Received 26 May 2024; accepted 23 June 2024. Date of publication 28 June 2024; date of current version 6 August 2024.
Digital Object Identifier 10.1109/OJCOMS.2024.3420440

Enhanced Cross Z-Complementary Set and Its Application in Generalized Spatial Modulation

ZHEN-MING HUANG (黃振銘) [1,2] **(Graduate Student Member, IEEE), CHENG-YU PAI (白承祐)** [1,3] **(Member, IEEE), ZILONG LIU (劉子龍)** [4] **(Senior Member, IEEE), AND CHAO-YU CHEN (陳昭羽)** [1,3] **(Senior Member, IEEE)**

[1] Institute of Computer and Communication Engineering, National Cheng Kung University, Tainan 701, Taiwan
[2] Department of Engineering Science, National Cheng Kung University, Tainan 701, Taiwan
[3] Department of Electrical Engineering, National Cheng Kung University, Tainan 701, Taiwan
[4] School of Computer Science and Electronic Engineering, University of Essex, CO4 3SQ Colchester, U.K.

CORRESPONDING AUTHOR: C.-Y. CHEN (e-mail: super@mail.ncku.edu.tw)

The work of Zhen-Ming Huang, Cheng-Yu Pai, and Chao-Yu Chen was supported by the National Science and Technology Council, Taiwan, under Grant NSTC 112–2221–E–006–122–MY3 and Grant NSTC 113–2927–I–006–501. The work of Zilong Liu was supported in part by the U.K. Engineering and Physical Sciences Research Council under Grant EP/X035352/1 and Grant EP/Y000986/1; in part by the Royal Society under Grant IEC\R3\223079; and in part by the Research Council of Norway under Grant 311646/O70.

ABSTRACT Generalized spatial modulation (GSM) is a novel multiple-antenna technique offering flexibility among spectral efficiency, energy efficiency, and the cost of RF chains. In this paper, a novel class of sequence sets, called enhanced cross Z-complementary set (E-CZCS), is proposed for efficient and feasible training sequence design in broadband GSM systems. Specifically, an E-CZCS consists of multiple CZCSs possessing front-end and tail-end zero-correlation zones (ZCZs), whereby any two distinct CZCSs have a tail-end ZCZ when a novel type of cross-channel aperiodic correlation sums is considered. The theoretical upper bound on the ZCZ width is first derived, upon which E-CZCSs with maximum ZCZ width and flexible parameters are constructed. For optimal channel estimation over frequency-selective channels, we introduce and evaluate a novel GSM training framework employing the proposed E-CZCSs. Numerical results demonstrate that the proposed E-CZCS-based training can achieve the minimum channel estimation mean square error (MSE) and outperform other classes of sequences.

INDEX TERMS Enhanced cross Z-complementary set (E-CZCS), cross Z-complementary set (CZCS), generalized spatial modulation (GSM), zero correlation zone (ZCZ), generalized Boolean function (GBF), training sequence.

I. INTRODUCTION

A. BACKGROUND

GOLAY complementary pair (GCP), found by Marcel J. E. Golay in the middle of the 20th century, is characterized by the property that the aperiodic autocorrelation sum of the two constituent sequences is zero at every non-zero time-shift [1]. In 1972, Tseng and Liu extended the concept of GCP to Golay complementary set (GCS), each consisting of more than two constituent sequences [2]. Additionally, a collection of GCSs was introduced, called the mutually orthogonal complementary set (MOCS), where any two distinct GCSs in an MOCS have zero aperiodic cross-correlation sums for all time-shifts. In [3], complete complementary code (CCC) was introduced as optimal MOCS with the maximum set size. Owing to these ideal autocorrelation and cross-correlation properties, complementary pairs/sets of sequences have been employed in many communications applications including synchronization [4], channel estimation [5], [6], interference suppression [7], [8], [9], [10], [11], peak-to-mean power control [12], [13], [14], [15], [16], cell search [17], and MIMO radar [18], [19]. As a generalization of MOCSs and CCCs, Z-complementary code set (ZCCS) having zero correlation zone (ZCZ) was proposed in [20].

On the other hand, spatial modulation (SM) has received tremendous research attention in recent years as a novel multiple-antenna technique. In SM, there is only one radio-frequency (RF) chain, whereby one transmit antenna (TA) is activated at each time-slot [21], [22], [23], [24], [25], [26], [27]. Because of this, SM enjoys zero inter-antenna interference (IAI), lower energy consumption, and reduced transceiver complexity. For a long time, efficient channel estimation schemes for SM systems in frequency selective channels were missing. In 2020, Liu et al. proposed cross Z-complementary pair (CZCP) for optimal sparse training matrix design in broadband SM systems [28]. Several constructions of CZCPs with larger ZCZ widths and more flexible lengths have been proposed in [28], [29], [30], [31], [32], [33], [34], [35]. However, the ZCZ width of every CZCP is theoretically upper bounded by a half of its sequence length. In [36], [37], cross Z-complementary set (CZCS) which can tolerate larger delay spreads was developed. Recently, CZCSs with more flexible lengths were presented in [38].

Unlike SM, generalized spatial modulation (GSM) system has been proposed for a higher spectral efficiency as it allows two or more active TAs at the same time [39], [40], [41], [42], [43]. To be specific, the transmitter of a GSM system is equipped with a few RF chains less than the number of TAs. During each transmission, a GSM symbol is modulated using two information parts. The message bits of the first information part are used to select the antenna activation patterns, whereas the second part carries message bits for selecting specific constellation points over those activated TAs. Therefore, GSM provides an excellent trade-off between the spectral efficiency and the cost of RF chains, while retaining most of the advantages of SM.

B. MOTIVATIONS AND CONTRIBUTIONS

Training sequence design for GSM is a more challenging task. First, dense training sequences designed for the traditional multiple-input multiple-output (MIMO) in [44], [45], [46] cannot be used since only a few GSM TAs are activated at each time-slot. Recently, symmetrical Z-complementary code set (SZCCS) was proposed in [47] for GSM training design. It is noted that SZCCS is a subclass of ZCCSs with symmetric ZCZ properties for its autocorrelation and cross-correlation sums. However, the proposed GSM training framework in [47] has an additional overhead for IAI mitigation incurred by zero-padding. Consequently, their approach suffers from a reduced training efficiency. For more efficient training design, CZCP and its mates were proposed as the training sequences in GSM [48]. However, the utilization of CZCP mates is restricted to GSM systems equipped with two active TAs. It is desirable to develop a more feasible training design for GSM systems, not limited to two active TAs. Motivated by [36], [48], we aim to go beyond the CZCP mates and CZCS by introducing new sequence properties for more feasible training design in GSM.

In this paper, we propose a novel family of CZCSs, called enhanced cross Z-complementary set (E-CZCS), each consisting of multiple CZCSs and any two distinct constituent CZCSs possess the following two correlation properties: 1) the aperiodic correlation sums have a front-end ZCZ and a tail-end ZCZ; 2) there is a tail-end ZCZ when a special type of cross-channel aperiodic correlation sums is considered. More specifically, such a tail-end ZCZ is required for the cross-correlation sums between the n-th constituent sequence of one CZCS and the $(n+1)_{\mathrm{mod}\,N}$-th constituent sequence of the other, where N refers to the total number of constituent sequences. Therefore, these additional aperiodic cross-correlation properties between distinct constituent CZCSs can eliminate the IAI caused by the multipath propagation. The major contributions of this paper are summarized as follows:

- We extend the concept of CZCS to E-CZCS by incorporating the aforementioned cross-channel aperiodic ZCZ property. Additionally, we derive an upper bound on the width of the ZCZ.
- Two constructions of E-CZCSs are proposed. The first construction is based on MOCSs, CCCs, and ZCCSs. The second construction is based on generalized Boolean functions. Both constructions can generate binary E-CZCSs with maximum ZCZ width and various set sizes.
- We present a novel training framework employing the proposed E-CZCSs for broadband GSM systems. The proposed GSM training framework can achieve optimal channel estimation over frequency-selective channels. Both IAI and ISI can be eliminated, thanks to the unique correlation properties of the proposed E-CZCSs.
- Simulations show that the proposed E-CZCS-based training scheme can achieve the minimum channel estimation mean square error (MSE) and outperform other classes of sequences, such as SZCCSs, ZCCSs, and Zadoff-Chu sequences.

C. ORGANIZATION OF THIS PAPER

The rest of this paper is organized as follows. In Section II, we first introduce some necessary notations, definitions, and the GSM system. In Section III, we define the E-CZCS and its correlation properties, and then propose two constructions of E-CZCSs. Section IV describes the requirements for training design in the GSM system and proposes a novel training framework based on E-CZCSs. The performance comparison is provided in Section V. Finally, concluding remarks are drawn in Section VI.

II. PRELIMINARIES AND DEFINITIONS

First, we introduce some notations which are used throughout this paper.

A. NOTATIONS

- "$\boldsymbol{a}\|\boldsymbol{b}$" denotes the concatenation of sequences $\boldsymbol{a}$ and $\boldsymbol{b}$;

- "+" and "−" denote 1 and −1, respectively;
- $\xi_q = e^{2\pi j/q}$ is a primitive complex qth root of unity;
- $\boldsymbol{X}^*$ denotes the complex conjugate of the matrix $\boldsymbol{X}$;
- $\boldsymbol{X}^T$ denotes the transpose of the matrix $\boldsymbol{X}$;
- $\boldsymbol{X}^H$ denotes the Hermitian of the matrix $\boldsymbol{X}$;
- $\lfloor\cdot\rfloor$ denotes the floor operation;
- $(\cdot)_{\bmod N}$ denotes the modulo operation with respect to a positive integer N;
- $\mathrm{Tr}(\boldsymbol{X})$ denotes the trace of the square matrix $\boldsymbol{X}$;
- $\boldsymbol{X}^{(L)}$ is the matrix where each row is the cyclic-shift (L elements to the right) of the corresponding row in $\boldsymbol{X}$.

Let $\boldsymbol{s}_0 = (s_{0,0}, s_{0,1}, \ldots, s_{0,L-1})$ and $\boldsymbol{s}_1 = (s_{1,0}, s_{1,1}, \ldots, s_{1,L-1})$ denote two complex-valued sequences of length L. For any integer displacement u, the *aperiodic cross-correlation function* (ACCF) of $\boldsymbol{s}_0$ and $\boldsymbol{s}_1$ is defined as

$$\rho(\boldsymbol{s}_0, \boldsymbol{s}_1; u) = \begin{cases} \sum_{k=0}^{L-1-u} s_{0,k+u} s_{1,k}^*, & 0 \le u \le L-1; \\ \sum_{k=0}^{L-1+u} s_{0,k} s_{1,k-u}^*, & -L+1 \le u < 0. \end{cases} \tag{1}$$

When $\boldsymbol{s}_0 = \boldsymbol{s}_1$, the function $\rho(\boldsymbol{s}_0, \boldsymbol{s}_1; u) = \rho(\boldsymbol{s}_0; u)$ is referred to as the *aperiodic autocorrelation function* (AACF) of $\boldsymbol{s}_0$. For periodic correlations, the *periodic cross-correlation function* (PCCF) of $\boldsymbol{s}_0$ and $\boldsymbol{s}_1$ at time-shift u is defined as

$$\phi(\boldsymbol{s}_0, \boldsymbol{s}_1; u) = \begin{cases} \sum_{k=0}^{L-1} s_{0,(k+u)_{\bmod L}} s_{1,k}^*, & 0 \le u \le L-1; \\ \sum_{k=0}^{L-1} s_{0,k} s_{1,(k-u)_{\bmod L}}^*, & -L+1 \le u < 0. \end{cases} \tag{2}$$

Accordingly, the *periodic autocorrelation function* (PACF) of $\boldsymbol{s}_0$ is denoted by $\phi(\boldsymbol{s}_0, \boldsymbol{s}_0; u) = \phi(\boldsymbol{s}_0; u)$.

Definition 1: For a set of N complex sequences $\mathcal{S} = \{\boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_{N-1}\}$ with length L, if

$$\phi(\boldsymbol{s}_i, \boldsymbol{s}_j; u) = \begin{cases} 0, \ 1 \le |u| \le Z, 0 \le i = j \le N-1; \\ 0, \ |u| \le Z, 0 \le i \ne j \le N-1, \end{cases} \tag{3}$$

then the set $\mathcal{S}$ is called an (N, L, Z)-ZCZ sequence set where Z is referred to as the width of ZCZ. The following lemma shows an upper bound among the parameters of the ZCZ sequence set.

Lemma 1 [49]: For an (N, L, Z)-ZCZ sequence set, there is a well-known theoretical upper bound, called *Tang-Fan-Matsufuji bound*, given as $Z \le L/N - 1$. A ZCZ sequence set is said to be optimal if the Tang-Fan-Matsufuji bound with equality is achieved. However, for binary case, the upper bound on ZCZ width is conjectured to be $Z \le L/(2N)$.

Consider a set of M sequence sets $\mathcal{S} = \{S^m | 0 \le m \le M-1\}$ where each constitute set $S^m = \{\boldsymbol{s}_n^m | 0 \le n \le N-1\}$ consists of N sequences of length L. For $S^{m_1}, S^{m_2} \in \mathcal{S}$ and $0 \le m_1, m_2 \le M-1$, we define

$$\rho(S^{m_1}, S^{m_2}; u) \triangleq \sum_{n=0}^{N-1} \rho(\boldsymbol{s}_n^{m_1}, \boldsymbol{s}_n^{m_2}; u). \tag{4}$$

Definition 2: A set of M sequence sets $\mathcal{S} = \{S^m | 0 \le m \le M-1\}$ is addressed as *Z-complementary code set*, denoted by (M, N, L, Z)-ZCCS, if

$$\rho(S^{m_1}, S^{m_2}; u) = \sum_{n=0}^{N-1} \rho(\boldsymbol{s}_n^{m_1}, \boldsymbol{s}_n^{m_2}; u) = \begin{cases} NL, \ u = 0, m_1 = m_2; \\ 0, \quad 0 < |u| < Z, m_1 = m_2; \\ 0, \quad |u| < Z, m_1 \ne m_2 \end{cases} \tag{5}$$

where M is the set size, N is the number of sequences in each S^m, L is the sequence length, and Z is the width of ZCZ. For an (M, N, L, Z)-ZCCS, a theoretical upper bound on the set size is given as $M \le N\lfloor L/Z \rfloor$. An (M, N, L, Z)-ZCCS is called optimal if the equality is achieved.

When $Z = L$, the set $\mathcal{S}$ is referred to as a *mutually orthogonal complementary set*, denoted by (M, N, L)-MOCS. Specifically, each constituent sequence set S^m reduces to a GCS and any two GCSs in set $\mathcal{S}$ are mutually orthogonal. Likewise, the upper bound on the set size for an MOCS satisfies the inequality $M \le N\lfloor L/L \rfloor = N$. If $M = N$, the set $\mathcal{S}$ is called a *complete complementary code*, denoted by (M, L)-CCC.

Definition 3 [47]: A set of M sequence sets $\mathcal{S} = \{S^m | 0 \le m \le M-1\}$ is called a *symmetrical Z-complementary code set*, denoted by (M, N, L, Z)-SZCCS, if

$$\rho(S^{m_1}, S^{m_2}; u) = \begin{cases} \sum_{n=0}^{N-1} \rho(\boldsymbol{s}_n^m; u) = 0, \\ \text{for } |u| \in \mathcal{T}_1 \cup \mathcal{T}_2, m_1 = m_2 = m; \\ \sum_{n=0}^{N-1} \rho(\boldsymbol{s}_n^{m_1}, \boldsymbol{s}_n^{m_2}; u) = 0, \\ \text{for } |u| \in \mathcal{T}_1 \cup \mathcal{T}_2 \cup \{0\}, m_1 \ne m_2 \end{cases} \tag{6}$$

where $\mathcal{T}_1 \triangleq \{1, 2, \ldots, Z\}$ and $\mathcal{T}_2 \triangleq \{L-Z, L-Z+1, \ldots, L-1\}$.

