

Received September 9, 2017, accepted October 3, 2017, date of publication October 23, 2017, date of current version November 14, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2765346

# Signature Design of Sparsely Spread Code Division Multiple Access Based on Superposed Constellation Distance Analysis

GUANGHUI SONG<sup>1</sup>, (Member, IEEE), XIANBIN WANG<sup>2</sup>, (Fellow, IEEE), AND JUN CHENG<sup>1</sup>, (Member, IEEE)

<sup>1</sup>Department of Intelligent Information Engineering and Sciences, Doshisha University, Kyoto 610-0321, Japan

<sup>2</sup>Department of Electrical and Computer Engineering, University of Western Ontario, London, ON N6A 5B9, Canada

Corresponding author: Guanghui Song (gsong2017@gmail.com)

This work was supported in part by the Japan Society for the Promotion of Science through the Grant-in-Aid for Scientific Research (C) under Grant 16K06373 and in part by the Ministry of Education, Culture, Sports, Science and Technology through the Strategic Research Foundation at Private Universities (2014–2018) under Grant S1411030.

**ABSTRACT** Sparsely spread code division multiple access (SCDMA) is a non-orthogonal superposition coding scheme that allows concurrent communications between a base station and multiple users over a common channel. However, the detection performance of an SCDMA system is mainly determined by its signature matrix, which should be sparse to facilitate the belief propagation (BP) detection. On the other hand, to guarantee good maximum likelihood (ML) detection performance, the minimum Euclidean distance for the equivalent signal constellation after multi-user superposition should be maximized. In this paper, a code distance analysis is proposed for SCDMA systems with a finite number of users and spreading lengths. Based on this analysis, good signature matrices whose factor graphs have very few short cycles and possess large superposed signal constellation distances are designed. The proposed signature matrices have both good BP and ML detection performances. Moreover, their BP detection performances exactly converge to their ML detection performances with few iterations. It is worth pointing out that the proposed signature matrix design could be directly applied to the 5G non-orthogonal multiple access systems.

**INDEX TERMS** Non-orthogonal multiple access, sparsely spread, signature design, code distance.

## I. INTRODUCTION

The future fifth generation (5G) mobile networks are expected to provide an unprecedented capacity in supporting the rapid growth of mobile data traffic with very limited spectrum resources. New multiple access techniques, i.e., non-orthogonal multiple access (NOMA), which allow multiple concurrent communications, have been recognized as one of most efficient solutions to fulfill these requirements [1]–[9]. Recently, NOMA is realized by employing non-orthogonal sparse codes for different users, such as in sparsely spread code division multiple access (SCDMA) [10], [11], low-density spreading [12], and sparse code multiple access [7]–[9], [13]. All of these techniques rely on sparse signature sequences and near-optimal joint multi-user belief propagation (BP) detections on sparse graphs. Therefore, we collectively call these techniques SCDMA. SCDMA yields many advantages with respect to the capacity load and detection complexity over conventional dense CDMA and other orthogonal multiple access schemes.

In the downlink of a general SCDMA system, a base station simultaneously communicates with multiple users. Data

streams for the multiple users are first spread (encoded) into vectors by multiplying their signature sequences, which are sparse and whose elements are usually selected from a given alphabet set. Multiple data streams after spreading are superimposed at the base station and broadcasted to the users over a common channel, i.e., using the common resources such as time slots and frequency channels. A multi-user BP detection is performed at each user to recover the data streams.

The performance of SCDMA detection is mainly determined by a signature matrix that consists of all the users' signature sequences as its row vectors. Generally, the signature matrix should have a good sparse property, i.e., without short cycles in the formed factor graph, to achieve a good BP detection performance. Theoretically, if its factor graph has no cycles, the BP detection converges to the maximum likelihood (ML) detection performance [14]. Moreover, the equivalent signal constellation after spreading and superposition should have large Euclidean distances which ultimately determine the performance bound of ML detection. This motivates us to design the elements in the signature matrix in SCDMA.

Signature design has been investigated for dense spreading in conventional CDMA [15]–[17], where an orthogonal or low-correlated sequence set is constructed to maximize an equivalent CDMA channel capacity. The problem becomes more complicated for sparse spreading in SCDMA since the design should be implemented under the sparsity constraint of the signature matrix. The problem becomes even more difficult when a two-dimensional modulation scheme is employed as in the scenarios of [7] and [13]. References [18] and [19] show that a user constellation rotation significantly affects detection performance of multi-user superposition codes. Convolutional codes are employed by each user in [18] and [19] and the multiple access scheme is referred to as trellis code multiple access (TCMA), which can be regarded as a spatial case of the scenarios in [7] and [13] with unitary spreading length. Reference [20] considers two-user TCMA and designs the user constellation rotation by maximizing an equivalent channel capacity. References [7] and [13] consider a general multi-user SCDMA with a non-trivial spreading length. For a given regular factor graph structure, References [7] and [13] show that a Latin-rectangular signature matrix significantly outperforms a randomly generated signature matrix due to a large minimum code distance property. However, most above works design good signature matrix based on practical system error rate simulation. Relatively few works propose efficient methods to find an optimal signature matrix with the maximum minimum code distance for a given SCDMA system. Moreover, improving the minimum code distance may only guarantee that the code has a good ML detection performance. How to design a code with both good BP and ML detection performances is also a challenging problem. Regarding to the performance analyses for SCDMA systems, the widely used methods, such as statistical mechanics-based analysis [10] and density evolution [11], rely on the assumption of a large number of users and a infinite spreading length. Relatively few analyses are proposed for practical SCDMA systems with finite number of users and spreading length.

In this paper, we consider SCDMA systems with a two-dimensional quadrature amplitude modulation (QAM) and develop a theoretical framework for signature design. We propose a code distance analysis for SCDMA systems with a finite number of users and spreading lengths. For a given sparse factor graph structure of an SCDMA code, we design the optimal signature matrix with the maximum minimum code distance. Since it is very difficult to give a general formula to find the optimal signature matrix, we propose some rules to simplify the design. Code designs for both cycle-free graph and graph with few cycles are proposed. We construct two SCDMA code families whose factor graphs have very few short cycles. The optimal signature labeling is hence proposed based on the graph structure. The constructed SCDMA codes outperform the existing codes in terms of both the word error rate (WER) performance and detection complexity. Our numerical results show that their BP detections exactly converge to their ML detection performances with

few iterations. Simulations for turbo-coded SCDMA systems with variety communication rates are given to validate our design in more practical applications.

The remainder of the paper is organized as follows. Section II describes the SCDMA system model and detection algorithms. Section III defines the SCDMA code distance and some properties on code distance are shown. Section IV gives the optimal signature matrix design. Section V gives two constructions of code families with few short cycles in their factor graph and large minimum code distance. Section VI gives simulations for our design in both uncoded and turbo-coded SCDMA systems. Section VII concludes this paper.

## II. SCDMA AND DETECTIONS

### A. SYSTEM MODEL

Figure 1 shows a  $K$ -user downlink SCDMA transmitter model at the base station. There are  $K$  data streams to be transmitted to the  $K$  mobile users. After forward error correction (FEC) encoding, each user’s data stream is modulated and spread by multiplying its signature sequence. Figure 1 illustrates the spread processing for an individual symbol of each user’s data stream. Here we consider QAM with  $x_k \in \mathcal{X} \triangleq \{\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}, \frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}, -\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}, -\frac{1}{\sqrt{2}} - \frac{i}{\sqrt{2}}\}$ , where  $i$  is the imaginaryy unit. The output after spreading is  $(s_{1,k}, \dots, s_{N,k})x_k$  for  $x_k$ , where  $(s_{1,k}, \dots, s_{N,k})$  with  $s_{n,k} = 0$  or  $s_{n,k} = e^{i\theta}$ ,  $\theta \in [0, 2\pi)$ ,  $n = 1, \dots, N$ , is called a signature sequence of user  $k$ . Here we considered unitary energy for each nonzero element of the signature sequence. It should be emphasized that the spreading vector is sparse, i.e., the majority of elements might be 0. The number of nonzero elements in a spreading vector is called an effective spreading length.

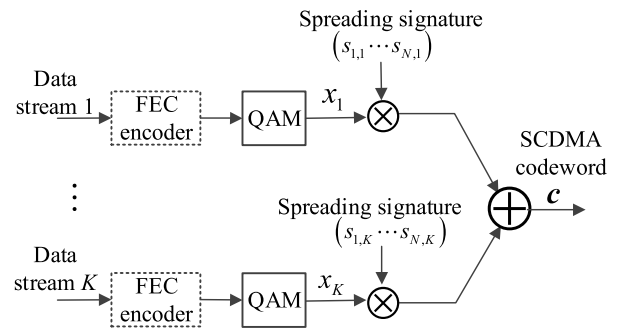


FIGURE 1. A  $K$ -user SCDMA transmitter model.

The  $K$  users’ data streams after spreading are superimposed and transmitted over  $N$  orthogonal channel resources, e.g., OFDMA tones or MIMO spatial layers. The transmitted vector is represented as

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,K} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ s_{N,1} & s_{N,2} & \cdots & s_{N,K} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} \quad (1)$$

which is referred to as an SCDMA codeword. Note that there is a total of  $4^K$  number of SCDMA codewords corresponding

to the  $4^K$  different variations of  $(x_1, \dots, x_K)^T$ , where  $()^T$  is the transpose of a matrix. Matrix  $S = [s_{n,k}]$  is sparse and is referred to as a signature matrix. By multiplexing  $K$  users over  $N$  channel resources, the load of the SCDMA code is  $K/N$ . Since for  $K/N \leq 1$ , we can use orthogonal spreading sequences to achieve near single-user performance, in this paper, we mainly consider overloaded SCDMA with  $K/N > 1$ .

