

# Robust and Energy Efficient Sparse-Coded OFDM-DCSK System via Matrix Recovery

Zhaofeng Liu<sup>ID</sup>, Hing Cheung So<sup>ID</sup>, *Fellow, IEEE*, Xiao Peng Li<sup>ID</sup>, Lin Zhang<sup>ID</sup>, *Senior Member, IEEE*, and Zhi-Yong Wang<sup>ID</sup>

**Abstract**—In this paper, we devise a sparse-coded orthogonal frequency division multiplexing (OFDM) differential chaos shift keying (DCSK) communication system based on low-rank matrix recovery which can handle Gaussian background noise and outlier-contaminated symbols simultaneously. As the noise-free OFDM-DCSK symbol matrix has rank 1, we exploit the vector outer product for its modeling, while sparse coding is also applied to reduce the transmission energy. To demodulate information bits from the sparse-coded signal, we formulate an objective function which consists of a sum of Frobenius norm for rank-1 matrix recovery and  $\ell_0$ -norm for identifying the possibly outlier-contaminated symbols, with a self-adaptive weight parameter. The resultant optimization problem is solved iteratively via block coordinate descent, and the Laplacian kernel with the Silverman's rule is adopted for outlier detection. Theoretical analysis including convergence of the objective function, bit error rate (BER), energy efficiency and computational complexity, are provided. Simulation results show that the proposed system has comparable mean square error and BER performance with the  $\ell_p$ -norm minimization based matrix recovery approach at  $p = 2$  in additive white Gaussian noise, and is superior to that of  $p = 1$  in Middleton class A noise, even when sparse coding is applied. Moreover, compared with other binary DCSK systems, our system achieves higher energy efficiency thanks to the sparse coding.

**Index Terms**—Differential chaos shift keying, sparse coding, low-rank matrix recovery, bit error rate, impulsive noise, energy efficiency.

Manuscript received 26 October 2022; revised 2 March 2023 and 24 April 2023; accepted 14 May 2023. Date of publication 24 May 2023; date of current version 16 August 2023. This work was supported in part by the National Science Fund for Distinguished Young Scholars under Grant 61925108, in part by the Key Project of International Cooperation and Exchanges of the National Natural Science Foundation of China under Grant 62220106009, and in part by the grant from the Research Grants Council of the Hong Kong SAR, China [Project No. CityU 11207922]. The associate editor coordinating the review of this article and approving it for publication was L. Ong. (*Corresponding author: Xiao Peng Li*)

Zhaofeng Liu, Hing Cheung So, and Zhi-Yong Wang are with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong, SAR, China (e-mail: zhaofeliu3-c@my.cityu.edu.hk; hcs@ee.cityu.edu.hk; z.y.wang@my.cityu.edu.hk).

Xiao Peng Li is with the College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: x.p.li@szu.edu.cn).

Lin Zhang is with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China (e-mail: isszl@mail.sysu.edu.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCOMM.2023.3279403>.

Digital Object Identifier 10.1109/TCOMM.2023.3279403

## I. INTRODUCTION

A WIRELESS communication system is easily attacked by malicious users for its inherent openness to natural environments and broadcast characteristics. To protect the system from being attacked, chaotic communication is a powerful encryption technique [1], thanks to the high-security properties of chaotic sequences [2] such as sensitivity of initial value, good auto-correlation, resistance to interference, pseudo randomness and ergodicity. Chaotic communication techniques are widely used in data transmission, including ultra-wide-band (UWB) [1], [3], power line communication (PLC) [4], and vehicle-to-vehicle (V2V) communication [5]. The chaotic sequences are used to modulate information bits, and this kind of techniques is referred to as chaotic modulation, which can be classified as coherent and non-coherent types. Their difference is that the former needs to regenerate the chaotic sequence at the receiver while this procedure is not required in the latter. Therefore, non-coherent modulation has less complexity than the coherent one, and has been widely discussed [6], [7].

Among various non-coherent chaotic modulation schemes, differential chaos shift keying (DCSK) [6] has attracted considerable attention. In DCSK, a reference chaotic sequence is transmitted, and the subsequent information-bearing chaotic sequence is generated by multiplying binary phase shift keying (BPSK) symbols by the reference chaotic sequence. Thanks to the reference sequence, the receiver does not need to regenerate the chaotic waveform, which reduces the need of chaos synchronization and significantly lowers the complexity. Moreover, in many communication scenarios such as UWB, PLC and V2V, DCSK can provide satisfactory bit error rate (BER) for reliable transmission [1]. However, since DCSK needs to transmit 2 sequences in different time slots for one symbol, its transmission efficiency is low, while highly-complex delay line circuits are involved. In order to enhance the efficiency, the  $M$ -ary modulation using Hilbert transform [8], [9], Walsh code [10], and index modulation [11] have been suggested but they cannot solve the delay line issue. In fact, delay line can be avoided by arranging different sequences in different subcarriers rather than in different time slots, which is the main idea of multi-carrier (MC) transmission. Since one reference sequence can be used for demodulating multiple information-bearing sequences, the MC based DCSK systems are able to

attain significantly higher efficiency than the traditional DCSK systems [12]. Furthermore, in order to accelerate the computational efficiency of the MC-DCSK, the orthogonal frequency division multiplexing (OFDM) technique is applied [13] via fast Fourier transform (FFT). Based on the OFDM-DCSK, numerous studies have been performed including multi-user extension [14], application in underwater acoustic communications [15], peak-to-average power ratio reduction [16], and security enhancement [17], [18].

Since impulsive noise occurs in broadband communications, there is a need for its suppression including broadband OFDM [19] and PLC [20], [21] based systems. When OFDM-DCSK is applied in PLC, the impulsive disturbance together with Gaussian noise in PLC channels can significantly degrade the BER performance [22], [23], [24]. However, handling of outlier-contaminated data has not been thoroughly investigated in the current DCSK-based systems. In [25], the reference sequence is repeated and moving average is employed at the receiver to suppress the Gaussian noise, while [26] also exploits average computation at the receiver, but it applies the iterative chaos generator. In addition, to reduce the noise for  $M$ -ary DCSK systems, [27] uses a divide-and-conquer strategy. For MC-DCSK systems, [28] designs an iterative receiver to suppress Gaussian noise, and [29] develops a feedback module of correlation coefficients to enhance the noise-tolerance performance of  $M$ -ary modulation. To deal with impulsive noise, [30] uses replica piecewise frame, and [31] develops a joint position and constellation mapping scheme. Moreover, [32] deploys bit-interleaved coded modulation for resisting impulsive noise in  $M$ -ary DCSK systems. However, they only focus on the DCSK in the time domain, and do not consider enhancing the transmission efficiency by the MC structure. Recently, we have proposed a rank-1 matrix approximation based OFDM-DCSK receiver using  $\ell_p$ -norm [33] to resist outliers. Nevertheless, the computational complexity is relatively high [34], [35], [36] because  $\ell_p$ -norm minimization requires an iterative updating procedure.

In this work, utilizing the rank-1 property of the noise-free OFDM-DCSK symbol matrix, we propose using  $\ell_0$ -norm to detect outlier-contaminated data to achieve robustness. We also suggest sparse coding for transmission energy reduction. To restore information symbols from the sparse-coded (SC) signal, we formulate an optimization problem which consists of a Frobenius norm for rank-1 matrix recovery and an  $\ell_0$ -norm for handling outliers. Moreover, matrix factorization is applied to obtain the solution for multiple variables, and we utilize block coordinate descent (BCD) [37], [38] to decompose the resultant problem into simpler subproblems. In one subproblem, only one variable is optimized at each iteration, while the remaining variables are fixed. Subsequently, all subproblems are alternatingly solved in a computationally efficient manner. Compared with [33], the proposed system effectively operates in the scenarios when both Gaussian background noise and impulsive disturbance occur at the same time without the need of determining an appropriate value of  $p$ . Furthermore, our system has higher

energy efficiency than other competing binary DCSK systems thanks to the sparse coding.

Our novelty and technical contributions are summarized as:

- 1) An SC-OFDM-DCSK system via  $\ell_0$ -norm minimization based rank-1 matrix recovery is proposed to suppress outliers. In our matrix completion formulation,  $\ell_0$ -norm minimization is proposed to detect the outliers or impulsive noise components. After excluding the outlier-contaminated entries,  $\ell_2$ -norm based matrix factorization solver can offer enhanced reliability even in impulsive noise environments. Furthermore, by exploiting the matrix recovery capability for an incomplete low-rank symbol matrix, sparse coding is devised at the transmitter to generate zero-valued symbols in the time domain, indicating less transmission energy is involved.
- 2) We derive the BER in additive white Gaussian noise (AWGN) and Middleton class A noise, which is a common impulsive noise model, of the proposed system, and the BER performance over fading channels is also evaluated. Moreover, the convergence, energy efficiency and computational complexity are analyzed. In particular, the proposed system has higher energy efficiency than competing binary DCSK systems, and is computationally simpler than [33] at  $p < 2$ .
- 3) We demonstrate that our solution can achieve comparable mean square error and BER performance with [33] at  $p = 2$  in AWGN, which is the optimum detector, and outperforms [33] at  $p \leq 2$  in the presence of Middleton class A noise with and without sparse coding. Note that the hyperparameter in our scheme is automatically adjusted to fit different noise environments while only a fixed value of  $p$  is allowed in [33].

The rest of the paper is organized as follows. Section II presents the structures of the transmitter and receiver of the proposed SC-OFDM-DCSK system, where impulsive noise suppression and symbol recovery from sparse-coded symbols are achieved by  $\ell_0$ -norm minimization and rank-1 matrix completion. The convergence, BER, energy efficiency and computational complexity, are analyzed in Section III. Numerical examples for assessing and corroborating the system performance are included in Section IV. Finally, Section V concludes the paper.

We use bold upper case and lower case letters to represent matrices and vectors, respectively. The superscripts  $(\cdot)^T$  and  $(\cdot)^H$  denote transpose and Hermitian transpose, respectively, and  $\circ$  is the element-wise product. Besides,  $\Re\{\cdot\}$  and  $\Im\{\cdot\}$  stand for the real and imaginary parts of a complex-valued number or vector. Moreover,  $\mathbb{E}\{\cdot\}$  denotes the expectation operator, while  $\|\cdot\|_F$  is the Frobenius norm and  $\|\cdot\|_0$  is the  $\ell_0$ -norm which returns the number of non-zero elements. Furthermore,  $|\cdot|$  means the modulus of a complex number or the cardinality of a set. Finally,  $\mathbb{R}$  and  $\mathbb{C}$  denote the sets of real and complex numbers, respectively.

## II. $\ell_0$ -NORM RANK-1 MATRIX RECOVERY BASED SC-OFDM-DCSK SYSTEM

In this section, we present the structure of our robust rank-1 matrix recovery based SC-OFDM-DCSK system.