B. GENERALIZED BOOLEAN FUNCTIONS

Let f be a function of m $\mathbb{Z}_2$-valued variables $x_1, x_2, \ldots, x_m$ mapping from $\mathbb{Z}_2^m$ to $\mathbb{Z}_q$, denoted by

$$f : (x_1, x_2, \ldots, x_m) \in \mathbb{Z}_2^m \to f(x_1, x_2, \ldots, x_m) \in \mathbb{Z}_q. \tag{7}$$

The function f is addressed as the generalized Boolean function. For a q-ary generalized Boolean function f, the associated sequence $\boldsymbol{f} \in \mathbb{Z}_q^{2^m}$ is given by

$$\boldsymbol{f} = (f_0, f_1, \ldots, f_{2^m-1}) \tag{8}$$

where $f_i = f(i_1, i_2, \ldots, i_m)$ and $(i_1, i_2, \ldots, i_m)$ is the binary representation of the integer $i = \sum_{k=1}^m i_k 2^{k-1}$. Note that i_1 is the least significant bit.

Also, we define the complex-valued sequence associated with a generalized Boolean function f to be

$$\zeta_q(\boldsymbol{f}) \triangleq \left(\xi_q^{f_0}, \xi_q^{f_1}, \ldots, \xi_q^{f_{2^m-1}}\right). \tag{9}$$

Let $m = 4$ and $q = 4$ as an example. The sequence $\boldsymbol{f}$ associated with $f = 2x_2 + x_1x_2$ is

$$\boldsymbol{f} = 2\boldsymbol{x}_2 + \boldsymbol{x}_1\boldsymbol{x}_2$$

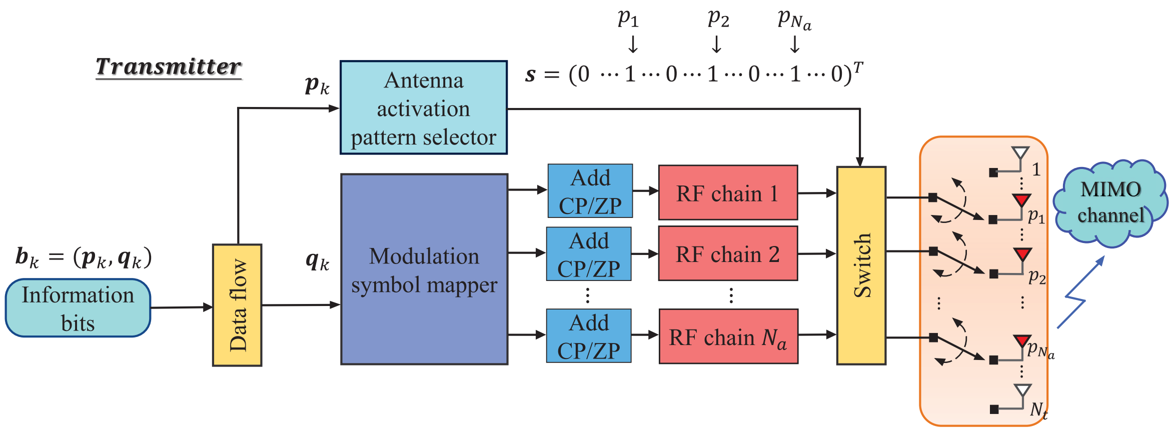

FIGURE 1. A generic transmitter structure of the SC-GSM system.

TABLE 1. An example of antenna activation patterns with $N_t = 4$ and $N_a = 2$.

GSM Mapping Rule		
Information bits	Antenna activation patterns	Note
00	$(1,1,0,0)^T$	Active antennas : 1, 2; Inactive antennas : 3, 4
01	$(0,1,1,0)^T$	Active antennas : 2, 3; Inactive antennas : 1, 4
10	$(1,0,0,1)^T$	Active antennas : 1, 4; Inactive antennas : 2, 3
11	$(0,0,1,1)^T$	Active antennas : 3, 4; Inactive antennas : 1, 2

$$= (f(0,0,0,0), f(1,0,0,0), \ldots, f(1,1,1,1))$$
$$= (0,0,2,3,0,0,2,3,0,0,2,3,0,0,2,3) \qquad (10)$$

and the complex modulated sequence

$$\zeta_4(\boldsymbol{f}) = \left(\xi_4^{f_0}, \xi_4^{f_1}, \ldots, \xi_4^{f_{15}}\right)$$
$$= (1,1,-1,-j,1,1,-1,-j,1,1,-1,-j,1,1,-1,-j). \qquad (11)$$

C. GSM

Consider a single-carrier GSM (SC-GSM) system over frequency-selective channels as depicted in Fig. 1. We denote N_t, N_r, and N_a as the number of TAs, receive antennas (RAs), and RF chains, respectively. An $N_a \times N_t$ switch is needed to connect the RF chains to the TAs. During each time-slot k, N_a of the N_t TAs are activated and the corresponding constellation symbols from quadrature amplitude modulation (QAM) or phase-shift keying (PSK) modulation $\mathcal{M}$ are transmitted on the activated TAs, while the remaining $N_t - N_a$ antennas are kept inactive. Specifically, $\lfloor \log_2 \binom{N_t}{N_a} \rfloor$ information bits, denoted by $\boldsymbol{p}_k$, are used for selecting N_a antennas based on activation pattern mapping. Additionally, we denote an $N_t \times 1$ vector $\boldsymbol{s} = (0 \cdots 1 \cdots 0 \cdots 1)^T$ as an antenna activation pattern where 1's correspond to the active antennas and 0's correspond to the silent antennas. On the other hand, $N_a \lfloor \log_2 |\mathcal{M}| \rfloor$ bits, denoted by $\boldsymbol{q}_k$, are mapped into a constellation from alphabet $\mathcal{M}$ through N_a active antennas. Therefore, the number of bits conveyed per symbol period is given by $\lfloor \log_2 \binom{N_t}{N_a} \rfloor + N_a \lfloor \log_2 |\mathcal{M}| \rfloor$.

Example 1: Let us consider an SC-GSM system with $N_t = 4$ and $N_a = 2$ using binary phase-shift keying (BPSK) modulation (i.e., $|\mathcal{M}| = 2$). The $\binom{4}{2} = 6$ possible activation patterns are shown as follows: $(1,1,0,0)^T$, $(1,0,1,0)^T$, $(1,0,0,1)^T$, $(0,1,1,0)^T$, $(0,1,0,1)^T$, $(0,0,1,1)^T$. However, only 4 activation patterns are selected since $\lfloor \log_2 \binom{4}{2} \rfloor = 2$. Then, the set of chosen activation patterns in this example is given by $\{(1,1,0,0)^T, (0,1,1,0)^T, (1,0,0,1)^T, (0,0,1,1)^T\}$. Table 1 shows a mapping from 2 information bits to the set of chosen activation patterns. Assume that the message bits (0101001110101111) are sent. There are 4 GSM symbols of which each symbol consists of $\lfloor \log_2 \binom{4}{2} \rfloor + 2 \lfloor \log_2 |2| \rfloor = 4$ bits. We have $\boldsymbol{b}_1 = (0101)$, $\boldsymbol{b}_2 = (0011)$, $\boldsymbol{b}_3 = (1010)$, and $\boldsymbol{b}_4 = (1111)$. Taking the first symbol for example, we have $\boldsymbol{b}_1 = (\boldsymbol{p}_1, \boldsymbol{q}_1)$ where $\boldsymbol{p}_1 = (01)$ and $\boldsymbol{q}_1 = (01)$. It indicates that TA 2 and TA 3 are activated to transmit the BPSK symbol "1" and "-1", respectively. Therefore, the first GSM symbol can be expressed as $(0,+,-,0)^T$. Then, the SC-GSM block for the 4 GSM symbols can be formulated as follows:

$$\left(\begin{array}{c|c|c|c} 0 & - & - & 0 \\ + & - & 0 & 0 \\ - & 0 & 0 & - \\ 0 & 0 & + & - \end{array}\right). \qquad (12)$$

III. ENHANCED CROSS Z-COMPLEMENTARY SETS

In this section, we will provide the definition and the optimality of the E-CZCS and then demonstrate two novel constructions.

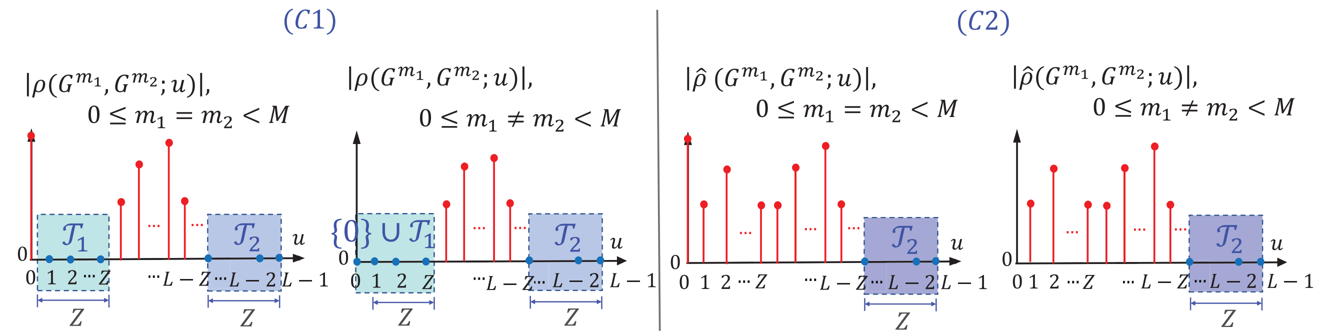

FIGURE 2. The correlation properties of E-CZCSs.

Consider a set of M sequence sets $\mathcal{G} = \{G^m | 0 \le m \le M-1\}$ where each constitute set $G^m = \{\mathbf{g}_n^m | 0 \le n \le N-1\}$ is composed of N sequences of length L. For $G^{m_1}, G^{m_2} \in \mathcal{G}$ with $0 \le m_1, m_2 \le M-1$, we define a special type of aperiodic cross-correlation sum as follows:

$$\hat{\rho}\left(G^{m_1}, G^{m_2}; u\right) \triangleq \sum_{n=0}^{N-1} \rho\left(\mathbf{g}_n^{m_1}, \mathbf{g}_{(n+1)_{\bmod N}}^{m_2}; u\right). \tag{13}$$

Note that (13) is different from the cross-correlation sum defined in (4). Taking $M = 2$ and $N = 4$ for example, we assume $\mathcal{G} = \{G^0, G^1\}$ where $G^0 = \{\mathbf{g}_0^0, \mathbf{g}_1^0, \mathbf{g}_2^0, \mathbf{g}_3^0\}$ and $G^1 = \{\mathbf{g}_0^1, \mathbf{g}_1^1, \mathbf{g}_2^1, \mathbf{g}_3^1\}$. Then, we have

$$\begin{aligned}\hat{\rho}\left(G^0, G^0; u\right) &= \sum_{n=0}^{3} \rho\left(\mathbf{g}_n^0, \mathbf{g}_{(n+1)_{\bmod 4}}^0; u\right) \\ &= \rho\left(\mathbf{g}_0^0, \mathbf{g}_1^0; u\right) + \rho\left(\mathbf{g}_1^0, \mathbf{g}_2^0; u\right) + \rho\left(\mathbf{g}_2^0, \mathbf{g}_3^0; u\right) \\ &\quad + \rho\left(\mathbf{g}_3^0, \mathbf{g}_0^0; u\right)\end{aligned} \tag{14}$$

and

$$\begin{aligned}\hat{\rho}\left(G^0, G^1; u\right) &= \sum_{n=0}^{3} \rho\left(\mathbf{g}_n^0, \mathbf{g}_{(n+1)_{\bmod 4}}^1; u\right) \\ &= \rho\left(\mathbf{g}_0^0, \mathbf{g}_1^1; u\right) + \rho\left(\mathbf{g}_1^0, \mathbf{g}_2^1; u\right) + \rho\left(\mathbf{g}_2^0, \mathbf{g}_3^1; u\right) \\ &\quad + \rho\left(\mathbf{g}_3^0, \mathbf{g}_0^1; u\right).\end{aligned} \tag{15}$$

Then we can define the E-CZCS based on the cross-correlation sums given in (4) and (13). This special type of aperiodic cross-correlation sum in (13) will be included in the definition of the E-CZCS.

Definition 4 (Enhanced Cross Z-Complementary Set): For positive integers M, N, L, and Z with $Z \le L$, we denote $\mathcal{T}_1 \triangleq \{1, 2, \ldots, Z\}$ and $\mathcal{T}_2 \triangleq \{L-Z, L-Z+1, \ldots, L-1\}$ as two distinct intervals. Let $\mathcal{G} = \{G^m | 0 \le m \le M-1\}$ be a set of M sequence sets and $G^m = \{\mathbf{g}_n^m | 0 \le n \le N-1\}$ where $\mathbf{g}_n^m$ is a sequence of length L. Then, the set $\mathcal{G}$ is addressed as an *enhanced cross Z-complementary set*, denoted by (M, N, L, Z)-E-CZCS, if it satisfies the following two conditions:

$$\text{(C1):}\quad \rho\left(G^{m_1}, G^{m_2}; u\right) = \sum_{n=0}^{N-1} \rho\left(\mathbf{g}_n^{m_1}, \mathbf{g}_n^{m_2}; u\right) = \begin{cases} 0, \text{ for all } |u| \in (\mathcal{T}_1 \cup \mathcal{T}_2) \cap \mathcal{T}^1, \\ \quad 0 \le m_1 = m_2 \le M-1; \\ 0, \text{ for all } |u| \in \mathcal{T}_1 \cup \mathcal{T}_2 \cup \{0\}, \\ \quad 0 \le m_1 \ne m_2 \le M-1; \end{cases}$$

$$\text{(C2):}\quad \hat{\rho}\left(G^{m_1}, G^{m_2}; u\right) = \sum_{n=0}^{N-1} \rho\left(\mathbf{g}_n^{m_1}, \mathbf{g}_{(n+1)_{\bmod N}}^{m_2}; u\right) = 0, \quad \text{for all } |u| \in \mathcal{T}_2 \text{ and any } m_1, m_2 \in \{0, 1, \ldots, M-1\} \tag{16}$$

where $\mathcal{T} = \{1, 2, \ldots, L-1\}$.

If $M = 1$, i.e., $m_1 = m_2 = 0$, then the E-CZCS reduces to the *cross Z-complementary set*, denoted by (N, L, Z)-CZCS. It also implies each G^{m_1} in an E-CZCS is a CZCS. Also, an E-CZCS is reduced to an (L, Z)-CZCP by taking $M = 1$ and $N = 2$. (C1) means the correlation sum $\rho(G^{m_1}, G^{m_2}; u)$ has symmetric ZCZs over $\mathcal{T}_1$ and $\mathcal{T}_2$. And (C2) indicates that the cross-correlation sum $\hat{\rho}(G^{m_1}, G^{m_2}; u)$ has a tail-end ZCZ for shifts over $\mathcal{T}_2$. We illustrate the correlation properties of E-CZCSs in Fig. 2. Besides, from condition (C1), the E-CZCS is also a ZCCS and a SZCCS. However, the ZCCS and SZCCS do not take the condition (C2) into account. Therefore, the E-CZCS can include CZCS, ZCCS, and SZCCS as special cases.

The aperiodic cross-correlation property (C2) of the E-CZCS can be utilized to eliminate the IAI when it is employed in the proposed training framework, which will be illustrated in Section IV-B.

Remark 1: For an (M, N, L, Z)-E-CZCS with $Z \ge L/2$, we have $\mathcal{T}_1 \cup \mathcal{T}_2 = \{1, 2, \ldots, L-1\}$. Therefore, (C1) in (16) implies that an (M, N, L, Z)-E-CZCS is also an (M, N, L)-MOCS.

