

# Transmit-Power Efficient Linear Precoding Utilizing Known Interference for the Multiantenna Downlink

S. Morteza Razavi, Tharmalingam Ratnarajah, *Senior Member, IEEE*, and  
Christos Masouros, *Senior Member, IEEE*

**Abstract**—It has been shown that the knowledge of both channel and data information at the base station prior to downlink transmission can help increase the received signal-to-noise ratio (SNR) of each user without the need to increase the transmitted power. Achievability is based on the idea of phase alignment (PA) precoding, where instead of nulling out the destructive interference, it judiciously rotates the phases of the transmitted symbols. In this way, they add up coherently at the intended user and yield higher received SNRs. In addition, it is well known that regularized channel inversion (RCI) precoding improves the performance of channel inversion (CI) in multiantenna downlink communications. In line with this and similar to the RCI precoding, in this paper, we propose the idea of regularized PA (RPA), which is shown to improve the performance of original PA precoding. To do this, we first rectify the original PA precoding, deriving a closed-form expression to evaluate the amount of transmit-power reduction achieved for the same average output SNR compared with CI precoding. We then use this new analysis to select the appropriate regularization factor for our proposed RPA scheme. It is shown by means of theoretical analysis and simulations that the proposed RPA precoding outperforms CI, RCI, and PA precoders from both symbol error rate (SER) and throughput perspectives and provides a more power-efficient alternative. This is particularly pronounced as the number of transmit antennas becomes larger, where up to a 50-times reduction in the transmit power is achieved by RPA (PA) compared with RCI (CI) precoding for a given performance.

**Index Terms**—Linear precoding, multiantenna downlink transmission, phase alignment (PA), power-efficient communications.

## I. INTRODUCTION

UTILIZING multiantenna base stations (BSs) is one of the most practical ways to achieve high performance in the wireless downlink transmission. In the downlink scenario, it makes sense to move the signal processing enhancements

Manuscript received May 9, 2013; revised September 4, 2013 and February 7, 2014; accepted April 8, 2014. Date of publication April 16, 2014; date of current version November 6, 2014. The work of S. M. Razavi and T. Ratnarajah was supported by the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission under FET-Open Grant HIATUS-265578. The work of C. Masouros was supported by the Royal Academy of Engineering, U.K. The review of this paper was coordinated by Dr. G. Bauch.

S. M. Razavi was with the Institute for Digital Communications, The University of Edinburgh, Edinburgh EH9 3JL, U.K. (e-mail: Morteza.Razavi@ed.ac.uk).

T. Ratnarajah is with the Institute for Digital Communications, The University of Edinburgh, Edinburgh EH9 3JL, U.K. (e-mail: T.Ratnarajah@ed.ac.uk).

C. Masouros is with the Department of Electrical and Electronic Engineering, University College London, London WC1E 7JE, U.K. (e-mail: C.Masouros@ucl.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2014.2317716

to the transmit side to keep mobile terminals (MTs) simple and low cost. Most of the related literature on this topic has focused on improving the system throughput. For example, in [1], it has been shown that dirty paper coding (DPC) can achieve the downlink capacity. However, since DPC is too complex for the contemporary systems, some less-complex *nonlinear* precoding techniques such as vector perturbation [2], [3] and Tomlinson–Harashima precoding [4], [5] have been proposed.

Although achieving less throughput, *linear* precoders are more practical due to reduced possessing complexity compared with their nonlinear counterparts. The least complex of the available techniques, i.e., channel inversion (CI) [6], is a linear precoding technique that yields reasonable performance in downlink communications. The generalization of CI precoding for multiantenna receivers has been investigated in [7]. In [8], it has been shown that the symbol error rate (SER) performance given by CI precoding becomes worse with the increase in the number of users. Regularized CI (RCI) proposed in [8] attains some performance with respect to the conventional CI in such a way that, with increasing the number of users, the SER performance of each user remains fixed at low SNRs and improves slightly at high SNRs.

In line with this, [9] presents a precoding technique based on phase rotation [hereafter, we call it phase alignment (PA)] for multiantenna downlink communications, where instead of removing the harmful symbol-to-symbol interference, it rotates the phases of the transmitted symbols such that the destructive interference becomes constructive, eventually leading to more received SNRs for fixed transmit power. Further, the superior performance of PA precoding of [9] compared with conventional linear precoders has been investigated in [10] and [11] for cognitive radio networks.

