the r-th position [5, 6].

To prove *Construction 2*, we need the following results.

Lemma 7 (**Theorem 1** of [6]). Consider a GCP $(\mathbf{a}, \mathbf{b})$ of length L. Let $(\mathbf{c}, \mathbf{d})$ be its Golay mate. If $x_1, y_1, x_2, y_2 \in \mathbb{U} = \{1, \xi_4^1, \xi_4^2, \xi_4^3\}$ and $\mathbf{e} = \mathfrak{I}(\mathbf{a}, 0, x_1)$, $\mathbf{f} = \mathfrak{I}(\mathbf{b}, 0, y_1)$, $\mathbf{g} = \mathfrak{I}(\mathbf{c}, L, x_2)$, $\mathbf{h} = \mathfrak{I}(\mathbf{d}, L, y_2)$. Then $\mathbf{A} = [\mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}]^T$ is a CSS of set size 4 and sequence length $L+1$ when $x_1 = y_2$ and $y_1 = -x_2$.

Lemma 8 (**Theorem 2** of [6]). Consider a GCP $(\mathbf{a}, \mathbf{b})$ of length L. Let $(\mathbf{c}, \mathbf{d})$ be its Golay mate. Also let $\mathbf{e} = \mathbf{a}||\mathbf{c}$ and $\mathbf{f} = \mathbf{b}||\mathbf{d}$ be sequences of length $2L$. If $x_1, y_1, x_2, y_2 \in \mathbb{U} = \{1, \xi_4^1, \xi_4^2, \xi_4^3\}$ and $\mathbf{p} = \mathfrak{I}(\mathbf{e}, L, x_1)$, $\mathbf{q} = \mathfrak{I}(\mathbf{f}, L, y_1)$, $\mathbf{r} = \mathfrak{I}(\mathbf{e}, L, x_2)$, $\mathbf{s} = \mathfrak{I}(\mathbf{f}, L, y_2)$. Then $\mathbf{A} = [\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}]^T$ is a CSS of set size 4 and sequence length $2L+1$ when the following conditions hold.

$$x_1^* - y_1 + x_2^* - y_2 = 0, \quad x_1 + y_1^* + x_2 + y_2^* = 0. \tag{C2}$$

Lemma 9 (**Theorem 3** of [6]). Consider a GCP $(\mathbf{a}, \mathbf{b})$ of length L. Let $(\mathbf{c}, \mathbf{d})$ be its Golay mate. Also, let $x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4 \in \mathbb{U} = \{1, \xi_4^1, \xi_4^2, \xi_4^3\}$ and $\mathbf{e} = \mathfrak{I}(\mathbf{a}, 0, x_1)$, $\mathbf{f} = \mathfrak{I}(\mathbf{b}, 0, y_1)$, $\mathbf{g} = \mathfrak{I}(\mathbf{c}, 0, x_3)$, $\mathbf{h} = \mathfrak{I}(\mathbf{d}, 0, y_3)$. If $\mathbf{p} = \mathfrak{I}(\mathbf{e}, L+1, x_2)$, $\mathbf{q} = \mathfrak{I}(\mathbf{f}, L+1, y_2)$, $\mathbf{r} = \mathfrak{I}(\mathbf{g}, L+1, x_4)$, $\mathbf{s} = \mathfrak{I}(\mathbf{h}, L+1, y_4)$. Then $\mathbf{A} = [\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}]^T$ is a CSS of set size 4 and sequence length $L+2$ when the following set of conditions hold

$$x_1 - y_4^* = 0; \; x_2^* - y_3 = 0; \; y_1 + x_4^* = 0; \; y_2^* + x_3 = 0. \tag{C3}$$

Appendix C.1 Proof of *Construction 2*

Proof.

