IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 18, 2017, accepted April 24, 2017, date of publication June 1, 2017, date of current version March 9, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2707396

Distributed Precoding for BER Minimization With PAPR Constraint in Uplink Massive MIMO Systems

WEI PENG, (Senior Member, IEEE), LU ZHENG, DA CHEN, (Member, IEEE), CHUNXING NI, AND TAO JIANG, (Senior Member, IEEE)

School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China

Corresponding author: Wei Peng (pengwei@hust.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61771214, in part by the National High Technology Development 863 Program of China under Grant 2014AA01A704, in part by the National Science Foundation for Distinguished Young Scholars of China under Grant 61325004, and in part by the Key Project of Hubei Province in China under Grant 2015BAA074.

ABSTRACT Precoding can effectively reduce the peak-to-average power ratio (PAPR) of the transmitted waveform but will usually degrade the bit error rate (BER) performance. In this paper, we investigate the PAPR-guaranteed BER minimization problem through precoding in uplink massive multiple input–multiple output (MIMO) systems. First, we formulate an optimization problem to minimize the BER via precoding design and derive the necessary condition of the optimal precoding matrix. Second, we discuss the BER minimization with PAPR constraint and propose a two-step distributed precoder to deal with this problem. Simulation results verify the effectiveness of the proposed precoding. Particularly, it is efficient for massive MIMO systems in terms of energy efficiency.

INDEX TERMS Precoding, peak-to-average power ratio, massive MIMO, uplink, energy efficiency.

I. INTRODUCTION

Recently, massive multiple input multiple output (MIMO) has attracted much attention from both academic and industrial communities [1]–[4]. In a massive MIMO system, the base station (BS) is equipped with tens or hundreds of antennas. Due to its great potential to improve the energy efficiency (EE) and spectral efficiency (SE), massive MIMO has been considered as a powerful candidate for future wireless communication systems [5]–[7]. To deal with the frequency selectivity of wireless channels, the orthogonal frequency division multiplexing (OFDM) technique is considered as a good waveform modulation scheme for massive MIMO systems [8]. However, as a multi-carrier modulation scheme, the OFDM suffers the high peak-to-average power ratio (PAPR). The high PAPR reduces the resolution of radio-frequency power amplifiers (PA) and decreases the system EE. When massive MIMO systems are considered, low-resolution PAs are generally employed to reduce the total cost. Consequently, the PAPR problem becomes serious in OFDM modulated massive MIMO systems.

Many techniques have been developed for PAPR reduction [9], such as the envelope clipping [10], tone reservation (TR) [11], active constellation extension (ACE) [12], selected mapping (SLM) [13] and partial transmission sequence (PTS) [14], etc. One effective method to reduce the PAPR is the precoding technique. Through precoding, the amplitudes and phases of signals on different sub-carriers can be scaled and rotated, respectively [15]–[18]. In this way, the cross correlation among the transmitted signals can be increased and the PAPR can be reduced. An extreme case is the constant-envelope precoding, by which the PAPR of the precoded waveform equals to one [19]–[23]. However, one main drawback of the precoding based PAPR reduction methods is that the bit error rate (BER) may be increased due to the loss of orthogonality among sub-carriers. An orthogonality-aware precoding is proposed for the single antenna system, by which the designing criterion is that the product of the precoding matrix's Hermitian transpose and itself should be an identity matrix [24]. As an extension of [24] and [25] proposes a precoding matrix with all-one singular values for the single antenna system. To deal with the BER degradation problem associated with PAPR reduction in MIMO systems, several improved precoding methods are proposed for the downlink transmission [26]–[28]. In the downlink, the full channel state information (CSI) is available at the transmitter. Thus, the multi-user interference (MUI) cancelation can be carried out simultaneously with PAPR reduction through precoding to improve the BER. However, in the non-cooperative multi-user uplink transmission, precoding is distributively carried out by each user without knowing the CSI of the other users. As a result, BER minimization has to be distributively considered by each user. To the best of our knowledge, the precoding method for PAPR-guaranteed BER minimization has not yet been reported in uplink MIMO systems.

In this paper, we consider the distributed uplink massive MIMO system and propose a novel precoding method to minimize the BER with guaranteed PAPR performance. Specifically, we first derive a closed-form BER expression with respect to the precoding matrix and then formulate an optimization problem that minimizes the BER in uplink massive MIMO systems. Then, we solve the optimization problem and derive a necessary condition of the precoding matrix. Similar to the single antenna case, the precoding matrix should have all-one singular values. We further consider the BER minimization problem with PAPR constraint and propose a two-step method. For the proposed method, a precoding matrix is selected to satisfy the PAPR constraint at the first step, and then, its singular values are adjusted to satisfy the derived necessary condition at the second step. Simulation results verify the effectiveness of the proposed precoding and particularly show its advantage for massive MIMO systems in terms of the EE performance.

The rest of the paper is organized as follows. The system model of the uplink massive MIMO is described in Section II. Then, the necessary condition of the precoding matrix for BER minimization is derived in Section III. In Section IV, the BER minimization problem with PAPR constraint is addressed and a low-complexity sub-optimal precoding method is proposed. The effectiveness of the proposed precoding is verified by simulation results in Section V. Finally, the paper is concluded in Section VI.

Notations: The superscripts $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ stand for the transpose, conjugate and the Hermitian operations, respectively. vec$\{\cdot\}$ represents the operation that stacks the columns of a matrix into a column vector. Pr$\{\cdot\}$ represents the probability. Re$\{\cdot\}$ stands for the real part of a complex number. tr$\{\cdot\}$ denotes the trace of a matrix. $\|\cdot\|$ stands for the norm of a vector. $E\{\cdot\}$, var$\{\cdot\}$ and max$\{\cdot\}$ denote the expectation, variance and maximum of a sequence, respectively.

II. SYSTEM MODEL

We consider an uplink massive MIMO system with K single-antenna users and an M-antenna base station (BS), as shown in Fig.1. Usually, it is assumed that $M \gg K$, where M takes value from several tens to hundreds. The transceiver of the precoding system is shown in Fig. 2. Since the users are non-cooperative, distributed precoding is carried out by each user to reduce the PAPR respectively. For the k^{th} user, the N-length information sequence $\mathbf{x}_k = [x_{k,0}, x_{k,1}, \cdots, x_{k,N-1}]^T$ with

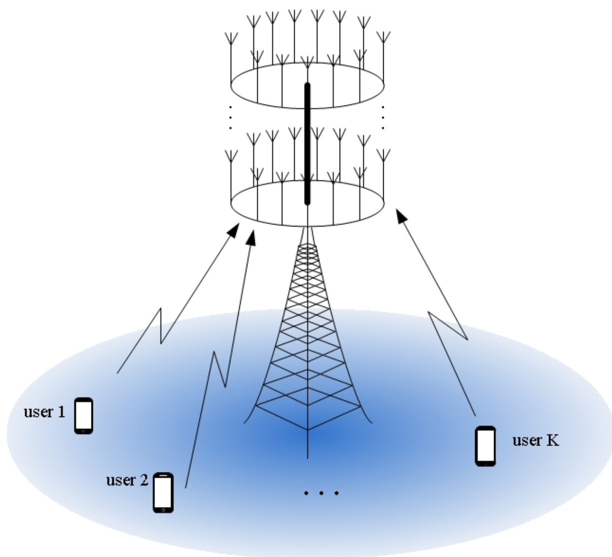

FIGURE 1. The uplink massive MIMO system.