Aside from increasing the throughput in downlink, designing power-efficient precoders has become important in recent years. The idea is to minimize the transmit power while securing the same quality of service for each user. In this paper, we focus on designing such a precoder that enables us to decrease the transmit power to achieve the same average output SNR for each user. Due to their practical complexity, in this paper, we focus on linear precoders, and the contributions of this paper can be summarized as follows.

- 1) We reformulate and enhance the performance analysis of PA precoding [9] so that, in line with the aims of “*green communications*” [12], the power efficiency (as opposed to received SNR) can be optimized. In addition, we complete the performance analysis of PA by analytically

calculating the scaling factors obtained, where in [9], only empirical scaling factors were used for the theoretical results.

- 2) We propose an enhanced PA technique, namely, regularized PA (RPA), where, based on the performance analysis of PA, we derive the required regularization factor for RPA.
- 3) We analytically derive the received SNR of the proposed RPA scheme. We also show that, to achieve the same average output SNR for each user, the transmit power reduction achieved by RPA compared with RCI precoding is the same as that of PA compared with CI precoding.
- 4) We show that the power gains of RPA compared with its counterparts PA, CI, and RCI magnify as the number of transmit antennas increases, which aligns the proposed scheme with the aims of *massive multiple-input multiple-output (MIMO)* systems [13]. In particular, we observe up to more than 50 times saving in the transmit power for RPA (PA) compared with RCI (CI) for systems with up to 100 transmit antennas.
- 5) We also consider the effect of channel estimation errors on the performance of the proposed scheme. We show that, with imperfect transmit-side channel state information (CSIT), the performance trend of the proposed RPA precoding follows the one of the conventional precoders, which further implies that the RPA precoding is as sensitive as the others to erroneous CSIT.

This paper is organized as follows. In Section II, the system model and conventional linear precoding are presented. In Section III, we reformulate and enhance the performance analysis of PA precoding. In Section IV, we propose the RPA precoding. In Section V, the power efficiency of PA and RPA precoding is evaluated. In Section VI, by using numerical simulations, we show that the proposed RPA precoding outperforms CI, RCI, and PA precoding and enables us to save more power at transmit side for a fixed average received SNR at each user. Finally, Section VII contains the conclusion.

## II. SYSTEM MODEL AND CONVENTIONAL LINEAR PRECODING

We consider a multiuser downlink scenario where an  $N$ -antenna transmitter communicating with MTs with  $M$  receive antennas in total. Since no signal processing treatment is going to be considered at each MT, the system configuration is irrelevant to whether the receive antennas cooperate or not. Therefore, the total number of receive antennas can belong to one user or be shared by several users; however, as purely transmitter-based precoders are most useful with single-antenna receivers, we consider single-antenna MTs for the remainder of this paper. We also assume that all the users are homogeneous and experience independent fading [8], [9], [14]. The received signals of all users can be expressed by

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{z} \quad (1)$$

where  $\mathbf{y} \in \mathbb{C}^{M \times 1}$ , and  $\mathbf{H} \in \mathbb{C}^{M \times N}$  denotes the channel from  $N$ -antenna transmitter to  $M$  single-antenna users such that the

absolute values of channel coefficients, i.e.,  $|h_{i,k}|$ , are bounded between a nonzero minimum value and a finite maximum value. We further assume that elements of  $\mathbf{H}$  can be modeled by independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and unit variance, i.e.,  $h_{i,k} \sim \mathcal{CN}(0, 1)$  for  $1 \leq i \leq M$ ,  $1 \leq k \leq N$ ,  $\mathbf{s} \in \mathbb{C}^{N \times 1}$  is the transmitted signal, and  $\mathbf{z} \in \mathbb{C}^{M \times 1}$  is the circularly symmetric additive white Gaussian noise with zero mean and variance  $\sigma^2$ , i.e.,  $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$ .