Each user receives a noise-corrupted codeword  $\mathbf{y} = h\mathbf{c} + \mathbf{z}$ , where  $\mathbf{c} = (c_1, \dots, c_N)^T$ ,  $\mathbf{y} = (y_1, \dots, y_N)^T$ ,  $h$  is a channel gain, and  $\mathbf{z} = (z_1, \dots, z_N)^T$  is a complex Gaussian noise vector with each element an independent and identically distributed (i.i.d.) mean-0 variance- $N_0$  complex Gaussian variable, i.e.,  $z_n \sim \mathcal{CN}(0, N_0)$ . A joint  $K$ -user SCDMA detection is performed to recover the data streams.

Each SCDMA code can be represented by a sparse factor graph. Figure 2 gives an example of factor graph representation for a ( $K = 6$ )-user ( $N = 4$ )-resource SCDMA code proposed in [7] and [8], where data nodes  $x_1, \dots, x_K$  denote data symbols of  $K$  users, and code nodes  $c_1, \dots, c_N$  denote  $N$  SCDMA coded symbols. There is an edge between  $c_n$  and  $x_k$ , denoted as  $e_{n,k}$ , if  $s_{n,k} \neq 0$ . Let  $E$  be the set that includes all the edges in the factor graph. Edge  $e_{n,k} \in E$  is labeled by signature element  $s_{n,k}$ . Each code node is a superposition of its neighboring data nodes, i.e.,  $c_n = \sum_{\{k|e_{n,k} \in E\}} s_{n,k}x_k$ . The degree of a node is the number of edges incident with the node. The graph is called code-node regular of degree  $d$  if all the code nodes have degree  $d$ . Figure 2 is code-node regular of degree 3. If a factor graph is cycle-free, we call it a tree graph, and the corresponding code is referred to as a tree SCDMA code.

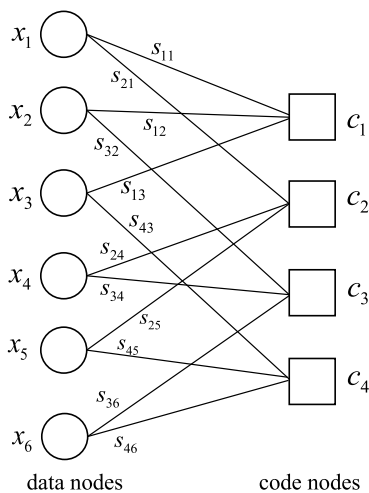


FIGURE 2. A factor graph representation of a 6-user 4-resource SCDMA code.

### B. SCDMA DETECTIONS

In this section, after we briefly review two detection algorithms of SCDMA, the ML and BP detections, we will give an approximate BP detection whose detection complexity

linearly increases with the user number. For all these three detections, we assume the receiver knows channel gain  $h$  perfectly.

#### 1) ML DETECTION

Based on  $\mathbf{y}$ , the ML detection is

$$\hat{\mathbf{c}} = \arg \max_{\mathbf{c}} \Pr(\mathbf{y}|\mathbf{c}) = \arg \min_{\mathbf{c}} \|\mathbf{y} - h\mathbf{c}\| \quad (2)$$

where the ML detection is reduced to the minimum distance detection due to the memoryless Gaussian channel. If more than one codeword satisfies (2), we randomly select one of them as our decision with equal probability. Here we assume that all SCDMA codewords are transmitted with equal probability, and thus, (2) is equivalent to the maximum *a posteriori* probability detection. Based on the estimated SCDMA codeword  $\hat{\mathbf{c}}$ , we can uniquely determine the transmitted data stream. The complexity of ML detection is  $O(4^K)$  for SCDMA with QAM.

#### 2) BP DETECTION

A suboptimal scheme with lower complexity is BP detection. This detection is performed on the factor graph. The whole detection is performed iteratively in a belief-propagation manner. In each iteration, each node of the factor graph performs a local processing and exchanges message with its neighboring nodes.

##### a: Code Node Processing

Consider the processing at code node  $c_n$ . Let  $P_{x_k, c_n}^{\ell-1}(x_k = \alpha)$ , a priori probability of  $x_k = \alpha$ ,  $\alpha \in \mathcal{X}$ , be the message output from data nodes  $x_k$ ,  $k \in \kappa = \{k|e_{n,k} \in E\}$  to code node  $c_n$  at the  $(\ell - 1)$ -th iteration. Let  $\kappa \setminus k$  be the set obtained by deleting  $k$  from  $\kappa$ . Based on this priori probability, and the channel observation  $y_n$ , code node  $c_n$  outputs a probability message of

$$\begin{aligned} P_{c_n, x_k}^{\ell}(x_k = \alpha) &= \sum_{\alpha_j \in \mathcal{X}, j \in \kappa \setminus k} \Pr(y_n | x_j = \alpha_j, j \in \kappa \setminus k, x_k = \alpha) \prod_{j \in \kappa \setminus k} P_{x_j, c_n}^{\ell-1}(x_j = \alpha_j) \\ &= \frac{1}{\pi N_0} \sum_{\alpha_j \in \mathcal{X}, j \in \kappa \setminus k} \exp\left(-\frac{|y_n - \sum_{j \in \kappa \setminus k} s_{n,j} \alpha_j - s_{n,k} \alpha|^2}{N_0}\right) \\ &\quad \times \prod_{j \in \kappa \setminus k} P_{x_j, c_n}^{\ell-1}(x_j = \alpha_j) \end{aligned}$$

to data node  $x_k$ ,  $k \in \kappa$ , which is a probability of  $x_k = \alpha$ ,  $\alpha \in \mathcal{X}$ .

##### b: Data Node Processing

Data node  $x_k$  combines the message obtained from the code nodes in its neighborhood and outputs a probability message of

$$P_{x_k, c_n}^{\ell}(x_k = \alpha) = \prod_{\{j|e_{j,k} \in E, j \neq n\}} P_{c_j, x_k}^{\ell}(x_k = \alpha), \alpha \in \mathcal{X}$$

to its neighboring code node  $c_n$ ,  $n \in \{n|e_{n,k} \in E\}$ .

*c: Hard Decision*

After a fixed number  $L$  of iterations, a hard decision is made for  $x_k$  as

$$\hat{x}_k = \arg \max_{\alpha \in \mathcal{X}} \prod_{\{j|e_{j,k} \in E\}} P_{c_j, x_k}^L(x_k = \alpha). \quad (3)$$

The complexity of BP detection is dominated by the complexity of code node processing, whose complexity is  $O(4^d)$ , where  $d$  is the maximum code node degree. The specific detection complexity is determined by the code node degree profile and the iteration number that is required for the detection.

Note that if the factor graph of an SCDMA code is a tree graph, its BP detection will converge to its ML detection with a finite number of iterations. Many works show that short cycles, such as length-4 cycles, significantly degrade the performance of BP detection [21]. Therefore, in this paper we only consider factor graphs without length-4 cycles.

3) APPROXIMATE BP DETECTION

In the BP detection, the processing at the code node is a MAP processing, where the accurate probability about the estimated data is calculated. In this section, we give an approximate BP detection with a simplified code node processing. We regard the summation of interferences for each data as complex Gaussian, so we only need to track a mean and variance message.

Take the processing at code node  $c_n$  as an example. The associated receive at this node is

$$y_n = \sum_{\{k|e_{n,k} \in E\}} s_{n,k} x_k + z_n = s_{n,k} x_k + \xi_k \quad (4)$$

where  $\xi_k = \sum_{j \in \mathcal{K} \setminus k} s_{n,j} x_j + z_n$  is the equivalent noise for  $x_k$ . We approximately regards  $\xi_k$  as a complex Gaussian variable, i.e.,  $\xi_k \sim \mathcal{CN}(\mu_k, N_k)$  with

$$\mu_k = E[\xi_k] = \sum_{j \in \mathcal{K} \setminus k} s_{n,j} E[x_j] = \sum_{j \in \mathcal{K} \setminus k} s_{n,j} \sum_{\alpha \in \mathcal{X}} P_{x_j, c_n}^{\ell-1}(x_j = \alpha) \alpha$$

$$\begin{aligned} N_k &= E[|\xi_k - \mu_k|^2] = \sum_{j \in \mathcal{K} \setminus k} E[|x_j - E[x_j]|^2] + N_0 \\ &= \sum_{j \in \mathcal{K} \setminus k} \left(1 - |E[x_j]|^2\right) + N_0 \\ &= \sum_{j \in \mathcal{K} \setminus k} \left(1 - \left|\sum_{\alpha \in \mathcal{X}} P_{x_j, c_n}^{\ell-1}(x_j = \alpha) \alpha\right|^2\right) + N_0 \end{aligned}$$

where  $E[*]$  takes the expectation of a random variable. Therefore, code node  $c_n$  outputs a probability message of

$$P_{c_n, x_k}^{\ell}(x_k = \alpha) = \frac{1}{\pi N_k} \exp\left(-\frac{|y_n - s_{n,k} \alpha - \mu_k|^2}{N_k}\right), \alpha \in \mathcal{X}$$

to data node  $x_k, k \in \mathcal{K}$ .

The approximate BP detection may work well when the code node degree is large or the noise level is high since at these two cases, interference term  $\xi_k$  is more like Gaussian.

The processing complexity of the code node reduces to  $O(d)$ , where  $d$  is the maximum code node degree.

III. SCDMA CODE DISTANCE AND PROPERTIES

In this section, we first define an SCDMA code distance and distance enumerator function, which is used to formulate a union bound for ML detection. Some properties about the SCDMA code distance enumerator function and the minimum code distance are derived.

*Definition 1:* Distance between two SCDMA codewords  $\mathbf{c}, \mathbf{c}' \in \mathcal{C}$  is

$$d(\mathbf{c}, \mathbf{c}') = \|\mathbf{c} - \mathbf{c}'\|$$

where  $\mathcal{C}$  is the SCDMA code set. □

*Definition 2:*

$$d_{\min} = \min_{\mathbf{c}, \mathbf{c}' \in \mathcal{C}, \mathbf{c} \neq \mathbf{c}'} d(\mathbf{c}, \mathbf{c}') \quad (5)$$

is called the minimum distance of SCDMA code  $\mathcal{C}$ . □

Applying (1) to (5), we obtain the following lemma straightforwardly.