1. For any $c \in \mathbb{Z}_4$, let $\mathbf{e} = \mathfrak{I}(\mathbf{a}, 0, \xi_q^c)$, $\mathbf{e}_1 = \mathfrak{I}(\mathbf{b}, 0, \xi_q^c)$, $\mathbf{e}_2 = \mathfrak{I}(\mathbf{c}, L, -\xi_q^c)$ and $\mathbf{e}_3 = \mathfrak{I}(\mathbf{d}, L, \xi_q^c)$, then according to **Lemma 7**, it can be obtained that $\{\mathbf{e}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is a CSS of length $L+1$. Hence, $\mathcal{E}_\Gamma$ is a $(4, \Gamma, L+1, 4)$-RSS, where $\Gamma = \{0\}$ and $\mathcal{E}_\Gamma$ is the RSS of $\mathbf{e}$ corresponding to Γ.

2. For any $c \in \mathbb{Z}_4$, let $\mathbf{h} = \mathfrak{I}(\mathbf{a}, L, \xi_q^c)$, $\mathbf{h}_1 = \mathfrak{I}(\mathbf{b}, L, \xi_q^c)$, $\mathbf{h}_2 = \mathfrak{I}(\mathbf{c}, 0, -\xi_q^c)$ and $\mathbf{h}_3 = \mathfrak{I}(\mathbf{d}, 0, \xi_q^c)$, then according to **Lemma 7**, we have $\{\mathbf{h}, \mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3\}$ is a CSS of length $L+1$. Hence, $\mathcal{H}_\Gamma$ is a $(4, \Gamma, L+1, 4)$-RSS, where $\Gamma = \{L\}$ and $\mathcal{H}_\Gamma$ is the RSS of $\mathbf{h}$ corresponding to Γ.

3. For any $c \in \mathbb{Z}_4$, let $\mathbf{k} = \mathfrak{I}(\mathbf{w}, L, \xi_q^c)$, $\bar{\mathbf{w}} = (\mathbf{b} \parallel \mathbf{d})$ and $\mathbf{k}_1 = \mathfrak{I}(\bar{\mathbf{w}}, L, \xi_q^{-c})$. Furthermore, let $\mathbf{k}_2 = \mathfrak{I}(\mathbf{w}, L, -\xi_q^c)$ and $\mathbf{k}_3 = \mathfrak{I}(\bar{\mathbf{w}}, L, -\xi_q^{-c})$, then according to **Lemma 8**, we have $\{\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3\}$ is a CSS of length $2L+1$. Hence, $\mathcal{K}_\Gamma$ is a $(4, \Gamma, 2L+1, 4)$-RSS, where $\Gamma = \{L\}$ and $\mathcal{K}_\Gamma$ is the RSS of $\mathbf{k}$ corresponding to Γ.

4. For any $c \in \mathbb{Z}_4$, let $\mathbf{e} = \mathfrak{I}(\mathbf{a}, 0, \xi_q^c)$, $\mathbf{e}_1 = \mathfrak{I}(\mathbf{b}, 0, -\xi_q^{-c})$, $\mathbf{e}_2 = \mathfrak{I}(\mathbf{c}, 0, \xi_q^c)$ and $\mathbf{e}_3 = \mathfrak{I}(\mathbf{d}, 0, \xi_q^{-c})$. Also, let $\mathbf{m} = \mathfrak{I}(\mathbf{e}, L+1, \xi_q^c)$, $\mathbf{m}_1 = \mathfrak{I}(\mathbf{e}_1, L+1, -\xi_q^{-c})$, $\mathbf{m}_2 = \mathfrak{I}(\mathbf{e}_2, L+1, \xi_q^c)$ and $\mathbf{m}_3 = \mathfrak{I}(\mathbf{e}_3, L+1, \xi_q^{-c})$, it can be verified from **Lemma 9** that $\{\mathbf{m}, \mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3\}$ is a CSS of length $L+2$. Hence, $\mathbf{m} = \mathfrak{I}(\mathbf{e}, L+1, \xi_4)$, $\mathcal{M}_\Gamma$ is a $(4, \Gamma, L+2, 4)$-RSS, where $\Gamma = \{0, L+1\}$ and $\mathcal{M}_\Gamma$ is the RSS of $\mathbf{m}$ corresponding to Γ.