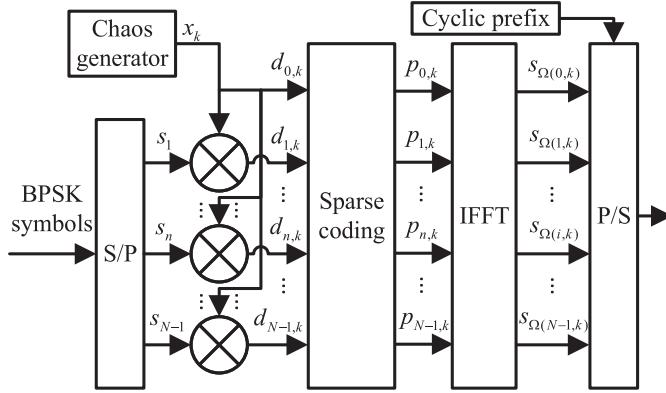


Fig. 1. Block diagram of SC-OFDM-DCSK transmitter.

### A. Transmitter Design

Figure 1 depicts the block diagram of the proposed SC-OFDM-DCSK transmitter. Compared with the conventional structure [13], we introduce a sparse coding module before inverse FFT (IFFT) for energy reduction. The chaos generator first produces chaotic sequences using the second-order Chebyshev polynomial function (CPF), which is

$$x_{k+1} = 1 - 2x_k^2, \quad x_k \in (-1, 0) \cup (0, 1) \quad (1)$$

where  $x_k$  represents the  $k$ th chip,  $0 \leq k \leq \beta - 1$ , and  $\beta$  is the sequence length. Then chaotic modulation is performed by multiplying the sequence with the BPSK data symbols which carry information bits after serial-to-parallel (S/P) conversion. The  $k$ th chip corresponding to the  $n$ th subcarrier is denoted as  $d_{n,k} = s_n x_k$ , where indices  $1 \leq n \leq N - 1$  refer to the information-bearing sequences, and  $n = 0$  corresponds to the reference sequence as  $d_{0,k} = x_k$ . They are then sent to sparse coding and IFFT modules to generate sparse chaotic OFDM signals. In Fig. 1, owing to the sparse coding, a portion of  $\{s_{\Omega_{i,k}}\}$  is 0 via transforming  $\{d_{n,k}\}$  to  $\{p_{n,k}\}$ . To introduce the sparse coding, we first express the IFFTs of  $\{d_{n,k}\}$  and  $\{p_{n,k}\}$  as:

$$\begin{aligned} s_{i,k} &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} d_{n,k} \exp\left(\frac{j2\pi ni}{N}\right) \\ &= \frac{x_k}{\sqrt{N}} \sum_{n=0}^{N-1} s_n \exp\left(\frac{j2\pi ni}{N}\right) = x_k \bar{s}_i \end{aligned} \quad (2)$$

$$s_{\Omega(i,k)} = s_{i,k} \Omega_{i,k} \quad (3)$$

where  $s_{i,k}$ ,  $s_{\Omega(i,k)}$  and  $\Omega_{i,k}$  are the  $(i + 1, k + 1)$  entries of  $\mathbf{S} \in \mathbb{C}^{N \times \beta}$ ,  $\mathbf{S}_{\Omega} \in \mathbb{C}^{N \times \beta}$ , and  $\Omega \in \mathbb{R}^{N \times \beta}$ , respectively,  $i = 0, \dots, N - 1$ ,  $k = 0, \dots, \beta - 1$ , and  $\bar{s}_i = \sum_{n=0}^{N-1} s_n \exp(j2\pi ni/N)/\sqrt{N}$ . According to (2), the rank of  $\mathbf{S}$  is 1, indicating that a subset of the matrix entries can fully characterize  $\mathbf{S}$ . As a result, we suggest only transmitting partial entries of  $\mathbf{S}$  in order to reduce the transmission energy, which is interpreted as sparse coding in the two-dimensional space. Given a pre-defined binary matrix  $\Omega$  which contains randomly-distributed values of either 0 or 1, the incomplete matrix  $\mathbf{S}_{\Omega}$  is formed by  $\mathbf{S}_{\Omega} = \mathbf{S} \circ \Omega$ , where its element-wise form is given in (3). According to (3), the transmitted  $s_{\Omega(i,k)}$

has zero value when  $\Omega_{i,k} = 0$ , hence sparse coding is achieved in the time domain, leading to energy consumption reduction. Rewriting (2) and (3), sparse coding corresponds to:

$$\begin{aligned} p_{n,k} &= \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} s_{\Omega(i,k)} \exp\left(-\frac{j2\pi ni}{N}\right) \\ &= \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} s_{i,k} \Omega_{i,k} \exp\left(-\frac{j2\pi ni}{N}\right) \\ &= \frac{1}{N} \sum_{i=0}^{N-1} \sum_{n'=0}^{N-1} d_{n',k} \exp\left(\frac{j2\pi n'i}{N}\right) \Omega_{i,k} \exp\left(-\frac{j2\pi ni}{N}\right) \\ &= \frac{1}{N} \sum_{n'=0}^{N-1} d_{n',k} \sum_{i=0}^{N-1} \exp\left(\frac{j2\pi(n' - n)i}{N}\right) \Omega_{i,k} \end{aligned} \quad (4)$$

where the first line is obtained by performing FFT on  $\{s_{\Omega(i,k)}\}$  and the relation between input  $d_{n',k}$  and output  $p_{n,k}$  is revealed. According to [39], when the number of observed entries is sufficiently large,  $\mathbf{S}$  can be uniquely recovered from  $\mathbf{S}_{\Omega}$  with high probability. Based on our empirical study, it is found that comparable performance is achieved when the number of zeros in  $\Omega$  is no more than  $0.2 N\beta$ . That is to say, we can remove up to 20% of the elements in  $\mathbf{S}$ .

It is worth mentioning that although our system and [40] have similar names, their transceiver structures and main focuses are different. In [40], sparse coding refers to the multiplication of sparse binary codes from a codebook with information symbols in the frequency domain. In doing so, multi-user access is achieved by allowing information symbols from different users to superimpose in the frequency domain. That is, the symbols are sparse prior to superimposing in the frequency domain, but they are not sparse in the time domain after the inverse Fourier transform. On the other hand, we transmit the sparse matrix  $\mathbf{S}_{\Omega}$ . Note that we cannot simply perform  $\mathbf{S}_{\Omega} = \mathbf{S} \circ \Omega$  because  $\mathbf{S}$  contains OFDM symbols in the time domain, and direct multiplication with  $\Omega$  will lead to inter-symbol interference (ISI). Consequently, our sparse coding is conducted in the frequency domain according to (4). It is worth noting that our proposed system can be also used in multi-user transmission with modifications like arranging the reference sequences of multiple users properly [14] or using the Walsh code [29]. Specifically, for  $P$  users, the rank of the transmit noise-free matrix is  $P$ , and hence rank- $P$  matrix completion is required at the receiver. The symbols from each user can then be recovered due to the orthogonality of chaos sequences or Walsh code.

After sparse coding, IFFT is performed on  $\{p_{n,k}\}$  to generate  $\{s_{\Omega(i,k)}\}$ . At the end of the transmitter, the parallel data streams are converted to serial format by the parallel-to-serial (P/S) converter, and the cyclic prefix (CP) is added to avoid the ISI and inter-carrier interference (ICI) due to frequency selective fading.

### B. Receiver Design

The structure of the proposed receiver with robust rank-1 matrix recovery is shown in Fig. 2. The received data first go through CP removal and serial-to-parallel (S/P) conversion.

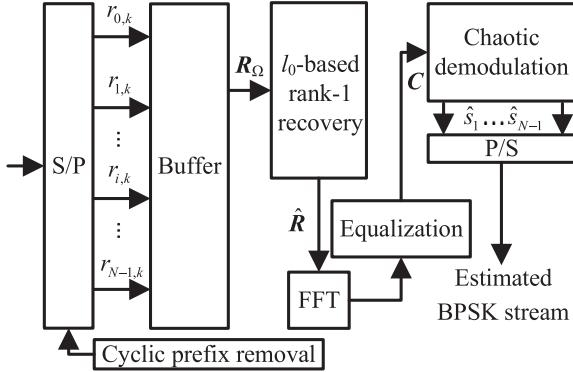


Fig. 2. Block diagram of SC-OFDM-DCSK receiver via rank-1 modeling and  $\ell_0$ -norm minimization.

The incomplete matrix  $\mathbf{R}_\Omega \in \mathbb{C}^{N \times \beta}$  bearing the information of the  $\beta$  OFDM symbols is then obtained at the buffer output, which is expressed as:

$$\mathbf{R}_\Omega = (\mathbf{H}\mathbf{S} + \mathbf{N}) \circ \Omega = (\mathbf{H}\mathbf{S})_\Omega + \mathbf{N}_\Omega = \bar{\mathbf{S}}_\Omega + \mathbf{N}_\Omega \quad (5)$$

where  $\mathbf{H} \in \mathbb{C}^{N \times N}$  is a Toeplitz matrix constituted by the multi-path impulse response in the time domain,  $\bar{\mathbf{S}} = \mathbf{H}\mathbf{S}$ , and  $\mathbf{N} \in \mathbb{C}^{N \times \beta}$  contains additive disturbances. The channel frequency response is derived as  $H_n = \sum_{i=0}^{N-1} h_i \exp(-j2\pi ni/N)/\sqrt{N}$ , where the multi-path impulse response  $h_i$  is equal to  $(a, (a-i) \bmod N)$  entry of  $\mathbf{H}$ ,  $a = 1, \dots, N$  and  $\bmod$  denotes the remainder. As the rank of  $\bar{\mathbf{S}}$  is equal to 1, its recovery, together with suppressing the noise component  $\mathbf{N}$ , can be achieved by means of low-rank matrix recovery [41]. In this research, we consider the general setting that  $\mathbf{N}_\Omega$  contains gross errors or outliers. That is,  $\mathbf{N}_\Omega$  can be written as

$$\mathbf{N}_\Omega = \mathbf{G}_\Omega + \mathbf{O}_\Omega \quad (6)$$

where  $\mathbf{G}_\Omega$  and  $\mathbf{O}_\Omega$  represent the Gaussian background noise and outlier components, respectively. Incorporating the rank-1 matrix constraint with the use of the outer product representation, we propose the following formulation to find  $\bar{\mathbf{S}}$ :

$$\min_{\mathbf{u}, \mathbf{v}, \mathbf{O}_\Omega} L(\mathbf{u}, \mathbf{v}, \mathbf{O}_\Omega) := \|\mathbf{R}_\Omega - (\mathbf{u}\mathbf{v}^H)_\Omega - \mathbf{O}_\Omega\|_F^2 + \mu \|\mathbf{O}_\Omega\|_0 \quad (7)$$

where  $\mathbf{u} \in \mathbb{C}^N$ ,  $\mathbf{v} \in \mathbb{C}^\beta$ , and the  $\ell_0$ -norm is defined as the number of non-zero elements of a matrix. Here, the Frobenius norm is adopted to suppress the Gaussian background noise  $\mathbf{G}_\Omega$  while the  $\ell_0$ -norm aims to detect the outlier components in  $\mathbf{O}_\Omega$ , which is modeled as a sparse matrix. The parameter  $\mu > 0$  controls the sparsity level of  $\mathbf{O}_\Omega$ , and ideally,  $\mathbf{O}_\Omega$  will store all gross errors. An adaptive rule for automatic determination of  $\mu$  will be provided shortly.