We further assume that the transmitted signal  $\mathbf{s}$  in (1) can be expressed as  $\mathbf{s} = g\mathbf{\Psi}\mathbf{c}$ . Similar to [8], [9], and [15], we consider  $g$  as the scaling factor that ensures transmit power constraint, i.e.,  $\mathbb{E}\{\|\mathbf{s}\|^2\} = 1$ , where  $\mathbb{E}\{\cdot\}$  is the expectation operator.  $\mathbf{\Psi}$  is the precoding matrix,  $\mathbf{c}$  represents the vector containing the symbols chosen from a desired constellation, and since we assume i.i.d. input signaling, we have  $\mathbb{E}\{\mathbf{c}\mathbf{c}^H\} = \mathbf{I}$ , where the superscript  $(\cdot)^H$  represents the Hermitian transpose.

Note that, throughout this paper, we consider the *input SNR* as  $1/\sigma^2$ , whereas *output SNR* is defined as the SNR received by each user.

Although the PA precoding (and, consequently, the proposed RPA precoding) is applicable when  $N \geq M$ ; hereafter, for notational and analytical simplicity, we assume that the number of transmit antennas is the same as the total number of receive antennas, i.e.,  $M = N = d$ , which is consistent with the same assumption in [8] and [9].

### A. Channel Inversion Precoding

Here, we briefly review the CI precoding. As is well known, the precoding matrix can be defined as  $\mathbf{\Psi}_{\text{CI}} = \mathbf{H}^H \mathbf{R}^{-1}$ , where  $\mathbf{R} = \mathbf{H}\mathbf{H}^H$  is the covariance matrix of the channel. Consequently, the transmitted signal in (1) is equal to

$$\mathbf{s}_{\text{CI}} = g_{\text{CI}} \mathbf{\Psi}_{\text{CI}} \mathbf{c} \quad (2)$$

where the scaling factor is equal to [8]

$$g_{\text{CI}} = \frac{1}{\sqrt{\text{Tr}[\mathbf{R}^{-1}]}} \quad (3)$$

such that  $\text{Tr}(\cdot)$  denotes “trace” operator.

For a given channel realization and with respect to the normalizing factor, the *unified* output SNR for each user is given by [8]

$$\eta'_{\text{CI}} = \frac{g_{\text{CI}}^2}{\sigma^2} = \frac{1}{\sigma^2 \text{Tr}[\mathbf{R}^{-1}]} \quad (4)$$

Since  $\text{Tr}[\mathbf{R}^{-1}] = \sum_{\ell=1}^d [\mathbf{R}^{-1}]_{\ell,\ell}$ , the output SNR for the  $\ell$ th user can be shown by [14]

$$\eta_{\text{CI}} = \frac{\text{snr}_{\text{CI}}}{[\mathbf{R}^{-1}]_{\ell,\ell}}, \quad 1 \leq \ell \leq d \quad (5)$$

where  $[\cdot]_{\ell,\ell}$  denotes the  $\ell$ th diagonal element, and we have

$$\text{snr}_{\text{CI}} = \frac{1}{d\sigma^2} \quad (6)$$

### B. Regularized Channel Inversion Precoding

Here, we briefly reintroduce the RCI precoding and derive a formula pertaining to its signal-to-interference-plus-noise ratio (SINR). There is a degradation in the performance of CI precoding in the multiple-input downlink communications since it requires inverting the square matrix  $\mathbf{R}$ , which is poorly conditioned with high probability when  $d$  is large [8]. One technique often used to regularize an inverse (and, consequently, to decrease the condition number of a matrix) is to add a multiple of the identity matrix before inverting. In [8], it has been shown that the optimum value for this regularization parameter of RCI precoding is equal to  $1/\text{snr}_{\text{CI}}$ , where  $\text{snr}_{\text{CI}}$  is defined in (6). For RCI precoding, the transmitted signal can thus be shown by

$$\mathbf{s}_{\text{RCI}} = g_{\text{RCI}} \mathbf{\Psi}_{\text{RCI}} \mathbf{c} \quad (7)$$

where  $\mathbf{\Psi}_{\text{RCI}} = \mathbf{H}^{\text{H}}(\mathbf{R} + (1/\text{snr}_{\text{CI}})\mathbf{I})^{-1}$  is the precoding matrix, and  $g_{\text{RCI}}$  is the scaling factor that can be represented as [15]

$$g_{\text{RCI}} = \frac{1}{\sqrt{\text{Tr} \left[ \mathbf{R} \left( \mathbf{R} + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-2} \right]}}. \quad (8)$$