a power of $\sigma_{x_k}^2$ is fed to the precoder, where $\mathbf{x}_k$ is mapped to an L-length sequence $\mathbf{s}_k$ by a precoding matrix $\mathbf{P}_k \triangleq [p_{k,l,n}] \in \mathbb{C}^{L\times N}$. The precoded sequence $\mathbf{s}_k$ is given as

$$\mathbf{s}_k = \mathbf{P}_k\mathbf{x}_k = [s_{k,0}, \cdots, s_{k,L-1}]^T, \tag{1}$$

where

$$s_{k,l} = \sum_{n=0}^{N-1} p_{k,l,n}x_{k,n}. \tag{2}$$

After the waveform modulation, the time-domain OFDM signal for the k^{th} user with a duration of T_s could be given as

$$\begin{aligned}\bar{x}_k(t) &= \text{IFFT}\{\mathbf{s}_k\} \\ &= \sum_{n=0}^{N-1} x_{k,n} \sum_{l=0}^{L-1} p_{k,l,n}e^{j2\pi l\frac{t}{T_s}},\end{aligned} \tag{3}$$

where IFFF$\{\cdot\}$ represents the inverse fast Fourier Transform. The transmitted sequence is then amplified by the PA and emitted by the transmitting antenna. The PAPR of the OFDM signal for the k^{th} user can be defined as

$$\text{PAPR}_k = \frac{\max\limits_{0\le t<T_s}\left\{|\bar{x}_k(t)|^2\right\}}{E\left\{|\bar{x}_k(t)|^2\right\}}. \tag{4}$$

At the receiver, the received signal is down-converted and de-modulated. The equivalent base-band received signal vector on the l^{th} sub-carrier could be given as

$$\mathbf{y}_l = \mathbf{H}_l\bar{\mathbf{s}}_l + \mathbf{z}_l, \tag{5}$$

where $\mathbf{H}_l = [\mathbf{h}_{0,l}, \cdots, \mathbf{h}_{K-1,l}]$ denotes the channel matrix. $\mathbf{h}_{k,l} = [h_{k,0,l}, \cdots, h_{k,M-1,l}]^T$ is the channel vector of the k^{th} user, and $h_{k,m,l}$ denotes the frequency domain channel response between the k^{th} user and the m^{th} antenna. The vector $\bar{\mathbf{s}}_l = [s_{0,l}, \cdots, s_{K-1,l}]^T$ consists of the transmitted signals of K users and $\mathbf{z}_l = [z_{0,l}, \cdots, z_{M-1,l}]^T$ is the vector of the complex additive white Gaussian noise with a power of σ_z^2.

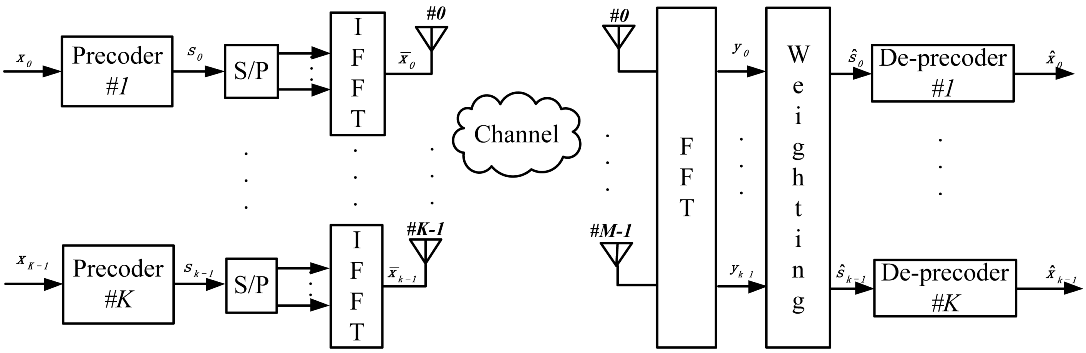

FIGURE 2. The block diagram of the precoded massive MIMO system.

It is assumed that the receiver has the full information of the channel as well as the precoding matrix. For the k^{th} user, the estimate of $s_{k,l}$ can be obtained by a weighted summation, given as

$$\hat{s}_{k,l} = \mathbf{w}_{k,l}^H \mathbf{y}_l, \tag{6}$$

where $\mathbf{w}_{k,l}$ denotes the weighting vector. If the minimum mean square error (MMSE) criterion is employed, the weighting vector could be obtained by

$$\mathbf{w}_{k,l} = (\mathbf{H}_l \mathbf{H}_l^H + \mu^2 \mathbf{I})^{-1} \mathbf{h}_{k,l}, \tag{7}$$

where $\mu^2 = \sigma_z^2/\sigma_{x_k}^2$. The estimate of the information sequence $\mathbf{x}_k$ is then obtained after de-precoding, given as

$$\hat{\mathbf{x}}_k = \mathbf{P}_k^H \hat{\mathbf{s}}_k, \tag{8}$$

where $\hat{\mathbf{x}}_k = [\hat{x}_{k,0}, \cdots, \hat{x}_{k,N-1}]^T$ and $\hat{\mathbf{s}}_k = [\hat{s}_{k,0}, \cdots, \hat{s}_{k,L-1}]^T$. Substituting (6) into (8), we have

$$\hat{\mathbf{x}}_k^T = \sum_{j=0}^{K-1} \sum_{l=0}^{L-1} \mathbf{w}_{k,l}^H \mathbf{h}_{j,l} \mathbf{x}_j \left(\mathbf{p}_{k,l}^H \mathbf{p}_{k,l}\right)^T + \sum_{l=0}^{L-1} \mathbf{w}_{k,l}^H \mathbf{z}_l \mathbf{p}_{k,l}^*, \tag{9}$$

where $\mathbf{p}_{k,l}$ represents the l^{th} row vector of $\mathbf{P}_k$.

III. NECESSARY CONDITION OF PRECODING DESIGN FOR BER MINIMIZATION

In this section, we derive the BER expression and formulate the optimization problem to minimize the BER. We start from the simple case when $K = 1$, then we extend the analysis to the general case when $K \geq 2$.

A. $K = 1$

Substituting (7) into (9), the estimated information sequence could be given as (since $K = 1$, the user index is omitted)

$$\hat{\mathbf{x}} = \mathbf{D} \sum_{m=0}^{M-1} \left(\mathbf{G}_m \mathbf{G}_m^* \mathbf{P} \mathbf{x} + \mathbf{G}_m \mathbf{z}_m\right), \tag{10}$$

where $\mathbf{D} = \mathbf{P}^H \text{inv}\left(\text{diag}[\|\mathbf{h}_0\|^2 + \mu^2, \cdots, \|\mathbf{h}_{L-1}\|^2 + \mu^2]\right)$ and $\mathbf{G}_m = \text{diag}[h_{m,0}^*, \cdots, h_{m,L-1}^*]$. It is clear that the i^{th} bit of $\hat{\mathbf{x}}$ is given as

$$\hat{x}_i = \mathbf{d}_i \sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{G}_m^* \mathbf{P} \mathbf{x} + \mathbf{d}_i \sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{z}_m, \tag{11}$$

where $\mathbf{d}_i$ is the i^{th} row of $\mathbf{D}$. In (11), the term $\mathbf{d}_i \sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{G}_m^* \mathbf{P}\mathbf{x}$ is the useful signal and $\mathbf{d}_i \sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{z}_m$ is the equivalent noise. Note that, $\mathbf{d}_i \sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{z}_m$ is a complex Gaussian random variable. Its expectation and variance can be obtained as

$$E\left\{\mathbf{d}_i \sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{z}_m\right\} = 0, \tag{12}$$

and

$$\text{var}\left\{\mathbf{d}_i \sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{z}_m\right\} = \sigma_i^2, \tag{13}$$

where $\sigma_i^2 \triangleq \sigma_z^2 \mathbf{d}_i \left(\sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{G}_m^H\right) \mathbf{d}_i^H$.