Before proceeding, we highlight the advantages of (7) over the  $\ell_p$ -norm minimization framework [35]:

$$\min_{\mathbf{u}, \mathbf{v}} \|(\mathbf{u}\mathbf{v}^H)_\Omega - \mathbf{R}_\Omega\|_p^p, \quad \|\mathbf{R}\|_p = \left( \sum_{i=1}^N \sum_{k=1}^\beta |r_{i,k}|^p \right)^{1/p}. \quad (8)$$

It is well known that the optimum value of  $p$  is 2 in the presence of AWGN, while  $0 < p < 2$  is required to operate in impulsive noise environments. Moreover, (8) applies  $\ell_p$ -norm minimization to all elements, which means that it treats all matrix entries as Gaussian noise contaminated when  $p = 2$  or outliers when  $p < 2$ . Hence (8) is unable to attain optimality when both small dense disturbances and gross errors appear. On the other hand, (7) utilizes two terms in the objective function to handle Gaussian background noise and outliers in a separate and simultaneous manner. Furthermore, the computational complexity of performing  $\ell_p$ -norm minimization can be high for  $p \neq 2$  as iterations are required such as using the iteratively reweighted least squares (IRLS), but there exists fast algorithms for tackling sparsity formulated in  $\ell_0$ -norm. These imply that (7) is superior to (8).

To handle (7), we adopt BCD to decompose it into three subproblems with alternate and iterative updates:

$$\mathbf{u}^{l+1} = \arg \min_{\mathbf{u}} \|\mathbf{R}_\Omega - (\mathbf{u}(\mathbf{v}^l)^H)_\Omega - \mathbf{O}_\Omega^l\|_F^2 \quad (9)$$

$$\mathbf{v}^{l+1} = \arg \min_{\mathbf{v}} \|\mathbf{R}_\Omega - (\mathbf{u}^{l+1}\mathbf{v}^H)_\Omega - \mathbf{O}_\Omega^l\|_F^2 \quad (10)$$

$$\begin{aligned} \mathbf{O}_\Omega^{l+1} = \arg \min_{\mathbf{O}_\Omega} & (\|\mathbf{R}_\Omega - (\mathbf{u}^{l+1}(\mathbf{v}^{l+1})^H)_\Omega - \mathbf{O}_\Omega\|_F^2 \\ & + \mu^{(l+1)} \|\mathbf{O}_\Omega\|_0) \end{aligned} \quad (11)$$

where  $l$  denotes the iteration number. Now  $\mu$  is also considered as an unknown parameter to be estimated adaptively, which is replaced by  $\mu^{(l+1)}$ . From (9)–(11), we see that BCD alternately finds optimum solutions among  $\mathbf{u}^{l+1}$ ,  $\mathbf{v}^{l+1}$  and  $\mathbf{O}_\Omega^{l+1}$  while the remaining variables are kept fixed. We first present the solutions for (9)–(11) as follows. It is easily observed that (9) can be decoupled in a row-by-row manner, facilitating independent estimation of each of its elements. Denote  $\mathbf{Y}_\Omega = \mathbf{R}_\Omega - \mathbf{O}_\Omega^l$  and let  $\mathbf{y}_{\Omega_i}^H$  be the vector storing the entries in the  $i$ th row of  $\mathbf{Y}_\Omega$  with indices in  $\Omega_i$  where  $\Omega_i$  is the index set of 1 in the  $i$ th row of  $\Omega$ . Similarly, we define  $\mathbf{v}_{\Omega_i}^l$  as the vector consisting of elements in  $\mathbf{v}^l$  with indices in  $\Omega_i$ . Then (9) is equivalent to:

$$u_i^{l+1} = \arg \min_{u_i} \|\mathbf{y}_{\Omega_i}^H - u_i(\mathbf{v}_{\Omega_i}^l)^H\|_2^2, \quad i = 1, \dots, N \quad (12)$$

which is a simple scalar linear least squares (LLS) problem. Analogously, considering column by column, each element of  $\mathbf{v}^{l+1}$  in (10) can be separately estimated from:

$$v_j^{l+1} = \arg \min_{v_j} \|\mathbf{y}_{\Omega_j} - \mathbf{u}_{\Omega_j}^{l+1} v_j\|_2^2, \quad j = 1, \dots, \beta \quad (13)$$

where  $\Omega_j$  is the index set of 1 in the  $j$ th column of  $\Omega$  while  $\mathbf{u}_{\Omega_j}^{l+1}$  and  $\mathbf{y}_{\Omega_j}$  contain elements of  $\mathbf{u}^{l+1}$  and  $j$ th column of  $\mathbf{Y}_\Omega$  with indices in  $\Omega_j$ , respectively. The LLS solutions to (12) and (13) are given by:

$$u_i^{l+1} = \frac{\mathbf{y}_{\Omega_i}^H \mathbf{v}_{\Omega_i}^l}{(\mathbf{v}_{\Omega_i}^l)^H \mathbf{v}_{\Omega_i}^l}, \quad (14)$$

$$v_j^{l+1} = \frac{(\mathbf{u}_{\Omega_j}^{l+1})^H \mathbf{y}_{\Omega_j}}{(\mathbf{u}_{\Omega_j}^{l+1})^H \mathbf{u}_{\Omega_j}^{l+1}}. \quad (15)$$

Defining  $\mathbf{E}_\Omega^{l+1} = \mathbf{R}_\Omega - (\mathbf{u}^{l+1}(\mathbf{v}^{l+1})^H)_\Omega$ , (11) becomes:

$$\mathbf{O}_\Omega^{l+1} = \arg \min_{\mathbf{O}_\Omega} \|\mathbf{E}_\Omega^{l+1} - \mathbf{O}_\Omega\|_F^2 + \mu^{(l+1)} \|\mathbf{O}_\Omega\|_0. \quad (16)$$

Vectorizing (16) yields:

$$\mathbf{o}^{l+1} = \arg \min_{\mathbf{o}} \|\mathbf{e}^{l+1} - \mathbf{o}\|_2^2 + \mu^{(l+1)} \|\mathbf{o}\|_0 \quad (17)$$

where  $\mathbf{e}^{l+1}$  and  $\mathbf{o}$  are vectors containing elements of  $\mathbf{E}_\Omega^{l+1}$  and  $\mathbf{O}_\Omega$  with indices in  $\Omega$  in the column-first manner, respectively. To facilitate real-valued processing, we decompose (17) into real and imaginary parts, resulting in:

$$\bar{\mathbf{o}}^{l+1} = \arg \min_{\bar{\mathbf{o}}} \|\bar{\mathbf{e}}^{l+1} - \bar{\mathbf{o}}\|_2^2 + \mu^{(l+1)} \|\bar{\mathbf{o}}\|_0 \quad (18)$$

where

$$\bar{\mathbf{e}}^{l+1} = \begin{bmatrix} \Re\{\mathbf{e}^{l+1}\} \\ \Im\{\mathbf{e}^{l+1}\} \end{bmatrix}, \quad \bar{\mathbf{o}} = \begin{bmatrix} \Re\{\mathbf{o}\} \\ \Im\{\mathbf{o}\} \end{bmatrix}. \quad (19)$$

Prior to solving (18), the value of  $\mu^{(l+1)}$ , which determines the sparsity of  $\mathbf{O}_\Omega^{l+1}$  or the number of outliers at each iteration, is required. One approach for outlier detection in  $\bar{\mathbf{e}}^{l+1}$  is to use the Laplacian kernel [42]:

$$k_{\sigma_s}(\bar{e}_i^{l+1}) = \frac{1}{2\sigma_s} \exp\left(-\frac{|\bar{e}_i^{l+1}|}{\sigma_s}\right) \quad (20)$$

where  $\bar{e}_i^{l+1}$  is the  $i$ th element of  $\bar{\mathbf{e}}^{l+1}$  and  $\sigma_s$  is the bandwidth. The principle is that the kernel operator (20) will return a small value when  $\bar{e}_i^{l+1}$  has a large magnitude. In kernel density estimation, the bandwidth can be calculated by the Silverman's rule as [43]:

$$\sigma_s = 1.06 \times \min\left\{\sigma_e, \frac{\text{IQR}_e}{1.34}\right\} \times \dim(\bar{\mathbf{e}}^{l+1})^{-0.2} \quad (21)$$

where  $\sigma_e$  and  $\text{IQR}_e$  are the standard derivation and interquartile range of  $\bar{\mathbf{e}}^{l+1}$ , respectively, and  $\dim(\cdot)$  denotes the vector length. When  $k_{\sigma_s}(\bar{e}_i^{l+1}) \leq \epsilon$  where  $\epsilon$  is a user-defined threshold, then  $\bar{e}_i^{l+1}$  is assigned as an outlier. Typically,  $\epsilon$  is set around  $10^{-20}$ . Defining  $\mathcal{I}_{l+1}$  as the set of indices  $i$  corresponding to outliers  $\{\bar{e}_i^{l+1}\}$ , the non-negative parameter  $\mu^{(l+1)}$  can be calculated as

$$\mu^{(l+1)} = \min \left\{ (\bar{e}_1^{l+1})^2, \dots, (\bar{e}_i^{l+1})^2, \dots, (\bar{e}_{|\mathcal{I}_{l+1}|}^{l+1})^2, \mu^{(l)} \right\} \quad (22)$$

where  $|\mathcal{I}_{l+1}|$  is the cardinality of  $\mathcal{I}_{l+1}$ .

With  $\mu^{(l+1)}$ , according to [44], (17) has a closed-form and global solution, which is:

$$\bar{o}_i^{l+1} = H_{\mu^{(l+1)}}(\bar{e}_i^{l+1}) = \begin{cases} \bar{e}_i^{l+1}, & \text{if } |\bar{e}_i^{l+1}| \geq \sqrt{\mu^{(l+1)}} \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

where  $\bar{o}_i^{l+1}$  is the  $i$ th element of  $\bar{\mathbf{o}}^{l+1}$  and  $H_{\mu^{(l+1)}}(\cdot)$  is the hard-thresholding operator which sets the values below  $\mu^{(l+1)}$  to 0. After all element-wise operations, the obtained  $\bar{\mathbf{o}}^{l+1}$  is reconstructed back to the complex-valued vector  $\mathbf{o}^{l+1}$ , and subsequently  $\mathbf{O}_\Omega^{l+1}$  according to  $\Omega$ . The overall procedure is summarized in Algorithm 1.

---

**Algorithm 1** Rank-1 Matrix Recovery With  $\ell_0$ -Norm Minimization

---

**Input:**  $\mathbf{R}_\Omega, \Omega, \epsilon = 10^{-20}$

**Initialize:** Randomly initialize  $\mathbf{u}^0$  and  $\mathbf{v}^0$ ,  $\mu^{(0)} = \infty$ .

**for**  $l = 0, 1, \dots$  **do**

**for**  $n = 0, 1, \dots, N - 1$  **do**  
    Update  $u_i^{l+1}$  based on (14).

**end for**  
    **for**  $k = 0, 1, \dots, \beta - 1$  **do**  
        Update  $v_j^{l+1}$  based on (15).

**end for**  
    Compute  $\mathbf{E}_\Omega^{l+1} = \mathbf{R}_\Omega - (\mathbf{u}^{l+1}(\mathbf{v}^{l+1})^H)_\Omega$ .

Transform  $\mathbf{E}_\Omega^{l+1}$  to  $\bar{\mathbf{e}}^{l+1}$ .