Fig. 3 illustrates the relationship between E-CZCSs and the related sequence sets, which include ZCCSs, SZCCSs, and MOCSs.

Next, we discuss the relationship among the ZCZ width Z, the set size M, and the number of sequences N.

[1]If Z equals to L, we have $\mathcal{T}_1 = \{1, 2, \ldots, L\}$ and $\mathcal{T}_2 = \{0, 1, \ldots, L-1\}$. Consequently, $\mathcal{T}_1 \cup \mathcal{T}_2 = \{0, 1, \ldots, L\}$. However, we have $\rho(G^{m_1}, G^{m_2}; 0) = NL$ for $0 \le m_1 = m_2 \le M-1$. So we have to exclude $u \ne 0$. Therefore, an extra condition of the intersection with $\mathcal{T} = \{1, 2, \ldots, L-1\}$ is needed.

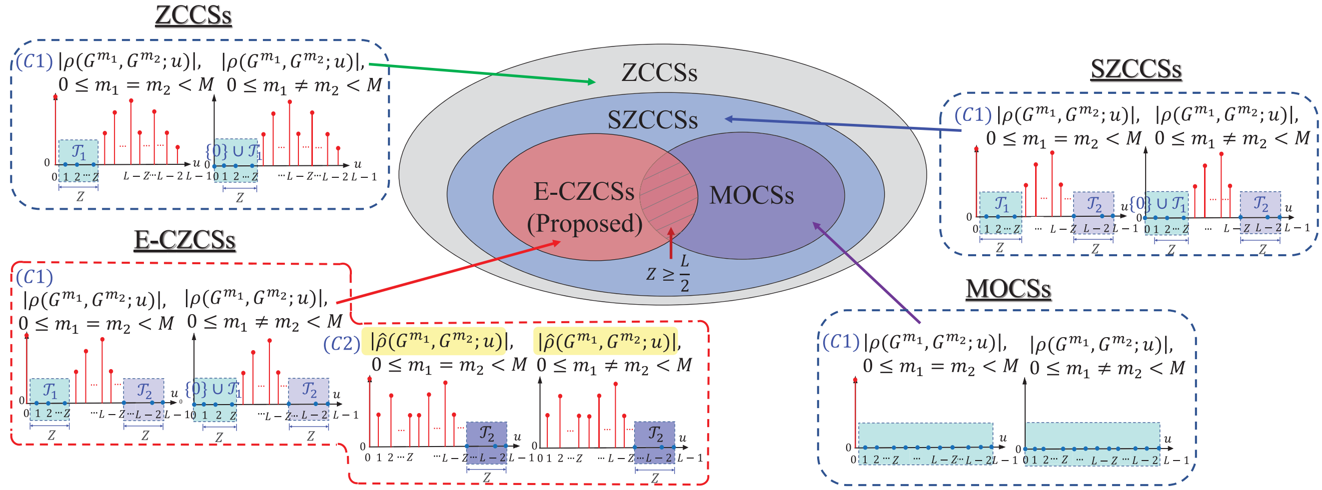

FIGURE 3. Relationship between E-CZCSs and the related sequence sets.

Theorem 1: For an (M, N, L, Z)-E-CZCS $\mathcal{G} = \{G^0, G^1, \ldots, G^{M-1}\}$, the upper bound on ZCZ width is given by

$$Z \leq \frac{NL}{M} - 1. \tag{17}$$

For the binary E-CZCS, we have

$$Z \leq \frac{NL}{2M}. \tag{18}$$

Proof: Let $G^m = \{\boldsymbol{g}_0^m, \boldsymbol{g}_1^m, \ldots, \boldsymbol{g}_{N-1}^m\}$ for $m = 0, 1, \ldots, M-1$ and also let

$$\begin{aligned}\boldsymbol{d}_0 &= \boldsymbol{g}_0^0 \| \boldsymbol{g}_1^0 \| \cdots \| \boldsymbol{g}_{N-1}^0,\\ \boldsymbol{d}_1 &= \boldsymbol{g}_0^1 \| \boldsymbol{g}_1^1 \| \cdots \| \boldsymbol{g}_{N-1}^1,\\ &\vdots\\ \boldsymbol{d}_{M-1} &= \boldsymbol{g}_0^{M-1} \| \boldsymbol{g}_1^{M-1} \| \cdots \| \boldsymbol{g}_{N-1}^{M-1}.\end{aligned} \tag{19}$$

For $m = 0, 1, \ldots, M-1$, we have

$$\begin{aligned}\phi(\boldsymbol{d}_m; u) &= \rho(\boldsymbol{g}_0^m; u) + \rho(\boldsymbol{g}_1^m; u) + \ldots + \rho(\boldsymbol{g}_{N-1}^m; u)\\ &\quad + \rho^*(\boldsymbol{g}_0^m, \boldsymbol{g}_1^m; L-u) + \rho^*(\boldsymbol{g}_1^m, \boldsymbol{g}_2^m; L-u) + \ldots\\ &\quad + \rho^*(\boldsymbol{g}_{N-1}^m, \boldsymbol{g}_0^m; L-u)\\ &= \sum_{n=0}^{N-1} \rho(\boldsymbol{g}_n^m; u) + \sum_{n=0}^{N-1} \rho^*\left(\boldsymbol{g}_n^m, \boldsymbol{g}_{(n+1)_{\bmod N}}^m; L-u\right)\\ &= \rho(G^m, G^m; u) + \hat{\rho}^*(G^m, G^m; L-u), \text{ for } 1 \leq u < L,\end{aligned} \tag{20}$$

and

$$\begin{aligned}&\phi(\boldsymbol{d}_m; L)\\ &= \rho^*(\boldsymbol{g}_0^m, \boldsymbol{g}_1^m; 0) + \rho^*(\boldsymbol{g}_1^m, \boldsymbol{g}_2^m; 0) + \ldots + \rho^*(\boldsymbol{g}_{N-1}^m, \boldsymbol{g}_0^m; 0)\\ &= \sum_{n=0}^{N-1} \rho^*\left(\boldsymbol{g}_n^m, \boldsymbol{g}_{(n+1)_{\bmod N}}^m; 0\right) = \hat{\rho}^*(G^m, G^m; 0).\end{aligned} \tag{21}$$

Therefore,

$$\phi(\boldsymbol{d}_m; u) = \begin{cases} \rho(G^m, G^m; u) + \hat{\rho}^*(G^m, G^m; L-u), \\ \text{for} \quad 1 \leq u < L, \quad 0 \leq m \leq M-1; \\ \hat{\rho}^*(G^m, G^m; 0), \\ \text{for} \quad u = L, \quad 0 \leq m \leq M-1. \end{cases} \tag{22}$$

Next, for two distinct integers m_1,m_2 with $0 \leq m_1, m_2 \leq M-1$, we have

$$\begin{aligned}\phi(\boldsymbol{d}_{m_1}, \boldsymbol{d}_{m_2}; u) &= \rho(\boldsymbol{g}_0^{m_1}, \boldsymbol{g}_0^{m_2}; u) + \rho(\boldsymbol{g}_1^{m_1}, \boldsymbol{g}_1^{m_2}; u) + \ldots\\ &\quad + \rho(\boldsymbol{g}_{N-1}^{m_1}, \boldsymbol{g}_{N-1}^{m_2}; u) + \rho^*(\boldsymbol{g}_0^{m_2}, \boldsymbol{g}_1^{m_1}; L-u)\\ &\quad + \rho^*(\boldsymbol{g}_1^{m_2}, \boldsymbol{g}_2^{m_1}; L-u) + \ldots + \rho^*(\boldsymbol{g}_{N-1}^{m_2}, \boldsymbol{g}_0^{m_1}; L-u)\\ &= \sum_{n=0}^{N-1} \rho(\boldsymbol{g}_n^{m_1}, \boldsymbol{g}_n^{m_2}; u) + \sum_{n=0}^{N-1} \rho^*\left(\boldsymbol{g}_n^{m_2}, \boldsymbol{g}_{(n+1)_{\bmod N}}^{m_1}; L-u\right)\\ &= \rho(G^{m_1}, G^{m_2}; u) + \hat{\rho}^*(G^{m_2}, G^{m_1}; L-u),\\ &\quad \text{for } 1 \leq u < L,\end{aligned} \tag{23}$$

$$\begin{aligned}\phi(\boldsymbol{d}_{m_1}, \boldsymbol{d}_{m_2}; 0) &= \rho(\boldsymbol{g}_0^{m_1}, \boldsymbol{g}_0^{m_2}; 0) + \rho(\boldsymbol{g}_1^{m_1}, \boldsymbol{g}_1^{m_2}; 0) + \ldots\\ &\quad + \rho(\boldsymbol{g}_{N-1}^{m_1}, \boldsymbol{g}_{N-1}^{m_2}; 0)\\ &= \sum_{n=0}^{N-1} \rho(\boldsymbol{g}_n^{m_1}, \boldsymbol{g}_n^{m_2}; 0) = \rho(G^{m_1}, G^{m_2}; 0),\end{aligned} \tag{24}$$

and

$$\begin{aligned}\phi(\boldsymbol{d}_{m_1}, \boldsymbol{d}_{m_2}; L) &= \rho^*(\boldsymbol{g}_0^{m_2}, \boldsymbol{g}_1^{m_1}; 0) + \rho^*(\boldsymbol{g}_1^{m_2}, \boldsymbol{g}_2^{m_1}; 0) + \ldots\\ &\quad + \rho^*(\boldsymbol{g}_{N-1}^{m_2}, \boldsymbol{g}_0^{m_1}; 0)\\ &= \sum_{n=0}^{N-1} \rho^*\left(\boldsymbol{g}_n^{m_2}, \boldsymbol{g}_{(n+1)_{\bmod N}}^{m_1}; 0\right) = \hat{\rho}^*(G^{m_2}, G^{m_1}; 0).\end{aligned} \tag{25}$$

Thus, we have

$$\phi(\boldsymbol{d}_{m_1}, \boldsymbol{d}_{m_2}; u) = \begin{cases} \rho(G^{m_1}, G^{m_2}; 0), \quad \text{for} \quad u = 0, \\ 0 \le m_1 \ne m_2 \le M-1; \\ \rho(G^{m_1}, G^{m_2}; u) \\ +\hat{\rho}^*(G^{m_2}, G^{m_1}; L-u), \\ \text{for} \quad 1 \le u < L, \\ \quad 0 \le m_1 \ne m_2 \le M-1; \\ \hat{\rho}^*(G^{m_2}, G^{m_1}; 0), \\ \text{for} \quad u = L, \ 0 \le m_1 \ne m_2 \le M-1. \end{cases} \tag{26}$$

From (16), (22), and (26), we have

$$\phi(\boldsymbol{d}_{m_1}, \boldsymbol{d}_{m_2}; u) = \begin{cases} 0, \text{ for } \quad 1 \le |u| \le Z \quad \text{and} \\ \quad 0 \le m_1 = m_2 \le M-1; \\ 0, \text{ for } \quad |u| \le Z \quad \text{and} \\ \quad 0 \le m_1 \ne m_2 \le M-1 \end{cases} \tag{27}$$

since $\mathcal{G}$ is an (M, N, L, Z)-E-CZCS. Therefore, the M sequences $\boldsymbol{d}_0$, $\boldsymbol{d}_1$, ..., $\boldsymbol{d}_{M-1}$ form an (M, NL, Z)-ZCZ sequence set. According to *Lemma 1*, the (M, NL, Z)-ZCZ sequence set satisfies that $Z \le (NL)/M - 1$ and $Z \le (NL)/(2M)$ for binary sequence sets. Therefore, we complete the proof. ■

Remark 2: According to *Theorem 1*, a q-ary (M, N, L, Z)-E-CZCS possesses the maximum ZCZ width if $Z = (NL)/M - 1$ for $q > 2$ or $Z = (NL)/(2M)$ for $q = 2$.

A. E-CZCSS BASED ON ZCCSS

We first present a construction of E-CZCSs based on ZCCSs.

Theorem 2: Given an $(M, N, L, Z+1)$-ZCCS $\mathcal{S} = \{S^m | 0 \le m \le M-1\}$ where each constitute set $S^m = \{\boldsymbol{s}_n^m | 0 \le n \le N-1\}$. Then, $\mathcal{G} = \{G^m | 0 \le m \le M-1\}$ is an $(M, N, 2L, Z)$-E-CZCS by letting

$$\begin{aligned} G^m = \big\{ & \boldsymbol{g}_0^m = \boldsymbol{s}_0^m \| \boldsymbol{s}_1^m, \\ & \boldsymbol{g}_1^m = \boldsymbol{s}_2^m \| \boldsymbol{s}_3^m, \\ & \vdots \\ & \boldsymbol{g}_{N/2-1}^m = \boldsymbol{s}_{N-2}^m \| \boldsymbol{s}_{N-1}^m, \\ & \boldsymbol{g}_{N/2}^m = \boldsymbol{s}_0^m \| (-\boldsymbol{s}_1^m), \\ & \boldsymbol{g}_{N/2+1}^m = \boldsymbol{s}_2^m \| (-\boldsymbol{s}_3^m), \\ & \vdots \\ & \boldsymbol{g}_{N-1}^m = \boldsymbol{s}_{N-2}^m \| (-\boldsymbol{s}_{N-1}^m) \big\} \end{aligned} \tag{28}$$

for $m = 0, 1, \ldots, M-1$. Furthermore, if $\mathcal{S}$ is an (M, N, L)-MOCS, then $\mathcal{G}$ is an $(M, N, 2L, L)$-E-CZCS.

Proof: We consider two cases to show that (C1) and (C2) in (16) are satisfied, respectively.

Case 1: Let $\mathcal{T}_1 = \{1, 2, \ldots, Z\}$. For $|u| \in \mathcal{T}_1 \cup \{0\}$, we have

$$\begin{aligned} & \rho(G^{m_1}, G^{m_2}; u) \\ & = \sum_{n=0}^{N-1} \rho(\boldsymbol{g}_n^{m_1}, \boldsymbol{g}_n^{m_2}; u) = 2\left(\sum_{n=0}^{N-1} \rho(\boldsymbol{s}_n^{m_1}, \boldsymbol{s}_n^{m_2}; u)\right) \\ & + \sum_{n=0}^{\frac{N}{2}-1} \rho^*(\boldsymbol{s}_{2n}^{m_2}, \boldsymbol{s}_{2n+1}^{m_1}; L-u) - \sum_{n=0}^{\frac{N}{2}-1} \rho^*(\boldsymbol{s}_{2n}^{m_2}, \boldsymbol{s}_{2n+1}^{m_1}; L-u) \\ & = 2\left(\sum_{n=0}^{N-1} \rho(\boldsymbol{s}_n^{m_1}, \boldsymbol{s}_n^{m_2}; u)\right) = \begin{cases} 2NL, \ u = 0, m_1 = m_2; \\ 0, \ 0 < |u| \le Z, m_1 = m_2; \\ 0, \ |u| \le Z, m_1 \ne m_2 \end{cases} \end{aligned} \tag{29}$$

since $\mathcal{S} = \{S^m | 0 \le m \le M-1\}$ is an $(M, N, L, Z+1)$-ZCCS. Let $\mathcal{T}_2 = \{2L-Z, 2L-Z+1, \ldots, 2L-1\}$. For $|u| \in \mathcal{T}_2$, we have

$$\begin{aligned} \rho(G^{m_1}, G^{m_2}; u) & = \sum_{n=0}^{N-1} \rho(\boldsymbol{g}_n^{m_1}, \boldsymbol{g}_n^{m_2}; u) \\ & = \sum_{n=0}^{\frac{N}{2}-1} \rho(\boldsymbol{s}_{2n+1}^{m_1}, \boldsymbol{s}_{2n}^{m_2}; u-L) - \sum_{n=0}^{\frac{N}{2}-1} \rho(\boldsymbol{s}_{2n+1}^{m_1}, \boldsymbol{s}_{2n}^{m_2}; u-L) \\ & = 0, \text{ for any } m_1, m_2 \in \{0, 1, \ldots, M-1\}. \end{aligned} \tag{30}$$

According to (29) and (30), we obtain that (C1) in (16) holds.