Assuming the constellation size of the modulated signal is C, the signal-to-noise ratio (SNR) of $\hat{x}_i$ could be obtained as

$$\text{SNR}_i = \frac{1}{A\sigma_i^2} \sum_{a=0}^{A-1} \left| \mathbf{d}_i \sum_{m=0}^{M-1} \mathbf{G}_m \mathbf{G}_m^* \mathbf{P} \dot{\mathbf{x}}_{i,a} \right|^2, \tag{14}$$

where $\dot{\mathbf{x}}_{i,a}, a = 0, \cdots, A-1$ belongs to the set $\mathbb{A}$. $\mathbb{A} \triangleq \left\{\mathbf{x}x_i | \mathbf{x} = [x_0, x_1, \cdots, x_{N-1}]^T, x_i \in \mathbb{C}^C\right\}$ is consisted of $A = C^{N-1}$ vectors.

The BER can then be obtained as

$$\text{BER} = \frac{1}{N} \sum_{i=0}^{N-1} f\left(\text{SNR}_i\right), \tag{15}$$

where $f\left(\text{SNR}_i\right) = \text{ber}_i$ is a function of SNR_i determined by the modulation scheme. For example, if the QPSK modulation is employed, $f(\text{SNR}_i) = \frac{1}{\sqrt{2\pi}} \int_{\text{SNR}_i}^{\infty} \exp(-\frac{x^2}{2}) dx$ [29]. To minimize the BER, the optimization problem can be formulated as follows

$$\begin{aligned} \min_{\mathbf{P}} \quad & \frac{1}{N} \sum_{i=0}^{N-1} \text{ber}_i \\ s.t. \quad & \|\text{vec}(\mathbf{P})\|^2 \leq N, \end{aligned} \tag{16}$$

where the constraint $\|\mathrm{vec}(\mathbf{P})\|^2 \leq N$ is the transmitting power limit. Similar to the work in [25], a Lagrange objective function is constructed to solve (16), given as

$$\lambda = \frac{1}{N}\sum_{i=0}^{N-1} \mathrm{ber}_i + \xi \left\|\mathrm{vec}(\mathbf{P})\right\|^2, \tag{17}$$

where ξ is a positive Lagrange multiplier. It can be proved that the optimization problem in (16) is equivalent to minimize (17) for a proper ξ. Therefore, we can obtain the solution by differentiating (17) with respect to $\mathbf{P}$ and set the result to zero, i.e.,

$$\frac{\partial\lambda}{\partial\mathbf{P}^*} = \frac{\partial(\mathbf{s}^H)}{\partial\mathrm{vec}(\mathbf{P}^*)}\cdot\frac{\partial\lambda}{\partial(\mathbf{s}^*)} = \mathbf{0}. \tag{18}$$

Substituting (14) (15) and (17) into (18), the necessary condition of the precoding matrix for BER minimization is derived and given in (19), as shown at the bottom of this page. Note that, the necessary condition in (19) is similar to its counterpart in the single antenna system, which is derived in [25]. Since (19) cannot be directly solved to obtain the optimal precoding matrix, in the following, we try to find a solution to (19).

Proposition 1: *The precoding matrix* $\mathbf{P}$ *with all-one singular values satisfies (19).*

Proof: By the principle of the singular value decomposition (SVD), the precoding matrix can be decomposed as

$$\mathbf{P} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}, \tag{20}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, containing the left and right singular vectors of $\mathbf{P}$, and $\boldsymbol{\Sigma}$ is a diagonal matrix containing the singular values of $\mathbf{P}$. When all the singular values of $\mathbf{P}$ equal to 1, it is easy to see that

$$\mathbf{P}^H\mathbf{P} = \mathbf{I}. \tag{21}$$

Given (21), the term $\dot{\mathbf{x}}_{i,a}\mathbf{d}_i\sum_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*$ in (19) could be rewritten as

$$\dot{\mathbf{x}}_{i,a}\mathbf{d}_i\sum_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^* = \dot{\mathbf{x}}_{i,a}\left[p_{i,0}^*,\cdots,p_{i,L-1}^*\right]. \tag{22}$$

Meanwhile, the term $e^{-\frac{1}{2\sigma_i^2}\left(\mathrm{Re}\left(\mathbf{d}_i\sum_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\mathbf{P}\dot{\mathbf{x}}_{i,a}\right)\right)^2}$ could be rewritten as

$$e^{-\frac{1}{2\sigma_i^2}\left(\mathrm{Re}\left(\mathbf{d}_i\sum\limits_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\mathbf{P}\dot{\mathbf{x}}_{i,a}\right)\right)^2} = e^{-\frac{1}{2\sigma_i^2}}. \tag{23}$$

Substituting (22) and (23) into (19), we have

$$\begin{aligned}
&\sum_{i=0}^{N-1}\sum_{a=0}^{A-1}\frac{1}{\sigma_i}e^{-\frac{1}{2\sigma_i^2}\left(\mathrm{Re}\left(\mathbf{d}_i\sum_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\mathbf{P}\dot{\mathbf{x}}_{i,a}\right)\right)^2} \\
&\quad\times\mathrm{vec}\left(\left(\dot{\mathbf{x}}_{i,a}\mathbf{d}_i\sum_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\right)^H\right) \\
&= \frac{1}{\sigma_i}e^{-\frac{1}{2\sigma_i^2}}\sum_{a=0}^{A-1}\Big[\mathrm{vec}\left(\left(\dot{\mathbf{x}}_{0,a}\left[p_{0,0}^*,\cdots,p_{0,L-1}^*\right]\right)^H\right) \\
&\quad+\mathrm{vec}\left(\left(\dot{\mathbf{x}}_{1,a}\left[p_{1,0}^*,\cdots,p_{1,L-1}^*\right]\right)^H\right) \\
&\quad+\cdots+\mathrm{vec}\left(\left(\dot{\mathbf{x}}_{N-1,a}\left[p_{N-1,0}^*,\cdots,p_{N-1,L-1}^*\right]\right)^H\right)\Big].
\end{aligned} \tag{24}$$

Since $\dot{\mathbf{x}}_{i,a}\in\mathbb{A}$, (24) can be rewritten as

$$\begin{aligned}
&\sum_{i=0}^{N-1}\sum_{a=0}^{A-1}\frac{1}{\sigma_i}e^{-\frac{1}{2\sigma_i^2}\left(\mathrm{Re}\left(\mathbf{d}_i\sum_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\mathbf{P}\dot{\mathbf{x}}_{i,a}\right)\right)^2} \\
&= \frac{A}{\sigma_i}e^{-\frac{1}{2\sigma_i^2}}\left(\mathrm{vec}[\mathbf{p}_0^T,\cdots,\mathbf{0}]+\cdots+\mathrm{vec}[\mathbf{0},\cdots,\mathbf{p}_{L-1}^T]\right) \\
&= \frac{A}{\sigma_i}e^{-\frac{1}{2\sigma_i^2}}\mathrm{vec}(\mathbf{P}).
\end{aligned} \tag{25}$$

Substituting (25) into (19), the denominator of the right hand side is given as $\frac{A}{\sigma_i}e^{-1/2\sigma_i^2}\|\mathrm{vec}(\mathbf{P})\|$. Since $\|\mathrm{vec}(\mathbf{P})\|^2 = N$, we have $\frac{A}{\sigma_i}e^{-1/2\sigma_i^2}\|\mathrm{vec}(\mathbf{P})\| = \frac{A}{\sigma_i}e^{-1/2\sigma_i^2}\sqrt{N}$. Thus, it is proved that both sides of (19) are equal.