Compute  $\sigma_s$  based on (21).

Compute  $k_{\sigma_s}(\bar{e}_i^{l+1})$  based on (20).

Construct  $\mathcal{I}_{l+1}$  which contains all indices  $i$  satisfying  $k_{\sigma_s}(\bar{e}_i^{l+1}) \leq \epsilon$ .

Compute  $\mu^{(l+1)}$  based on (22).

Compute  $\bar{\mathbf{o}}^{l+1} = H_{\mu^{(l+1)}}(\bar{\mathbf{e}}^{l+1})$  based on (23).

Construct  $\mathbf{O}_\Omega^{l+1}$  with  $\bar{\mathbf{o}}^{l+1}$  and  $\Omega$ .

**Stop** when termination condition is met.

**end for**

**Output:**  $\hat{\mathbf{R}} = \mathbf{u}^{l+1}(\mathbf{v}^{l+1})^H$

---

After rank-1 matrix recovery, FFT is performed on  $\hat{\mathbf{R}}$  for OFDM demodulation, followed by the frequency domain equalizer (FDE) [45] to yield the rank-1 matrix  $\mathbf{C}$  [33]:

$$\mathbf{C} = [\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_n, \dots, \mathbf{c}_{N-1}]^H \quad (24)$$

That is, every row vector of  $\mathbf{C}$  represents a chaotic sequence, with the first one,  $\mathbf{c}_0$ , being the reference sequence while the rest are information-bearing sequences. Utilizing the real part of  $\mathbf{C}$ , the detected BPSK symbol  $\hat{s}_n$  is determined via chaotic demodulation as:

$$\hat{s}_n = \text{sgn}(\Re\{\mathbf{c}_0^T\} \cdot \Re\{\mathbf{c}_n\}), \quad 1 \leq n \leq N - 1 \quad (25)$$

where  $\text{sgn}(\cdot)$  is the sign function. Finally, the information bits are straightforwardly decoded from the BPSK symbols.

### III. THEORETICAL ANALYSIS

#### A. Convergence

We add superscripts  $\mu^{(l)}$  and  $\mu^{(l+1)}$  in the objective function of (7) to denote its values at the  $l$ th and  $(l+1)$ th iterations, respectively. The objective function difference in two adjacent iterations can then be written as:

$$\begin{aligned} & L_{\mu^{(l+1)}}(\mathbf{u}^{l+1}, \mathbf{v}^{l+1}, \mathbf{O}_\Omega^{l+1}) - L_{\mu^{(l)}}(\mathbf{u}^l, \mathbf{v}^l, \mathbf{O}_\Omega^l) \\ &= \underbrace{L_{\mu^{(l)}}(\mathbf{u}^{l+1}, \mathbf{v}^l, \mathbf{O}_\Omega^l) - L_{\mu^{(l)}}(\mathbf{u}^l, \mathbf{v}^l, \mathbf{O}_\Omega^l)}_{L_1} \\ &\quad + \underbrace{L_{\mu^{(l)}}(\mathbf{u}^{l+1}, \mathbf{v}^{l+1}, \mathbf{O}_\Omega^l) - L_{\mu^{(l)}}(\mathbf{u}^{l+1}, \mathbf{v}^l, \mathbf{O}_\Omega^l)}_{L_2} \\ &\quad + \underbrace{L_{\mu^{(l+1)}}(\mathbf{u}^{l+1}, \mathbf{v}^{l+1}, \mathbf{O}_\Omega^l) - L_{\mu^{(l)}}(\mathbf{u}^{l+1}, \mathbf{v}^{l+1}, \mathbf{O}_\Omega^l)}_{L_3} \end{aligned}$$

$$+ \underbrace{L_{\mu^{(l+1)}}(\mathbf{u}^{l+1}, \mathbf{v}^{l+1}, \mathbf{O}_{\Omega}^{l+1}) - L_{\mu^{(l+1)}}(\mathbf{u}^{l+1}, \mathbf{v}^{l+1}, \mathbf{O}_{\Omega}^l)}_{L_4} \quad (26)$$

where  $L_1$ ,  $L_2$ ,  $L_3$  and  $L_4$  refer to the objective function differences after computing (9), (10), (22) and (11), respectively. Since  $\mathbf{u}^{l+1}$ ,  $\mathbf{v}^{l+1}$ , and  $\mathbf{O}_{\Omega}^{l+1}$  are the optimal solutions of (9), (10), and (11), respectively, as well as (22) indicates  $\mu^{(l+1)} \leq \mu^{(l)}$ , all  $L_1$ ,  $L_2$ ,  $L_3$  and  $L_4$  are not greater than 0, implying that  $L(\mathbf{u}, \mathbf{v}, \mathbf{O}_{\Omega})$  is non-increasing. Moreover, the non-convex function  $L(\mathbf{u}, \mathbf{v}, \mathbf{O}_{\Omega})$  is upper bounded by  $L(\mathbf{u}^0, \mathbf{v}^0, \mathbf{O}_{\Omega}^0)$  and lower bounded by 0, thus Algorithm 1 makes the objective function value converge.

### B. Bit Error Rate

We first derive the BER in AWGN over frequency selective fading channel. Subsequently, as a special case, the BERs over flat fading and no fading channel are derived. Moreover, we provide the BER in Middleton class A noise. In the BER analysis, we assume that perfect channel state information is obtained and used in FDE. We also assume that the SC-OFDM-DCSK symbol duration is smaller than the channel coherent time, which means that the slow fading remains constant in SC-OFDM-DCSK symbol duration, and the Doppler spread and the Doppler-induced ICI are ignored.

*1) BER in AWGN Over Frequency Selective Fading Channel:* According to Fig. 2, BPSK demodulation is processed after equalization in the frequency domain. We write the matrix used for BPSK demodulation as  $\Re\{\mathbf{C}\} \approx \tilde{\mathbf{u}}\tilde{\mathbf{v}}^T$  where  $\tilde{\mathbf{u}} = [\tilde{u}_0, \tilde{u}_1, \dots, \tilde{u}_n, \dots, \tilde{u}_{N-1}]^T \in \mathbb{R}^N$  and  $\tilde{\mathbf{v}} \in \mathbb{R}^\beta$ . Because of BPSK modulation, only the signs of the demodulated symbols affect the demodulation results. Based on [33], we have  $\hat{s}_n = \text{sgn}(\tilde{u}_0) \cdot \text{sgn}(\tilde{u}_n)$  and the error probability of each element in  $\tilde{\mathbf{u}}$  can be expressed as:

$$\begin{aligned} P_{b_n} &= Q\left(\frac{d_{\mathbf{x}_n}\tilde{\mathbf{v}}}{\sigma_g}\right) = Q\left(|H_n| \cos \theta \sqrt{\frac{\beta/2}{N_0/2}}\right) \\ &= Q\left(|H_n| \cos \theta \sqrt{\frac{\beta}{N_0}}\right). \end{aligned} \quad (27)$$

where  $N_0 = 2\sigma_g^2$  with  $\sigma_g^2$  being the AWGN variance,  $Q(x) = \int_x^\infty \exp(-t^2/2) dt$  is the Q-function,  $\theta$  is the numerically-determined angle between the transmitted chaotic sequence vector  $\mathbf{x}_n$  and  $\tilde{\mathbf{v}}$ ,  $\mathbf{x}_n = |H_n|\mathbf{x}$  and the length of the projection of  $\mathbf{x}_n$  onto  $\tilde{\mathbf{v}}$  is:

$$\begin{aligned} d_{\mathbf{x}_n}\tilde{\mathbf{v}} &= |H_n|d_{\mathbf{x}} \cos \theta = |H_n| \sqrt{\beta \mathbb{E}\{x_k^2\}} \cos \theta \\ &= |H_n| \sqrt{\frac{\beta}{2}} \left| \frac{\tilde{\mathbf{v}}^T \mathbf{x}}{\|\tilde{\mathbf{v}}\|_2 \|\mathbf{x}\|_2} \right| \end{aligned} \quad (28)$$

where  $\cos \theta = \tilde{\mathbf{v}}^T \mathbf{x} / (\|\tilde{\mathbf{v}}\|_2 \|\mathbf{x}\|_2)$ . Subsequently, since chaos demodulation is non-coherent, the BER of one BPSK symbol is calculated based on the product of  $\tilde{u}_0$  and  $\tilde{u}_n$ . Noting that even when both values have incorrect polarities, the resultant BPSK symbol is still correctly demodulated, while error appears only when one of  $\tilde{u}_0$  and  $\tilde{u}_n$  has wrong sign.

Therefore, the error rate of BPSK symbol under certain  $\theta$  and  $\{H_0, \dots, H_{N-1}\}$  can be expressed as

$$P_e = \frac{1}{N-1} \sum_{n=1}^{N-1} [P_{b_0}(1-P_{b_n}) + (1-P_{b_0})P_{b_n}]. \quad (29)$$

Finally, the BER in AWGN over frequency selective fading channel is numerically calculated by taking the expected value of  $P_e$  under the admissible range of  $\theta$  [46], [47] and  $\{H_0, \dots, H_{N-1}\}$ , which can be expressed as:

$$\begin{aligned} \text{BER}_{\text{FSF}} &= \mathbb{E}\{P_e|\theta; H_0, \dots, H_{N-1}\} \\ &= \mathbb{E}\left\{ \sum_{n=1}^{N-1} Q\left(|H_0| \cos \theta \sqrt{\frac{\beta}{N_0}}\right) \left(1 - Q\left(|H_n| \cos \theta \sqrt{\frac{\beta}{N_0}}\right)\right) \right. \\ &\quad \left. + \left(1 - Q\left(|H_0| \cos \theta \sqrt{\frac{\beta}{N_0}}\right)\right) Q\left(|H_n| \cos \theta \sqrt{\frac{\beta}{N_0}}\right) \right. \\ &\quad \left. \left| \theta; H_0, \dots, H_{N-1} \right. \right\}. \end{aligned} \quad (30)$$

When the frequency selective fading is originated from the complex Gaussian distributed multi-path fading with unit power spectral density [17],  $|H_n|$  is independent and identically Rayleigh distributed with the scale parameter  $\sigma_r = \sqrt{2}/2$ . Therefore, assuming that the number of received samples is large, the BER over multi-path fading channel is calculated by a multi-variable integral of  $|H_0|, \dots, |H_{N-1}|$  as [11]:

$$\begin{aligned} \text{BER}_{\text{MPF}} &= \int_0^\infty \dots \int_0^\infty \mathbb{E}\{P_e|\theta\} \prod_{n=0}^{N-1} f(|H_n|) d|H_0|, \dots, d|H_{N-1}| \end{aligned} \quad (31)$$

where  $f(|H_n|)$  is the probability density function (PDF) of  $|H_n|$  over multi-path fading channel, denoted as  $f_{\text{MPF}}(|H_n|) = 2|H_n| \exp(-|H_n|^2)$ . To compute the BER values, the Monte Carlo simulation with  $T$  independent trials and  $N$  subcarriers is applied. In the simulation,  $T \times (N-1) > 10^5$  bits are generated, indicating that the precision of the empirical BER can achieve  $10^{-5}$ .