*Lemma 1:* Let  $\Delta \mathcal{X} \triangleq \{0, \pm\sqrt{2}, \pm\sqrt{2}i, \sqrt{2} \pm \sqrt{2}i, -\sqrt{2} \pm \sqrt{2}i\}$  and  $\Delta \mathcal{X}^K$  be the universal set of length  $K$  vectors over  $\Delta \mathcal{X}$ . The minimum distance of the SCDMA code with spreading signature matrix  $S$  is

$$\begin{aligned} d_{\min}(S) &= \min_{\mathbf{u} \in \Delta \mathcal{X}^K, \mathbf{u} \neq \mathbf{0}} F(S, \mathbf{u}) \\ F(S, \mathbf{u}) &\triangleq \sqrt{\sum_{n=1}^N \left| \sum_{k=1}^K s_{n,k} u_k \right|^2} \end{aligned}$$

where  $\mathbf{u} = (u_1, \dots, u_K)$  and  $\mathbf{0} = (0, \dots, 0)$ . □

To give a global description of the code distance spectrum of an SCDMA code, we have the following definition.

*Definition 3:* The distance enumerator function for an SCDMA code with signature matrix  $S$  is

$$\begin{aligned} A(S, Z) &= \frac{1}{|\mathcal{C}|} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{c}' \in \mathcal{C}, \mathbf{c}' \neq \mathbf{c}} Z^{d(\mathbf{c}, \mathbf{c}')} \\ &= \frac{1}{4^K} \sum_{\mathbf{u} \in \mathcal{X}^K} \sum_{\mathbf{u}' \in \mathcal{X}^K, \mathbf{u}' \neq \mathbf{u}} Z^{F(S, \mathbf{u} - \mathbf{u}')} \quad (6) \end{aligned}$$

where  $Z$  is a dummy variable, and  $|\mathcal{C}| = 4^K$  for QAM is the cardinality of the code set. □

Equation (6) in fact gives an average distance spectrum for all the codewords in the code set.

The distance enumerator function of SCDMA code can be used to calculate a multi-user union bound developed in [22]. It is an WER upper bound on ML detection.

Let  $A(S, Z) = \sum_d A(d) Z^d$  be the distance enumerator function of an SCDMA code with signature matrix  $S$ , where  $A(d)$  can be regarded as the average number of codeword pairs with distance  $d$ . The WER  $P_W$  under ML detection is upper bounded by

Union Bound [22]:

$$P_W \leq A(0) + \sum_{d>0} A(d)Q\left(\frac{d}{\sqrt{2N_0}}\right). \quad (7)$$

Note that (7) has a form different from that in [22] since the definition of code distance in this work has a form different from that in [22].

We give the following properties for the SCDMA code distance enumerator function and minimum code distance.

*Lemma 2 (Row Rotation Invariance):*

$$A(S', Z) = A(S, Z) \quad (8)$$

holds for  $s'_{n,k} = e^{i\theta_n} s_{n,k}$ ,  $\theta_n \in [0, 2\pi)$ ,  $k = 1, \dots, K$ ,  $n = 1, \dots, N$ .  $\square$

*Proof:* Equation (8) holds because

$$\begin{aligned} F(S', \mathbf{u}) &= \sqrt{\sum_{n=1}^N \left| \sum_{k=1}^K e^{i\theta_n} s_{n,k} u_k \right|^2} = \sqrt{\sum_{n=1}^N \left| \sum_{k=1}^K s_{n,k} u_k \right|^2} \\ &= F(S, \mathbf{u}) \end{aligned}$$

holds for any  $\mathbf{u} \in \Delta\mathcal{X}^K$ .  $\square$

*Lemma 3 (Column Rotation Invariance):*

$$A(S^*, Z) = A(S, Z) \quad (9)$$

holds for any  $s^*_{n,k} = e^{im_k\pi/2} s_{n,k}$ ,  $m_k \in \mathbb{Z}$ ,  $k = 1, \dots, K$ ,  $n = 1, \dots, N$ , where  $Z$  is a set of integer numbers.  $\square$

*Proof:* Equation (9) holds because

$$\begin{aligned} F(S^*, \mathbf{u}) &= \sqrt{\sum_{n=1}^N \left| \sum_{k=1}^K e^{i\frac{m_k\pi}{2}} s_{n,k} u_k \right|^2} = \sqrt{\sum_{n=1}^N \left| \sum_{k=1}^K s_{n,k} u_k^* \right|^2} \\ &= F(S, \mathbf{u}^*) \end{aligned}$$

where  $u_k^* = e^{i\frac{m_k\pi}{2}} u_k \in \Delta\mathcal{X}$ , holds for any  $\mathbf{u} \in \Delta\mathcal{X}^K$ .  $\square$

*Lemma 4 (Add a User or Resource):* For a give signature matrix  $S$ , it holds that

$$d_{\min}(\bar{S}_c) \leq d_{\min}(S) \leq d_{\min}(\bar{S}_r)$$

where  $\bar{S}_r$  and  $\bar{S}_c$  are signature matrices obtained by adding a row (resource) and column (user) to  $S$ , respectively.  $\square$

Similarly, we can obtain an opposite proposition of Lemma 4 by deleting a user or resource.

*Corollary 1:* For a give signature matrix  $S$ ,

$$d_{\min}(S) \leq \sqrt{2w} \quad (10)$$

where  $w$  is the minimum effective spreading length.  $\square$

*Proof:* Equation (10) is obtained from the fact that  $\sqrt{2w}$  is the minimum distance achieved by the matrix obtained by deleting all the columns of  $S$  except the one with the minimum effective spreading length.  $\square$

*Lemma 5 (Concatenation of Signature Matrices):*

$$d_{\min}(S) \geq \sqrt{\sum_{j=1}^n d_{\min}(S_j)^2}$$

where  $S = [S_1^T, S_2^T, \dots, S_n^T]^T$  is a concatenation of  $S_j, j = 1, \dots, n$ .  $\square$

#### IV. OPTIMAL SIGNATURE MATRIX

For a given factor graph structure, we design the optimal signature matrix with the maximum minimum SCDMA code distance.

For factor graph  $G$ , we have an infinite number of SCDMA codes by varying its edge labels, i.e., the phases of nonzero elements of the signature matrix. Let  $\mathcal{S}_G$  be the universal set that includes all the possible signature matrices, edge labels, associated with  $G$ . We aim to find the SCDMA code with the maximum minimum code distance.

Since a disconnected factor graph can be considered as multiple independent SCDMA codes, in the following we only consider connected factor graphs.

*Definition 4:*

$$S^{\text{opt}} = \arg \max_{S \in \mathcal{S}_G} d_{\min}(S)$$

is called an optimal signature matrix of  $G$ .  $\square$

For a factor graph, there are an infinite number of optimal signature matrices due to Lemmas 2 and 3.

For a given factor graph  $G$ , finding  $S^{\text{opt}}$  in  $\mathcal{S}_G$  is a non-convex problem with high complexity. However, if  $G$  is cycle-free, the problem can be simplified. We first give the following theorem.

*Theorem 1:* If  $G$  is a cycle-free factor graph, for each signature matrix  $S \in \mathcal{S}_G$ , there exists a matrix  $S^* \in \mathcal{S}_G^* = \{S \mid s_{n,k} = e^{i\theta_k}, \theta_1 = 0, \theta_2, \dots, \theta_K \in [0, \frac{\pi}{2}), \text{ for } e_{n,k} \in E\}$  with  $A(S^*, Z) = A(S, Z)$ .  $\square$

*Proof:* We first prove that for each  $S \in \mathcal{S}_G$  there exists  $S' \in \mathcal{S}'_G = \{S \mid s_{n,k} = e^{i\theta_k}, \theta_1 = 0, \theta_2, \dots, \theta_K \in (-\infty, \infty), \text{ for } e_{n,k} \in E\}$  with  $A(S', Z) = A(S, Z)$ . Using the row rotation invariance property of Lemma 2, we just need to show that for a given  $S \in \mathcal{S}_G$ , there exists an  $S' \in \mathcal{S}'_G$  which is a row rotation of  $S$ . Assume that  $S$  is given. We determine  $S'$  as follows. Since zero elements in  $S'$  are predetermined by the factor graph  $G$ , we only determine the nonzero elements in  $S'$  in the following steps:

i. For each  $n \in \{n \mid e_{n,1} \in E\}$ , the  $n$ -th row of  $S'$  is a rotation of the  $n$ -th row of  $S$ , i.e.,  $s'_{n,k} = s_{n,k}/s_{n,1}$  for  $e_{n,k} \in E$ .

ii. Find a column of  $S'$  that has only one determined nonzero element and at least one undetermined element. Assume that the  $m$ -th column is found and the only determined nonzero element is  $s'_{j,m}$ . For each  $n \in \{n \mid e_{n,m} \in E, n \neq j\}$ , the  $n$ -th row of  $S'$  is a rotation of the  $n$ -th row of  $S$ , i.e.,  $s'_{n,k} = s'_{j,m} s_{n,k}/s_{n,m}$  for  $e_{n,k} \in E$ .

iii. If all elements of  $S'$  are determined, terminate the procedure, otherwise, repeat step ii.