The results can be derived from the argument above along with the condition that $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and $\mathbf{d}$ have the same energy.
This proof is completed.

Appendix D General Methods for Constructing q-th RSSs from CSSs

The proof of *Construction 3* is straightforward from the following result, which is direct from the Rudin-Shapiro construction in [7].

Lemma 10. For M, N are positive integer with $M = 2N$, let $\mathcal{S}^0 = \{\mathbf{s}_1^0, \mathbf{s}_2^0, \cdots, \mathbf{s}_M^0\}$ be a CSS of length L. Then, a CSS $\mathcal{S}^k = \{\mathbf{s}_1^k, \mathbf{s}_2^k, \cdots, \mathbf{s}_M^k\}$ of length $2^k L$ with size M can be constructed using the following two recursive formulas:

1. $\mathbf{s}_{2n-1}^k = \mathbf{s}_{2n-1}^{k-1} \Diamond \mathbf{s}_{2n}^{k-1}$, $\mathbf{s}_{2n}^k = \mathbf{s}_{2n-1}^{k-1} \Diamond - \mathbf{s}_{2n}^{k-1}$
2. $\mathbf{s}_{2n-1}^k = \left(\mathbf{s}_{2n-1}^{k-1} \parallel \mathbf{s}_{2n}^{k-1}\right)$, $\mathbf{s}_{2n}^k = \left(\mathbf{s}_{2n-1}^{k-1} \parallel -\mathbf{s}_{2n}^{k-1}\right)$

, where $n \in \{1, 2, ..., N\}$, "$\Diamond$" denotes the bit-interleaved operation and "$\|$" denotes the concatenation operation.

Remark 2. If the given CSS satisfies the condition in *Construction 3*, new RSSs can be generated by all the general methods of CSSs based on existing CSSs using the given CSSs. The CSSs in **Lemmas 2,3,4** and **Lemma 7** can be used as the seed CSS of *Construction 3*, which can enlarge the sequence length along with larger rotatable index set size. Specially, some classes of RSSs in *Construction 1* can be generated from the second method in *Construction 3* where the parameters of seed CSSs satisfy the following conditions.

1. With the same notation as **Lemma 2**, for $M = 2$, let $\mathbf{s}_1 = \psi(\mathbf{a})$, $\mathbf{s}_2 = \psi(\mathbf{c})$. By the second method in *Construction 3*, the corresponding generalized Boolean function of $\mathbf{s}_1^k$ is

$$\frac{q}{2} \sum_{w=1}^{m-1} x_{\pi(w)} x_{\pi(w+1)} + \frac{q}{2} x_{\pi(m)} x_{m+1} + \frac{q}{2} \sum_{w=1}^{k-1} x_{m+w} x_{m+w+1} + \sum_{w=1}^{m} c_w x_w + c$$

, and it can generate a $(q, \Gamma_1^k, 2^{m+k}, 2)$-RSS, where $\Gamma_{k'} = \{i : i_{k'} = 1\}$, $1 \leqslant k' \leqslant m+k$, $0 \leqslant i \leqslant 2^{m+k} - 1$ and $i_{k'}$ is the k'-th element of the binary representation of index i.

2. With the same notation as **Lemma 3**, for $M = 4$, let $\mathbf{s}_1 = \psi(\mathbf{a})$, $\mathbf{s}_2 = \psi(\mathbf{c})$, $\mathbf{s}_3 = \psi(\mathbf{b})$, $\mathbf{s}_4 = \psi(\mathbf{d})$, where $t = m - 1$. By the second method in *Construction 3*, the corresponding generalized Boolean function of $\mathbf{s}_1^k$ is

$$\frac{q}{2}\sum_{w=1}^{m-1} x_{\pi(w)}x_{\pi(w+1)} + \frac{q}{2}x_{\pi(t)}x_{m+1} + \frac{q}{2}\sum_{w=1}^{k-1} x_{m+w}x_{m+w+1} + \lambda x_{\pi(t)}x_{\pi(t+1)} + \sum_{w=1}^{m} c_w x_w + c$$