Case 2: For $|u| \in \mathcal{T}_2$, we have

$$\begin{aligned} \hat{\rho}(G^{m_1}, G^{m_2}; u) & = \sum_{n=0}^{N-1} \rho\left(\boldsymbol{g}_n^{m_1}, \boldsymbol{g}_{(n+1)_{\bmod N}}^{m_2}; u\right) \\ & = \sum_{n=0}^{\frac{N}{2}-1} \rho\left(\boldsymbol{s}_{2n+1}^{m_1}, \boldsymbol{s}_{(2n+2)_{\bmod N}}^{m_2}; u-L\right) \\ & \quad - \sum_{n=0}^{\frac{N}{2}-1} \rho\left(\boldsymbol{s}_{2n+1}^{m_1}, \boldsymbol{s}_{(2n+2)_{\bmod N}}^{m_2}; u-L\right) \\ & = 0, \text{ for any } m_1, m_2 \in \{0, 1, \ldots, M-1\} \end{aligned} \tag{31}$$

which means (C2) in (16) holds. Therefore, $\mathcal{G}$ is an $(M, N, 2L, Z)$-E-CZCS.

Moreover, if $\mathcal{S}$ is an (M, N, L)-MOCS, we substitute Z by L in *Case 1* and *Case 2* and hence have $\mathcal{T}_1 \cup \mathcal{T}_2 = \{1, 2, \ldots, 2L-1\}$. Therefore, $\mathcal{G}$ is an $(M, N, 2L, L)$-E-CZCS. ■

Remark 3: To possess the largest ZCZ width, we consider a binary (M, M, L)-MOCS $\mathcal{S}$, i.e., (M, L)-CCC, in *Theorem 2*. Then the ZCZ width of the constructed $(M, M, 2L, L)$-E-CZCS satisfies the equality given in (18). Therefore, a binary $(M, M, 2L, L)$-E-CZCS with maximum ZCZ width can be obtained.

An example of the binary $(4, 4, 14, 7)$-E-CZCS is provided below.

Example 2: Consider a $(4, 7)$-CCC $\mathcal{S} = \{S^0, S^1, S^2, S^3\}$ obtained from [50] as shown in Table 2. We let

$$\begin{aligned} G^0 = \Big\{ & \boldsymbol{g}_0^0 = \boldsymbol{s}_0^0 \| \boldsymbol{s}_1^0 = (+++--+++-+----), \\ & \boldsymbol{g}_1^0 = \boldsymbol{s}_2^0 \| \boldsymbol{s}_3^0 = (+-+----+--++-+), \\ & \boldsymbol{g}_2^0 = \boldsymbol{s}_0^0 \| \left(-\boldsymbol{s}_1^0\right) = (+++--++-+-++++), \\ & \boldsymbol{g}_3^0 = \boldsymbol{s}_2^0 \| \left(-\boldsymbol{s}_3^0\right) = (+-+------++--+-) \Big\}, \end{aligned} \tag{32}$$

TABLE 2. A binary (4, 7)-CCC $\mathcal{S} = \{S^0, S^1, S^2, S^3\}$.

S^0	$\left\{\begin{array}{l} s_0^0 = (+++--++), \\ s_1^0 = (+-+----), \\ s_2^0 = (+-+----), \\ s_3^0 = (+--++-+) \end{array}\right\}$	S^2	$\left\{\begin{array}{l} s_0^2 = (+-++--+), \\ s_1^2 = (+-+----), \\ s_2^2 = (-+-++++), \\ s_3^2 = (--++---) \end{array}\right\}$
S^1	$\left\{\begin{array}{l} s_0^1 = (++++-+-), \\ s_1^1 = (++--+++), \\ s_2^1 = (-+--++-), \\ s_3^1 = (----+-+) \end{array}\right\}$	S^3	$\left\{\begin{array}{l} s_0^3 = (----+-+), \\ s_1^3 = (-++--+-), \\ s_2^3 = (---++--), \\ s_3^3 = (----+-+) \end{array}\right\}$

$$\begin{aligned} G^1 = \Big\{ & \boldsymbol{g}_0^1 = \boldsymbol{s}_0^1 \| \boldsymbol{s}_1^1 = (++++-+-++--+++), \\ & \boldsymbol{g}_1^1 = \boldsymbol{s}_2^1 \| \boldsymbol{s}_3^1 = (-+--++------+-+), \\ & \boldsymbol{g}_2^1 = \boldsymbol{s}_0^1 \| \left(-\boldsymbol{s}_1^1\right) = (++++-+---++---), \\ & \boldsymbol{g}_3^1 = \boldsymbol{s}_2^1 \| \left(-\boldsymbol{s}_3^1\right) = (-+--++-++++-+-) \Big\}, \end{aligned} \tag{33}$$

$$\begin{aligned} G^2 = \Big\{ & \boldsymbol{g}_0^2 = \boldsymbol{s}_0^2 \| \boldsymbol{s}_1^2 = (+-++--++-+----), \\ & \boldsymbol{g}_1^2 = \boldsymbol{s}_2^2 \| \boldsymbol{s}_3^2 = (-+-++++--++---), \\ & \boldsymbol{g}_2^2 = \boldsymbol{s}_0^2 \| \left(-\boldsymbol{s}_1^2\right) = (+-++--+-+-++++), \\ & \boldsymbol{g}_3^2 = \boldsymbol{s}_2^2 \| \left(-\boldsymbol{s}_3^2\right) = (-+-++++++--+++) \Big\}, \end{aligned} \tag{34}$$

and

$$\begin{aligned} G^3 = \Big\{ & \boldsymbol{g}_0^3 = \boldsymbol{s}_0^3 \| \boldsymbol{s}_1^3 = (----+-+-++--+-), \\ & \boldsymbol{g}_1^3 = \boldsymbol{s}_2^3 \| \boldsymbol{s}_3^3 = (---++------+-+), \\ & \boldsymbol{g}_2^3 = \boldsymbol{s}_0^3 \| \left(-\boldsymbol{s}_1^3\right) = (----+-++--++-+), \\ & \boldsymbol{g}_3^3 = \boldsymbol{s}_2^3 \| \left(-\boldsymbol{s}_3^3\right) = (---++--++++-+-) \Big\}. \end{aligned} \tag{35}$$

A (4, 4, 14, 7)-E-CZCS $\mathcal{G} = \{G^0, G^1, G^2, G^3\}$ can be obtained by *Theorem* 2 in Table 3. We list the correlation sums $\rho(G^0, G^0; u)$ and $\hat{\rho}(G^0, G^1; u)$ as follows:

$$\begin{aligned} \left|\rho\left(G^0, G^0; u\right)\right|_{u=0\sim13} &= (56, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0); \\ \left|\hat{\rho}\left(G^0, G^1; u\right)\right|_{u=0\sim13} &= (4, 16, 4, 8, 4, 0, 4, 0, 0, 0, 0, 0, 0, 0). \end{aligned} \tag{36}$$

According to *Theorem 2*, E-CZCSs can be constructed based on the MOCSs, CCCs, and ZCCSs. Since MOCSs, CCCs, and ZCCSs with various lengths can be obtained from [50], [51], [52], [53], the lengths of the constructed E-CZCSs from *Theorem 2* are flexible.

B. E-CZCSS BASED ON GENERALIZED BOOLEAN FUNCTIONS

In this subsection, we will present a direct construction of E-CZCSs based on generalized Boolean functions. The proposed construction can generate E-CZCSs with various set sizes and large ZCZ widths including binary E-CZCS with the maximum ZCZ width.

Theorem 3: For nonnegative integers m, v, k with $v \le k$, we let nonempty sets $U_1, U_2, \ldots, U_k$ be a partition of $\{1, 2, \ldots, m\}$. Also let m_α be the order of U_α and π_α be a bijection from $\{1, 2, \ldots, m_\alpha\}$ to U_α for $\alpha = 1, 2, \ldots, k$. The generalized Boolean function f is given as

$$f = \frac{q}{2} \sum_{\alpha=1}^{k} \sum_{\beta=1}^{m_\alpha - 1} x_{\pi_\alpha(\beta)} x_{\pi_\alpha(\beta+1)} + \sum_{i=1}^{m} \eta_i x_i + \eta_0 \tag{37}$$

where $\eta_i \in \mathbb{Z}_q$ for $i = 0, 1, \ldots, m$. If $v < k$, we set $\pi_{v+\gamma}(1) = m - \gamma + 1$ for $\gamma = 1, 2, \ldots, k - v$. For $p = 0, 1, \ldots, 2^k - 1$, we let $G^p = \{\zeta_q(\boldsymbol{g}_0^p), \zeta_q(\boldsymbol{g}_1^p), \ldots, \zeta_q(\boldsymbol{g}_{2^v-1}^p)\}$ where

$$\boldsymbol{g}_n^p = \boldsymbol{f} + \frac{q}{2}\left(\sum_{\alpha=1}^{v} n_{v-\alpha+1} \boldsymbol{x}_{\pi_\alpha(1)} + \sum_{\alpha=1}^{k} p_\alpha \boldsymbol{x}_{\pi_\alpha(m_\alpha)}\right) \tag{38}$$

for $n = 0, 1, \ldots, 2^v - 1$; $(n_1, n_2, \ldots, n_v)$ and $(p_1, p_2, \ldots, p_k)$ are binary representations of n and p, respectively. Then the set $\mathcal{G} = \{G^0, G^1, \ldots, G^{2^k-1}\}$ is a $(2^k, 2^v, 2^m, 2^{\pi_1(1)-1})$-E-CZCS.

Proof: The proof is given in Appendix. ■

Remark 4: To obtain the largest ZCZ width, we set $\pi_1(1) = m - k + v$ in *Theorem* 3 to construct the $(2^k, 2^v, 2^m, 2^{m-k+v-1})$-E-CZCS which can achieve the upper bound on the ZCZ width in (18). Therefore, binary $(2^k, 2^v, 2^m, 2^{m-k+v-1})$-E-CZCSs with maximum ZCZ width can be obtained.

Example 3: Let us consider $q = 2$, $m = 5$, $k = 2$, and $v = 1$. We let a partition of $\{1, 2, 3, 4, 5\}$ by $U_1 = \{1, 2, 4\}$ and $U_2 = \{3, 5\}$ with $m_1 = 3$ and $m_2 = 2$, respectively. We also let bijections $\pi_1 = (4, 1, 2)$ and $\pi_2 = (5, 3)$. Then, the generalized Boolean function f in (37) can be written as $f = x_4x_1 + x_1x_2 + x_5x_3$ by setting $\eta_i = 0$ for all i. Following *Theorem 3*, a binary (4, 2, 32, 8)-E-CZCS can be constructed as $\mathcal{G} = \{G^p = \{\zeta_2(\boldsymbol{g}_0^p), \zeta_2(\boldsymbol{g}_1^p)\} : p \in \{0, 1, 2, 3\}\}$ where $\boldsymbol{g}_n^p = \boldsymbol{f} + n_1\boldsymbol{x}_4 + p_1\boldsymbol{x}_2 + p_2\boldsymbol{x}_3$. For example, the sequence $\boldsymbol{g}_0^0$ in G^0 can be constructed as follows:

$$\begin{aligned} \boldsymbol{g}_0^0 &= \boldsymbol{f} = \boldsymbol{x}_4\boldsymbol{x}_1 + \boldsymbol{x}_1\boldsymbol{x}_2 + \boldsymbol{x}_5\boldsymbol{x}_3 \\ &= (f(0,0,0,0,0), f(1,0,0,0,0), \ldots, f(1,1,1,1,1)) \\ &= (0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, \\ &\qquad 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1). \end{aligned} \tag{39}$$

Subsequently, we can obtain the corresponding modulated sequence

$$\begin{aligned} \zeta_2\left(\boldsymbol{g}_0^0\right) = (&1, 1, 1, -1, 1, 1, 1, -1, 1, -1, 1, 1, 1, -1, 1, 1, \\ &1, 1, 1, -1, -1, -1, -1, 1, 1, -1, 1, 1, -1, 1, -1, -1). \end{aligned} \tag{40}$$

We list the constituent sequence sets G^0, G^1, G^2, and G^3 in Table 4. The correlation sums $\rho(G^0, G^2; u)$ and $\hat{\rho}(G^0, G^2; u)$ are given as follows:

$$\begin{aligned} \left|\rho\left(G^0, G^2; u\right)\right|_{u=0\sim31} = (&0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 16, \\ &0, 0, 0, 32, 0, 0, 0, 16, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0); \\ \left|\hat{\rho}\left(G^0, G^2; u\right)\right|_{u=0\sim31} = (&0, 0, 0, 0, 0, 0, 0, 0, 0, 4, 0, 4, \\ &0, 12, 0, 12, 0, 12, 0, 12, 0, 4, 0, 4, 0, 0, 0, 0, 0, 0, 0, 0). \end{aligned} \tag{41}$$

TABLE 3. Binary (4, 4, 14, 7)-E-CZCS in Example 2.

(4, 4, 14, 7)-E-CZCS $\mathcal{G}=\{G^0,G^1,G^2,G^3\}$			
G^0	$g_0^0=(+++--+++-+----)$, $g_1^0=(+-+----+--++-+)$, $g_2^0=(+++--++-+-++++)$, $g_3^0=(+-+-----++--+-)$	G^2	$g_0^2=(+-++--++-+----)$, $g_1^2=(-+-++++--++---)$, $g_2^2=(+-++--+-+-++++)$, $g_3^2=(-+-++++++--+++)$
G^1	$g_0^1=(++++-+-++--+++)$, $g_1^1=(-+--++-----+-+)$, $g_2^1=(++++-+---++---)$, $g_3^1=(-+--++-++++-+-)$	G^3	$g_0^3=(----+-+-++--+-)$, $g_1^3=(---++------+-+)$, $g_2^3=(----+-++--++-+)$, $g_3^3=(---++--++++-+-)$

TABLE 4. Binary (4, 2, 32, 8)-E-CZCS in Example 3.

(4, 2, 32, 8)-E-CZCS $\mathcal{G}=\{G^0,G^1,G^2,G^3\}$		
	G^0	$\zeta_2(g_0^0)=(+++-+++-+-+++-+++++----++-++-+--)$, $\zeta_2(g_1^0)=(+++-+++--+---+--+++----+-+--+-++)$
	G^1	$\zeta_2(g_0^1)=(++-+++-++---+---++-+--+-+----+++)$, $\zeta_2(g_1^1)=(++-+++-+-+++-+++++-+--+--++++---)$
	G^2	$\zeta_2(g_0^2)=(+++----++-++-+--+++-+++-+-+++-++)$, $\zeta_2(g_1^2)=(+++----+-+--+-+++++-+++--+---+--)$
	G^3	$\zeta_2(g_0^3)=(++-+--+-+----+++++-+++-++---+---)$, $\zeta_2(g_1^3)=(++-+--+--++++---++-+++-+-+++-+++)$

IV. PROPOSED TRAINING FRAMEWORK FOR BROADBAND GSM SYSTEMS

A. TRAINING DESIGN

In this subsection, we formulate the system model and the requirements for training design in the GSM system.