We show that Step ii can always be successfully carried out if there exist undetermined elements in  $S'$ . First, if there exists an undetermined element in  $S'$ , we can always find a column with both determined and undetermined nonzero elements since  $G$  is connected and the nonzero elements are determined in a row-by-row manner according to the procedure. Moreover, if a column has both determined and undetermined nonzero elements, the number of determined nonzero element must be one. Suppose that there are more

than one determined nonzero elements in column  $m$  in Step ii, i.e., elements  $s'_{j,m} \neq 0$  and  $s'_{j',m} \neq 0, j' \neq j$ , are determined. Since the labeling begins from the edges incident with data node  $x_1$  (in Step i), there should exist two paths from data node  $x_1$  to both code nodes  $c_j$  and  $c_{j'}$ , i.e., there exists a path between  $c_j$  and  $c_{j'}$  that goes through  $x_1$ . Since there exists another path of  $c_j \rightarrow x_m \rightarrow c_{j'}$  between  $c_j$  and  $c_{j'}$ , which results in a cycle in  $G$ . This conflicts with the fact that  $G$  is cycle-free. Therefore, Step ii can always be successful.

It also holds that  $S' \in \mathcal{S}_G^*$ , since Step i guarantees that the nonzero elements in the first column are 1, and Step ii guarantees that nonzero elements in each column are the same. Note that each column except the first with more than one nonzero elements will be found in Step ii, otherwise, we can show that a cycle exists in the graph similarly. Since  $S'$  is a row rotation of  $S$ ,  $A(S', Z) = A(S, Z)$ .

Directly applying the column rotation invariance property of Lemma 3, we can get a matrix  $S^* \in \mathcal{S}_G^*$  through column rotations from  $S'$  with  $A(S^*, Z) = A(S', Z) = A(S, Z)$ . Thus, the theorem is proved.  $\square$

*Corollary 2:* If  $G$  is a cycle-free factor graph, there exists an optimal signature matrix with  $S^{\text{opt}} \in \mathcal{S}_G^*$ .  $\square$

We consider a special tree factor graph with  $K$  data nodes and one code node with load  $K$  in which the signature matrix becomes a vector. This is the scenario considered in TCMA [18], [19]. We first consider the simplest case of  $K = 2$ .

*Theorem 2:* For the SCMA factor graph with user number  $K = 2$  and resource number  $N = 1$ ,  $S^{\text{opt}} = [1, e^{-i\pi/6}]$  is an optimal signature matrix with  $d_{\min}(S^{\text{opt}}) = \sqrt{3} - 1$ .  $\square$

*Proof:* See Appendix A.  $\square$

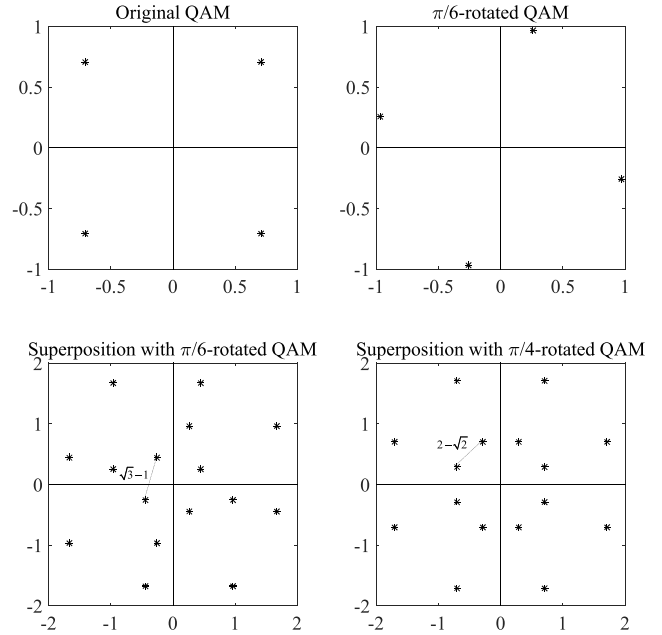
Using (6), the distance enumerator function of the two-user SCMA code with the optimal signature matrix is calculated as:  $2Z\sqrt{3}-1 + \frac{1}{4}Z\sqrt{6}-\sqrt{2} + 5Z\sqrt{2} + \frac{9}{4}Z^2 + Z\sqrt{8-2\sqrt{3}} + Z\sqrt{6} + 2Z\sqrt{3}+1 + Z\sqrt{8+2\sqrt{3}} + \frac{1}{4}Z^2\sqrt{3} + \frac{1}{4}Z^2\sqrt{2+\sqrt{3}}$ , which could be used to estimate the WER performance based on the union bound (7).

Signature matrix  $[1, e^{i\pi/4}]$  is used in [18] and [19] for a two-user TCMA system. It is suboptimal since it gives a smaller minimum distance of  $2 - \sqrt{2}$ . We illustrate the superimposed constellations for both the cases in Fig. 3.

It is difficult to formulate the optimal signature matrix for  $N = 1, K > 2$ . We obtain the optimal signature matrix and minimum distance, denoted as  $\delta_K$ , for  $K \leq 6$  in TABLE 1 by numerical search. The minimum distance decreases as the number of users  $K$  increases.

In general, a signature matrix of a multi-resource SCMA system with each row an optimal signature vector for a single-resource SCMA system may not be optimal. We only have the following lower bound on the minimum code distance of such a signature matrix.

*Definition 5:* Let  $\alpha \subset \mathfrak{N} \triangleq \{1, 2, \dots, N\}$  be an index subset. Subgraph  $G_{\setminus\alpha}$  is obtained by deleting all the data nodes that are adjacent to the code nodes with index in  $\alpha$  and edges induced from these data nodes.  $\square$



**FIGURE 3.** Constellation diagram of two SCMA code sets for  $K = 2, N = 1$ . One is for the optimal signature matrix  $[1, e^{i\pi/6}]$  and the other is for signature matrix  $[1, e^{i\pi/4}]$  used in [18] and [19].

**TABLE 1.** Optimal signature matrix and minimum distance for  $N = 1, K \leq 6$ .

$K$	Optimal signature vector	$\delta_K$
1	[1]	$\sqrt{2}$
2	$[1, e^{i\pi/6}]$	$\sqrt{3} - 1$
3	$[1, e^{i0.0974\pi}, e^{i0.4026\pi}]$	0.4310
4	$[1, e^{i0.0477\pi}, e^{i0.0947\pi}, e^{i0.1965\pi}]$	0.2086
5	$[1, e^{i0.0851\pi}, e^{i0.1368\pi}, e^{i0.1631\pi}, e^{i0.1894\pi}]$	0.1142
6	$[1, e^{i0.0266\pi}, e^{i0.0664\pi}, e^{i0.1696\pi}, e^{i0.473\pi}, e^{i0.4866\pi}]$	0.0595

*Lemma 6:* Let factor graph  $G$  be code-node regular degree of  $q, q > 1$ . Let  $S \in \mathcal{S}_G$  be a signature matrix whose rows are optimal signature vectors of single-resource SCMA systems. It holds that

$$d_{\min}(S) \geq \min_{\alpha \subset \mathfrak{N}} \sqrt{n_1(G_{\setminus\alpha})\delta_1^2 + n_2(G_{\setminus\alpha})\delta_q^2}$$

where  $n_1(G_{\setminus\alpha})$  and  $n_2(G_{\setminus\alpha})$  are the numbers of code nodes in  $G_{\setminus\alpha}$  with degree one and larger than one, respectively.  $\square$

*Proof:* See Appendix B.  $\square$

Note that Lemma 6 can be extended to the case of irregular code-node degree factor graph with a more complex expression.

*Example 1 (An Optimal  $K$ -User,  $(K - 1)$ -Resource Tree SCMA Code):* We have the following  $K$ -user  $(K - 1)$ -resource tree SCMA code with each row an optimal two-user single-resource signature vector,

$$S_{K-1,K}^{\text{opt}} = \begin{bmatrix} 1 & e^{\pi/6} & & & \\ & e^{\pi/6} & 1 & & \\ & & 1 & e^{\pi/6} & \\ & & & e^{\pi/6} & 1 \\ & & & & \dots \end{bmatrix}$$

This SCDMA code has a load of  $K/(K-1)$ . Using Lemma 6 we have  $d_{\min}(S_{K-1,K}^{\text{opt}}) \geq \min\{\sqrt{K-1}(\sqrt{3}-1), \sqrt{2}\}$ . Also based on the proof of Theorem 2 in Appendix A, we know that  $d_{\min}(S_{K-1,K}^{\text{opt}}) = \min\{\sqrt{K-1}(\sqrt{3}-1), \sqrt{2}\}$  exactly holds. For  $K \geq 5$ , the minimum distance is  $\sqrt{2}$  which achieves the upper bound of Corollary 1. In fact, we can further show that  $S_{K-1,K}^{\text{opt}}$  is also an optimal signature labeling for  $K < 5$ .  $\square$

For a factor graph with cycles, we can find an edge set so that after deleting the edges in the set, the remaining graph is a tree. We simplify its labeling based on the following theorem.