, and it can generate a $(q, \Gamma_1^k, 2^{m+k}, 4)$-RSS, where $\Gamma_1^k = \{i : i_{\pi(t)}i_{\pi(t+1)} = 1\}$, $1 \leqslant t \leqslant m-1$, $0 \leqslant i \leqslant 2^{m+k} - 1$ and i_t is the t-th element of the binary representation of index i.

3. With the same notation as **Lemma 4**, for $M = 2^r$, let $\mathbf{s}_{2n-1} = \psi(\mathbf{S}_n)$, $\mathbf{s}_{2n} = \psi(\mathbf{S}_{n+2^{r-1}})$ for $1 \leqslant n \leqslant 2^{r-1}$. By the second method in *Construction 3*, the corresponding generalized Boolean function of $\mathbf{s}_1^k$ is

$$\frac{q}{2}\sum_{w=1}^{m-1} x_{\pi(w)}x_{\pi(w+1)} + \frac{q}{2}x_{\pi(m)}x_{m+1} + \frac{q}{2}\sum_{w=1}^{k-1} x_{m+w}x_{m+w+1} + \lambda x_{\pi(m)}\prod_{w\in\mathcal{S}} x_{\pi(w)} + \sum_{w=1}^{m} c_w x_w + c$$

, and it can generate a $(q, \Gamma, 2^{m+k}, 2^r)$-RSS where $\Gamma = \{i : i_{\pi(m)}\prod_{w\in\mathcal{S}} i_{\pi(w)} = 1\}$, $0 \leqslant i \leqslant 2^{m+k} - 1$ and i_w is the w-th element of the binary representation of index i.

Example 3. Taking $\{(1,1),(1,-1)\}$ as the GCP with $\{(-1,1),(-1,-1)\}$ its Golay mate. Let $\mathbf{e} = (1,1,1)$, $\mathbf{f} = (-1,1,-1)$, $\mathbf{g} = (1,1,-1)$, $\mathbf{h} = (-1,-1,-1)$, and $\Gamma_1 = \Gamma_2 = \{0\}$, $\Gamma_3 = \Gamma_4 = \{2\}$, then it can be easily checked that for any $p \in \mathbb{Z}_4$, $\{\mathbf{e}_p^{\Gamma_1}, \mathbf{f}_p^{\Gamma_2}, \mathbf{g}_p^{\Gamma_3}, \mathbf{h}_p^{\Gamma_4}\}$ is a CSS, where $\mathbf{e}_p^{\Gamma_1} = (\xi_4^p, 1, 1)$, $\mathbf{f}_p^{\Gamma_2} = (-\xi_4^p, 1, -1)$, $\mathbf{g}_p^{\Gamma_3} = (1, 1, -\xi_4^p)$, $\mathbf{h}_p^{\Gamma_4} = (-1, -1, -\xi_4^p)$. Consider $\mathbf{a} = \mathbf{e}\Diamond\mathbf{f} = (1, -1, 1, 1, 1, -1)$ and $\tilde{\mathbf{a}} = \mathbf{e} \parallel \mathbf{f} = (1, 1, 1, -1, 1, -1)$, Tables D1 and D2 show the PMEPR of $\mathbf{a}_p^\Gamma$ and $\tilde{\mathbf{a}}_p^\Gamma$, it can be observed that $\{\mathbf{a}^\Gamma, \mathbf{a}_1^\Gamma, \mathbf{a}_2^\Gamma, \mathbf{a}_3^\Gamma\}$ is a $(4, \{0,1\}, 6, 2.6667)$-RSS and $\{\tilde{\mathbf{a}}^\Gamma, \tilde{\mathbf{a}}_1^\Gamma, \tilde{\mathbf{a}}_2^\Gamma, \tilde{\mathbf{a}}_3^\Gamma\}$ is a $(4, \{0,3\}, 6, 2.6667)$-RSS.