Note that, although the SINR of the RCI precoder can be related to the eigenvalues of  $\mathbf{R}$  [8], here, we derive the SINR of RCI precoding in a different way since it also facilitates the SINR analysis of the proposed RPA precoding, as will be discussed later. To further proceed, first, we rewrite the precoding matrix as

$$\mathbf{\Psi}_{\text{RCI}} = \mathbf{H}^{\text{H}} \left( \mathbf{R} + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} = \left( \mathbf{H}^{\text{H}} \mathbf{H} + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{H}^{\text{H}}. \quad (9)$$

Therefore, the received signal can be now represented as

$$\mathbf{y} = g_{\text{RCI}} \mathbf{H} \left( \mathbf{H}^{\text{H}} \mathbf{H} + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{H}^{\text{H}} \mathbf{c} + \mathbf{z}. \quad (10)$$

Let  $\mathbf{h}_\ell \in \mathbb{C}^{1 \times d}$  denote the  $\ell$ th row of  $\mathbf{H}$  and  $\mathbf{H}_\ell$  designate the submatrix obtained by striking  $\mathbf{h}_\ell$  out of  $\mathbf{H}$ . The received signal at the  $\ell$ th user is then given by

$$\begin{aligned} y_\ell &= g_{\text{RCI}} \mathbf{h}_\ell \left( \mathbf{H}^{\text{H}} \mathbf{H} + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{H}^{\text{H}} \mathbf{c} + z_\ell \\ &= g_{\text{RCI}} \mathbf{h}_\ell \left( \mathbf{H}^{\text{H}} \mathbf{H} + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^{\text{H}} c_\ell \\ &\quad + \sum_{x \neq \ell}^d g_{\text{RCI}} \mathbf{h}_\ell \left( \mathbf{H}^{\text{H}} \mathbf{H} + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{h}_x^{\text{H}} c_x + z_\ell. \end{aligned} \quad (11)$$

Note that, in the given equation, the term  $g_{\text{RCI}} \mathbf{h}_\ell (\mathbf{H}^{\text{H}} \mathbf{H} + (1/\text{snr}_{\text{CI}})\mathbf{I})^{-1} \mathbf{h}_\ell^{\text{H}} c_\ell$  is the desired signal of user  $\ell$ . By observing that  $(\mathbf{H}^{\text{H}} \mathbf{H} + (1/\text{snr}_{\text{CI}})\mathbf{I}) = (\mathbf{H}_\ell^{\text{H}} \mathbf{H}_\ell + (1/\text{snr}_{\text{CI}})\mathbf{I} + \mathbf{h}_\ell^{\text{H}} \mathbf{h}_\ell)$  and based on the matrix inverse lemma,<sup>1</sup> we have

$$\left( \mathbf{H}^{\text{H}} \mathbf{H} + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^{\text{H}} = \frac{\left( \mathbf{H}_\ell^{\text{H}} \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^{\text{H}}}{1 + \mathbf{h}_\ell \left( \mathbf{H}_\ell^{\text{H}} \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^{\text{H}}}. \quad (12)$$

<sup>1</sup>If  $\mathbf{x}$  is a row vector, then  $(\mathbf{A} + \mathbf{x}^{\text{H}} \mathbf{x})^{-1} \mathbf{x}^{\text{H}} = (\mathbf{A}^{-1} \mathbf{x}^{\text{H}} / (1 + \mathbf{x} \mathbf{A}^{-1} \mathbf{x}^{\text{H}}))$  [16].

Therefore, by considering i.i.d. input signaling, i.e.,  $\mathbb{E}\{\mathbf{c}\mathbf{c}^{\text{H}}\} = \mathbf{I}$ , the desired signal energy of the  $\ell$ th user can be represented as

$$\kappa_\ell = g_{\text{RCI}}^2 \left( \frac{A_\ell}{1 + A_\ell} \right)^2 \quad (13)$$

where  $A_\ell = \mathbf{h}_\ell (\mathbf{H}_\ell^{\text{H}} \mathbf{H}_\ell + (1/\text{snr}_{\text{CI}})\mathbf{I})^{-1} \mathbf{h}_\ell^{\text{H}}$ .