The proof of Proposition 1 is completed. Next, we extend the proposition for $K = 1$ to the general case when $K \geq 2$. ■

B. $K \geq 2$

Proposition 2: *The precoding matrix* $\mathbf{P}$ *with all-one singular values can minimize the BER in the multi-user systems.*

Proof: According to (9), the i^{th} bit of the estimated information sequence $\hat{\mathbf{x}}_k^T$ can be obtained as

$$\begin{aligned}
\hat{x}_{k,i} &= \sum_{l=0}^{L-1}\mathbf{w}_{k,l}^H\mathbf{h}_{k,l}x_{k,i}(\bar{\mathbf{p}}_{k,i}^*)^H\bar{\mathbf{p}}_{k,i}^* \\
&\quad+\sum_{j=0,j\neq k}^{K-1}\sum_{l=0}^{L-1}\mathbf{w}_{k,l}^H\mathbf{h}_{j,l}x_{j,i}(\bar{\mathbf{p}}_{j,i}^*)^H\bar{\mathbf{p}}_{j,i}^*+\sum_{l=0}^{L-1}\mathbf{w}_{k,l}^H\mathbf{z}_l\bar{\mathbf{p}}_{k,i}^*,
\end{aligned} \tag{26}$$

$$\mathrm{vec}(\mathbf{P}) = \frac{\sqrt{N}\sum\limits_{i=0}^{N-1}\sum\limits_{a=0}^{A-1}\frac{1}{\sigma_i}e^{-\frac{1}{2\sigma_i^2}\left(\mathrm{Re}\left(\mathbf{d}_i\sum\limits_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\mathbf{P}\dot{\mathbf{x}}_{i,a}\right)\right)^2}\times\mathrm{vec}\left(\left(\dot{\mathbf{x}}_{i,a}\mathbf{d}_i\sum\limits_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\right)^H\right)}{\left\|\sum\limits_{i=0}^{N-1}\sum\limits_{a=0}^{A-1}\frac{1}{\sigma_i}e^{-\frac{1}{2\sigma_i^2}\left(\mathrm{Re}\left(\mathbf{d}_i\sum\limits_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\mathbf{P}\dot{\mathbf{x}}_{i,a}\right)\right)^2}\times\mathrm{vec}\left(\left(\dot{\mathbf{x}}_{i,a}\mathbf{d}_i\sum\limits_{m=0}^{M-1}\mathbf{G}_m\mathbf{G}_m^*\right)^H\right)\right\|}. \tag{19}$$

where $\bar{\mathbf{p}}_{k,i} = [p_{0,i}, \cdots, p_{L-1,i}]^T$ is the i^{th} column vector of $\mathbf{P}_k$. It is clear that $\hat{x}_{k,i}$ in (26) consists of two components: the desired signal $\sum_{l=0}^{L-1} \mathbf{w}_{k,l}^H \mathbf{h}_{k,l} x_{k,i} (\bar{\mathbf{p}}_{k,i}^*)^H \bar{\mathbf{p}}_{k,i}^*$, and the interference plus noise $\sum_{j=0,j\neq k}^{K-1} \sum_{l=0}^{L-1} \mathbf{w}_{k,l}^H \mathbf{h}_{j,l} x_{j,i} (\bar{\mathbf{p}}_{j,i}^*)^H \bar{\mathbf{p}}_{j,i}^* + \sum_{l=0}^{L-1} \mathbf{w}_k^H \mathbf{z}_l \bar{\mathbf{p}}_{k,i}^*$. The signal-to-interference-plus-noise-ratio (SINR) can then be calculated as

$$\Gamma_k = \frac{\sum_{l=0}^{L-1} |\mathbf{w}_{k,l}^H \mathbf{h}_{k,l} \mathbf{P}_k^H \mathbf{P}_k|^2 \sigma_{x_k}^2}{\sum_{j=0,j\neq k}^{K-1} \sum_{l=0}^{L-1} |\mathbf{w}_{j,l}^H \mathbf{h}_{j,l} \mathbf{P}_j^H \mathbf{P}_j|^2 \sigma_{x_j}^2 + \sigma_z^2 \sum_{l=0}^{L-1} \mathbf{w}_{k,l}^H \mathbf{P}_k^H \mathbf{P}_k \mathbf{w}_{k,l}}. \tag{27}$$

For a precoding matrix with all-one singular values, we have $\mathbf{P}_k^H \mathbf{P}_k = \mathbf{I}$ and $(\bar{\mathbf{p}}_{k,i}^*)^H \bar{\mathbf{p}}_{k,i}^* = 1$. Equations (26) and (27) could respectively be rewritten as

$$\hat{x}_{k,i} = \sum_{l=0}^{L-1} \mathbf{w}_{k,l}^H \mathbf{h}_{k,l} x_{k,i} + \sum_{l=0}^{L-1} \sum_{j=0,j\neq k}^{K-1} \mathbf{w}_{j,l}^H \mathbf{h}_{j,l} x_{j,i} + \sum_{l=0}^{L-1} \mathbf{w}_{k,l}^H \mathbf{z}_l \bar{\mathbf{p}}_{k,i}^*, \tag{28}$$

and

$$\Gamma_k = \frac{\sum_{l=0}^{L-1} |\mathbf{w}_{k,l}^H \mathbf{h}_{k,l}|^2 \sigma_{x_k}^2}{\sum_{j=0,j\neq k}^{K-1} \sum_{l=0}^{L-1} |\mathbf{w}_{j,l}^H \mathbf{h}_{j,l}|^2 \sigma_{x_j}^2 + \sigma_z^2 \sum_{l=0}^{L-1} \mathbf{w}_{k,l}^H \mathbf{w}_{k,l}}. \tag{29}$$

Note that, no cooperation could be obtained through the distributed precoder. Therefore, the distributed precoding could not improve the BER performance without increasing the transmitting power. In other words, the minimum BER that can be achieved is the BER of the un-precoded system [24]. Next, we will prove that the precoding matrix $\mathbf{P}$ with all-one singular values can achieve the same BER as the un-precoded system.

As we know, the un-precoded system is equivalent to the case when an identity matrix $\mathbf{I}$ is used as the pre-coding matrix. Replacing the precoding matrix $\mathbf{P}_k$ in (27) by $\mathbf{I}$, it is easy to see that the SINR of the un-precoded system has the same value as that in (29). Therefore, the precoded system has the same BER as the un-precoded system.