Consider a generic training-based multiple-antenna transmission structure depicted in Fig. 4. Prior to data payload transmission, the training sequences $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_{N_t}$ transmitted from the N_t TAs are used to estimate the channel state information. The cyclic prefix (CP) is inserted before each training sequence for ISI suppression in dispersive channels. Let $\boldsymbol{\Psi}$ denote the training matrix given by

$$\boldsymbol{\Psi}=\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \vdots\\ \boldsymbol{x}_{N_t}\end{bmatrix}=\begin{bmatrix}x_{1,0} & x_{1,1} & \cdots & x_{1,L'-1}\\ x_{2,0} & x_{2,1} & \cdots & x_{2,L'-1}\\ \vdots & \vdots & \ddots & \vdots\\ x_{N_t,0} & x_{N_t,1} & \cdots & x_{N_t,L'-1}\end{bmatrix}_{N_t\times L'} \quad (42)$$

where $\boldsymbol{x}_p=(x_{p,0},x_{p,1},\ldots,x_{p,L'-1})$ stands for the training sequence conveyed over the p-th TA for $p=1,2,\ldots,N_t$. Note that all the training sequences with identical energy $E=\sum_{t=0}^{L'-1}|x_{p,t}|^2$. In addition, we consider a quasi-static frequency-selective channel with the delay spread λ. Assume that the channel impulse response from the p-th TA to the receiver is $\boldsymbol{h}_p=[h_{p,0},h_{p,1},\ldots,h_{p,\lambda}]$. To formulate the model in matrix form, we let

$$\boldsymbol{X}=\left[\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_{N_t}\right]_{L'\times N_t(\lambda+1)} \quad (43)$$

where

$$\boldsymbol{X}_p=\begin{bmatrix}x_{p,0} & x_{p,L'-1} & \cdots & x_{p,L'-\lambda}\\ x_{p,1} & x_{p,0} & \cdots & x_{p,L'-\lambda+1}\\ \vdots & \vdots & \ddots & \vdots\\ x_{p,L'-1} & x_{p,L'-2} & \cdots & x_{p,L'-\lambda-1}\end{bmatrix}_{L'\times(\lambda+1)} \quad (44)$$

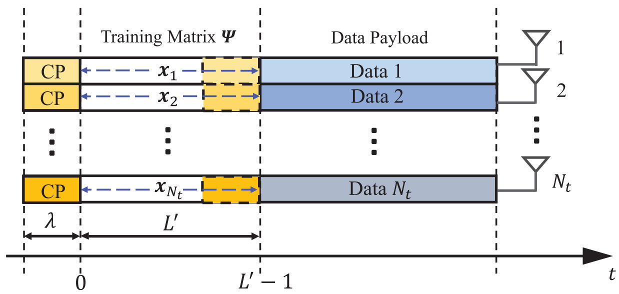

FIGURE 4. A generic training-based SC-MIMO transmission structure.

for $p=1,2,\ldots,N_t$. Then, the $L'\times 1$ complex received signal vector at a RA can be expressed as

$$\boldsymbol{y}=\boldsymbol{X}\boldsymbol{h}+\boldsymbol{w} \quad (45)$$

where $\boldsymbol{h}=[\boldsymbol{h}_1,\boldsymbol{h}_2,\ldots,\boldsymbol{h}_{N_t}]^T$ stands for the channel matrix and $\boldsymbol{w}=[w_0,w_1,\ldots,w_{L'-1}]^T$ stands for the complex additive white Gaussian noise (AWGN) with zero mean and variance σ^2. By using the LS channel estimator [28], [44], the normalized mean square error can be derived as

$$\text{MSE}=\frac{\sigma^2}{N_t\lambda+N_t}\text{Tr}\left(\left(\boldsymbol{X}^H\boldsymbol{X}\right)^{-1}\right). \quad (46)$$

Therefore, the minimum MSE can be achieved as σ^2/E [28] if and only if

$$\phi\left(\boldsymbol{x}_i,\boldsymbol{x}_j;u\right)=\begin{cases}E, & \text{if } i=j, u=0;\\ 0, & \text{if } i\neq j, 0\leq u\leq\lambda,\\ & \text{or } i=j, 1\leq u\leq\lambda.\end{cases} \quad (47)$$

Remark 5: Since the training-based multiple-antenna transmission incorporates the GSM transmission scheme in Fig. 1 as a special case with a particular focus on the training matrix design, (47) is referred to as the optimal condition for GSM training sequences under the LS channel estimator.

Furthermore, it should be noted that the training matrix $\boldsymbol{\Psi}$ needs to be sparse since every GSM system only activates a few antennas at each time-slot.

Hence, the following design criterion provides the optimal channel estimation conditions for the GSM system.

Design criterion: A training matrix $\boldsymbol{\Psi}$ for the GSM system can achieve the optimal channel estimation over the frequency-selective channel with delay spread λ, if it satisfies the following two conditions.

(1) Each column of the training matrix $\boldsymbol{\Psi}$ has exactly N_a non-zero entries since N_a TAs are activated over each time-slot in the GSM system.

(2) The training matrix $\boldsymbol{\Psi}$ needs to meet the condition in (47).

B. PROPOSED GSM TRAINING MATRIX

Based on the design criteria outlined in Section IV-A, we generate the training matrix employing the proposed E-CZCSs for the broadband GSM system.

For positive integers N_t and N_a, we let $V = \lceil \frac{N_t}{N_a} \rceil$ where N_t is the number of transmit antennas and N_a is the number of RF chains. Let $\boldsymbol{\Psi}_1, \boldsymbol{\Psi}_2, \ldots, \boldsymbol{\Psi}_V$ be the training blocks as follows:

$$\boldsymbol{\Psi}_1 = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_{N_a} \end{bmatrix}, \boldsymbol{\Psi}_2 = \begin{bmatrix} \boldsymbol{x}_{N_a+1} \\ \boldsymbol{x}_{N_a+2} \\ \vdots \\ \boldsymbol{x}_{2N_a} \end{bmatrix}, \ldots, \boldsymbol{\Psi}_V = \begin{bmatrix} \boldsymbol{x}_{(V-1)N_a+1} \\ \boldsymbol{x}_{(V-1)N_a+2} \\ \vdots \\ \boldsymbol{x}_{VN_a} \end{bmatrix}. \tag{48}$$

Choosing an (M, N, L, Z)-E-CZCS $\mathcal{G}$ with the condition $M \geq N_a$, we let $\mathcal{X}_0, \mathcal{X}_1, \ldots, \mathcal{X}_{N-1}$ be the training sub-blocks of size $N_a \times VL$ as follows:

$$\mathcal{X}_0 = \begin{bmatrix} \boldsymbol{g}_0^0 & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{g}_0^1 & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{g}_0^{N_a-1} & \boldsymbol{0} & \ldots & \boldsymbol{0} \end{bmatrix}, \mathcal{X}_1 = \begin{bmatrix} \boldsymbol{g}_1^0 & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{g}_1^1 & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{g}_1^{N_a-1} & \boldsymbol{0} & \ldots & \boldsymbol{0} \end{bmatrix},$$
$$\ldots, \mathcal{X}_{N-1} = \begin{bmatrix} \boldsymbol{g}_{N-1}^0 & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{g}_{N-1}^1 & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{g}_{N-1}^{N_a-1} & \boldsymbol{0} & \ldots & \boldsymbol{0} \end{bmatrix} \tag{49}$$

where $\{\boldsymbol{g}_0^0, \boldsymbol{g}_1^0, \ldots, \boldsymbol{g}_{N-1}^0\}$, $\{\boldsymbol{g}_0^1, \boldsymbol{g}_1^1, \ldots, \boldsymbol{g}_{N-1}^1\}$, $\ldots$, $\{\boldsymbol{g}_0^{N_a-1}, \boldsymbol{g}_1^{N_a-1}, \ldots, \boldsymbol{g}_{N-1}^{N_a-1}\}$ are N_a constituent sequence sets from the (M, N, L, Z)-E-CZCS $\mathcal{G}$ and $\boldsymbol{0}$ represents all-zero vector of length L. Then, a $VN_a \times NVL$ GSM training matrix (N_t, N_a, V, N, L)-$\boldsymbol{\Psi}$ is provided as

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\Psi}_1 \\ \boldsymbol{\Psi}_2 \\ \vdots \\ \boldsymbol{\Psi}_V \end{bmatrix} = \begin{bmatrix} \mathcal{X}_0 & \mathcal{X}_1 & \ldots & \mathcal{X}_{N-1} \\ \mathcal{X}_0^{(L)} & \mathcal{X}_1^{(L)} & \ldots & \mathcal{X}_{N-1}^{(L)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}_0^{((V-1)L)} & \mathcal{X}_1^{((V-1)L)} & \ldots & \mathcal{X}_{N-1}^{((V-1)L)} \end{bmatrix}. \tag{50}$$

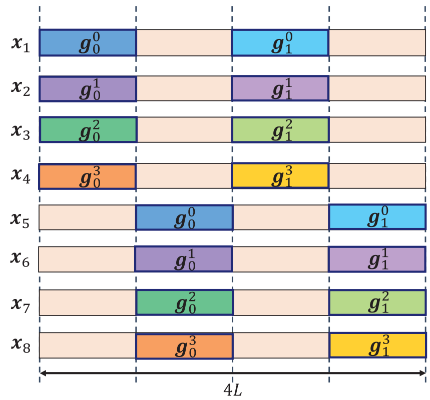

FIGURE 5. The GSM training matrix (8, 4, 2, 2, L)-Ψ based on a (4, 2, L, Z)-E-CZCS in the Example 4.

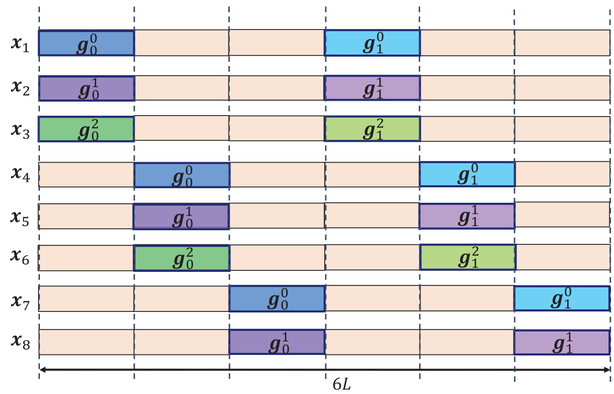

FIGURE 6. The GSM training matrix (8, 3, 3, 2, L)-Ψ based on a (4, 2, L, Z)-E-CZCS in the Example 4.

It is noted that in the scenario where $N_t < VN_a$, the first N_t rows of $\boldsymbol{\Psi}$ are selected as training sequences for N_t transmit antennas.

Example 4: Consider a GSM system equipped with $N_t = 8$ TAs and $N_a = 4$ RF chains. We have $V = \lceil \frac{N_t}{N_a} \rceil = 2$. The $(8, 4, 2, 2, L)$-$\boldsymbol{\Psi}$ GSM training matrix based on a $(4, 2, L, Z)$-E-CZCS is shown in Fig. 5. If we consider another scenario for the GSM system with $N_t = 8$ TAs and $N_a = 3$ RF chains (i.e., $V = \lceil \frac{N_t}{N_a} \rceil = 3$), the $(8, 3, 3, 2, L)$-$\boldsymbol{\Psi}$ GSM training matrix based on a $(4, 2, L, Z)$-E-CZCS can be expressed as shown in Fig. 6.

We then demonstrate that the training matrix $\boldsymbol{\Psi}$ employing the proposed E-CZCSs satisfies the design criteria mentioned in Section IV-A Firstly, there are N_a non-zero entries and $(V-1)N_a$ zeros in each column of the training matrix $\boldsymbol{\Psi}$ as shown in (50). Secondly, we prove that the training sequences $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{N_t}$, i.e., the first N_t rows of $\boldsymbol{\Psi}$, meet the condition in (47) if $Z \geq \lambda$. Here, we consider an (M, N, L, Z)-E-CZCS $\mathcal{G} = \{G^0, G^1, \ldots, G^{M-1}\}$ with $M \geq N_a$ and $Z \geq \lambda$ where λ is the delay spread. The number of transmit antennas N_t is divided into V antennas groups $B_1, B_2, \ldots, B_V$ where each B_v consists of N_a antennas where $B_v = \{1 + (v-1)N_a, 2 + (v-1)N_a, \ldots, vN_a\}$ for $v = 1, 2, \ldots, V$. We consider three cases below to show that the training matrix $\boldsymbol{\Psi}$ satisfies the condition in (47).

Case 1: Since the sets $G^0 = \{\boldsymbol{g}_0^0, \boldsymbol{g}_1^0, \ldots, \boldsymbol{g}_{N-1}^0\}$, $G^1 = \{\boldsymbol{g}_0^1, \boldsymbol{g}_1^1, \ldots, \boldsymbol{g}_{N-1}^1\}, \ldots,$ and $G^{N_a-1} = \{\boldsymbol{g}_0^{N_a-1}, \boldsymbol{g}_1^{N_a-1}, \ldots, \boldsymbol{g}_{N-1}^{N_a-1}\}$ in $\mathcal{G}$ satisfy the condition (C1) in (16), we have

$$\begin{aligned}
&\phi(\boldsymbol{x}_k, \boldsymbol{x}_l; u) \\
&= \sum_{j=0}^{N-1} \rho\left(\boldsymbol{g}_j^{m_k}, \boldsymbol{g}_j^{m_l}; u\right) \\
&= \begin{cases} \sum_{j=0}^{N-1} \rho\left(\boldsymbol{g}_j^{m_l}, \boldsymbol{g}_j^{m_l}; u\right), \text{ for } k = l, 0 \le u \le Z; \\ \sum_{j=0}^{N-1} \rho\left(\boldsymbol{g}_j^{m_k}, \boldsymbol{g}_j^{m_l}; u\right), \text{ for } k \ne l, 0 \le u \le Z; \end{cases} \\
&= \begin{cases} \rho(G^{m_l}, G^{m_l}; u), \text{ for } k = l, 0 \le u \le Z; \\ \rho(G^{m_k}, G^{m_l}; u), \text{ for } k \ne l, 0 \le u \le Z; \end{cases} \\
&= \begin{cases} NL, \text{ for } k = l, u = 0; \\ 0, \text{ for } k = l, 1 \le u \le Z; \\ 0, \text{ for } k \ne l, 0 \le u \le Z, \end{cases}
\end{aligned} \tag{51}$$

for any $k, l \in B_v$ and $v = 1, 2, \ldots, V$ where $m_k = (k-1)_{\bmod N_a}$ and $m_l = (l-1)_{\bmod N_a}$. Over each training block $\boldsymbol{\Psi}_v$, the ISI at each TA and the IAI between the k-th and the l-th TAs caused by multipath delay can be eliminated.

Case 2: For $k \in B_v$, $l \in B_{v+1}$, and $v = 1, 2, \ldots, V-1$, we have

$$\begin{aligned}
&\phi(\boldsymbol{x}_l, \boldsymbol{x}_k; u) \\
&= \sum_{j=0}^{N-1} \rho^*\left(\boldsymbol{g}_j^{m_k}, \boldsymbol{g}_j^{m_l}; L-u\right) \\
&= \begin{cases} \sum_{j=0}^{N-1} \rho^*\left(\boldsymbol{g}_j^{m_k}, \boldsymbol{g}_j^{m_k}; L-u\right), \\ \text{if } l = k + N_a; \\ \sum_{j=0}^{N-1} \rho^*\left(\boldsymbol{g}_j^{m_k}, \boldsymbol{g}_j^{m_l}; L-u\right), \\ \text{otherwise}; \end{cases} \\
&= \begin{cases} \rho^*(G^{m_k}, G^{m_k}; L-u), \text{ if } l = k + N_a; \\ \rho^*(G^{m_k}, G^{m_l}; L-u), \text{ otherwise}; \end{cases} \\
&= 0
\end{aligned} \tag{52}$$

where $m_k = (k-1)_{\bmod N_a}$, $m_l = (l-1)_{\bmod N_a}$, and $1 \le u \le Z$. The IAI between the k-th TA in $\boldsymbol{\Psi}_v$ and the l-th TA in $\boldsymbol{\Psi}_{v+1}$ is eliminated.