**Theorem 3:** Let  $G$  be a factor graph with cycles and  $\varphi \subset E$  be an edge subset that after deleting the edges  $\varphi$ , the remaining graph is a tree. For each signature matrix  $S \in \mathcal{S}_G$ , there exists a matrix  $S^* \in \mathcal{S}_G^*(\varphi) = \{S | s_{n,k} = e^{i\theta_k}, \theta_1 = 0, \theta_2, \dots, \theta_K \in [0, \frac{\pi}{2}), \text{ for } e_{n,k} \in \bar{\varphi}, \text{ and } s_{n,k} = e^{i\theta_{n,k}}, \theta_{n,k} \in [0, 2\pi), \text{ for } e_{n,k} \in \varphi\}$  with  $A(S^*, Z) = A(S, Z)$ .  $\square$

*Proof:* See Appendix C.  $\square$

**Corollary 3:** For factor graph  $G$  with cycles and  $\varphi$  defined in Theorem 3, there exists an optimal signature matrix with  $S^{\text{opt}} \in \mathcal{S}_G^*(\varphi)$ .  $\square$

**Example 2 (Optimal Labeling for Factor Graph in Fig. 2):** Based on Corollary 3, we can simplify the labeling for the 6-user and 4-resource SCDMA factor graph in Fig. 2 proposed in [7] and [8] as

$$S_{4,6} = \begin{bmatrix} 1 & e^{i\theta_2} & e^{i\theta_3} & 0 & 0 & 0 \\ 1 & 0 & 0 & e^{i\theta_4} & e^{i\theta_5} & 0 \\ 0 & e^{i\theta_2} & 0 & e^{i\theta_{3,4}} & 0 & e^{i\theta_6} \\ 0 & 0 & e^{i\theta_3} & 0 & e^{i\theta_{4,5}} & e^{i\theta_{4,6}} \end{bmatrix} \quad (11)$$

where  $\theta_j \in [0, \pi/2), j = 2, \dots, 6$  and  $\theta_{3,4}, \theta_{4,5}, \theta_{4,6} \in [0, 2\pi)$  since after deleting the edges in set  $\varphi = \{e_{3,4}, e_{4,5}, e_{4,6}\}$ , the factor graph will be a tree. Its load is 1.5. We can find 4 length-6 cycles in its factor graph. Through an exhaustive search based on (11), we obtain the following optimal signature matrix

$$S_{4,6}^{\text{opt}} = \begin{bmatrix} 1 & e^{i0.1431\pi} & e^{i0.2021\pi} & 0 & 0 & 0 \\ 1 & 0 & 0 & e^{i0.3127\pi} & e^{i0.3765\pi} & 0 \\ 0 & e^{i0.1431\pi} & 0 & e^{i0.5736\pi} & 0 & e^{i0.2667\pi} \\ 0 & 0 & e^{i0.2021\pi} & 0 & e^{i0.3935\pi} & e^{i0.3078\pi} \end{bmatrix}$$

which has the minimum distance  $d_{\min}(S_{4,6}^{\text{opt}}) = 1.3726$ . Note that the Latin-rectangular labeling proposed in [7] and [13], which uses the elements  $\{1, e^{i\pi/6}, e^{i\pi/3}\}$  with permutation for each row, only gives the minimum distance of 1.1658.  $\square$

## V. TWO CONSTRUCTIONS OF CODE FAMILIES

In this section, we give two constructions of code families whose factor graph has very few short cycles. We give examples to obtain the optimal signature labeling for these two constructions.

**Construction 1 (A  $Kq$ -User,  $(K-1)q$ -Resource SCDMA Code Family):**

$$S_{(K-1)q, Kq} = \begin{bmatrix} I & \mathbf{v}_1 I & & & & \mathbf{v}_{K-1} P \\ & \mathbf{v}_1 I & \mathbf{v}_2 I & & & \\ & & \mathbf{v}_2 I & \ddots & & \\ & & & \ddots & \mathbf{v}_{K-2} I & \\ & & & & \mathbf{v}_{K-2} I & \mathbf{v}_K I \end{bmatrix}$$

where  $\mathbf{v}_k = [e^{i\theta_1^k} \dots e^{i\theta_q^k}]$ ,  $\theta_j^1, \dots, \theta_j^{K-1} \in [0, \pi/2)$  for  $j = 1, \dots, q$ , and  $\theta_j^K = \theta_j^{K-1} \in [0, \pi/2)$ , for  $j = 1, \dots, q-1$ ,  $\theta_j^K \in [0, 2\pi)$ ,  $I$  is a  $q \times q$  identity matrix, and  $P$  is the following  $q \times q$  permutation matrix

$$P = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}. \quad (12)$$

This SCDMA code family has a load of  $K/(K-1)$ . Vectors  $\mathbf{v}_j, j = 1, \dots, K$ , should be optimized to achieve the maximum minimum distance. It is easy to see that there is only one length- $2(K-1)q$  cycle in its corresponding factor graph. Since if we delete one edge in  $\varphi = \{e_{(K-1)q, Kq}\}$  the graph will be cycle-free, we have used the simplified labeling for the remaining tree graph based on Corollary 3.  $\square$

**Example 3 (An Optimal 6-User, 4-Resource SCDMA Code):** Consider  $K = 3, q = 2$  in Construction 1. The graph will be cycle-free by deleting  $\varphi = \{e_{4,6}\}$ . Through a full search, we obtain the following optimal 6-user, 4-resource SCDMA code:

$$\bar{S}_{4,6}^{\text{opt}} = \begin{bmatrix} 1 & 0 & e^{i\pi/6} & 0 & 0 & e^{i\pi/6} \\ 0 & 1 & 0 & e^{i\pi/6} & e^{i\pi/3} & 0 \\ 0 & 0 & e^{i\pi/6} & 0 & e^{i\pi/3} & 0 \\ 0 & 0 & 0 & e^{i\pi/6} & 0 & -1 \end{bmatrix}$$

with the minimum distance  $d_{\min}(\bar{S}_{4,6}^{\text{opt}}) = 1.2679$ .  $\square$

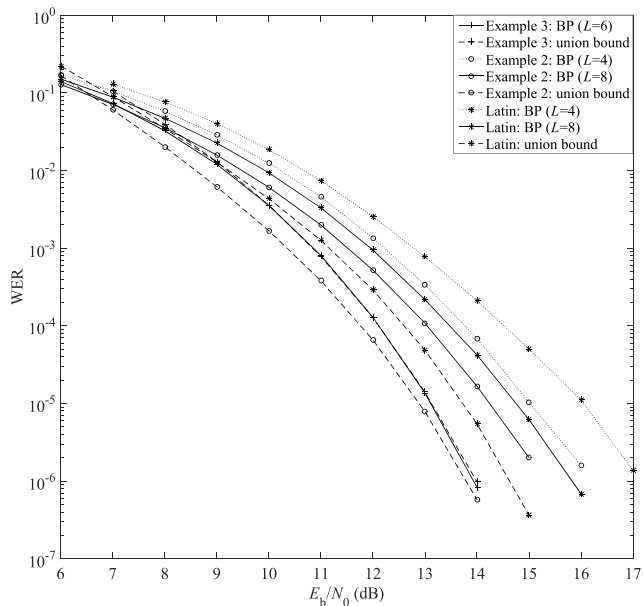
**Remark 1** If we delete the edges corresponding to  $\mathbf{v}_{K-1} P$  in Construction 1, the graph will become a tree, and the maximum minimum code distance will reduce to  $\min\{\sqrt{K-1}(\sqrt{3}-1), \sqrt{2}\}$ , which is achieved by allocating single-resource optimal signature vector to each row as in Example 1. Introducing the part of  $\mathbf{v}_{K-1} P$  in Construction 1 increases the minimum code distance for  $K < 5$ , and thus, improves the performance of ML detection.  $\square$

**Example 4 (An Optimal 8-User, 6-Resource SCDMA Code):** Similarly, by considering  $K = 4, q = 2$  in Construction 1, we obtain the following optimal 8-user, 6-resource SCDMA code:

$$S_{6,8}^{\text{opt}} = \begin{bmatrix} 1 & 0 & e^{i\pi/6} & 0 & 0 & 0 & 0 & e^{i\pi/6} \\ 0 & e^{i\pi/3} & 0 & e^{i\pi/6} & 0 & 0 & e^{i\pi/6} & 0 \\ 0 & 0 & e^{i\pi/6} & 0 & e^{i\pi/3} & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{i\pi/6} & 0 & e^{i\pi/3} & 0 & 0 \\ 0 & 0 & 0 & 0 & e^{i\pi/3} & 0 & e^{i\pi/6} & 0 \\ 0 & 0 & 0 & 0 & 0 & e^{i\pi/3} & 0 & e^{i\pi/6} \end{bmatrix}.$$





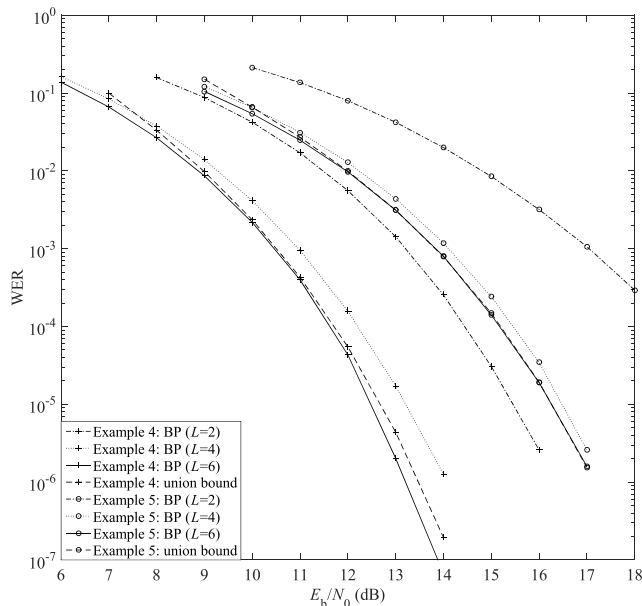


**FIGURE 5.** WER of 6-user 4-resource SCDMA codes (the optimal codes obtained in Examples 2, 3, and the code with Latin-rectangular labeling proposed in [7] and [13]) under BP detections with  $L = 4, 6, 8$  iterations and their union bounds.

Latin-rectangular labeling. However, both of their WER under BP detection with 8 iterations are more than 1 dB worse than their union bounds due to too many short cycles in their factor graph. Although the code designed in Example 3 of Construction 1 has a slightly worse union bound than that of the code obtained in Example 2, its BP detection with 6 iterations converges to its union bound, which means that its BP detection may converge to its ML detection. It has a performance gain of more than 1 dB under BP detection over the optimal code obtained in Example 2 with even fewer iterations.

Figure 6 illustrates the WER of the optimal 8-user 6-resource and 8-user 4-resource SCDMA codes designed in Examples 4 and 5 under BP detections with  $L = 2, 4, 6$  iterations and their union bounds. With 6 iterations, their BP detections converge to or even exceed their union bounds which means that their BP detections may converge to their ML detections.