The proof of Proposition 2 is completed. It is concluded from Propositions 1 and 2 that, the precoding matrix should have all-one singular values to minimize the BER. ■

IV. PRECODING DESIGN FOR BER MINIMIZATION WITH PAPR CONSTRAINT

When the PAPR constraint is considered, the precoding matrix must simultaneously satisfy the necessary condition in (19) and the PAPR requirement. Mathematically, the problem is formulated as

$$\begin{aligned} \min_{\mathbf{P}} \quad & \frac{1}{N} \sum_{i=0}^{N-1} \text{ber}_i \\ s.t. \quad & \|\text{vec}(\mathbf{P})\|^2 \leq N, \\ & \text{PAPR} \leq \gamma_0. \end{aligned} \tag{30}$$

Directly solving (30) requires huge computational complexity. Enlightened by [25], we propose a two-step method to deal with the problem. At the first step, a PAPR-reducing precoding matrix is selected to meet the PAPR requirement. At the second step, the singular values of the selected precoding matrix are adjusted to satisfy (19) for BER minimization.

A. MEETING THE PAPR CONSTRAINT

The PAPR in (4) is generally upper bounded by the sum of pulse shaping functions of N sub-carriers. To reduce the PAPR, pulse shaping functions of N symbols could be designed so that their peak amplitudes do not occur simultaneously within the symbol duration [24]. For the n^{th} symbol of the k^{th} user, a feasible way to generate the pulse shaping functions is given as

$$p_{k,n}(t) = \begin{cases} p_k(t - \frac{n}{N}T_s + T_s), & 0 \leq t < \frac{n}{N}T_s, \\ p_k(t - \frac{n}{N}T_s), & \frac{n}{N}T_s \leq t < T_s, \end{cases} \tag{31}$$

where $p_k(t)$ is the pulse generation function of the k^{th} user. There are several commonly used pulse generation functions such as the square root of the raised cosine function and the trapezoidal function. Their roll-off factors can be adjusted to satisfy the PAPR requirement. A PAPR-reducing precoding matrix can then be generated, denoted by $\mathbf{P}_T$, where its $(l, n)^{\text{th}}$ element $p_{k,l,n}$ is given by

$$p_{k,l,n} = p_{k,n}\left(\frac{l}{L}T_s\right). \tag{32}$$

B. MEETING THE NECESSARY CONDITION FOR BER MINIMIZATION

The singular values of $\mathbf{P}_T$ can be obtained by the SVD of $\mathbf{P}_T$, given as

$$\mathbf{P}_T = \mathbf{U}_{\mathbf{P}_T} \mathbf{\Sigma}_{\mathbf{P}_T} \mathbf{V}_{\mathbf{P}_T}, \tag{33}$$

where $\mathbf{U}_{\mathbf{P}_T}$ is an $L \times L$ unitary matrix, $\mathbf{V}_{\mathbf{P}_T}$ is an $N \times N$ unitary matrix, and $\mathbf{\Sigma}_{\mathbf{P}_T}$ is an $L \times N$ diagonal matrix made up of the singular values of $\mathbf{P}_T$.

To meet the necessary condition for BER minimization, all the singular values of $\mathbf{P}_T$ are set to one. The newly generated precoding matrix could be given as

$$\mathbf{P}_{\text{CT}} = \mathbf{U}_{\mathbf{P}_T} \begin{bmatrix} \mathbf{I}_{N\times N} \\ \mathbf{0}_{(L-N)\times N} \end{bmatrix} \mathbf{V}_{\mathbf{P}_T}. \tag{34}$$

Then, $\mathbf{P}_{\text{CT}}$ will be used as the precoding matrix to minimize the BER with PAPR constraint.

V. ENERGY EFFICIENCY OF THE PRECODED MASSIVE MIMO SYSTEM

Since the EE is an important issue for PAPR-reducing techniques, in this section, we analyze the EE of a massive MIMO system with precoding. The EE is defined as

$$\eta_{\text{EE}} = \frac{\sum_{k=0}^{K-1} R_k}{\sum_{k=0}^{K-1} P_{k,\text{total}}}, \tag{35}$$

where R_k and $P_{k,\text{total}}$ are the achievable data rate and total power consumption of the k^{th} user, respectively. For simplicity, it is assumed that all the users are well balanced so that $R_k = R$ and $P_{k,\text{total}} = P_{\text{total}}$. To obtain η_{EE}, R and P_{total} will be evaluated respectively as follows.

A. EVALUATION OF R

The achievable data rate can be calculated by the Shannon's theory, given as

$$R = B \log_2 (1 + \Gamma), \tag{36}$$

where B is the bandwidth of the transmitted signal and Γ is the SNR.

According to the theory of high-frequency circuits, the output signal of the PA could be formulated as

$$s_{\text{out}} = \alpha(s_{\text{in}}) s_{\text{in}} + n_{\text{PA}}, \tag{37}$$

where s_{in} is the input signal. In the radio frequency circuit, the OFDM modulated sequence $\bar{x}_k(t)$ is designed as the input signal of the PA. Therefore, we have $s_{\text{in}} = \bar{x}_k$ and $\text{var}\{s_{\text{in}}\} = \text{var}\{\bar{x}_k\} = \frac{L}{N}\sigma_{x_k}^2$. The parameter $\alpha(s_{\text{in}})$ is the scaling factor, and n_{PA} is the noise caused by the non-linear amplifying. The scaling factor $\alpha(s_{\text{in}})$ is related to the input signal and the maximum output voltage $u_{\max}$ of the PA. It could be calculated as [30]

$$\alpha(s_{\text{in}}) = 1 - e^{-\frac{u_{\max}^2}{\sigma_{\text{in}}^2}} + \frac{\sqrt{\pi}\, u_{\max}}{2\sigma_{\text{in}}} \text{erfc}\left(\frac{u_{\max}}{\sigma_{\text{in}}}\right), \tag{38}$$

where $\sigma_{\text{in}}^2 = \text{var}\{s_{\text{in}}\} = \frac{L}{N}\sigma_{x_k}^2$, $\text{erfc}(u_{\max}/\sigma_{\text{in}}) = \frac{2}{\sqrt{\pi}}\int_{u_{\max}/\sigma_{\text{in}}}^{\infty} e^{-u^2} du$. The non-linear amplifying noise n_{PA} could be approximated as

$$\sigma_{\text{PA}}^2 \approx \left(1 - e^{-\frac{u_{\max}^2}{\sigma_{\text{in}}^2}} - \alpha^2(\sigma_{\text{in}})\right)\sigma_{\text{in}}^2. \tag{39}$$

Taking the non-linear amplifying noise of the PA and the noise at the receiver into consideration, the equivalent SNR could be calculated as

$$\Gamma = \frac{\alpha^2(\sigma_{\text{in}})\sigma_{x_k}^2}{\sigma_{\text{PA}}^2 + \sigma_i^2}. \tag{40}$$

Then, the achievable data rate R is obtained by substituting (40) into (36).