The term  $\sum_{x \neq \ell}^d g_{\text{RCI}} \mathbf{h}_\ell (\mathbf{H}^{\text{H}} \mathbf{H} + (1/\text{snr}_{\text{CI}})\mathbf{I})^{-1} \mathbf{h}_x^{\text{H}} c_x$  in (11) is the interference caused to the  $\ell$ th user. Similarly and by doing some straightforward matrix manipulation, the power of interference induced to the  $\ell$ th user can be shown as

$$\varrho_\ell = \frac{g_{\text{RCI}}^2 B_\ell}{(1 + A_\ell)^2} \quad (14)$$

where

$$B_\ell = \mathbf{h}_\ell \left( \mathbf{H}_\ell^{\text{H}} \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{H}_\ell^{\text{H}} \mathbf{H}_\ell \left( \mathbf{H}_\ell^{\text{H}} \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{CI}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^{\text{H}}. \quad (15)$$

Therefore, the SINR for the  $\ell$ th user can be shown as

$$\eta_{\text{RCI}} = \frac{\kappa_\ell}{\varrho_\ell + \sigma^2} = \frac{g_{\text{RCI}}^2 A_\ell^2}{g_{\text{RCI}}^2 B_\ell + (1 + A_\ell)^2 \sigma^2}. \quad (16)$$

### III. PHASE ALIGNMENT REVISITED

Here, we represent the basic idea of the PA precoding. Note that, although the PA precoding was defined in [9], it needs to be redefined in a relatively different way. The reason is that there is no closed-form expression for average output SNR of PA precoding. However, in this paper, we derive a closed-form expression for this average output SNR, which eventually enables us to calculate the amount of transmit-power reduction of PA precoding compared with CI precoding for a fixed average output SNR at each user. This also facilitates the selection of an optimized regularization parameter for the proposed RPA precoding.

We note that the concept of PA is most beneficial in high-interference scenarios where more gains are to be gleaned by exploiting interference. In these scenarios, typically, low-order modulation is employed to secure low error rates. Therefore, while the benefits of the proposed scheme extend to high-order quadrature amplitude modulations, here, we focus on low-order phase-shift keying (PSK).

With PA precoding, instead of nulling out the destructive symbol-to-symbol (or cochannel) interference (as being done by  $\mathbf{R}^{-1}$  for CI precoding), the knowledge of the data's and channels' covariance matrices at transmit side can be used to make the harmful interference constructive. Fig. 1 shows how the PA precoding works for QPSK input signaling. If we consider the signal of interest as  $c_\ell = (1 + j)/\sqrt{2}$  and the interfering symbol as  $c_x = (-1 + j)/\sqrt{2}$ , the symbol-to-symbol cochannel interference  $\rho_{\ell,x} c_x$  resulted from  $c_x$  to  $c_\ell$  through the  $(\ell, x)$ th element of channel's covariance matrix  $\mathbf{R}$ , i.e.,  $\rho_{\ell,x}$ , is denoted by the dashed red arrow in the figure. The phase of the interference  $\rho_{\ell,x} c_x$  with respect to the signal of interest  $c_\ell$  is denoted by  $\theta_{\ell,x}$ . For QPSK constellation, the real and imaginary axes are decision thresholds. It is clear that, for the interfering symbol  $c_x$ , the resulting interference  $\rho_{\ell,x} c_x$

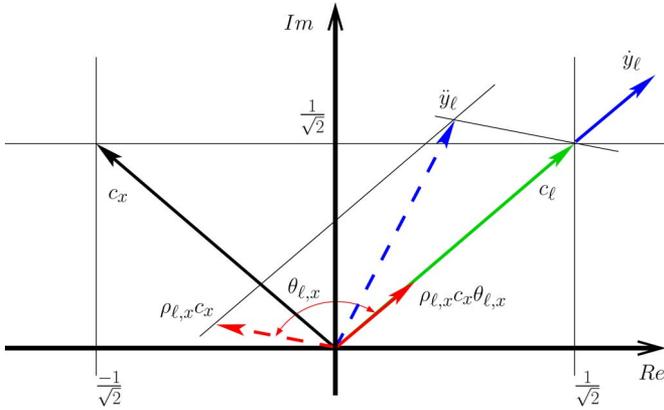


Fig. 1. PA for QPSK constellation.  $\hat{y}_\ell$  is the received symbol without PA, whereas  $\tilde{y}_\ell$  is the received symbol with PA.