Let's consider their BP detection complexities. As mentioned in Section II-B.2, the complexity of an SCDMA decoding is mainly determined by the code node degrees and the iteration numbers. The factor graph of the code in Example 2 has 4 degree-3 code nodes, and its BP detection requires 8 iterations to converge (according to our simulation observations). The factor graphs of the codes designed in Examples 3–5 have 2, 2, and 4 degree-3 code nodes, respectively, and their BP detections only require 6 iterations to converge. Let  $C_{Eg,j}$  denote the detection complexity of the code designed in Example  $j$ . Based on a full consideration of their code node degree profile and iteration number, we can rank their complexities as  $C_{Eg,2} > C_{Eg,5} > C_{Eg,4} > C_{Eg,3}$ .



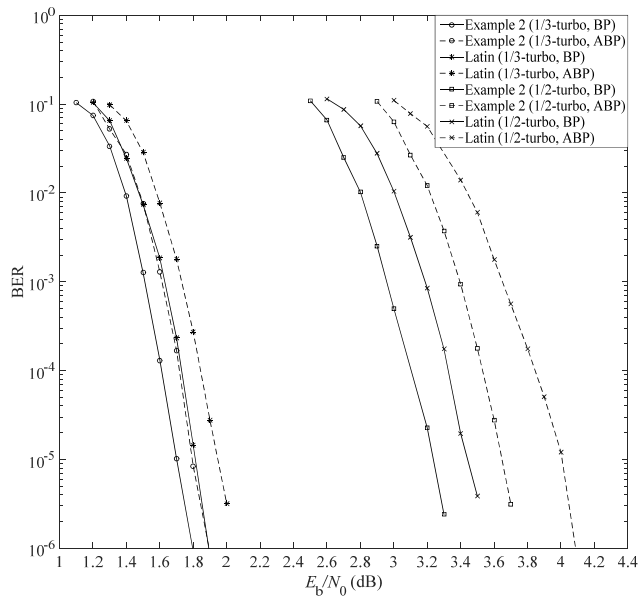
**FIGURE 6.** WER of the optimal 8-user 6-resource and 8-user 4-resource SCDMA codes obtained in Examples 4, 5 under BP detections with  $L = 2, 4, 6$  iterations and their union bounds.

### B. TURBO-CODED SCDMA

In this section, we give bit-error-rate (BER) simulations of turbo-coded SCDMA systems with QAM modulation, where the FEC code in Fig. 1 is realized by a turbo code. Here the turbo code we considered consists two 8-state parallel concatenated convolutional codes with generator matrix  $[1, \frac{1+D+D^3}{1+D^2+D^3}]$ , which is used in 3GPP LTE networks. By puncturing its parity bits, we can obtain different turbo encoding rates: 1/3, 1/2, 2/3, 4/5. For all the simulations, the data stream length for turbo encoding of each user is 1024. For both BP and approximate BP detections, the global decoding iteration (each global iteration includes a turbo decoding iteration and an SCDMA iteration) number is 30, which is enough for all the considered decodings converge to their best performances.

Moreover, codes given by Example 1 and Constructions 1, 2 have irregular effective spreading profiles, i.e., effective spreading lengths for symbols of different users may be different. To realize user fairness, we alternately use column permutations of a signature matrix so that each user's symbol is spread with equal effective spreading length in average sense. For example, the signature matrix given in Example 3 has effective spreading length profile (1, 1, 2, 2, 2, 2) for the six users. In our simulations, we divide the modulated symbol stream within a turbo codeword of each user into three sub-streams with equal length. The first sub-streams of the six users are spread based on signature matrix  $\bar{S}_{4,6}^{opt}$  in Example 3. For the second and third sub-streams we use permuted matrices  $\bar{S}_{4,6}^{opt}P(1, 3)P(2, 4)$  and  $\bar{S}_{4,6}^{opt}P(1, 5)P(2, 6)$ , respectively, where  $P(i, j)$  is a  $6 \times 6$  column permutation matrix that swaps columns  $i$  and  $j$ . The permuted

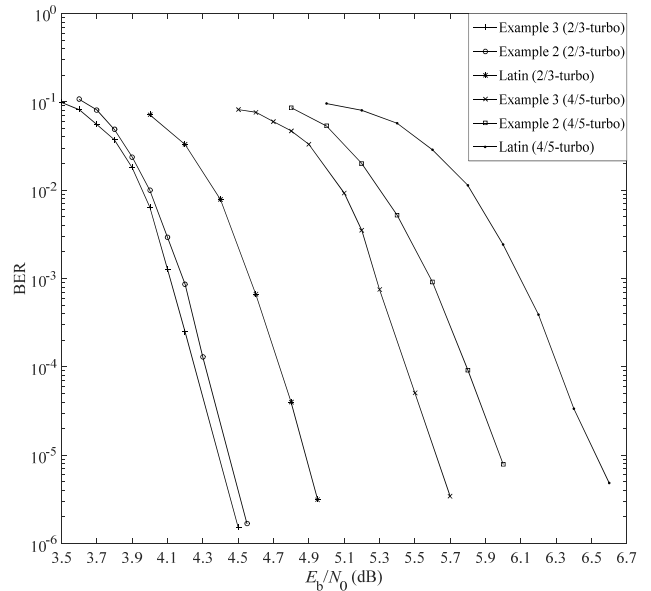
signature matrices have the same distance property with the original matrix but have the effective spreading length profiles (2, 2, 1, 1, 2, 2) and (2, 2, 2, 2, 1, 1), respectively. By doing this, the average effective spreading length for each symbol, which is the same for each user, becomes  $1/3 + 2 \cdot 1/3 + 2 \cdot 1/3 = 5/3$ . Therefore, the detection error rate of each user will also be the same.



**FIGURE 7.** BER of turbo-coded (with turbo coding rates 1/3 and 1/2) 6-user 4-resource SCDMA codes (the optimal code obtained in Example 2 and the code with Latin-rectangular labeling proposed in [7] and [13]) under BP and approximate BP (ABP) detections with 30 iterations.

Figure 7 illustrates BER of rate-1/3 and 1/2 turbo-coded 6-user 4-resource SCDMA systems under BP and approximate BP (ABP) detections, where the optimal SCDMA code obtained in Example 2 and the code with Latin-rectangular labeling proposed in [7] and [13] are considered. The sum communication rates of these two turbo-coded SCDMA systems are  $1/3 \cdot 3/2 \cdot 2 = 1$  bit/resource and  $1/2 \cdot 3/2 \cdot 2 = 3/2$  bit/resource. The rate-1/3 turbo-coded SCDMA system with the optimal SCDMA code designed in Example 2 has a performance gain of about 0.1 dB over the same rate SCDMA system with Latin-rectangular labeling under both BP and ABP decodings. This gain increases if we considered a higher rate turbo code, which works at a higher  $E_b/N_0$  regime, i.e., the gain increases to 0.2 ~ 0.4 dB for the rate-1/2 turbo-coded SCDMA system. Comparing with BP decoding, the performance loss of the ABP is about 0.1 dB for rate-1/3 turbo-coded SCDMA system in the low  $E_b/N_0$  regime since the interference term in (4) is very similar to Gaussian. This performance loss increases to 0.4 ~ 0.55 dB in the high  $E_b/N_0$  regime for the rate-1/2 turbo coded SCDMA system.

Figure 8 illustrates the BER of rate-2/3 and 4/5 turbo-coded SCDMA systems under BP detection, where the optimal codes obtained in Examples 2, 3 and the code with Latin-rectangular labeling are considered. The rate-2/3



**FIGURE 8.** BER of turbo-coded (with turbo coding rates 2/3 and 4/5) SCDMA codes (the optimal codes obtained in Examples 2, 3 and the code with Latin-rectangular labeling) under BP detection with 30 iterations.

turbo-coded SCDMA system with the optimal SCDMA codes designed in Examples 3 has a slightly better BER than the code designed in Examples 2 and has a performance gain of about 0.5 dB over the code with the Latin-rectangular labeling. They have the sum communication rate of 2 bit/resource. For an even higher encoding rate, i.e., a rate-4/5 turbo-coded SCDMA system that works in the higher  $E_b/N_0$  regime, this gain increases and the code in Example 3 has larger performance gains of about 0.4 dB and 1 dB over the code in Example 2 and the code with the Latin-rectangular labeling. In this case, the sum communication rate reaches 12/5 bit per channel use.

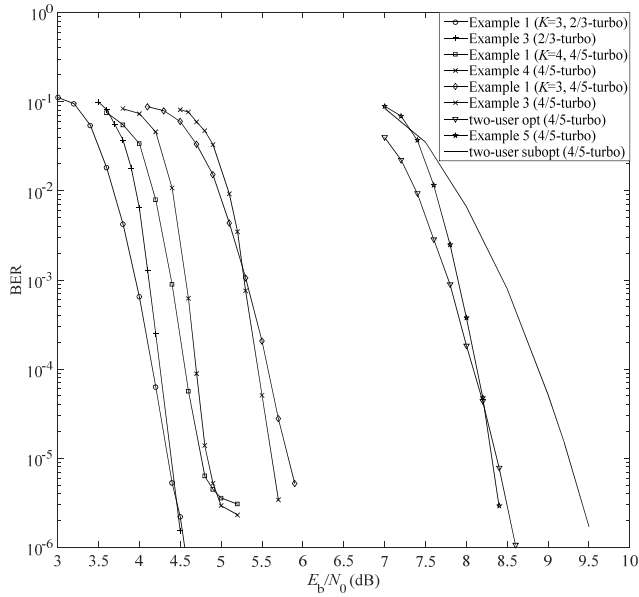
Figure 9 compares four pairs of turbo-coded SCDMA systems:

a) Rate-2/3 turbo-coded SCDMA systems with the ( $K = 3$ )-user 2-resource optimal tree SCDMA code constructed in Example 1 and ( $K = 6$ )-user 4-resource SCDMA code designed in Example 3. Their communication rate is 2 bit/resource.

b) Rate-4/5 turbo-coded SCDMA systems with the ( $K = 4$ )-user 3-resource optimal tree SCDMA code constructed in Example 1 and ( $K = 8$ )-user 6-resource SCDMA code with the optimal SCDMA code designed in Example 4. Their communication rate is 32/15 bit/resource.

c) Rate-4/5 turbo-coded SCDMA systems with the ( $K = 3$ )-user 2-resource optimal tree SCDMA code constructed in Example 1 and ( $K = 6$ )-user 4-resource SCDMA code designed in Example 3. Their communication rate is 12/5 bit/resource.

d) Rate-4/5 turbo-coded SCDMA systems with the two-user single-resource optimal SCDMA code obtained in Theorem 2 and ( $K = 8$ )-user 4-resource SCDMA code designed in Example 5. Their communication rate is 16/5 bit/resource.