B. EVALUATION OF P_{total}

Generally speaking, the total power consumption consists of three parts, given as

$$P_{\text{total}} = P_{\text{PA}} + P_{\text{ADC}} + P_{\text{LNA}}, \tag{41}$$

where P_{PA}, P_{ADC} and P_{LNA} denote the power consumption of the PA, the analog-to-digital converter and the low noise amplifier, respectively. Since the PA consumes most of the power consumption, P_{total} could be approximated as

$$P_{\text{total}} \approx \theta P_{\text{PA}}, \tag{42}$$

where $\theta \geq 1$ is a constant. For this reason, we only need to evaluate P_{PA}, which could be obtained as

$$P_{\text{PA}} = \frac{1}{\eta_{\text{PA}}}\left(\alpha^2(\sigma_{\text{in}})\sigma_{\text{in}}^2 + \sigma_{\text{PA}}^2\right), \tag{43}$$

where η_{PA} is the efficiency of the PA. It could be approximated as [31]

$$\eta_{\text{PA}} \approx \frac{1}{8}\pi^{\frac{3}{2}} \text{erfc}\left(\frac{u_{\max}}{\sigma_{\text{in}}}\right) + \left(1 - \frac{\pi}{4}\right)\left(e^{-\frac{u_{\max}^2}{\sigma_{\text{in}}^2}} + \int_{-\infty}^{-\frac{u_{\max}^2}{\sigma_{\text{in}}^2}} \frac{e^u}{u} du\right). \tag{44}$$

Substituting (39), (43) and (44) into (42), the total power consumption is obtained as

$$P_{\text{total}} = \frac{\theta}{\eta_{\text{PA}}}\left(1 - e^{-\frac{u_{\max}^2}{\sigma_{\text{in}}^2}}\right)\sigma_{\text{in}}^2. \tag{45}$$

Next, the EE can be obtained by substituting (45) and (40) into (35).

VI. SIMULATION RESULTS

Simulations are conduced with Monte Carlo method and some representative results are shown in this section. The length of the transmitted sequence is set to be $N = 128$ and BPSK modulation is used. The single-user scenario with $K = 1$ and multi-user scenario with $k \geq 2$ are simulated. Note that the proposed precoding matrix is applicable to massive MIMO systems with large number of antennas. However, to obtain smooth resulting curves, the sampling size of Monte Carlo trails becomes extremely large when the ratio $\frac{M}{K}$ is large. For example, the BER performance of a 1×16 system is around 10^{-8} when SNR $= 6$ dB, which takes 10^9 Monte Carlo trails to make the BER curve smooth. For this reason, a moderate number of antennas is used in the single-user scenario and a comparatively large number of antennas is used in the multi-user scenario. The trapezoidal function [32] is selected as the pulse generation function. Its Fourier transform is given as

$$S_T(f) = \begin{cases} -\frac{t_s^{3/2}}{\beta}\left(f - \frac{1+\beta}{2t_s}\right), & \frac{1-\beta}{2t_s} < f \leq \frac{1+\beta}{2t_s}, \\ \sqrt{t_s}, & 0 < f \leq \frac{1-\beta}{2t_s}, \end{cases} \tag{46}$$

where $t_s = T_s/N$ and $\beta = (L - N)/N$ is the roll-off factor.

A. PAPR PERFORMANCE

The PAPR performance of the proposed precoding is shown in Fig. 3. The PAPR of the system without precoding is also shown as a reference, which is marked as w/o precoding hereafter. For comparison, the PAPR with the PAPR-reducing-only matrix $\mathbf{P}_T$ is also presented. The roll-off factor β is set to be 0.1, 0.3, and 0.7, respectively. It can be observed that, the proposed precoded system can effectively reduce the PAPR. Actually, the proposed precoding matrix has a PAPR

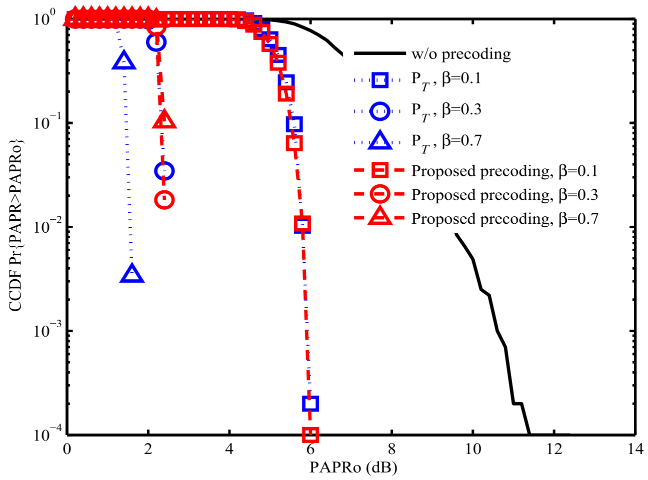

FIGURE 3. **The PAPR of the transmitted waveform.**

performance degradation compared with the PAPR-reducing-only matrix $\mathbf{P}_T$. However, only a slight degradation occurs when a big value of roll-off factor (e.g. when $\beta = 0.7$) is used. For small roll-off factors such as $\beta = 0.1$ and $\beta = 0.3$, the PAPR of the proposed precoding matrix is almost the same as that with $\mathbf{P}_T$. In the practical systems, the roll-off factor takes the value within [0.15, 0.5] and the PAPR performance of the proposed precoding will be acceptable.

B. BER PERFORMANCE

When the BER performance is considered, a sixteen-path channel model is used in simulations. I.e., the channel fading between each user and each antenna consists of sixteen paths, and the channel gain on each path follows the Rayleigh distribution with zero mean and a variance of $1/16$.

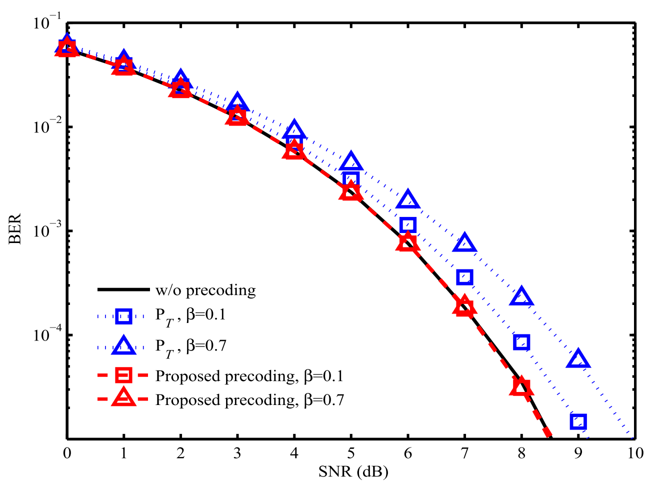

FIGURE 4. **The BER of the single-user scenario, $M = 4$.**

Firstly, the single-user scenario is simulated, where $M = 4$ is used. The BER performance using the proposed precoding is shown in Fig. 4. It is observed that, the proposed precoded system has the same BER as the un-precoded system. In other words, the minimum BER can be achieved. In contrast, the BER performance with the PAPR-reducing-only precoding matrix $\mathbf{P}_T$ is obviously worse than the un-precoded system. The reason lies in the fact that $\mathbf{P}_T$ can not satisfy the necessary condition for BER minimization in (19). It can be further observed that, the BER performance of the proposed precoding is not sensitive to the value of the roll-off factor. That is to say, the proposed precoding can obtain BER minimization as long as the necessary condition in (19) is satisfied. Thus, system designers can select the value of the roll-off factor according to the PAPR requirement without worrying about degrading the BER performance.