Case 3: For $k \in B_1$ and $l \in B_V$, according to (C2) in (16), we have

$$\begin{aligned}
&\phi(\boldsymbol{x}_k, \boldsymbol{x}_l; u) \\
&= \sum_{j=0}^{N-1} \rho^*\left(\boldsymbol{g}_j^{m_l}, \boldsymbol{g}_{(j+1)_{\bmod N}}^{m_k}; L-u\right) \\
&= \begin{cases} \sum_{j=0}^{N-1} \rho^*\left(\boldsymbol{g}_j^{m_k}, \boldsymbol{g}_{(j+1)_{\bmod N}}^{m_k}; L-u\right), \\ \text{if } l = k + (V-1)N_a; \\ \sum_{j=0}^{N-1} \rho^*\left(\boldsymbol{g}_j^{m_l}, \boldsymbol{g}_{(j+1)_{\bmod N}}^{m_k}; L-u\right), \\ \text{otherwise}; \end{cases} \\
&= \begin{cases} \hat{\rho}^*(G^{m_k}, G^{m_k}; L-u), \\ \text{if } l = k + (V-1)N_a; \\ \hat{\rho}^*(G^{m_l}, G^{m_k}; L-u), \\ \text{otherwise}; \end{cases} \\
&= 0
\end{aligned} \tag{53}$$

where $m_k = (k-1)_{\bmod N_a}$, $m_l = (l-1)_{\bmod N_a}$, and $1 \le u \le Z$. It means that the IAI between the k-th TA in $\boldsymbol{\Psi}_1$ and the l-th TA in $\boldsymbol{\Psi}_V$ is eliminated. Note that the last equality in (53) follows from the aperiodic cross-correlation property (C2) of the E-CZCS. From the above three cases, we can conclude that the training matrix $\boldsymbol{\Psi}$ employing the proposed (M, N, L, Z)-E-CZCS $\mathcal{G}$ achieves the condition in (47) if $Z \ge \lambda$.

V. SIMULATIONS

In this section, we examine the channel estimation performance of the proposed E-CZCS-based training for the GSM system over the frequency-selective channel. We consider a $(\lambda+1)$-path channel separated by integer symbol durations as $h[t] = \sum_{i=0}^{\lambda} h_i \delta[t - iT_s]$ where h_i's are complex Gaussian random variables with zero mean and $E(|h_i|^2) = 1/(\lambda+1)$ for all i. We evaluate the channel estimation performance of the GSM training matrices based on our proposed E-CZCSs and other classes of sequence sets including the SZCCS, ZCCS, CZCPs, binary random sequences, and Zadoff-Chu sequences. Our first simulation setup consists of $N_t = 8$ TAs, $N_a = 4$ RF chains, and $N_r = 1$ RA. We employ the binary $(4, 2, 32, 8)$-E-CZCS from *Example 3* to generate the $(8, 4, 2, 2, 32)$-$\boldsymbol{\Psi}$ as depicted in Fig. 5. For the ZCCS and the SZCCS, the training matrix is given by

$$\begin{bmatrix} \boldsymbol{s}_0^0 & \boldsymbol{0} & \boldsymbol{s}_1^0 & \boldsymbol{0} \\ \boldsymbol{s}_0^1 & \boldsymbol{0} & \boldsymbol{s}_1^1 & \boldsymbol{0} \\ \boldsymbol{s}_0^2 & \boldsymbol{0} & \boldsymbol{s}_1^2 & \boldsymbol{0} \\ \boldsymbol{s}_0^3 & \boldsymbol{0} & \boldsymbol{s}_1^3 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{s}_0^0 & \boldsymbol{0} & \boldsymbol{s}_1^0 \\ \boldsymbol{0} & \boldsymbol{s}_0^1 & \boldsymbol{0} & \boldsymbol{s}_1^1 \\ \boldsymbol{0} & \boldsymbol{s}_0^2 & \boldsymbol{0} & \boldsymbol{s}_1^2 \\ \boldsymbol{0} & \boldsymbol{s}_0^3 & \boldsymbol{0} & \boldsymbol{s}_1^3 \end{bmatrix}_{8\times 128} \tag{54}$$

where $\{\boldsymbol{s}_0^0, \boldsymbol{s}_1^0\}$, $\{\boldsymbol{s}_0^1, \boldsymbol{s}_1^1\}$, $\{\boldsymbol{s}_0^2, \boldsymbol{s}_1^2\}$, and $\{\boldsymbol{s}_0^3, \boldsymbol{s}_1^3\}$ are the constituent sets of a $(4, 2, 32, 16)$-ZCCS and the first four sequence sets of the $(8, 2, 32, 7)$-SZCCS from [47], respectively. For the training matrix based on CZCPs, the pairs $(\boldsymbol{s}_0^0, \boldsymbol{s}_1^0)$, $(\boldsymbol{s}_0^1, \boldsymbol{s}_1^1)$, $(\boldsymbol{s}_0^2, \boldsymbol{s}_1^2)$, and $(\boldsymbol{s}_0^3, \boldsymbol{s}_1^3)$ in (54) are 4 distinct $(32, 16)$-CZCPs from [28]. These CZCPs satisfy the conditions (C1) and (C2) in (16) only when $m_1 = m_2$. For binary random sequences, the elements of $\boldsymbol{s}_0^0$, $\boldsymbol{s}_1^0$, $\boldsymbol{s}_0^1$, $\boldsymbol{s}_1^1$, $\boldsymbol{s}_0^2$, $\boldsymbol{s}_1^2$, $\boldsymbol{s}_0^3$, and $\boldsymbol{s}_1^3$ in (54) are randomly generated from "+1" or "−1". For the training matrix based on Zadoff-Chu sequences, the sequences $\boldsymbol{s}_0^0$, $\boldsymbol{s}_1^0$, $\boldsymbol{s}_0^1$, $\boldsymbol{s}_1^1$, $\boldsymbol{s}_0^2$, $\boldsymbol{s}_1^2$, $\boldsymbol{s}_0^3$, and $\boldsymbol{s}_1^3$ are assigned by 8 distinct Zadoff-Chu sequences of length 32 with low cross-correlations. The MSE performances of the channel estimation based on different training matrices as shown in Figs. 7 and 8. Fig. 7 shows that the training matrix based on

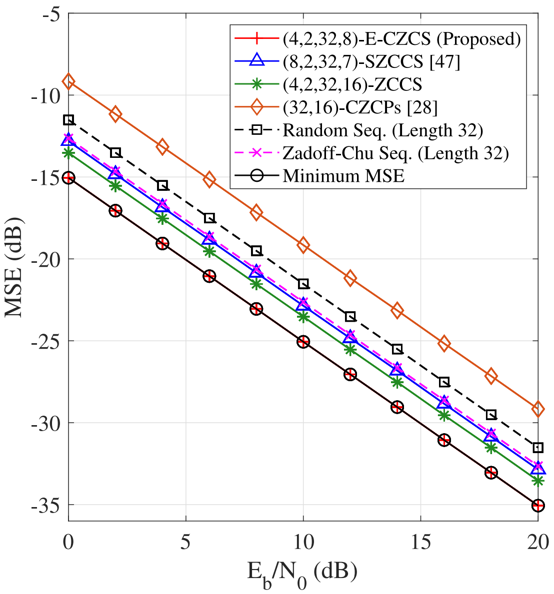

FIGURE 7. MSE comparison of GSM training based on different sequences with 8 TAs and 4 active TAs.

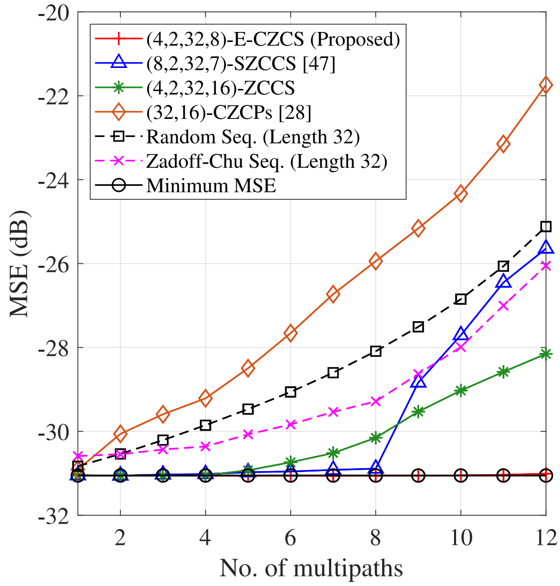

FIGURE 8. MSE comparison of GSM training based on different sequences with 8 TAs and 4 active TAs.

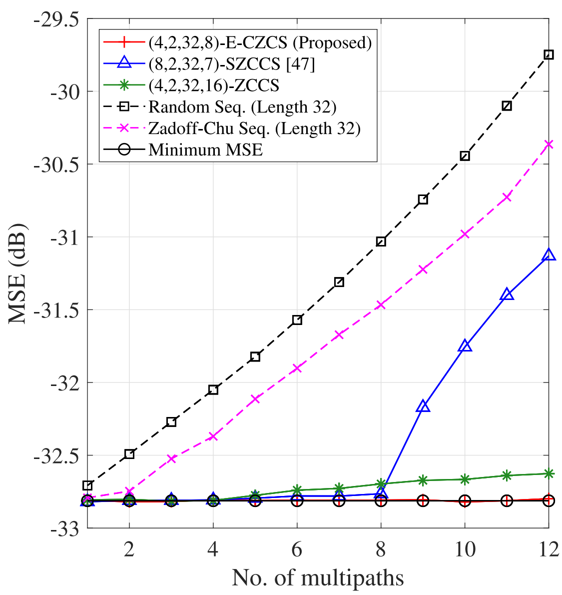

FIGURE 9. MSE comparison of GSM training based on different sequences with 8 TAs and 3 active TAs.

the $(4, 2, 32, 8)$-E-CZCS achieves the minimum MSE with 2.2 and 1.5 dB gains over the SZCCS and ZCCS based training, respectively, when the number of multipaths is 9. In Fig. 8, we consider different numbers of multipaths at $E_b/N_0 = 16$ dB. When the number of multipaths is less than or equal to 9, i.e., $\lambda = 8$, our proposed E-CZCS-based training matrix can achieve the minimum MSE since the ZCZ width is 8. Also, we observe that the $(4, 2, 32, 8)$-E-CZCS outperforms others and still performs close to the MSE lower bound even the number of multipaths is larger than 9. This is because the out-of-zone correlations of the proposed $(4, 2, 32, 8)$-E-CZCS are small. For the SZCCS-based GSM, the performance is worse when the number of multipaths is larger than 8. This is because the SZCCS only consider the condition (C1) in (16) and the ZCZ width is only 7. When the number of multipaths is larger than 8, the out-of-zone correlations of the SZCCS degrade the channel estimation performance. For the ZCCS-based training, the condition (C2) in (16) is not met, thus leading nonzero IAI. For the CZCP-based training, the performance is worse since the conditions (C1) and (C2) are not taken into consideration when $m_1 \neq m_2$ and hence the IAI is introduced.

Furthermore, we discuss the comparison with the GSM training in [47] regarding the training efficiency. The training efficiency is modeled by T/T' where T stands for the length of the interval during which the training sequences are transmitted and T' stands for the length of the total training interval. This metric indicates the effectiveness of the training framework. When the training efficiency is 1, it implies that the training sequences are transmitted on every time slot during the training interval. Compared to the GSM training in [47], the training efficiency of our proposed training framework is $NVL/NVL = 1$, whereas that of the training framework in [47] is $NVL/(NVL + N\lambda) < 1$ where λ is the delay spread. In the case we consider in Fig. 8, we have $V = \lceil \frac{N_t}{N_a} \rceil = \lceil \frac{8}{4} \rceil = 2$, $N = 2$, and $L = 32$. As a result, the training efficiency of the training framework in [47] is only 0.85 with $\lambda = 11$.

In Fig. 9, we consider the GSM system with $N_t = 8$ TAs, $N_a = 3$ RF chains, and $N_r = 1$ RA. We use the GSM training matrix $(8, 3, 3, 2, 32)$-$\mathbf{\Psi}$ as depicted in Fig. 6 based on the binary $(4, 2, 32, 8)$-E-CZCS from *Example 3*. For comparison, we take the first three sequence sets of the $(8, 2, 32, 7)$-SZCCS from [47] into consideration, as well as the $(4, 2, 32, 16)$-ZCCS, binary random sequences, and

Zadoff-Chu sequences. The training matrix is given by

$$\begin{bmatrix} s_0^0 & 0 & 0 & s_1^0 & 0 & 0 \\ s_0^1 & 0 & 0 & s_1^1 & 0 & 0 \\ s_0^2 & 0 & 0 & s_1^2 & 0 & 0 \\ 0 & s_0^0 & 0 & 0 & s_1^0 & 0 \\ 0 & s_0^1 & 0 & 0 & s_1^1 & 0 \\ 0 & s_0^2 & 0 & 0 & s_1^2 & 0 \\ 0 & 0 & s_0^0 & 0 & 0 & s_1^0 \\ 0 & 0 & s_0^1 & 0 & 0 & s_1^1 \end{bmatrix}_{8\times 192} \tag{55}$$

where the component sequences $\boldsymbol{s}_n^m$'s are assigned in a similar manner as in the previous simulation. For example, $\{\boldsymbol{s}_0^0, \boldsymbol{s}_1^0\}$, $\{\boldsymbol{s}_0^1, \boldsymbol{s}_1^1\}$, and $\{\boldsymbol{s}_0^2, \boldsymbol{s}_1^2\}$ are the three constituent sets of the $(4, 2, 32, 16)$-ZCCS if the training matrix is based on the ZCCS. For Fig. 9, the E_b/N_0 is fixed at 16 dB. We observe that the $(4, 2, 32, 8)$-E-CZCS outperforms others and still performs close to the MSE lower bound even the number of multipaths is larger than 9. This is because the out-of-zone correlations of the proposed $(4, 2, 32, 8)$-E-CZCS are small. For the SZCCS based training, the performance degrades significantly when the number of multipaths is larger than 8.

VI. CONCLUSION

This paper is focused on a novel class of sequence sets called E-CZCSs, each consisting of a collection of CZCSs and with an additional cross-channel aperiodic correlation sum property. We have proposed two systematic constructions of E-CZCSs including one based on ZCCSs, MOCSs, and CCCs (*Theorem 2*) and the other based on generalized Boolean functions (*Theorem 3*). Both constructions can generate the binary E-CZCSs with maximum ZCZ width.

Furthermore, a novel GSM training framework has been proposed based on E-CZCSs. It is shown that the proposed training design can achieve the optimal channel estimation performance in frequency-selective channels, by fully exploiting the correlation properties of E-CZCS.

Although *Theorem 3* can generate E-CZCSs with various set sizes and ZCZ widths, the lengths are currently limited to powers of two. Therefore, a potential future direction is to construct E-CZCSs with non-power-two sequence lengths.

APPENDIX

PROOF OF THEOREM 3

Before proving *Theorem 3*, we introduce the following lemma which can be used to prove our main theorem.

Lemma 2 [54]: For any integer m and k with $0 < k \le m$, let nonempty sets $U_1, U_2, \ldots, U_k$ be a partition of $\{1, 2, \ldots, m\}$. Also let π_α be a bijection from $\{1, 2, \ldots, m_\alpha\}$ to U_α where m_α is the order of U_α for $\alpha = 1, 2, \ldots, k$. Given an even positive integer q and the generalized Boolean function f

$$f = \frac{q}{2}\sum_{\alpha=1}^{k}\sum_{\beta=1}^{m_\alpha-1} x_{\pi_\alpha(\beta)}x_{\pi_\alpha(\beta+1)} + \sum_{i=1}^{m}\eta_i x_i + \eta_0 \tag{56}$$

where η_i's $\in \mathbb{Z}_q$. For $0 \le \kappa, \nu \le 2^k - 1$, the set $C^\nu = \{\boldsymbol{c}_0^\nu, \boldsymbol{c}_1^\nu, \ldots, \boldsymbol{c}_{2^k-1}^\nu\}$ can be constructed as follows:

$$\boldsymbol{c}_\kappa^\nu = \boldsymbol{f} + \frac{q}{2}\sum_{\alpha=1}^{k}\kappa_\alpha \boldsymbol{x}_{\pi_\alpha(1)} + \frac{q}{2}\sum_{\alpha=1}^{k}\nu_\alpha \boldsymbol{x}_{\pi_\alpha(m_\alpha)} \tag{57}$$

where $(\kappa_1, \kappa_2, \ldots, \kappa_k)$ and $(\nu_1, \nu_2, \ldots, \nu_k)$ are binary representations of κ and ν, respectively. Then, $C^0, C^1, \ldots, C^{2^k-1}$ form a $(2^k, 2^m)$-CCC.