*2) BER in AWGN Over Flat Fading Channel:* Since in flat fading channel, the frequency responses over multiple subcarriers remain constant, we have  $|H_0| = |H_1| = \dots = |H_{N-1}| = |H|$ . Therefore (27) is rewritten as:

$$P_b = Q\left(|H| \cos \theta \sqrt{\frac{\beta}{N_0}}\right). \quad (32)$$

While (29) is:

$$\bar{P}_e = 2P_b(1-P_b). \quad (33)$$

Subsequently, similar to (31), the BER in AWGN over flat fading channel is derived as:

$$\begin{aligned} \text{BER}_{\text{flat}} &= \int_0^\infty \mathbb{E}\{\bar{P}_e|\theta\} f(|H|) d|H| \\ &= \int_0^\infty \mathbb{E}\left\{2Q\left(|H| \cos \theta \sqrt{\frac{\beta}{N_0}}\right) \left(1 - Q\left(|H| \cos \theta \sqrt{\frac{\beta}{N_0}}\right)\right)\right\} d|H|. \end{aligned}$$

$$2|H| \exp(-|H|^2) d|H| \quad (34)$$

where the PDF of  $|H|$  can be also modeled by a Rayleigh distribution with scale parameter  $\sigma_r = \sqrt{2}/2$ , denoted as  $f(|H|) = 2|H| \exp(-|H|^2)$ .

3) *BER in AWGN Without Fading*: When there is no fading, we have  $H = 1$ . Therefore, according to (34), the BER in AWGN is:

$$\begin{aligned} \text{BER}_{\text{AWGN}} &= \mathbb{E}\{\bar{P}_e|\theta\} \\ &= \mathbb{E}\left\{2Q\left(\cos\theta\sqrt{\frac{\beta}{N_0}}\right)\left(1-Q\left(\cos\theta\sqrt{\frac{\beta}{N_0}}\right)\right)|\theta\right\}. \end{aligned} \quad (35)$$

4) *BER Lower Bound in Middleton Class A Noise*: We use the Middleton class A model to depict the impulsive noise. Middleton class A noise is originated from the widespread electromagnetic interference in telecommunication systems [48]. The PDF of the Middleton class A noise is expressed as:

$$p(\xi) = \exp(-A) \sum_{m=0}^{\infty} \frac{A^m}{m! \sqrt{2\pi\sigma_m^2}} e^{-\frac{\xi^2}{2\sigma_m^2}} \quad (36)$$

where  $A$  is the impulsive index and  $\Gamma$  denotes the ratio of Gaussian noise power to impulsive noise power [22], while the variance of the  $m$ th Gaussian component is

$$\sigma_m^2 = \sigma_g^2 \frac{\frac{m}{A} + \Gamma}{1 + \Gamma}. \quad (37)$$

We set  $A$  and  $\Gamma$  as  $A = \Gamma = 0.01$ . To generate the complex-valued version, the zero-mean real and imaginary parts are independently realized using (36) with the same power.

In (36), when  $m$  becomes larger, the variance of the  $m$ th Gaussian component becomes larger, and the probability becomes smaller. Outlier resistance aims at suppressing the noise from Gaussian components with large  $m$ . Since the variance of the  $m$ th Gaussian component obeys (37), we obtain  $\sigma_m^2 \gg \sigma_0^2$  when  $m > 0$ ,  $A = 0.01$  and  $\Gamma = 0.01$ , which means that the noise from Gaussian components with  $m \geq 1$  is considered as outlier. Therefore, for the best outlier resistance performance, the Gaussian components with  $m > 0$  are suppressed, and only the Gaussian component with  $m = 0$  is left. For the BER calculation with the best outlier resistance, we use  $\sigma_0$  to replace  $\sigma_g$ . According to (37), we know that  $\sigma_0^2 = \sigma_g^2 \Gamma / (1 + \Gamma) = \Gamma \sqrt{N_0/2} / (1 + \Gamma)$ , and hence the BER lower bound in Middleton class A noise can be derived without fading according to (35):

$$\begin{aligned} \text{BER}_{\text{MCA,ideal}} &= \mathbb{E}\left\{2Q\left(\frac{(1+\Gamma)\cos\theta}{\Gamma}\sqrt{\frac{\beta}{N_0}}\right)\right. \\ &\quad \left.\left(1-Q\left(\frac{(1+\Gamma)\cos\theta}{\Gamma}\sqrt{\frac{\beta}{N_0}}\right)\right)|\theta\right\}. \end{aligned} \quad (38)$$

For cases of Middleton class A noise plus fading channels, the lower bound can be derived by replacing  $\sigma_g^2 = N_0/2$  to  $\sigma_0^2$  in (30), (31) and (34).

TABLE I  
ENERGY EFFICIENCY COMPARISON

System	Energy efficiency
DCSK [6]	$\frac{1}{2}$
OFDM-DCSK [13]	$\frac{N-1}{N}$
$\ell_p$ -norm minimization based OFDM-DCSK [33]	$\frac{N-1}{N}$
Sparse code spreading MC-DCSK [40], $M = 2$	$\frac{\lambda(N-1)}{\lambda(N-1)+1}, \lambda > 0$
Carrier index MC-DCSK [49]	$\frac{N-1+\log_2 N}{N+1}$
EECI-OFDM-DCSK [50]	$\frac{N-1+(\log_2 N)-1}{N}$
Proposed SC-OFDM-DCSK	$\frac{N-1}{N(1-R)}$

### C. Energy Efficiency

Without loss of generality, we assume that the chips in chaotic sequences are transmitted in unit time, and the energy cost of CP is ignored [11]. In the SC-OFDM-DCSK system, each entry of  $\Omega$  is an independent Bernoulli variable with  $R$  as the probability of being 0. That is,  $R$  refers to the percentage of missing data. Therefore, we can adjust the energy cost by adjusting  $R$ . According to the Parseval's theorem, the energy cost of one SC-OFDM-DCSK symbol is  $E_{sc} = N\beta\mathbb{E}\{x_k^2\}(1-R)$ . Moreover, as one SC-OFDM-DCSK symbol carries  $N-1$  bits, hence the average energy distributed per bit is  $E_b = N\beta\mathbb{E}\{x_k^2\}(1-R)/(N-1)$ . Since the energy cost is on sequence-by-sequence basis in chaotic systems using chaotic sequences, the energy efficiency can be measured by the reference-to-bit-energy ratio (RBR) [11], denoted as RBR =  $E_{ref}/E_b = (N-1)/(N(1-R))$  where the energy of the reference chaotic sequence is  $E_{ref} = \beta\mathbb{E}\{x_k^2\}$ .

Table I shows the comparison of energy efficiency between the proposed system and other binary DCSK systems, where  $M$  and  $\lambda$  are parameters in [40]. We see that if  $R$  is set as a large value, the energy efficiency of our system can be significantly larger than that of other competing systems.

### D. Computational Complexity

For (9) and (10), it is clear from (14) and (15) that calculating  $u_i^{l+1}$  has the complexity of  $\mathcal{O}(|\Omega_i|)$ , and that of  $v_j^{l+1}$  is  $\mathcal{O}|\Omega_j|$ . Therefore, updates of both  $u^{l+1}$  and  $v^{l+1}$  are of  $\mathcal{O}(|\Omega|)$  because  $\sum_{i=0}^{N-1} |\Omega_i| = \sum_{j=0}^{\beta-1} |\Omega_j| = |\Omega|$ . As linear computation is performed on all elements with indices in  $\Omega$  in updating (11), its complexity is also  $\mathcal{O}(|\Omega|)$ . It is worth pointing out that even without sparse coding or  $|\Omega| = N\beta$ , the complexity becomes  $\mathcal{O}(N\beta)$ , which is smaller than that of [33], namely,  $\mathcal{O}(N\beta N_{\text{IRLS}})$  where  $N_{\text{IRLS}}$  is the number of iterations required in the IRLS. Moreover, the complexity of directly performing truncated singular value decomposition, corresponding to [33] at  $p = 2$ , is  $\mathcal{O}(N\beta)$ , which is equal to that of the proposed method.

## IV. SIMULATION RESULTS

Computer simulations have been carried out to evaluate the properties and performance of the proposed SC-OFDM-DCSK system. Second-order CPF is used for chaotic sequence generation. The number of subcarriers is  $N = 128$ , and the length of chaotic sequence is  $\beta = 50$ .

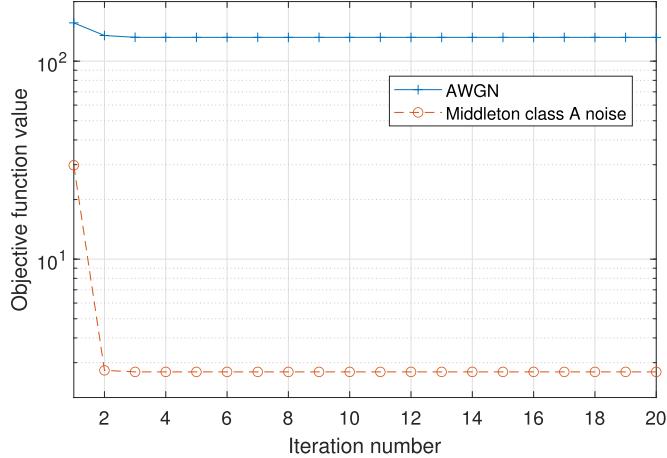


Fig. 3. Objective function value versus iteration number.

#### A. Convergence of Objective Function Value

Figure 3 plots the objective function value in (7) versus number of iterations for a single trial in the presence of AWGN and Middleton class A noise at  $E_b/N_0 = 10$  dB. Furthermore, no sparse coding is applied, that is,  $R = 0$ . We see that the analysis in Section III-A is validated, and the objective function converges very fast, namely, at the third iteration. Moreover, the objective function value over Middleton class A noise channel is smaller than the objective function value over AWGN channel. This is because in impulsive noise channels, most noise power is allocated to outliers. When  $E_b/N_0$  is fixed, outlier suppression means significant reduction of the power of impulsive noise. However, since no outliers are in AWGN channel, noise suppression effect is not obvious. Therefore, if the outliers from impulsive noise channels are suppressed properly, the power of remaining errors is smaller than AWGN power, which makes objective function value smaller.

#### B. Steady-State Error

We utilize the normalized root square error (NRSE) to determine the convergence of the algorithm, which is defined as:

$$\text{NRSE} = \frac{\|\hat{\mathbf{R}} - \mathbf{S}\|_F}{\|\mathbf{S}\|_F} \quad (39)$$

Apparently, the NRSE measures the difference between the estimated rank-1 matrix  $\hat{\mathbf{R}}$  and  $\mathbf{S}$  in the time domain, where it is assumed that the channel response matrix is  $\mathbf{H} = \mathbf{I}$  for simplicity. Lower NRSE leads to more effective low-rank matrix recovery and better BER performance for low-rank matrix approximation based OFDM-DCSK systems. In our study, Algorithm 1 is terminated when the NRSE difference between two successive iterations is smaller than  $\delta = 10^{-4}$ . Employing a PC with Intel i7-9700 3.00GHz CPU, the convergence time of Algorithm 1 varies between 0.03s and 0.16s for various  $E_b/N_0$  and input symbol values, outperforming the  $\ell_p$ -norm method [33] formulated in (8) for  $p < 2$ , which needs 0.2s to 0.6s to converge.