**FIGURE 9.** BER of turbo-coded (with turbo coding rates 2/3 and 4/5) SCDMA codes (the optimal codes obtained in Examples 1, 3-5, and two-user single-resource SCDMA with optimal [Theorem 2] and suboptimal [18], [19] labeling) under BP detection with 30 iterations.

Each pair has the same communication rate but the code with more users has a steeper BER curve, better asymptotic BER performance, due to the joint multi-user processing gain. The rate-4/5 turbo-coded two-user single-resource optimal SCDMA code still has a 1 dB performance gain over the same rate turbo-coded suboptimal SCDMA code used in [18] and [19].

## VII. CONCLUDING REMARKS

We proposed a code distance analysis and signature optimization for practical SCDMA systems with a finite number of users and spreading lengths. Good SCDMA codes that work well under both BP and ML detections with low detection complexities are constructed. The constructed codes can support very diverse high-rate services and can be applied to the future 5G NOMA systems.

In our code distance analysis, we only considered uncoded SCDMA systems, i.e., without FEC code, and SCDMA with QAM and equal power for each user. One possible extension is to do distance analysis for coded SCDMA systems, which leads to a joint FEC and SCDMA code design. The new system can be treated as a concatenated code. Some works related to a concatenated code are given in [22]–[24]. Another possible extension is to consider a more general modulation and unequal-power user transmission.

Although we focused on SCDMA systems, our design is also applicable to several similar well-documented system proposals, such as TCMA, low density spreading, and superposition modulation.

## APPENDIX A PROOF OF THEOREM 2

*Proof:* We first prove that there exists an optimal signature matrix in  $[1, e^{i\theta}]$ ,  $\theta \in [0, \pi/4)$ . According to Corollary 2, optimal signature matrix exists in  $[1, e^{i\theta}]$ ,  $\theta \in [0, \pi/2)$ . Moreover, signature matrices  $S = [1, e^{i(\pi/4+\theta)}]$ ,  $\theta \in [0, \pi/4)$  and  $S^* = [1, e^{i(\pi/4-\theta)}]$  give the same distance enumerator function since for any  $\mathbf{u} = [u_1, u_2] \in \Delta\mathcal{X}^2$ ,

$$\begin{aligned} F(S^*, \mathbf{u}) &= |u_1 + e^{i(\pi/4-\theta)}u_2| = |\bar{u}_1 + e^{-i(\pi/4-\theta)}\bar{u}_2| \\ &= |\bar{u}_1 + e^{i(\pi/4+\theta)}\bar{u}_2 e^{-i\pi/2}| = F(S, \mathbf{u}^*) \end{aligned}$$

holds with  $\mathbf{u}^* = [\bar{u}_1, \bar{u}_2 e^{-i\pi/2}] \in \Delta\mathcal{X}^2$ , where  $\bar{u}$  is the complex conjugate of  $u$ .

To continue prove Theorem 2, we simplify the expression of minimum distance as

$$\begin{aligned} \min_{\mathbf{u} \in \Delta\mathcal{X}, \mathbf{u} \neq \mathbf{0}} |u_1 + s_2 u_2| \\ = \min \left\{ \min_{u_2 \in \Delta\mathcal{X}} |\sqrt{2} + s_2 u_2|, \min_{u_2 \in \Delta\mathcal{X}} |\sqrt{2}(1+i) + s_2 u_2| \right\} \end{aligned}$$

due to the following facts:

$$\text{for } u_1 = 0, \min_{u_2 \in \Delta\mathcal{X}, u_2 \neq 0} |u_1 + s_2 u_2| = \sqrt{2} \geq \min_{u_2 \in \Delta\mathcal{X}} |\sqrt{2} + s_2 u_2|,$$

$$\text{for } u_1 = \pm\sqrt{2}, \pm\sqrt{2}i, \min_{u_2 \in \Delta\mathcal{X}} |u_1 + s_2 u_2| = \min_{u_2 \in \Delta\mathcal{X}} |\sqrt{2} + s_2 u_2|,$$

$$\text{for } u_1 = \pm\sqrt{2}(1+i), \pm\sqrt{2}(1-i),$$

$$\min_{u_2 \in \Delta\mathcal{X}} |u_1 + s_2 u_2| = \min_{u_2 \in \Delta\mathcal{X}} |\sqrt{2}(1+i) + s_2 u_2|.$$

Since for any  $s_2 = e^{i\theta}$ ,  $\theta \in [0, \pi/4)$ , the following holds

$$\begin{aligned} \min_{u_2 \in \Delta\mathcal{X}} |\sqrt{2} + s_2 u_2| \\ = \sqrt{2} \min \{|1 - s_2|, |1 + (-1+i)s_2|\}, \\ \min_{u_2 \in \Delta\mathcal{X}} |\sqrt{2} + \sqrt{2}i + s_2 u_2| \\ = \sqrt{2} \min \{|1 + i - s_2|, |1 + i - (1+i)s_2|\}, \\ |1 - s_2| \leq |1 + i - (1+i)s_2|, \\ |1 + (-1+i)s_2| \\ = | -s_2(1 + (-1+i)s_2) | = |1 + i - s_2| \end{aligned}$$

we obtain the final expression of minimum distance as

$$d_{\min}(S) = \sqrt{2} \min \{|1 - s_2|, |1 + i - s_2|\}.$$

Since for  $s_2 = e^{i\theta}$ ,  $\theta \in [0, \pi/4)$ ,  $|1 - s_2|$  increases as  $\theta$  increases, and  $|1 + i - s_2|$  decreases as  $\theta$  increases, the optimal  $s_2$  should satisfy  $|1 - s_2| = |1 + i - s_2|$ , which leads to  $s_2 = e^{i\pi/6}$ , i.e.,  $S^{\text{opt}} = [1, e^{i\pi/6}]$  with  $d_{\min}(S^{\text{opt}}) = \sqrt{3} - 1$ . The theorem is proved.  $\square$

## APPENDIX B PROOF OF LEMMA 6

*Proof:* Let  $\beta \subseteq \mathcal{K} \triangleq \{1, \dots, K\}$  be an index subset. Let  $\mathcal{U}(\beta) = \{\mathbf{u} \in \Delta\mathcal{X}^K, u_k \neq 0 \text{ for } k \in \beta, u_k = 0 \text{ for } k \in \bar{\beta}\}$ ,

where  $\bar{\beta}$  is the complementary set of  $\beta$ . From Lemma 1,

$$d_{\min}(S) = \min_{\beta \subseteq \mathcal{K}, \beta \neq \emptyset} \min_{\mathbf{u} \in \mathcal{U}(\beta)} F(S, \mathbf{u})$$

$$= \min_{\beta \subseteq \mathcal{K}, \beta \neq \emptyset} \min_{\mathbf{u} \in \mathcal{U}(\beta)} \sqrt{\sum_{n \in \mathcal{Y}} \left| \sum_{k \in \beta} s_{n,k} u_k \right|^2}$$

where  $\mathcal{Y} = \{n | n \in \mathfrak{N}, s_{k,n} \neq 0 \text{ for some } k \in \beta\}$ . Since each row of  $S$  is a length- $q$  optimal signature vector for a single-resource SCDMA system, for a given  $n \in \mathcal{Y}$ ,

$$\left| \sum_{k \in \beta} s_{n,k} u_k \right| \geq 1_{\{|\beta \cap \{k | s_{n,k} \neq 0\}|=1\}} \delta_1 + 1_{\{|\beta \cap \{k | s_{n,k} \neq 0\}|>1\}} \delta_q$$

hold for any  $\mathbf{u} \in \mathcal{U}(\beta)$ , where  $1_{\{E\}} = 1$  if  $E$  holds, otherwise,  $1_{\{E\}} = 0$ . Since  $\delta_1 > \delta_q$ ,  $\left| \sum_{k \in \beta} s_{n,k} u_k \right|$  decreases as  $|\beta \cap \{k | s_{n,k} \neq 0\}|$  increases. Moreover, since  $|\beta \cap \{k | s_{n,k} \neq 0\}| \leq d_n$ , where  $d_n$  is the degree of the  $n$ -th code node in  $G_{\setminus \bar{\mathcal{Y}}}$ , we have

$$\left| \sum_{k \in \beta} s_{n,k} u_k \right| \geq 1_{\{d_n=1\}} \delta_1 + 1_{\{d_n>1\}} \delta_q$$

$$\sqrt{\sum_{n \in \mathcal{Y}} \left| \sum_{k \in \beta} s_{n,k} u_k \right|^2} \geq \sqrt{n_1(G_{\setminus \bar{\mathcal{Y}}}) \delta_1^2 + n_2(G_{\setminus \bar{\mathcal{Y}}}) \delta_q^2}$$

for any  $\mathbf{u} \in \mathcal{U}(\beta)$ . Since by varying  $\beta$ ,  $\bar{\mathcal{Y}}$  can be any proper subset of  $\mathfrak{N}$ , by denoting  $\alpha = \bar{\mathcal{Y}}$  we obtain

$$\min_{\beta \subseteq \mathcal{K}, \beta \neq \emptyset} \min_{\mathbf{u} \in \mathcal{U}(\beta)} \sqrt{\sum_{n \in \mathcal{Y}} \left| \sum_{k \in \beta} s_{n,k} u_k \right|^2}$$

$$\geq \min_{\alpha \subset \mathfrak{N}} \sqrt{n_1(G_{\setminus \alpha}) \delta_1^2 + n_2(G_{\setminus \alpha}) \delta_q^2}.$$