Table D1 PMEPR($\mathbf{a}_p^\Gamma$) in **Example 3** for $p = 1, 2, 3$

	$p=0$	$p=1$	$p=2$	$p=3$
$\Gamma = \{0,1\}$	2.667	1.911	2.000	1.911

Table D2 PMEPR($\tilde{\mathbf{a}}_p^\Gamma$) in **Example 3** for $p = 1, 2, 3$

	$p=0$	$p=1$	$p=2$	$p=3$
$\Gamma = \{0,3\}$	2.667	2.488	2.000	2.488

Appendix E Comparisons

In this section, first, we compare the PMEPRs of the proposed sequences with those of the well-known Zadoff-Chu (ZC) sequences [8] and m-sequences [9], which have been widely used in MIMO-OFDM systems [10–14]. Then we compare the code rates of the proposed sequences with those of the ZC sequences, the m-sequences, and the complementary sequences proposed in [2, 15].

Appendix E.1 Comparison of PMEPRs

Let us consider sequences of lengths 32 and 64. The m-sequence of lengths 32 and 64 is obtained by adding an "1" at the end of the m-sequence of lengths 31 and 64, respectively. Tables E1 and E2 give some numerical results of PMEPRs for difference sequences of lengths 32 and 64, it is obvious that the proposed sequences have much lower PMEPRs than the m-sequences when the rotation is considered as per the given rotatable index set when the sequence length is 32. Although the PMEPRs of our proposed sequences are slightly higher than those of the ZC sequences when $q = 2$ with $\Gamma = \{i : i_1i_2 = 1\}$ and $\Gamma = \{i : i_1i_2i_5 = 1\}$ and when $q = 4$ with $\Gamma = \{i : i_1i_2 = 1\}$, our sequences perform much better in other cases. As a whole, our proposed sequences outperform than the m-sequences and the ZC sequences because of their low and steady PMEPRs for varies q.

Appendix E.2 Comparison of Code Rates

For the proposed sequences in this paper and complementary sequences in [2, 15], we only consider the sequences whose PMEPRs are at most 4. For the ZC sequences and the m-sequences, we compute all the non-equivalent sequences, where the m-sequences of length 2^m is obtained by adding an "1" at the end of the m-sequences of length $2^m - 1$. As shown in Table E3, the proposed sequences have much larger code rates than the m-sequences, the ZC sequences, and the GCSs proposed in [15]. Although the code rate of our proposed sequences is slightly lower than that of the CSs proposed in [2] when $L = 256$, our sequences have larger code rates in other cases.

References

1 Tseng C-C, Liu C. Complementary sets of sequences. IEEE Trans Inf Theory, 1972, 18: 644-652

2 Paterson K G. Generalized Reed-Muller codes and power control in OFDM modulation. IEEE Trans Inf Theory, 2000, 46: 104-120

3 Liu Z, Guan Y L. 16-QAM almost-complementary sequences with low PMEPR. IEEE Trans Commun, 2016, 64: 668-679

4 Chen C-Y, Wang C-H, Chao C-C. Complete complementary codes and generalized Reed-Muller codes. IEEE Commun Lett, 2008, 12: 849-851

Table E1 Comparisons of PMEPR($\mathcal{A}_\Gamma$) for different sequences of length 32

PMEPR / Seq. \ Γ, q	$\Gamma=\{i: i_1=1\}$			$\Gamma=\{i: i_1 i_2=1\}$			$\Gamma=\{i: i_1 i_2 i_5=1\}$		
	$q=2$	$q=4$	$q=6$	$q=2$	$q=4$	$q=6$	$q=2$	$q=4$	$q=6$
Proposed sequence	2.000	2.000	2.000	3.633	3.633	3.633	3.444	3.444	3.444
ZC sequence	3.027	3.027	3.424	3.434	3.434	3.702	3.345	5.027	4.860
m-sequence	3.385	3.385	3.744	6.485	6.485	6.485	3.701	4.526	4.598