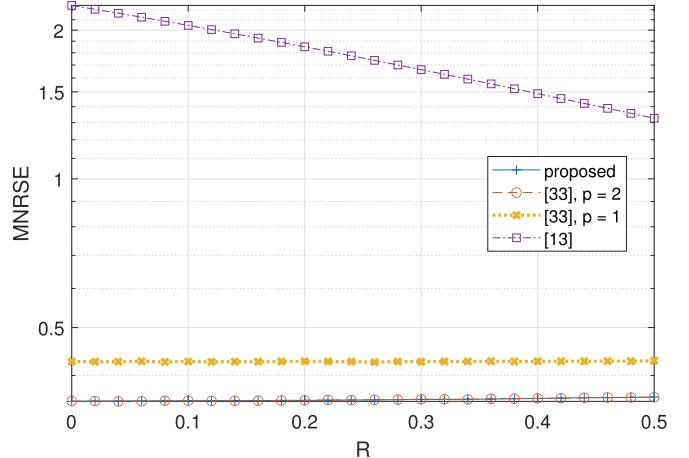


Fig. 4. MNRSE versus missing percentage in AWGN.

The steady-state mean value of the NRSE, denoted by  $\text{MNRSE} = \mathbb{E}\{\text{NRSE}\}$ , is examined for different values of  $R$  and  $E_b/N_0$  based on the average of 1000 independent runs. We evaluate the proposed system with [33] at  $p = 1$  and  $p = 2$ , and comparison with sparse coding aided OFDM-DCSK [13] is also conducted. Fig. 4 plots the MNRSE versus missing percentage  $R \in [0, 0.5]$  in AWGN at  $E_b/N_0 = 10$  dB. It is observed that the proposed scheme and  $\ell_2$ -norm method have similar MNRSE values for the whole range of  $R$ , and are much smaller than  $\ell_1$ -norm method and [13]. As  $\ell_2$ -norm minimization corresponds to the optimum estimator in AWGN, indicating that Algorithm 1 is able to yield the best rank-1 matrix recovery. This also implies that (7) is approximately reduced to (8) at  $p = 2$ , and only a few or even no entries have been identified as outliers. On the other hand, the MNRSE performance versus  $E_b/N_0 \in [-20, 15]$  dB is plotted in Fig. 5 for  $R = 0, 0.2$  and  $0.4$ . We see that the results align with those of Fig. 4. That is, both the proposed system and [33] with  $p = 2$  perform comparably, and are superior to the  $\ell_1$ -norm method and [13]. In addition, the proposed system and [33] are insensitive to missing data even when  $R \leq 0.4$ , while [13] has slightly smaller MNRSE when  $R$  is larger. This is because the sparse coding reduces the signal energy, and the noise power is also reduced when  $E_b/N_0$  is fixed. Although [13] has smaller MNRSE when  $R$  is larger, the value is still higher than other systems with low-rank approximation.

We replace the AWGN by an impulsive process, namely, Middleton class A noise [48], to evaluate the outlier suppression performance via repeating the above tests. The results are plotted in Figs. 6 and 7, while the  $E_b/N_0$  in Fig. 6 is 10 dB. As [13] does not perform the low-rank matrix approximation, and  $p = 2$  is not robust to outliers, [13] and the  $\ell_2$ -norm minimization method have very large MNRSEs for different values of  $R$  and  $E_b/N_0$ . Moreover, we see that the proposed system produces the minimum MNRSE in all these settings, and outperforms the  $\ell_1$ -norm scheme. This means that the formulation in (7) which simultaneously handles both Gaussian background noise and impulsive components is more effective than (8) because the  $\ell_1$ -norm treats all entries as outlier-contaminated terms, particularly in more noisy conditions. From Figs. 4 to 7, the advantage of automatic

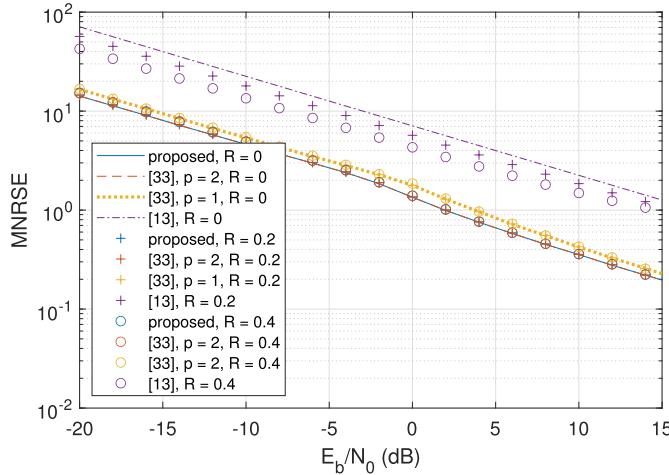
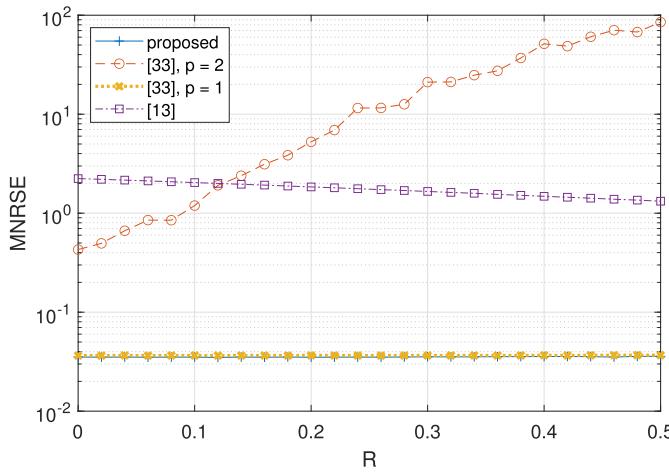
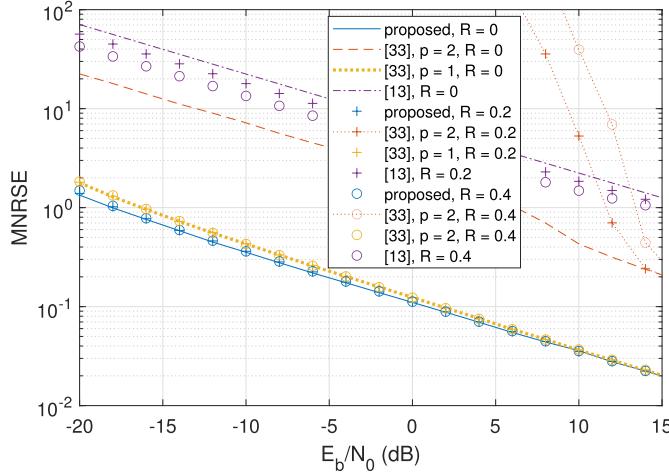
Fig. 5. MNRSE versus  $E_b/N_0$  in AWGN.

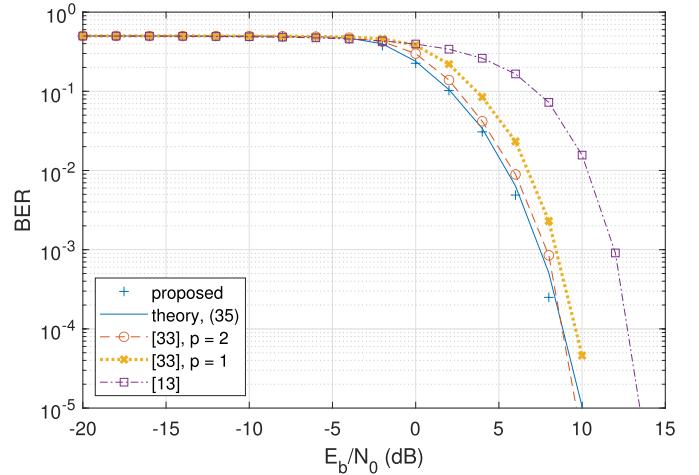
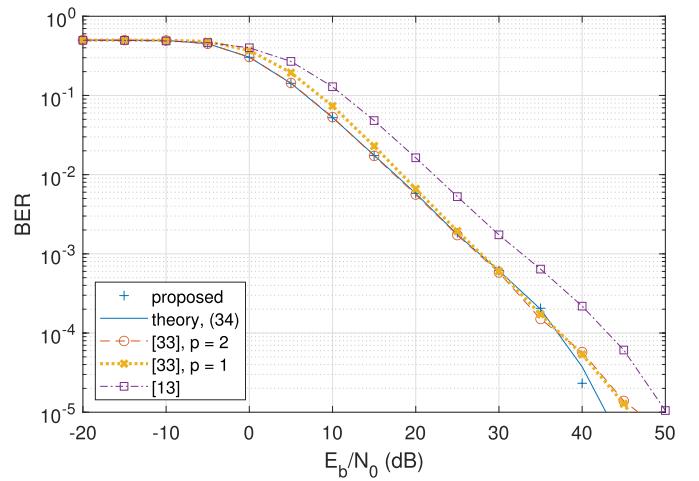
Fig. 6. MNRSE versus missing percentage in Middleton class A noise.

Fig. 7. MNRSE versus  $E_b/N_0$  in Middleton class A noise.

parameter adjustment to achieve the best performance of (7) in different noise models over the  $\ell_p$ -norm approach with a fixed value of  $p$  is also demonstrated.

### C. Bit Error Rate

Finally, the BER performance of the proposed system is evaluated by comparing with [33]. For each value of  $E_b/N_0$ , 3938 independent trials, corresponding to  $3938 \times (N - 1) =$

Fig. 8. BER versus  $E_b/N_0$  in AWGN at  $R = 0$ .Fig. 9. BER versus  $E_b/N_0$  in AWGN over flat fading channel at  $R = 0$ .

500126 bits, are performed. To show the results over fading channels, the average power of flat fading is 1, and the multi-path fading has three paths with average path power  $1/3$  and  $[0, 2, 4]$  delayed symbols for each path. Moreover, we assume that perfect channel estimation information is obtained for frequency equalization [45] at the receiver.

Figures 8–11 plot the BER in AWGN versus  $E_b/N_0$  for the proposed scheme, [33] at  $p = 1$  and  $p = 2$ , and [13]. The  $E_b/N_0$  range in Figs. 8 and 11 is set as  $[-20, 15]$  dB, while Figs. 9 and 10 extend the range of  $E_b/N_0$  to  $[-20, 50]$  dB. The theoretical BERs in AWGN, given by (35), which even holds for incomplete data but with accurate matrix recovery, are also included. It is observed that the proposed system and the optimum  $\ell_2$ -norm method produce similar BERs and are superior to the  $\ell_1$ -norm scheme and [13], while (35) is validated. Figs. 9 and 10 show that similar results are obtained over flat fading channel and multi-path fading channel. The theoretical BERs in (34) and (31) are also validated. Moreover, we see from Fig. 11 that losing 20% of the transmitted symbols basically will not degrade the BER performance, which aligns with the findings in Figs. 4 and 5.