Proof of Theorem 3: We consider three parts to illustrate that $\mathcal{G}$ satisfies (C1) and (C2) in (16) where $\mathcal{T}_1 = \{1, 2, \ldots, 2^{\pi_1(1)-1}\}$ and $\mathcal{T}_2 = \{2^m - 2^{\pi_1(1)-1}, 2^m - 2^{\pi_1(1)-1} + 1, \ldots, 2^m - 1\}$. Let $\boldsymbol{g}_n^p = (g_{n,0}^p, g_{n,1}^p, \ldots, g_{n,L-1}^p)$ for $p = 0, 1, \ldots, 2^\nu - 1$ and $n = 0, 1, \ldots, 2^k - 1$.

In the first part, we have to demonstrate that

$$\begin{aligned}\rho(G^p, G^p; u) &= \sum_{n=0}^{2^\nu-1}\rho(\zeta_q(\boldsymbol{g}_n^p), \zeta_q(\boldsymbol{g}_n^p); u) \\ &= \sum_{n=0}^{2^\nu-1}\sum_{i=0}^{2^m-1-u}\xi_q^{g_{n,i+u}^p - g_{n,i}^p} = \sum_{i=0}^{2^m-1-u}\sum_{n=0}^{2^\nu-1}\xi_q^{g_{n,i+u}^p - g_{n,i}^p} = 0,\end{aligned} \tag{58}$$

for $|u| \in \mathcal{T}_1 \cup \mathcal{T}_2$. If $\nu = k$, the sequences $\boldsymbol{g}_n^p$ in (38) can be rewritten as

$$\boldsymbol{g}_n^p = \boldsymbol{f} + \frac{q}{2}\left(\sum_{\alpha=1}^{k} n_{k-\alpha+1}\boldsymbol{x}_{\pi_\alpha(1)} + \sum_{\alpha=1}^{k} p_\alpha \boldsymbol{x}_{\pi_\alpha(m_\alpha)}\right) \tag{59}$$

implying $G^p \in \mathcal{G}$ is a GCS as given in *Lemma 2*. Hence, we have $\rho(G^p, G^p; u) = 0$, $|u| \in \{1, 2, \ldots, 2^m - 1\}$.

If $\nu < k$, we consider two cases to show that $\rho(G^p, G^p; u) = \sum_{n=0}^{2^\nu-1}\rho(\zeta_q(\boldsymbol{g}_n^p); u) = 0$ when $|u| \in \mathcal{T}_1$ and $|u| \in \mathcal{T}_2$, respectively. For a nonnegative integer i with binary representation $(i_1, i_2, \ldots, i_m)$, we let $j = i + u$ with binary representation $(j_1, j_2, \ldots, j_m)$.

Case 1-A: We assume $i_{\pi_1(1)} \neq j_{\pi_1(1)}$ in this case. For any sequence $\boldsymbol{g}_n^p \in G^p$ where $0 \le p \le 2^k - 1$ and $0 \le n \le 2^\nu - 1$, there exists a sequence $\boldsymbol{g}_s^p = \boldsymbol{g}_n^p + (q/2)\boldsymbol{x}_{\pi_1(1)} \in G^p$ such that

$$g_{n,j}^p - g_{n,i}^p - g_{s,j}^p + g_{s,i}^p = \frac{q}{2}\left(i_{\pi_1(1)} - j_{\pi_1(1)}\right) \equiv \frac{q}{2} \pmod{q}. \tag{60}$$

Since $i_{\pi_1(1)} \neq j_{\pi_1(1)}$, we can obtain

$$\xi_q^{g_{n,j}^p - g_{n,i}^p} / \xi_q^{g_{s,j}^p - g_{s,i}^p} = \xi_q^{\frac{q}{2}\left(i_{\pi_1(1)} - j_{\pi_1(1)}\right)} = e^{\frac{j2\pi}{q}\frac{q}{2}} = -1 \tag{61}$$

implying $\xi^{g_{n,j}^p - g_{n,i}^p} + \xi^{g_{s,j}^p - g_{s,i}^p} = 0$. Therefore, we have $\sum_{n=0}^{2^\nu-1}\xi_q^{g_{n,j}^p - g_{n,i}^p} = 0$.

Case 1-B: In this case we have $i_{\pi_1(1)} = j_{\pi_1(1)}$ and we can deduce that $i_{\pi_{\nu+\gamma}(1)} = j_{\pi_{\nu+\gamma}(1)}$ for $\gamma = 1, 2, \ldots, k - \nu$. Suppose not, let α' be the smallest integer satisfying $i_{\pi_{\nu+\alpha'}(1)} \neq j_{\pi_{\nu+\alpha'}(1)}$. Therefore, $i_m = j_m, i_{m-1} = j_{m-1}, \ldots, i_{m-\alpha'+2} = j_{m-\alpha'+2}$. Then,

$$u = j - i = 2^{m-\alpha'} + \sum_{s=1,\ s\neq\pi_1(1)}^{m-\alpha'}(j_s - i_s)2^{s-1}$$

$$\geq 2^{m-\alpha'} - \sum_{s=1}^{m-\alpha'} 2^{s-1} + 2^{\pi_1(1)-1} = 2^{\pi_1(1)-1} + 1 \quad (62)$$

which contradicts the assumption that $u \leq 2^{\pi_1(1)-1}$. So we have $i_{\pi_{v+1}(1)} = j_{\pi_{v+1}(1)}, i_{\pi_{v+2}(1)} = j_{\pi_{v+2}(1)}, \ldots, i_{\pi_k(1)} = j_{\pi_k(1)}$ here. Then we consider two subcases below.

Case 1-B (i): We assume $i_{\pi_\alpha(1)} \neq j_{\pi_\alpha(1)}$ for some $\alpha = 2, 3, \ldots, v$. For any sequence $\boldsymbol{g}_n^p \in G^p$, there exists another sequence $\boldsymbol{g}_s^p = \boldsymbol{g}_n^p + (q/2)\boldsymbol{x}_{\pi_\alpha(1)} \in G^p$ such that $\xi_q^{g_{n,j}^p - g_{n,i}^p} + \xi_q^{g_{s,j}^p - g_{s,i}^p} = 0$.

Case 1-B (ii): Following the above case, we have $i_{\pi_\alpha(1)} = j_{\pi_\alpha(1)}$ for all $\alpha = 1, 2, \ldots, k$. We assume $i_{\pi_\alpha(\beta)} = j_{\pi_\alpha(\beta)}$ for $\alpha = 1, 2 \ldots, \hat{\alpha} - 1$ with $\hat{\alpha} \leq k$ and $\beta = 1, 2, \ldots, m_\alpha$. Then we suppose that $\hat{\beta}$ is the smallest integer such that $i_{\pi_{\hat{\alpha}}(\hat{\beta})} \neq j_{\pi_{\hat{\alpha}}(\hat{\beta})}$. Let i' and j' be two integers which are distinct from i and j, respectively, only in one position $\pi_{\hat{\alpha}}(\hat{\beta} - 1)$. That is, $i'_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} = 1 - i_{\pi_{\hat{\alpha}}(\hat{\beta}-1)}$ and $j'_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} = 1 - j_{\pi_{\hat{\alpha}}(\hat{\beta}-1)}$. Hence, we have

$$\begin{aligned} g_{n,i'}^p - g_{n,i}^p &= \frac{q}{2}\Big(i_{\pi_{\hat{\alpha}}(\hat{\beta}-2)} i'_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} - i_{\pi_{\hat{\alpha}}(\hat{\beta}-2)} i_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} \\ &\quad + i'_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} i_{\pi_{\hat{\alpha}}(\hat{\beta})} - i_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} i_{\pi_{\hat{\alpha}}(\hat{\beta})}\Big) \\ &\quad + \eta_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} i'_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} + \eta_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} i_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} \\ &\equiv \frac{q}{2}\Big(i_{\pi_{\hat{\alpha}}(\hat{\beta}-2)} + i_{\pi_{\hat{\alpha}}(\hat{\beta})}\Big) + \eta_{\pi_{\hat{\alpha}}(\hat{\beta}-1)}\Big(1 - 2i_{\pi_{\hat{\alpha}}(\hat{\beta}-1)}\Big) \\ &\quad \pmod{q}. \end{aligned} \quad (63)$$

Since $i_{\pi_{\hat{\alpha}}(\hat{\beta}-2)} = j_{\pi_{\hat{\alpha}}(\hat{\beta}-2)}$ and $i_{\pi_{\hat{\alpha}}(\hat{\beta}-1)} = j_{\pi_{\hat{\alpha}}(\hat{\beta}-1)}$, we have

$$\begin{aligned} g_{n,j}^p - g_{n,i}^p - g_{n,j'}^p + g_{n,i'}^p &\equiv \frac{q}{2}\Big(i_{\pi_{\hat{\alpha}}(\hat{\beta})} - j_{\pi_{\hat{\alpha}}(\hat{\beta})}\Big) \\ &\equiv \frac{q}{2} \pmod{q}. \end{aligned} \quad (64)$$

Then, we can obtain

$$\xi_q^{g_{n,j}^p - g_{n,i}^p} / \xi_q^{g_{n,j'}^p - g_{n,i'}^p} = \xi_q^{\frac{q}{2}\left(i_{\pi_{\hat{\alpha}}(\hat{\beta})} - j_{\pi_{\hat{\alpha}}(\hat{\beta})}\right)} = e^{\frac{j2\pi}{q}\frac{q}{2}} = -1. \quad (65)$$

Therefore,

$$\xi_q^{g_{n,j}^p - g_{n,i}^p} + \xi_q^{g_{n,j'}^p - g_{n,i'}^p} = 0. \quad (66)$$

From Case 1-A and Case 1-B, we can conclude that $\rho(G^p, G^p; u) = 0, \text{ for } |u| \in \mathcal{T}_1$.

Case 2: Then, let us consider $|u| \in \mathcal{T}_2$, i.e., $2^m - 2^{\pi_1(1)-1} \leq |u| \leq 2^m - 1$. In this case, we should have $i_{\pi_1(1)} \neq j_{\pi_1(1)}$. Suppose not. If $i_{\pi_1(1)} = j_{\pi_1(1)}$, then we have

$$u = j - i = \sum_{s=1, s\neq\pi_1(1)}^{m} (j_s - i_s)2^{s-1} \leq 2^m - 2^{\pi_1(1)-1} - 1 \quad (67)$$

which contradicts the assumption $2^m - 2^{\pi_1(1)-1} \leq |u| \leq 2^m - 1$. Hence, we must have $i_{\pi_1(1)} \neq j_{\pi_1(1)}$ here. Following the similar arguments as given in *Case 1-A*, we can also obtain $\xi_q^{g_{n,j}^p - g_{n,i}^p} + \xi_q^{g_{s,j}^p - g_{s,i}^p} = 0$ where $\boldsymbol{g}_s^p = \boldsymbol{g}_n^p + (q/2)\boldsymbol{x}_{\pi_1(1)} \in G^p$. Therefore,

$$\rho\big(G^p, G^p; u\big) = \sum_{n=0}^{2^v-1} \sum_{i=0}^{2^m-1-u} \xi_q^{g_{n,j}^p - g_{n,i}^p} = 0, \quad \text{for } |u| \in \mathcal{T}_2. \quad (68)$$

In the second part, we will demonstrate that any two distinct constituent sets G^p and G^l, where $0 \leq p \neq l \leq 2^k - 1$, have zero cross-correlation sum for $|u| \in \mathcal{T}_1 \cup \mathcal{T}_2$, i.e.,

$$\begin{aligned} \rho\big(G^p, G^l; u\big) &= \sum_{n=0}^{2^v-1} \rho\big(\zeta_q(\boldsymbol{g}_n^p), \zeta_q(\boldsymbol{g}_n^l); u\big) \\ &= \sum_{n=0}^{2^v-1} \sum_{i=0}^{2^m-1-u} \xi_q^{g_{n,i+u}^p - g_{n,i}^l} = \sum_{i=0}^{2^m-1-u} \sum_{n=0}^{2^v-1} \xi_q^{g_{n,i+u}^p - g_{n,i}^l} = 0. \end{aligned} \quad (69)$$

For $v = k$, similar to the first part, we can obtain that G^p and G^l are mutually orthogonal GCSs, i.e., $\rho(G^p, G^l; u) = 0, \ |u| \in \{0, 1, 2, \ldots, 2^m - 1\}$. For $v < k$, by following the similar arguments in the first part, we can obtain that

$$\sum_{n=0}^{2^v-1} \rho\big(\zeta_q(\boldsymbol{g}_n^p), \zeta_q(\boldsymbol{g}_n^l); u\big) = 0, \quad \text{for } |u| \in \mathcal{T}_1 \cup \mathcal{T}_2. \quad (70)$$

Now, it only suffices to show that

$$\rho\big(G^p, G^l; 0\big) = \sum_{n=0}^{2^v-1} \sum_{i=0}^{2^m-1} \xi_q^{g_{n,i}^p - g_{n,i}^l} = 0. \quad (71)$$

We denote p_α and l_α as the α-th bits of the binary representations of p and l, respectively. Also, let $i_{\pi_\alpha(m_\alpha)}$ be the $\pi_\alpha(m_\alpha)$-th bit of the binary representation of i. According to (38), we have

$$g_{n,i}^p - g_{n,i}^l \equiv \frac{q}{2}\sum_{\alpha=1}^{k}(p_\alpha - l_\alpha) i_{\pi_\alpha(m_\alpha)} \pmod{q}. \quad (72)$$

It can be observed that (72) is the linear combination of the term $i_{\pi_\alpha(m_\alpha)}$. For i ranging from 0 to $2^m - 1$, there are 2^{m-1} i's such that $\xi_q^{g_{n,i}^p - g_{n,i}^l} = \xi_q^{q/2} = -1$ and 2^{m-1} i's such that $\xi_q^{g_{n,i}^p - g_{n,i}^l} = \xi_q^0 = 1$. Therefore, we can obtain $\sum_{i=0}^{2^m-1} \xi_q^{g_{n,i}^p - g_{n,i}^l} = 0$.