Table E2 Comparisons of PMEPR($\mathcal{A}_\Gamma$) for different sequences of length 64

PMEPR / Seq. \ Γ, q	$\Gamma=\{i: i_2=1\}$			$\Gamma=\{i: i_2 i_5=1\}$			$\Gamma=\{i: i_2 i_5 i_6=1\}$		
	$q=2$	$q=4$	$q=6$	$q=2$	$q=4$	$q=6$	$q=2$	$q=4$	$q=6$
Proposed sequence	1.983	2.00	1.997	2.867	3.314	3.473	4.582	4.582	4.582
ZC sequence	2.907	4.304	4.321	7.888	7.888	7.888	6.487	6.487	6.753
m-sequence	3.812	4.405	4.548	3.419	3.943	3.806	3.203	3.423	3.452

Table E3 Comparisons of code rate for binary sequences with various lengths

$R(\mathcal{C})$ / Seq. \ Length	$L=32$	$L=64$	$L=128$	$L=256$
Proposed sequence	0.500	0.297	0.188	0.110
GCS [15]	0.375	0.250	0.156	0.094
CS [2]	0.344	0.266	0.180	0.113
ZC sequence	0.125	0.078	0.047	0.027
m-sequence	0.094	0.047	0.039	0.016

5 Adhikary A R, Majhi S, Liu Z, et al. New sets of even-length binary Z-complementary pairs with asymptotic ZCZ ratio of 3/4. IEEE Signal Process Lett, 2018, 25: 970-973

6 Adhikary A R, Majhi S. New constructions of complementary sets of sequences of lengths non-power-of-two. IEEE Commun Lett, 2019, 23: 1119-1122

7 Stanczak S, Boche H. Aperiodic properties of generalized binary Rudin-Shapiro sequences and some recent results on sequences with a quadratic phase function. In Proceedings of 2000 International Zurich Seminar on Broadband Communications, Accessing, Transmission, Networking. (Cat. No.00TH8475), Zurich, Switzerland, 2000, 279-286

8 Hyder M, Mahata K. Zadoff-Chu sequence design for random access initial uplink synchronization in LTE-like systems. IEEE Trans Wireless Commun, 2017, 16: 503-511

9 Fan P, Darnell M. Sequence design for communications applications. Electronic & electrical engineering research studies: Communications systems, techniques, and applications series. Research Studies Press, 1996.

10 Li Z, Xia X. PAPR reduction for repetition space-time-frequency coded MIMO-OFDM systems using Chu sequences. IEEE Trans Wireless Commun, 2008, 7: 1195-1202

11 Xian J, Gong Y. A novel method in SLM to produce m sequence to reduce the PAPR in MIMO-OFDM system. In: Proceedings of 2009 3rd IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, Beijing, China, 2009, 733-736

12 Wu Y, Attallah S, Bergmans J W M. Carrier frequency offset estimation for multi-user MIMO-OFDM uplink using CAZAC sequences. In: Proceedings of 2009 IEEE Wireless Communications and Networking Conference, Budapest, Hungary, 2009, 1-5

13 Kang J, Whang Y, Lee H, et al. Optimal pilot sequence design for multi-cell MIMO-OFDM systems. IEEE Trans Wireless Commun, 2011, 10: 3354-3367

14 Li L, Zhou P. Synchronization for B3G MIMO-OFDM in DL initial acquisition by CAZAC sequence. In Proceedings of 2006 International Conference on Communications, Circuits and Systems, Tirrenia, Italy, 2006, 2: 1035-1039

15 Davis J A, Jedwab J. Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed-Muller codes. IEEE Trans Inf Theory, 1999, 45: 2397-2417