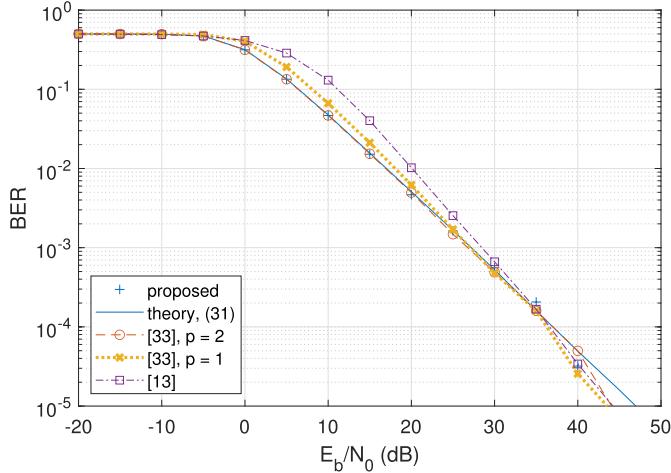


Fig. 10. BER versus  $E_b/N_0$  in AWGN over multi-path fading channel at  $R = 0$ .

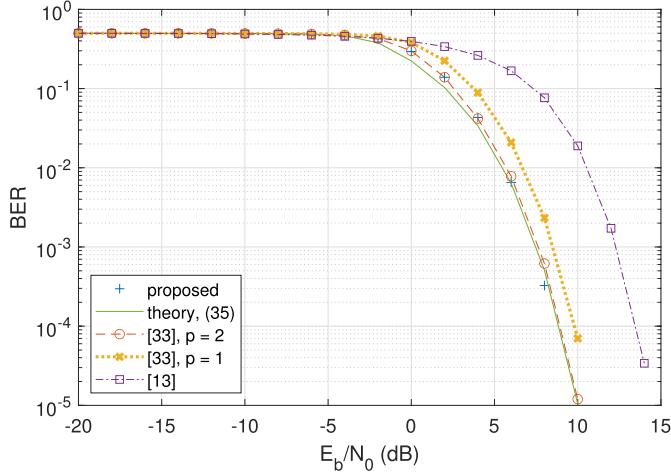


Fig. 11. BER versus  $E_b/N_0$  in AWGN at  $R = 0.2$ .

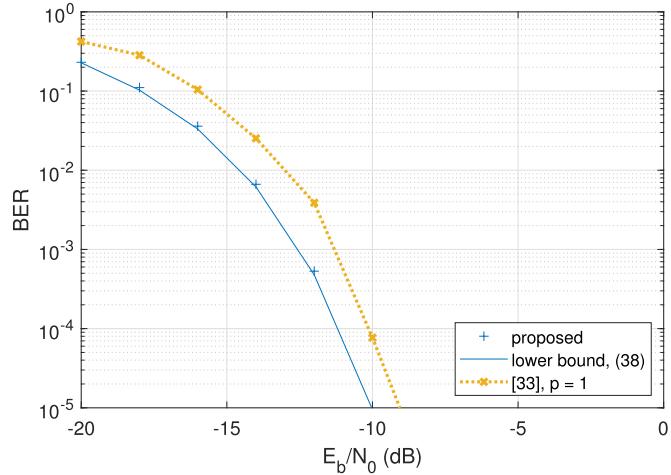


Fig. 12. BER versus  $E_b/N_0$  in Middleton class A noise at  $R = 0$ .

The above tests are repeated for the Middleton class A noise, and the results are shown in Figs. 12-15. As the  $\ell_2$ -norm method and [13] seriously deteriorate in impulsive noise environments, their results are not included. To calculate the lower bound in Figs. 13 and 14,  $\cos \theta$  in (38) should be first calculated according to (28) with the transmitted chaos sequence  $\mathbf{x}$  and estimated chaos sequence  $\tilde{\mathbf{v}}$ . Subsequently,

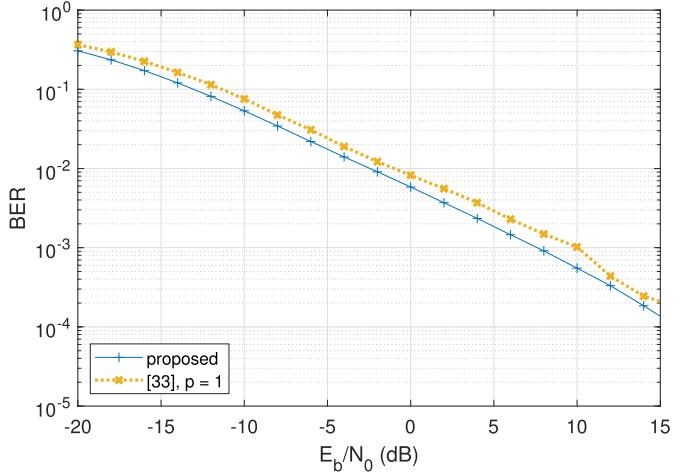


Fig. 13. BER versus  $E_b/N_0$  in Middleton class A noise over flat fading channel at  $R = 0$ .

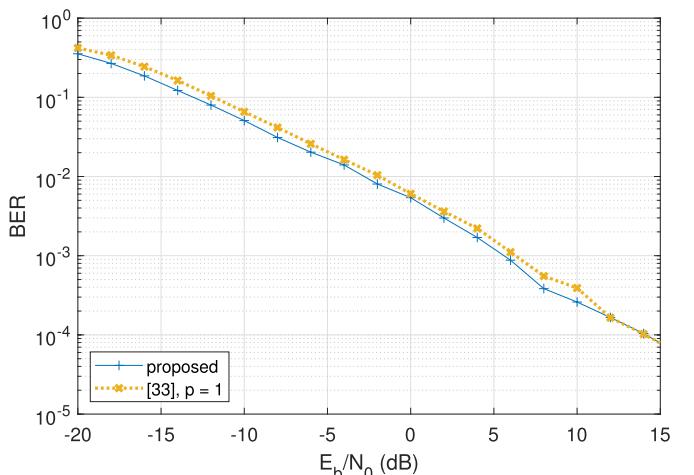


Fig. 14. BER versus  $E_b/N_0$  in Middleton class A noise over multi-path fading channel at  $R = 0$ .

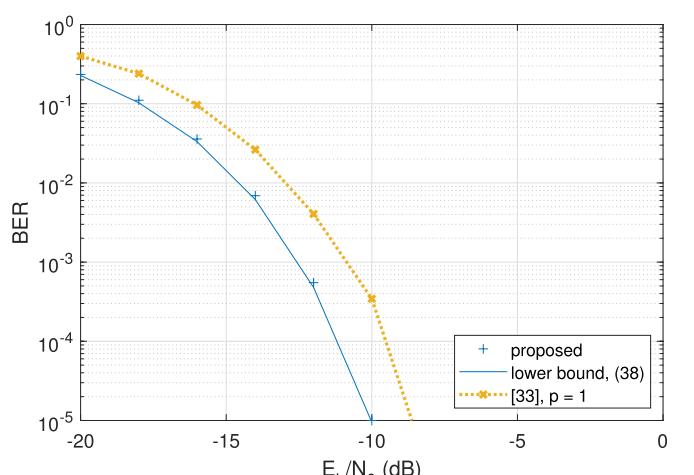


Fig. 15. BER versus  $E_b/N_0$  in Middleton class A noise at  $R = 0.2$ .

the conditional BER of a specific  $\theta$  is calculated by the equation inside the expectation in (38). Finally, the value of (38) is calculated by taking the expectation of different conditional BERs with different  $\theta$ . We clearly see that the BER superiority of the proposed system over the  $\ell_1$ -norm scheme in

the absence and presence of missing transmitted data, and the BER of the proposed system matches the BER lower bound in (38) well, which means that the outliers are suppressed well. Moreover, in Figs. 13 and 14, the proposed system has better BER performance than the  $\ell_1$ -norm scheme in Middleton class A noise over flat fading channel and multi-path fading channel, respectively. Furthermore, their BER performance basically remains unchanged at  $R = 0.2$ . These findings also agree with Figs. 6 and 7.

## V. CONCLUSION

A robust and energy efficient SC-OFDM-DCSK system is devised to tackle possibly outlier-contaminated received symbols and enhance energy efficiency. By making use of the rank-1 property of the noise-free OFDM-DCSK symbol matrix, the outer product representation is adopted as its model, whose estimation is formulated in a least squares setting, together with an  $\ell_0$ -norm term for outlier detection. In doing so, both Gaussian background noise and impulsive components can be handled simultaneously by matrix recovery. Due to the rank-1 redundancy, we also suggest sparse coding to reduce the transmission energy, and the original signal can be recovered from the sparse-coded signal. BCD is utilized as the fast solver with automatic parameter control and the Laplacian kernel for identifying the outliers. Compared with the  $\ell_p$ -norm minimization based rank-1 matrix recovery approach [33] in the AWGN environments, our solution is comparable to that of  $p = 2$  which is the optimum estimator, and the derived BER is validated even when the missing symbol ratio is up to 20%. In the presence of Middleton class A noise, the proposed system outperforms [33] with  $p = 1$  in both computational complexity and BER, which are also theoretically justified. Furthermore, our scheme can achieve higher energy efficiency than competing binary DCSK based systems.