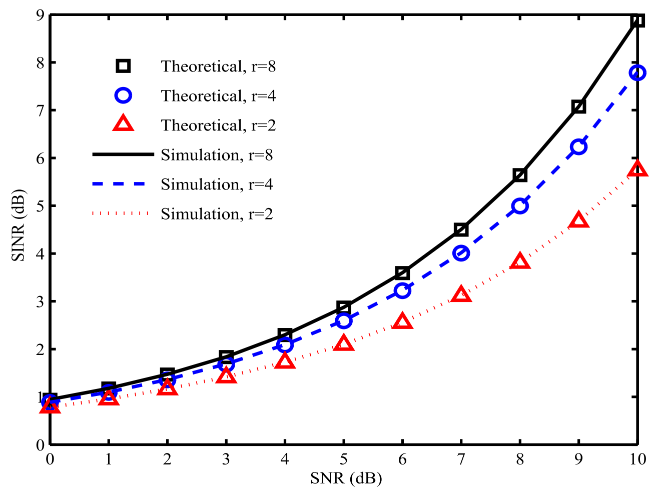

FIGURE 5. **The SINR of the multi-user system, $M = 128$, $r = M/K$.**

Then, the multi-user scenario is considered, where $M = 128$ is used. In order to verify the theoretical analysis for the multi-user scenario, we calculate the SINR in (29) and compare the theoretical and simulation results, as shown in Fig. 5. By varying the values of $r = \frac{M}{K}$, we observe the SINRs of the proposed precoding method. It is shown that the theoretical results well match the simulation results. The BER performance of the precoded multi-user system is then shown in Fig. 6, where the roll-off factor $\beta = 0.5$ is used. It is

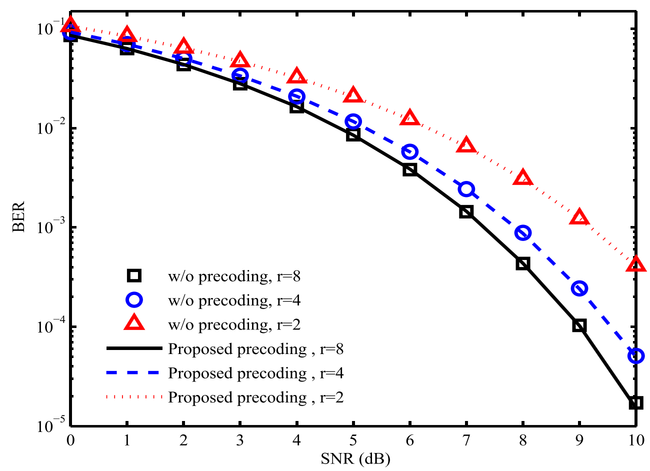

FIGURE 6. **The BER of the multi-user scenario, $M = 128$, $r = M/K$.**

observed that, as expected, the precoded multi-user system has the same BER performance as the un-precoded system. Thus, it is verified that the proposed precoding can obtain the minimum BER.

Discussion: When the number of antennas becomes very large, the numerical calculation of SINR will be computational complex. In order to obtain an efficient performance evaluation, we derive the SINR estimation method when the number of antennas goes infinity as follows.

The weighting vector in (7) can be rewritten as

$$\mathbf{w}_{k,l} = \mathbf{B}_{k,l}^{-1}\mathbf{h}_{k,l} = \frac{\mathbf{B}_{\bar{k},l}^{-1}\mathbf{h}_{k,l}}{1+\mathbf{h}_{k,l}^{H}\mathbf{B}_{\bar{k},l}^{-1}\mathbf{h}_{k,l}}, \tag{47}$$

where $\mathbf{B}_{k,l} = \mathbf{H}_l\mathbf{H}_l^H + \mu^2\mathbf{I}$, and $\mathbf{B}_{\bar{k},l} = \mathbf{B}_{k,l} - \mathbf{h}_{k,l}\mathbf{h}_{k,l}^H$ are used for notation simplicity, recalling that μ^2 was defined as $\mu^2 = \sigma_z^2/\sigma_{x_k}^2$. The SINR in (29) is then given as

$$\Gamma_k = \frac{1}{L}\sum_{l=0}^{L-1}\mathbf{h}_{k,l}^{H}\mathbf{B}_{\bar{k},l}^{-1}\mathbf{h}_{k,l}. \tag{48}$$

When the number of receiving antennas M grows to infinity while the ratio $\frac{M}{K} = r$ is fixed, we have the limit

$$\lim_{M\to\infty}\mathbf{h}_{k,l}^{H}\mathbf{B}_{k,l}^{-1}\mathbf{h}_{k,l} = \frac{1}{M}\mathrm{tr}(\mathbf{B}_{k,l}^{-1}). \tag{49}$$

According to the M-P law [33],

$$\lim_{M\to\infty}\left(\frac{1}{M}\mathrm{tr}\left(\mathbf{B}_{k,l}^{-1}\right)\right) = \int\frac{1}{x+\mu^2}dF_r(x), \tag{50}$$

where $F_r(x)$ is the empirical spectrum distribution of the random matrix $\mathbf{H}_l\mathbf{H}_l^H$. If a unit transmitting power is allocated for each user, the SINR in (29) can be obtained by substituting (49) and (50) in to (48), given as

$$\Gamma_k = \frac{\sqrt{(\mu^2+r+1)^2-4r}-\mu^2-r+1}{2\mu^2}. \tag{51}$$

C. EE PERFORMANCE

The numerical results of the EE are calculated and shown in Fig. 7, where the single-user scenario is used and the number of the receiving antennas is 1, 4, and 16. For simplicity, θ in (42) is approximated by 1. The operating point of the HPA is set by the input back-off (IBO) (dB), which is defined as

$$\mathrm{IBO} = 10\log\frac{u_{\max}^2}{\sigma_{\mathrm{in}}^2}. \tag{52}$$

It can be observed that, in the single antenna system, the EE with the proposed precoding is worse than that without precoding. As the number of antennas increases, the EE of both systems with and without precoding increase. However, the EE of the proposed precoding increases more significantly than that without precoding. Interestingly, when the number of antennas is large, a higher EE can be obtained by the proposed precoding compared with that without precoding. The EE performance in the multi-user scenario is similar to the single-user scenario. It can be affirmed that, an EE gain over the un-precoded system can be obtained when the number of antennas is large. Therefore, the proposed precoding is highly efficient for massive MIMO systems in terms of the EE performance.

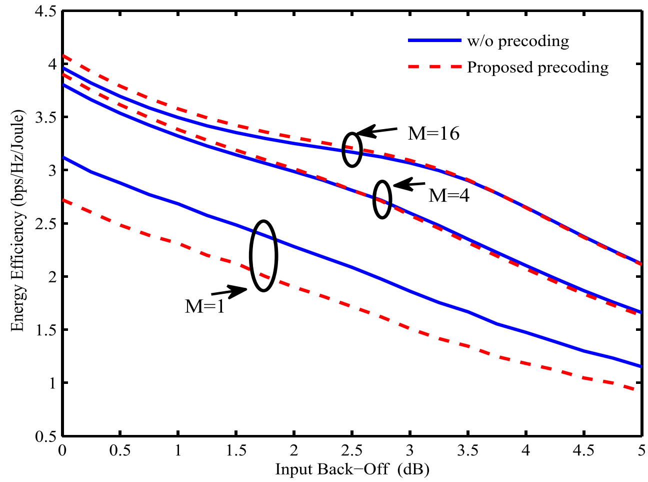

FIGURE 7. The energy efficiency.

According to the simulation results, it can be summarized that: 1) BER minimization with PAPR constraint can be obtained by the proposed precoding in both the single-user and multi-user scenarios. 2) The BER of the proposed precoded system is not sensitive to the roll-off factor. 3) With a large number of antennas, the EE of the proposed precoded system outperforms the un-precoded system considering the EE performance.