In the last part, we will prove the condition (C2) in (16) holds for $\mathcal{G}$, i.e.,

$$\begin{aligned} \hat{\rho}\big(G^p, G^l; u\big) &= \sum_{n=0}^{N-1} \rho\big(\zeta_q(\boldsymbol{g}_n^p), \zeta_q(\boldsymbol{g}_{(n+1)_{\bmod N}}^l); u\big) \\ &= \sum_{i=0}^{2^m-1-u} \sum_{n=0}^{N-1} \xi_q^{g_{n,i+u}^p - g_{(n+1)_{\bmod N},i}^l} = 0 \end{aligned} \quad (73)$$

where $2^m - 2^{\pi_1(1)-1} \leq |u| \leq 2^m - 1$ and $N = 2^v$. Similarly, we let $j = i + u$ for any integer i. From *Case 2* in the first part, we deduce that $i_{\pi_1(1)} \neq j_{\pi_1(1)}$. Let $(n_1, n_2, \ldots, n_v)$ and $(h_1, h_2, \ldots, h_v)$ be the binary representations of n and $h = (n+1)_{\bmod N}$, respectively. We also let n' and h' be the integers that are distinct from n and h, respectively, in only one position, i.e., $n'_v = 1 - n_v$ and $h'_v = 1 - h_v$. We can obtain

$$g^p_{n,j} - g^p_{n',j} = \frac{q}{2}\big(n_v j_{\pi_1(1)} - (1-n_v) j_{\pi_1(1)}\big) = -\frac{q}{2} j_{\pi_1(1)} + q n_v j_{\pi_1(1)} \equiv -\frac{q}{2} j_{\pi_1(1)} \pmod q \tag{74}$$

and

$$g^l_{h,i} - g^l_{h',i} = \frac{q}{2}\big(h_v i_{\pi_1(1)} - (1-h_v) i_{\pi_1(1)}\big) = -\frac{q}{2} i_{\pi_1(1)} + q h_v i_{\pi_1(1)} \equiv -\frac{q}{2} i_{\pi_1(1)} \pmod q. \tag{75}$$

Then, we have

$$g^p_{n,j} - g^l_{h,i} - g^p_{n',j} + g^l_{h',i} \equiv \frac{q}{2}\big(i_{\pi_1(1)} - j_{\pi_1(1)}\big) \equiv \frac{q}{2} \pmod q \tag{76}$$

since $i_{\pi_1(1)} \neq j_{\pi_1(1)}$. Therefore, $\xi_q^{g^p_{n,j} - g^l_{h,i}} + \xi_q^{g^p_{n',j} - g^l_{h',i}} = 0$ and (73) is proved. From the above three parts, we can conclude that $\mathcal{G}$ is a $(2^k, 2^v, 2^m, 2^{\pi_1(1)-1})$-E-CZCS. ∎

REFERENCES

[1] M. J. E. Golay, "Complementary series," *IRE Trans. Inf. Theory*, vol. 7, no. 2, pp. 82–87, Apr. 1961.
[2] C.-C. Tseng and C. L. Liu, "Complementary sets of sequences," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 5, pp. 644–652, Sep. 1972.
[3] N. Suehiro and M. Hatori, "*N*-shift cross-orthogonal sequences," *IEEE Trans. Inf. Theory*, vol. 34, no. 1, pp. 143–146, Jan. 1988.
[4] J. M. Groenewald and B. T. Maharaj, "MIMO channel synchronization using Golay complementary pairs," in *Proc. AFRICON*, Windhoek, South Africa, Sep. 2007, pp. 1–5.
[5] P. Spasojevic and C. N. Georghiades, "Complementary sequences for ISI channel estimation," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1145–1152, Mar. 2001.
[6] S. Wang and A. Abdi, "MIMO ISI channel estimation using uncorrelated Golay complementary sets," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 3024–3039, Sep. 2007.
[7] S.-M. Tseng and M. R. Bell, "Asynchronous multicarrier DS-CDMA using mutually orthogonal complementary sets of sequences," *IEEE Trans. Commun.*, vol. 48, pp. 53–59, Jan. 2000.
[8] H.-H. Chen, J.-F. Yeh, and N. Suehiro, "A multicarrier CDMA architecture based on orthogonal complete complementary codes for new generations of wideband wireless communications," *IEEE Commun. Mag.*, vol. 39, no. 10, pp. 126–134, Oct. 2001.
[9] H.-H. Chen, *The Next Generation CDMA Technologies*. Hoboken, NJ, USA: Wiley, 2007.
[10] Z. Liu, Y. L. Guan, and U. Parampalli, "New complete complementary codes for peak-to-mean power control in multi-carrier CDMA," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 1105–1113, Mar. 2014.
[11] Z. Liu, Y. L. Guan, and H.-H. Chen, "Fractional-delay-resilient receiver design for interference-free MC-CDMA communications based on complete complementary codes," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1226–1236, Mar. 2015.
[12] R. D. J. van Nee, "OFDM codes for peak-to-average power reduction and error correction," in *Proc. IEEE Global Telecommun. Conf.*, London, U.K., Nov. 1996, pp. 740–744.
[13] J. A. Davis and J. Jedwab, "Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed–Müller codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2397–2417, Nov. 1999.
[14] K. G. Paterson, "Generalized Reed-Müller codes and power control in OFDM modulation," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 104–120, Jan. 2000.
[15] C.-Y. Chen, "Complementary sets of non-power-of-two length for peak-to-average power ratio reduction in OFDM," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7538–7545, Dec. 2016.
[16] C.-Y. Chen, "A novel construction of complementary sets with flexible lengths based on boolean functions," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 260–263, Feb. 2018.
[17] C.-Y. Chen, Y.-J. Min, K.-Y. Lu, and C.-C. Chao, "Cell search for cell-based OFDM systems using quasi complete complementary codes," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, May 2008, pp. 4840–4844.
[18] S. F. Li, J. Chen, and L. Q. Zhang, "Optimisation of complete complementary codes in MIMO radar system," *Electron. Lett.*, vol. 46, no. 16, pp. 1157–1159, Aug. 2010.
[19] J. Tang, N. Zhang, Z. Ma, and B. Tang, "Construction of Doppler resilient complete complementary code in MIMO radar," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4704–4712, Sep. 2014.
[20] P. Fan, W. Yuan, and Y. Tu, "Z-complementary binary sequences," *IEEE Signal Process. Lett.*, vol. 14, no. 8, pp. 509–512, Aug. 2007.
[21] R. Y. Mesleh, H. Haas, Y. Lee, and S. Yun, "Interchannel interference avoidance in MIMO transmission by exploiting spatial information," in *Proc. IEEE 16th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Berlin, Germany, Sep. 2005, pp. 141–145.
[22] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.
[23] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, Dec. 2011.
[24] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.
[25] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 6–26, 1st Quart., 2015.
[26] P. Yang et al., "Single-carrier SM-MIMO: A promising design for broadband large-scale antenna systems," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1687–1716, 3rd Quart., 2016.
[27] M. Wen et al., "A survey on spatial modulation in emerging wireless systems: Research progresses and applications," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 1949–1972, Sep. 2019.
[28] Z. Liu, P. Yang, Y. L. Guan, and P. Xiao, "Cross Z-complementary pairs for optimal training in spatial modulation over frequency selective channels," *IEEE Trans. Signal Process.*, vol. 68, pp. 1529–1543, Feb. 2020.
[29] C. Fan, D. Zhang, and A. R. Adhikary, "New sets of binary cross Z-complementary sequence pairs," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1616–1620, Aug. 2020.
[30] A. R. Adhikary, Z. Zhou, Y. Yang, and P. Fan, "Constructions of cross Z-complementary pairs with new lengths," *IEEE Trans. Signal Process.*, vol. 68, pp. 4700–4712, Aug. 2020.
[31] Z.-M. Huang, C.-Y. Pai, and C.-Y. Chen, "Binary cross Z-complementary pairs with flexible lengths from Boolean functions," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1057–1061, Apr. 2021.
[32] M. Yang, S. Tian, N. Li, and A. R. Adhikary, "New sets of quadriphase cross Z-complementary pairs for preamble design in spatial modulation," *IEEE Signal Process. Lett.*, vol. 28, pp. 1240–1244, May 2021.
[33] F. Zeng, X. He, Z. Zhang, and L. Yan, "Quadriphase cross Z-complementary pairs for pilot sequence design in spatial modulation systems," *IEEE Signal Process. Lett.*, vol. 29, pp. 508–512, Jan. 2022.
[34] S. Das, A. Banerjee, and Z. Liu, "New family of cross Z-complementary sequences with large ZCZ width," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Espoo, Finland, Jun. 2022, pp. 522–527.
[35] H. Zhang, C. Fan, Y. Yang, and S. Mesnager, "New binary cross Z-complementary pairs with large CZC ratio," *IEEE Trans. Inf. Theory*, vol. 69, no. 2, pp. 1328–1336, Feb. 2023.
[36] Z.-M. Huang, C.-Y. Pai, and C.-Y. Chen, "Cross Z-complementary sets for training design in spatial modulation," *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5030–5045, Aug. 2022.
[37] Z.-M. Huang, C.-Y. Pai, and C.-Y. Chen, "A novel construction of optimal cross Z-complementary sets based on generalized Boolean functions," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Espoo, Finland, Jun. 2022, pp. 1725–1730.

[38] P. Kumar, S. Majhi, and S. Paul, "Direct construction of CZCS with flexible parameters for channel estimation in SM-MIMO system," *IEEE Commun. Lett.*, vol. 27, no. 12, pp. 3325–3329, Dec. 2023.

[39] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. Conf. Rec. 44th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2010, pp. 1498–1502.

[40] J. Fu, C. Hou, W. Xiang, L. Yan, and Y. Hou, "Generalised spatial modulation with multiple active transmit antennas," in *Proc. IEEE Globecom Workshops*, Miami, FL, USA, Dec. 2010, pp. 839–844.

[41] J. Wang, S. Jia, and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1605–1615, Apr. 2012.

[42] L. He, J. Wang, and J. Song, "Spatial modulation for more spatial multiplexing: RF-chain-limited generalized spatial modulation aided MM-wave MIMO with hybrid precoding," *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 986–998, Mar. 2018.

[43] S. Guo, H. Zhang, P. Zhang, S. Dang, C. Liang, and M.-S. Alouini, "Signal shaping for generalized spatial modulation and generalized quadrature spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4047–4059, Aug. 2019.

[44] S.-A. Yang and J. Wu, "Optimal binary training sequence design for multiple-antenna systems over dispersive fading channels," *IEEE Trans. Veh. Technol.*, vol. 51, no. 5, pp. 1271–1276, Sep. 2002.

[45] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple-antenna broadband transmissions," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 384–391, Mar. 2003.

[46] P. Fan and W. H. Mow, "On optimal training sequence design for multiple-antenna systems over dispersive fading channels and its extensions," *IEEE Trans. Veh. Technol.*, vol. 53, no. 5, pp. 1623–1626, Sep. 2004.

[47] Y. Zhou, Z. Zhou, Z. Liu, Y. Yang, P. Yang, and P. Fan, "Symmetrical Z-complementary code sets for optimal training in generalized spatial modulation," *Signal Process.*, vol. 208, Jul. 2023, Art. no. 108990.

[48] Z.-M. Huang, C. Hu, Z. Liu, and C.-Y. Chen, "On the mates of cross Z-complementary pairs for training sequence design in generalized spatial modulation," *IEEE Access*, vol. 11, pp. 145231–145237, Dec. 2023.

[49] X. H. Tang, P. Z. Fan, and S. Matsufuji, "Lower bounds on correlation of spreading sequence set with low or zero correlation zone," *Electron. Lett.*, vol. 36, no. 13, pp. 551–552, Mar. 2000.

[50] T. Yu, A. R. Adhikary, Y. Wang, and Y. Yang, "New class of optimal Z-complementary code sets," *IEEE Signal Process. Lett.*, vol. 29, pp. 1477–1481, Jun. 2022.

[51] S.-W. Wu, A. Şahin, Z.-M. Huang, and C.-Y. Chen, "Z-complementary code sets with flexible lengths from generalized boolean functions," *IEEE Access*, vol. 9, pp. 4642–4652, 2021.

[52] S.-W. Wu, C.-Y. Chen, and Z. Liu, "How to construct mutually orthogonal complementary sets with non-power-of-two lengths?" *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 3464–3472, Jun. 2021.

[53] B. Shen, Y. Yang, Y. Feng, and Z. Zhou, "A generalized construction of mutually orthogonal complementary sequence sets with non-power-of-two lengths," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4247–4253, Jul. 2021.

[54] C.-Y. Chen, C.-H. Wang, and C.-C. Chao, "Complete complementary codes and generalized Reed-Muller codes," *IEEE Commun. Lett.*, vol. 12, pp. 849–851, Nov. 2008.

ZHEN-MING HUANG (黃振銘) (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from the National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan, in 2019, and the M.S. degree in engineering science from National Cheng Kung University, Tainan, Taiwan, in 2021, where he is currently pursuing the Ph.D. degree in engineering science and communications engineering. From October 2023 to November 2023, he was a visiting Ph.D. student with the University of Essex, Colchester, U.K. (with Prof. Z. Liu). His research interests include sequence design and its applications in communications.

CHENG-YU PAI (白承祐) (Member, IEEE) received the B.S. and Ph.D. degrees in engineering science from National Cheng Kung University, Tainan, Taiwan, in 2018 and 2023, respectively, where he has been serving as a Postdoctoral Research Fellow with the Institute of Computer and Communication Engineering since January 2024. From July 2022 to June 2023, he was a visiting Ph.D. student with the University of Essex, Colchester, U.K. (with Prof. Z. Liu). His research interest includes sequence design and its applications in communications. He was a recipient of the 2nd Hon Hai Technology Award administered by Hon Hai Education Foundation, Taiwan, in 2022.

ZILONG LIU (劉子龍) (Senior Member, IEEE) received the bachelor's degree from the School of Electronics and Information Engineering, Huazhong University of Science and Technology, China, in 2004, the master's degree from the Department of Electronic Engineering, Tsinghua University, China, in 2007, the Ph.D. from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2014. He has been with the School of Computer Science and Electronic Engineering, University of Essex, since December 2019, first as a Lecturer and then a Senior Lecturer since October 2023. From January 2018 to November 2019, he was a Senior Research Fellow with the Institute for Communication Systems, Home of the 5G Innovation Centre, University of Surrey, during which he studied the air-interface design of 5G communication networks (e.g., machine-type communications, V2X communications, and 5G new radio). Prior to his career in U.K., he spent nine and half years in NTU, first as a Research Associate from July 2008 to October 2014 and then a Research Fellow from November 2014 to December 2017. His Ph.D. thesis "Perfect- and Quasi- Complementary Sequences", focusing on fundamental limits, algebraic constructions, and applications of complementary sequences in wireless communications, has settled a few long-standing open problems in the field. He was a Consultant to the Japanese Government on 6G Assisted Autonomous Driving in 2023. His research lies in the interplay of coding, signal processing, and communications, with a major objective of bridging theory and practice as much as possible. Recently, he has developed an interest in advanced 6G V2X communication technologies for future connected autonomous vehicles as well as machine learning for enhanced communications and networking.

Dr. Liu is an Associate Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE ACCESS, *Frontiers in Communications and Networks*, *Frontiers in Signal Processing*, and *Advances in Mathematics of Communications*. He was the Track Co-Chair on Networking and MAC in IEEE PIMRC'2023. He was the Hosting General Co-Chair of the 10th IEEE International Workshop on Signal Design and its Applications in Communications in 2022 and the TPC Co-Chair of the 2020 IEEE International Conference on Advanced Networks and Telecommunications Systems. He was a Tutorial Speaker of VTC-Fall'2021 and APCC'2021 on code-domain NOMA. Details of his research can be found at: https://sites.google.com/site/zilongliu2357.

CHAO-YU CHEN (陳昭羽) (Senior Member, IEEE) received the B.S. degree in electrical engineering and the M.S. and Ph.D. degrees in communications engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2000, 2002 and 2009, respectively, under the supervision of Prof. C.-c. Chao.

He was a visiting Ph.D. student with the University of California at Davis, Davis, CA, USA, from 2008 to 2009 (with Prof. S. Lin). From 2009 to 2016, he was a Technical Manager with the Communication System Design Division, Mediatek Inc., Hsinchu. From July 2018 to August 2018, he was with the University of California at Davis as a Visiting Scholar (with Prof. S. Lin). Since February 2016, he has been a Faculty Member with the National Cheng Kung University, Tainan, Taiwan. From February 2016 to July 2022, he was with the Department of Engineering Science. He is currently a Professor with the Department of Electrical Engineering and the Director of the Institute of Computer and Communication Engineering. His current research interests include sequence design, error-correcting codes, digital communications, and wireless networks.

Dr. Chen was a recipient of the 15th and 20th Y. Z. Hsu Science Paper Award administered by Far Eastern Y. Z. Hsu Science and Technology Memorial Foundation, Taiwan, in 2017 and 2022, and the Best Paper Award for Young Scholars by the IEEE Information Theory Society Taipei/Tainan Chapter and the IEEE Communications Society Taipei/Tainan Chapter in 2018. From January 2019 to December 2023, he served as the Vice Chair for the IEEE Information Theory Society Tainan Chapter. Since January 2024, he has been serving as the Vice Chair for the IEEE Communications Society Tainan Chapter.