## REFERENCES

- [1] Y. Fang, G. Han, P. Chen, F. C. M. Lau, G. Chen, and L. Wang, "A survey on DCSK-based communication systems and their application to UWB scenarios," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1804–1837, 3rd Quart., 2016.
- [2] F. C. M. Lau and C. K. Tse, *Chaos-Based Digital Communication Systems*. New York, NY, USA: Springer, 2003.
- [3] H. Ma, G. Cai, Y. Fang, J. Wen, P. Chen, and S. Akhtar, "A new enhanced energy-detector-based FM-DCSK UWB system for tactile internet," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 3028–3039, May 2019.
- [4] G. Kaddoum and N. Tadayon, "Differential chaos shift keying: A robust modulation scheme for power-line communications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 1, pp. 31–35, Jan. 2017.
- [5] Z. Chen, L. Zhang, J. Zhang, Z. Wu, and D. Luobu, "An OFDM-based pre-coded chaos shift keying transceiver for reliable V2V transmission," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6710–6715, Jun. 2022.
- [6] G. Kolumbán, B. Vizvári, W. Schwarz, and A. Abel, "Differential chaos shift keying: A robust coding for chaos communication," in *Proc. Nonlinear Dyn. Electron. Syst.*, Seville, Spain, Jun. 1996, pp. 87–92.
- [7] H. Cao, Y. T. Chan, and H. C. So, "Compressive TDOA estimation: Cramér–Rao bound and incoherent processing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 4, pp. 3326–3331, Aug. 2020.
- [8] Z. Galias and G. M. Maggio, "Quadrature chaos-shift keying: Theory and performance analysis," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 12, pp. 1510–1519, Dec. 2001.
- [9] L. Wang, G. Cai, and G. R. Chen, "Design and performance analysis of a new multiresolution  $M$ -ary differential chaos shift keying communication system," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5197–5208, Sep. 2015.
- [10] N. Li, J. Martínez-Ortega, V. H. Díaz, and J. M. M. Chaus, "A new high-efficiency multilevel frequency-modulation different chaos shift keying communication system," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3334–3345, Dec. 2018.
- [11] Z. Liu, L. Zhang, Z. Wu, and Y. Jiang, "Energy efficient parallel concatenated index modulation and  $M$ -ary PSK aided OFDM-DCSK communications with QoS consideration," *IEEE Trans. Vehi. Technol.*, vol. 69, no. 9, pp. 9469–9482, Sep. 2020.
- [12] G. Kaddoum, F.-D. Richardson, and F. Gagnon, "Design and analysis of a multi-carrier differential chaos shift keying communication system," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3281–3291, Aug. 2013.
- [13] S. Li, Y. Zhao, and Z. Wu, "Design and analysis of an OFDM-based differential chaos shift keying communication system," *J. Commun.*, vol. 10, no. 3, pp. 199–205, 2015.
- [14] G. Kaddoum, "Design and performance analysis of a multiuser OFDM based differential chaos shift keying communication system," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 249–260, Jan. 2016.
- [15] X. Cai, W. Xu, L. Wang, and G. Kaddoum, "Joint energy and correlation detection assisted non-coherent OFDM-DCSK system for underwater acoustic communications," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 3742–3759, Jun. 2022.
- [16] Z. Liu, L. Zhang, and Z. Chen, "Low PAPR OFDM-based DCSK design with carrier interferometry spreading codes," *IEEE Commun. Lett.*, vol. 22, no. 8, pp. 1588–1591, Aug. 2018.
- [17] Z. Liu, L. Zhang, Z. Wu, and J. Bian, "A secure and robust frequency and time diversity aided OFDM-DCSK modulation system not requiring channel state information," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1684–1697, Mar. 2020.
- [18] Z. Liu, L. Zhang, and Z. Wu, "Reliable and secure pre-coding OFDM-DCSK design for practical cognitive radio systems with the carrier frequency offset," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 1, pp. 189–200, Mar. 2020.
- [19] M. Mirahmadi, A. Al-Dweik, and A. Shami, "BER reduction of OFDM based broadband communication systems over multipath channels with impulsive noise," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4602–4615, Nov. 2013.
- [20] K. M. Rabie, B. Adebisi, A. M. Tonello, and G. Nauryzbayev, "For more energy-efficient dual-hop DF relaying power-line communication systems," *IEEE Syst. J.*, vol. 12, no. 2, pp. 2005–2016, Jun. 2018.
- [21] T. Bai et al., "Fifty years of noise modeling and mitigation in power-line communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 41–69, 1st Quart., 2021.
- [22] K. C. Wiklundh, P. F. Stenumgaard, and H. M. Tullberg, "Channel capacity of Middleton's class A interference channel," *Electron. Lett.*, vol. 45, no. 24, pp. 1227–1229, Nov. 2009.
- [23] G. Laguna-Sánchez and M. Lopez-Guerrero, "On the use of alpha-stable distributions in noise modeling for PLC," *IEEE Trans. Power Del.*, vol. 30, no. 4, pp. 1863–1870, Aug. 2015.
- [24] B. Van Nguyen, M. Le, H. Jung, and K. Kim, "Antijamming receiver with hybrid blanking—Clipping for DCSK system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 5, pp. 3751–3761, Oct. 2020.
- [25] G. Kaddoum and E. Soujeri, "NR-DCSK: A noise reduction differential chaos shift keying system," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 63, no. 7, pp. 648–652, Jul. 2016.
- [26] W. Rao, L. Zhang, Z. Zhang, and Z. Wu, "Noise-suppressing chaos generator to improve BER for DCSK systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.
- [27] X. Cai, W. Xu, and L. Wang, "Design of divide-and-conquer noise decontamination strategy for  $M$ -ary DCSK: From remodulation to denoising," *IEEE Commun. Lett.*, vol. 26, no. 7, pp. 1673–1677, Jul. 2022.
- [28] B. Chen, L. Zhang, and Z. Wu, "General iterative receiver design for enhanced reliability in multi-carrier differential chaos shift keying systems," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7824–7839, Nov. 2019.
- [29] X. Cai, W. Xu, L. Wang, and G. Kolumbán, "Design and performance analysis of a robust multi-carrier  $M$ -ary DCSK system: A noise suppression perspective," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 1623–1637, Mar. 2022.

- [30] M. Miao, L. Wang, G. Chen, and W. Xu, "Design and analysis of replica piecewise  $M$ -ary DCSK scheme for power line communications with asynchronous impulsive noise," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 12, pp. 5443–5453, Dec. 2020.
- [31] X. Cai, W. Xu, L. Wang, and G. Chen, "Design of joint position and constellation mapping assisted DCSK scheme subject to Laplacian impulsive noise," *IEEE Commun. Lett.*, vol. 26, no. 2, pp. 463–467, Feb. 2022.
- [32] M. Miao, L. Wang, and G. Chen, "Performance and capacity analysis of MDCSK-BICM for impulsive noise in PLC," *IEEE Trans. Power Del.*, vol. 37, no. 4, pp. 3164–3175, Aug. 2022.
- [33] Z. Liu, H. C. So, L. Zhang, and X. P. Li, "Robust receiver for OFDM-DCSK modulation via rank-1 modeling and  $\ell_p$ -minimization," *Signal Process.*, vol. 188, Nov. 2021, Art. no. 108219.
- [34] W.-J. Zeng, H. C. So, and L. Huang, " $\ell_p$ -MUSIC: Robust direction-of-arrival estimator for impulsive noise environments," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4296–4308, Sep. 2013.
- [35] W.-J. Zeng and H. C. So, "Outlier-robust matrix completion via  $\ell_p$ -minimization," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1125–1140, Mar. 2018.
- [36] X. P. Li, Q. Liu, and H. C. So, "Rank-one matrix approximation with  $\ell_p$ -norm for image inpainting," *IEEE Signal Process. Lett.*, vol. 27, pp. 680–684, 2020.
- [37] Z. Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *J. Optim. Theory Appl.*, vol. 72, no. 1, pp. 7–35, Jan. 1992.
- [38] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1758–1789, Jan. 2013.
- [39] E. J. Candès and Y. Plan, "Matrix completion with noise," *Proc. IEEE*, vol. 98, no. 6, pp. 925–936, Jun. 2010.
- [40] Z. Chen, L. Zhang, Z. Wu, L. Wang, and W. Xu, "Reliable and efficient sparse code spreading aided MC-DCSK transceiver design for multiuser transmissions," *IEEE Trans. Commun.*, vol. 69, no. 3, pp. 1480–1495, Mar. 2021.
- [41] Q. Liu, Y. Gu, and H. C. So, "DOA estimation in impulsive noise via low-rank matrix approximation and weakly convex optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 6, pp. 3603–3616, Dec. 2019.
- [42] J. Kim and C. D. Scott, "Robust kernel density estimation," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 2529–2565, Jan. 2012.
- [43] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Evanston, IL, USA: Routledge, 2018.
- [44] A. Beck, *First-Order Methods in Optimization*. Philadelphia, PA, USA: SIAM, 2017.
- [45] X. Ouyang and J. Zhao, "Single-tap equalization for fast OFDM signals under generic linear channels," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1319–1322, Aug. 2014.
- [46] M. Barkat, *Signal Detection and Estimation* (Artech House Radar Library). Norwood, MA, USA: Artech House, 2005.
- [47] L. Zhang, J. Zheng, B. Chen, and Z. Wu, "Reliable low-rank approximation of matrices detection aided multicarrier DCSK receiver design," *IEEE Syst. J.*, vol. 15, no. 4, pp. 5277–5288, Dec. 2021.
- [48] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: New methods and results for class a and class b noise models," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1129–1149, May 1999.
- [49] G. Cheng, L. Wang, W. Xu, and G. Chen, "Carrier index differential chaos shift keying modulation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 8, pp. 907–911, Aug. 2017.
- [50] Z. Liu, L. Zhang, Z. Wu, and J. Bian, "Carrier interferometry code index modulation aided OFDM-based DCSK communications," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Honolulu, HI, USA, Sep. 2019, pp. 1–5.



**Zhaofeng Liu** received the B.S. and M.S. degrees from the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China, in 2018 and 2020, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, City University of Hong Kong. His research interests include chaotic communication, multi-carrier modulation, robust signal processing, and MIMO radar.



**Hing Cheung So** (Fellow, IEEE) was born in Hong Kong. He received the B.Eng. degree in electronic engineering from the City University of Hong Kong, Hong Kong, in 1990, and the Ph.D. degree in electronic engineering from The Chinese University of Hong Kong, Hong Kong, in 1995.

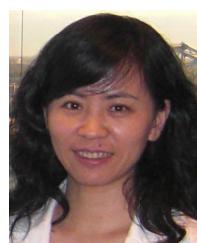
From 1990 to 1991, he was an Electronic Engineer with the Research and Development Division, Everex Systems Engineering Ltd., Hong Kong. From 1995 to 1996, he was a Post-Doctoral Fellow with The Chinese University of Hong Kong.

From 1996 to 1999, he was a Research Assistant Professor with the Department of Electronic Engineering, City University of Hong Kong, where he is currently a Professor. His research interests include detection and estimation, fast and adaptive algorithms, multidimensional harmonic retrieval, robust signal processing, source localization, and sparse approximation. He has been on the editorial boards of *IEEE Signal Processing Magazine* from 2014 to 2017, *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 2010 to 2014, *Signal Processing* since 2010, and *Digital Signal Processing* since 2011. He was also the Lead Guest Editor of *IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING*, Special Issue on "Advances in Time/Frequency Modulated Array Signal Processing" in 2017. In addition, he was an Elected Member in Signal Processing Theory and Methods Technical Committee, IEEE Signal Processing Society, from 2011 to 2016, where he was the chair in the awards subcommittee from 2015 to 2016.



**Xiao Peng Li** received the B.Eng. degree in electronic science and technology from Yanshan University, Qinhuangdao, China, in 2015, and the M.Sc. degree (Hons.) in electronic information engineering and the Ph.D. degree in electrical engineering from the City University of Hong Kong, Hong Kong, SAR, China, in 2018 and 2022, respectively.

He was a Research Assistant with the Department of Information Engineering, Shenzhen University, Shenzhen, China, from 2018 to 2019, and a Post-Doctoral Fellow with the Department of Electrical Engineering, City University of Hong Kong, from 2022 to 2023. He is currently an Assistant Professor with the College of Electronics and Information Engineering, Shenzhen University. His research interests include optimization methods, machine learning, sparse recovery, matrix processing, and tensor processing, with applications in target estimation, image recovery, video restoration, hyperspectral unmixing, and stock market analysis.



**Lin Zhang** (Senior Member, IEEE) received the Ph.D. degree in information and communication engineering from Sun Yat-sen University, Guangzhou, China, in 2003. In 2003, she joined the School of Information Science and Technology, Sun Yat-sen University, where she has been an Associate Professor since 2007. From 2008 to 2009, she was a Visiting Researcher with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA. In 2016, she joined the School of Electronics and Information Technology, Sun Yat-sen University. She is also with the Southern Marine Science and Engineering Guangdong Laboratory. Her research has been supported by the National Natural Science Foundation of China and the Science and Technology Program Project of Guangdong Province. Her current research interests include signal processing and their applications to wireless communication systems.



**Zhi-Yong Wang** received the B.S. degree in mechanical engineering from Zhengzhou University, Zhengzhou, China, in 2017, and the M.S. degree in mechanical engineering from Xi'an Jiaotong University, Xi'an, China, in 2020. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong, SAR, China. His research interests include nonconvex optimization, sparse recovery, matrix/tensor completion, and robust signal processing.