VII. CONCLUSIONS

In this paper, the problem of minimizing the BER with PAPR constraint has been considered for uplink massive MIMO systems. By formulating the optimizing problem, the necessary condition of the precoding matrix for BER minimization has been derived. To satisfy the derived necessary condition, it has been proved that a matrix with all-one singular values could make a possible solution. In order to efficiently minimize the BER with PAPR constraint, we proposed a sub-optimal precoding, which takes two steps to generate the precoding matrix: 1) Generating a PAPR-reducing matrix to meet the PAPR requirement. 2) Decomposing the selected precoding matrix and replacing its singular values with all ones. Simulations have been carried out to testify the proposed precoding. It has been shown that, the proposed precoding could obtain BER minimization with PAPR constraint for both the single-user and multi-user scenarios. In addition, the proposed precoded system is highly energy efficient, especially with a large number of antennas.

REFERENCES

[1] J. Zhang, C.-K. Wen, S. Jin, X. Gao, and K.-K. Wong, "On capacity of large-scale MIMO multiple access channels with distributed sets of correlated antennas," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 133–148, Feb. 2013.

[2] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[3] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving 'massive MIMO' spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sep. 2012.

[4] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[6] Y. Huang, S. He, S. Jin, and W. Chen, "Decentralized energy-efficient coordinated beamforming for multicell systems," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4302–4314, Nov. 2014.

[7] L. Zhao, K. Li, K. Zheng, and M. O. Ahmad, "An analysis of the tradeoff between the energy and spectrum efficiencies in an uplink massive MIMO-OFDM system," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 62, no. 3, pp. 291–295, Mar. 2015.

[8] A. Pitarokoilis, E. Björnson, and E. G. Larsson, "Performance of the massive MIMO uplink with OFDM and phase noise," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1595–1598, Aug. 2016.

[9] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 257–268, Jun. 2008.

[10] Y. Wang and A. N. Akansu, "Low-complexity peak-to-average power ratio reduction method for orthogonal frequency-division multiplexing communications," *IET Commun.*, vol. 9, no. 17, pp. 2153–2159, Sep. 2015.

[11] C. Ni, Y. Ma, and T. Jiang, "A novel adaptive tone reservation scheme for PAPR reduction in large-scale multi-user MIMO-OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 5, no. 5, pp. 480–483, Oct. 2016.

[12] S.-H. Wang, W.-L. Lin, B.-R. Huang, and C.-P. Li, "PAPR reduction in OFDM systems using active constellation extension and subcarrier grouping techniques," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2378–2381, Dec. 2016.

[13] K. Bae, C. Shin, and E. J. Powers, "Performance analysis of OFDM systems with selected mapping in the presence of nonlinearity," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2314–2322, May 2013.

[14] L. Li and D. Qu, "Joint decoding of LDPC code and phase factors for OFDM systems with PTS PAPR reduction," *IEEE Trans. Veh. Technol.*, vol. 62, no. 1, pp. 444–449, Jan. 2013.

[15] Z. Sharifian, M. J. Omidi, H. Saeedi-Sourck, and A. Farhang, "Linear precoding for PAPR reduction of GFDMA," *IEEE Wireless Commun. Lett.*, vol. 5, no. 5, pp. 520–523, Oct. 2016.

[16] J. Fang and I.-T. Lu, "Precoder designs for jointly suppressing out-of-band emission and peak-to-average power ratio in an orthogonal frequency division multiplexing system," *IET Commun.*, vol. 8, no. 10, pp. 1705–1713, Jul. 2014.

[17] J. Gao, X. Zhu, and A. K. Nandi, "Non-redundant precoding and PAPR reduction in MIMO OFDM systems with ICA based blind equalization," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3038–3049, Jun. 2009.

[18] D. Qiao, Y. Wu, and D. Chen, "Is precoding for massive MIMO systems well-analyzed?" in *Proc. 11th Int. Symp. Wireless Commun. Syst.*, Aug. 2014, pp. 198–202.

[19] S. Khademi and A. van der Veen, "Constant modulus algorithm for peak-to-average power ratio (PAPR) reduction in MIMO OFDM/A," *IEEE Signal Process. Lett.*, vol. 20, no. 5, pp. 531–534, May 2013.

[20] J. Zhang, Y. Huang, J. Wang, B. Ottersten, and L. Yang, "Per-antenna constant envelope precoding and antenna subset selection: A geometric approach," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6089–6104, Dec. 2016.

[21] J. Pan and W.-K. Ma, "Constant envelope precoding for single-user large-scale MISO channels: Efficient precoding and optimal designs," *IEEE J. Sel. Areas Signal Process.*, vol. 8, no. 5, pp. 982–995, Oct. 2014.

[22] J.-C. Chen, C.-K. Wen, and K.-K. Wong, "Improved constant envelope multiuser precoding for massive MIMO systems," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1311–1314, Aug. 2014.

[23] A. Liu and V. K. N. Lau, "Two-stage constant-envelope precoding for low-cost massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 485–494, Jan. 2016.

[24] S. B. Slimane, "Reducing the peak-to-average power ratio of OFDM signals through precoding," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 686–695, Mar. 2007.

[25] M.-J. Hao and C.-H. Lai, "Precoding for PAPR reduction of OFDM signals with minimum error probability," *IEEE Trans. Broadcast.*, vol. 56, no. 1, pp. 120–128, Mar. 2010.

[26] S. H. Wang, C. P. Li, K. C. Lee, and H. J. Su, "A novel low-complexity precoded OFDM system with reduced PAPR," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1366–1376, Mar. 2015.

[27] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 303–313, Feb. 2013.

[28] S. K. Mohammed and E. G. Larsson, "Single-user beamforming in large-scale MISO systems with per-antenna constant-envelope constraints: The doughnut channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 3992–4005, Nov. 2012.

[29] J. G. Proakis, *Digital Communications*, 3rd ed. New York, NY, USA: McGraw-Hill, 1995.

[30] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IRE Trans. Inf. Theory*, vol. 4, no. 2, pp. 69–72, Jun. 1958.

[31] T. Jiang, C. Li, and C. Ni, "Effect of PAPR reduction on spectrum and energy efficiencies in OFDM systems with class-A HPA over AWGN channel," *IEEE Trans. Broadcast.*, vol. 59, no. 3, pp. 513–519, Sep. 2013.

[32] Z. You, I.-T. Lu, R. Yang, and J. Li, "Flexible companding design for PAPR reduction in OFDM and FBMC systems," in *Proc. Int. Conf. Comput., Netw. Commun.*, 2013, pp. 408–412.

[33] M. L. Mehta, *Random Matrices*, 3rd ed. Amsterdam, The Netherlands: Elsevier, 2004.

WEI PENG received the Ph.D. degree from Hong Kong University. She was an Assistant Professor with Tohoku University, Japan, from 2009 to 2013. She is currently an Associate Professor with the School of Electronics and Communications, Huazhong University of Science and Technology, China. Her research interests include wireless signal processing, network design, and optimization.

LU ZHENG received the master's degree from Huazhong University of Science and Technology, China. He is currently a Senior Engineer in telecommunications with China Mobile Communications Corporation.

DA CHEN received the Ph.D. degree from Huazhong University of Science and Technology, China. He is currently an Assistant Professor with the School of Electronics and Communications, Huazhong University of Science and Technology. His research interests include waveform designing and network optimization.

CHUNXING NI is currently pursuing the Ph.D. degree with Huazhong University of Science and Technology, China. His research interests include PAPR-reducing techniques and waveform designing.

TAO JIANG is currently a Full Professor with the School of Electronics and Communications, Huazhong University of Science and Technology, China. His current research interests include cognitive wireless access, waveform designing, co-operative networks, and wireless sensor networks.