The lemma is proved.  $\square$

### APPENDIX C PROOF OF THEOREM 3

*Proof:* We first prove that for each  $S \in \mathcal{S}_G$  there exists  $S' \in \mathcal{S}'_G(\varphi) = \{S | s_{n,k} = e^{i\theta_k}, \theta_1 = 0, \theta_2, \dots, \theta_K \in (-\infty, \infty), \text{ for } e_{n,k} \in \bar{\varphi}, \text{ and } s_{n,k} = e^{i\theta_{n,k}}, \theta_{n,k} \in (-\infty, \infty), \text{ for } e_{n,k} \in \varphi\}$  with  $A(S', Z) = A(S, Z)$ . We just need to show that for a given  $S \in \mathcal{S}_G$ , there exists an  $S' \in \mathcal{S}'_G$  which is a row rotation of  $S$ . Assume that  $S$  is given. We determine  $S'$  as follows. Since the zero elements in  $S'$  are predetermined by the factor graph  $G$ , we only determine the nonzero elements in  $S'$ . The procedure is similar as that in the proof of Theorem 1 except some modifications.

i. For each  $n \in \{n | e_{n,1} \in \bar{\varphi}\}$ , the  $n$ -th row of  $S'$  is a rotation of the  $n$ -th row of  $S$ , i.e.,  $s'_{n,k} = s_{n,k}/s_{n,1}$  for  $e_{n,k} \in E$ .

ii. Find a column of  $S'$  that satisfies the following two conditions. a) The column has a determined labeling for an edge in  $\bar{\varphi}$ . b) The column has at least one undetermined labeling for an edge in  $\varphi$ . If the  $m$ -th column is found and

$s'_{j,m}$  is the determined labeling for  $e_{j,m} \in \bar{\varphi}$ , For each  $n \in \{n | e_{n,m} \in \bar{\varphi}, n \neq j\}$ , the  $n$ -th row of  $S'$  is a rotation of the  $n$ -th row of  $S$ , i.e.,  $s'_{n,k} = s'_{j,m} s_{n,k}/s_{n,m}$  for  $e_{n,k} \in E$ .

iii. If all the labelings for the edges in  $\bar{\varphi}$  are determined, terminate the procedure, otherwise, repeat step ii.

Note that if there are still undetermined labelings for edges in  $\varphi$  at the end of the procedure, we simply use the same labeling as in  $S$ . Moreover, the above procedure only applies to the case that the remaining graph after deleting edges in  $\varphi$  is connected, otherwise, if it contain multiple trees, we can label each of them independently in a similar way.

Step ii can always be successful since the remaining graph with edges in  $\bar{\varphi}$  is a tree. Step ii guarantees that labeling in each column for the edges in  $\bar{\varphi}$  are the same.

Applying the column rotation invariance property of Lemma 3, we can get a matrix  $S^* \in \mathcal{S}_G^*(\varphi) = \{S | s_{n,k} = e^{i\theta_k}, \theta_1 = 0, \theta_2, \dots, \theta_K \in [0, \frac{\pi}{2}), \text{ for } e_{n,k} \in \bar{\varphi}, \text{ and } s_{n,k} = e^{i\theta'_{n,k}}, \theta'_{n,k} \in [0, 2\pi), \text{ for } e_{n,k} \in \varphi\}$  through column rotations from  $S'$  with  $A(S^*, Z) = A(S', Z) = A(S, Z)$ . Thus, the theorem is proved.  $\square$

### REFERENCES

- [1] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE VTC-Spring*, Jun. 2013, pp. 1–5.
- [2] Z. Ding, X. Lei, G. Karagiannidis, R. Schober, J. Yuan, and V. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017, doi: 10.1109/JSAC.2017.2725519.
- [3] Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, "Non-orthogonal multiple access in large-scale underlay cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10152–10157, Dec. 2016.
- [4] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 65, no. 6, pp. 6010–6023, Aug. 2016.
- [5] Z. Ding et al., "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.
- [6] P. D. Diamantoulakis, K. N. Pappi, Z. Ding, and G. K. Karagiannidis, "Wireless powered communications with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8422–8436, Dec. 2016.
- [7] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE PIMRC*, Sep. 2013, pp. 332–336.
- [8] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA codebook design," in *Proc. IEEE VTC Fall*, Sep. 2014, pp. 1–5.
- [9] H. Nikopour et al., "SCMA for downlink multiple access of 5G wireless networks," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 3940–3945.
- [10] M. Yoshida and T. Tanaka, "Analysis of sparsely-spread CDMA via statistical mechanics," in *Proc. IEEE ISIT*, Jul. 2006, pp. 2378–23826.
- [11] D. Guo and C.-C. Wang, "Multiuser detection of sparsely spread CDMA," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 421–431, Apr. 2008.
- [12] R. Hoshyari, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616–1626, Apr. 2008.
- [13] Y. Zhou, Q. Yu, W. Meng, and C. Li, "SCMA codebook design based on constellation rotation," in *Proc. IEEE ICC*, May 2017, pp. 1–6.
- [14] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
- [15] M. Rupf and J. L. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1261–1266, Jul. 1994.
- [16] S. Ulukus and R. D. Yates, "Iterative construction of optimum signature sequence sets in synchronous CDMA systems," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1989–1998, Jul. 2001.

- [17] K. Alishahi, S. Dashmiz, P. Pad, and F. Marvasti, "Design of signature sequences for overloaded CDMA and bounds on the sum capacity with arbitrary symbol alphabets," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1441–1469, Mar. 2012.
- [18] T. Aulin and R. Espineira, "Trellis coded multiple access (TCMA)," in *Proc. IEEE ICC*, May 1999, pp. 1177–1181.
- [19] F. Brannstrom, T. M. Aulin, and L. K. Rasmussen, "Iterative detectors for trellis-code multiple-access," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1478–1485, Sep. 2002.
- [20] J. Harshan and B. S. Rajan, "On two-user Gaussian multiple access channels with finite input constellations," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1299–1327, Mar. 2011.
- [21] W. E. Ryan and S. Lin, *Channel Codes: Classical and Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [22] G. Song and J. Cheng, "Distance enumerator analysis for interleaved-division multi-user codes," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 4039–4053, Jul. 2016.
- [23] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.
- [24] H. Jin and R. J. McEliece, "Coding theorems for turbo code ensembles," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1451–1461, Jun. 2002.



**GUANGHUI SONG** (M'13) received the B.S. degree in communication and engineering from Henan Normal University, Xinxiang, China, in 2006, the M.S. degree in telecommunications engineering from Xidian University, Xi'an, China, in 2009, and the Ph.D. degree from the Department of Intelligent Information Engineering and Sciences, Doshisha University, Kyoto, Japan, in 2012. From 2013 to 2015, he was a Research Assistant with Doshisha University. From 2015 to 2016, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Western Ontario, London, Canada. From 2016 to 2017, he was a Post-Doctoral Researcher with the Singapore University of Technology and Design, Singapore. He is currently a Post-Doctoral Researcher with Doshisha University. His research interests are in the areas of channel coding theory, multiuser coding, and coding for storage systems.



**XIANBIN WANG** (S'98–M'99–SM'06–F'17) received the Ph.D. degree in electrical and computer engineering from the National University of Singapore in 2001. He is currently a Professor and the Canada Research Chair with Western University, Canada.

He was a Research Scientist/Senior Research Scientist with the Communications Research Centre Canada (CRC) between 2002 and 2007. From 2001 to 2002, he was a System Designer with STMicroelectronics, where he was responsible for the system design of DSL and gigabit ethernet chipsets. He has authored or co-authored more than 300 peer-reviewed journals and conference papers, in addition to 26 granted and pending patents and several standard contributions. His current research interests include 5G technologies, Internet-of-Things, communications security, and locationing technologies. Dr. Wang has received many awards and recognition, including the Canada Research Chair, the CRC President's Excellence Award, the Canadian Federal Government Public Service Award, the Ontario Early Researcher Award, and five IEEE Best Paper Awards. He was involved in a number of IEEE conferences, including GLOBECOM, ICC, VTC, PIMRC, WCNC, and CWIT, in different roles such as the Symposium Chair, a Tutorial Instructor, the Track Chair, the Session Chair, and the TPC Co-Chair. He is an IEEE Distinguished Lecturer. He was also an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS between 2007 and 2011 and the IEEE WIRELESS COMMUNICATIONS LETTERS between 2011 and 2016. He currently serves as an Editor/Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON BROADCASTING, and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.



**JUN CHENG** (M'00) received the B.S. and M.S. degrees in telecommunications engineering from Xidian University, Xi'an, China, in 1984 and 1987, respectively, and the Ph.D. degree in electrical engineering from Doshisha University, Kyoto, Japan, in 2000. From 1987 to 1994, he was an Assistant Professor and a Lecturer with the Department of Information Engineering, Xidian University. From 1995 to 1996, he was an Associate Professor with the National Key Laboratory on Integrated Service Network, Xidian University. In 2000, he joined ATR Adaptive Communications Research Laboratories, Kyoto, Japan, where he was a Visiting Researcher. From 2002 to 2003, he was a Staff Engineer with the Research and Development Center, Panasonic Mobile Communications Co., Ltd. (formerly Wireless Solution Labs., Matsushita Communication Industrial Co., Ltd.), Yokosuka, Japan, where he was a Staff Engineer with the Next-Generation Mobile Communications Development Center from 2003 to 2004. In 2004, he joined Doshisha University. He is currently a Professor with the Department of Intelligent Information Engineering and Sciences, Doshisha University. His research interests are in the areas of communications theory, information theory, coding theory, array signal processing, and radio communication systems.

...