is harmful since its accumulation with the signal of interest moves the received symbol  $\hat{y}_\ell$  closer to the QPSK decision thresholds. The goal of the PA precoding is to correct the phase of all transmitted symbols and to rotate the angle of correlation between them such that the resulting symbols after precoding are aligned to the signal of interest  $c_\ell$ . The desired symbol  $c_\ell$  and the aligned interference  $\rho_{\ell,x}c_x\theta_{\ell,x}$ , respectively, are shown by the solid green and red arrows in Fig. 1, which add up to  $\tilde{y}_\ell$  denoted by solid blue one. With respect to the fact that  $|c_\ell| = |c_x| = 1$ , the relative phase  $\theta_{\ell,x}$  can be expressed as

$$\rho_{\ell,x}c_x\theta_{\ell,x} \propto c_\ell \rightarrow \theta_{\ell,x} = \frac{(\rho_{\ell,x}c_x)^H}{|\rho_{\ell,x}|}c_\ell \quad (17)$$

where  $\propto$  means linear proportionality. From (17), it is evident that  $|\theta_{\ell,x}| = 1$ ; therefore, the amplitude of the rotated correlations remains unchanged. Now, matrix  $\mathbf{R}_\theta$ , which contains the phase-rotated correlation elements, can be shown as

$$[\mathbf{R}_\theta]_{\ell,x} = \rho_{\ell,x}\theta_{\ell,x} = \rho_{\ell,x}\frac{(\rho_{\ell,x}c_x)^H}{|\rho_{\ell,x}|}c_\ell = |\rho_{\ell,x}|c_\ell c_x^H \quad (18)$$

where  $\rho_{\ell,x}$  is the  $(\ell, x)$ th element of the channel's covariance matrix  $\mathbf{R}$ . From the matricial notation perspective, (18) is equivalent to

$$\mathbf{R}_\theta = |\mathbf{R}| \odot \mathbf{Q} \quad (19)$$

where  $\mathbf{Q} = \mathbf{c}\mathbf{c}^H$  is the covariance matrix of the input data vector  $\mathbf{c}$ ,  $\odot$  denotes the Hadamard (element-wise) matrix product, and  $|\cdot|$  represents the element-wise absolute value.

*Remark 1:* The PA precoding (and, consequently, the proposed RPA precoding as will be discussed later) is linear, as stated in [9]. However, this can be also deduced from (19) since the Hadamard product is a linear operator.

Now, the precoding matrix can be shown as

$$\Psi_{\text{PA}} = \mathbf{H}^H \mathbf{R}^{-1} \mathbf{R}_\theta \quad (20)$$

which yields

$$\mathbf{S}_{\text{PA}} = g_{\text{PA}} \Psi_{\text{PA}} \mathbf{C} \quad (21)$$

where the scaling factor is equal to [9]

$$g_{\text{PA}} = \frac{1}{\sqrt{\text{Tr}[\mathbf{R}_\theta^2 \mathbf{R}^{-1}]}}. \quad (22)$$

### A. Output SNR

After going through the channel, the received signal related to the  $\ell$ th user ( $\ell = 1, \dots, d$ ) can be shown as

$$\begin{aligned} y_\ell &= g_{\text{PA}} [\mathbf{R}_\theta]_{\ell,\star} \mathbf{c} + z_\ell = g_{\text{PA}} \sum_{x=1}^d |\rho_{\ell,x}| c_\ell c_x^H c_x + z_\ell \\ &= g_{\text{PA}} c_\ell \sum_{x=1}^d |\rho_{\ell,x}| + z_\ell \\ &= \underbrace{g_{\text{PA}} c_\ell |\rho_{\ell,\ell}|}_{\text{desired signal}} + \underbrace{g_{\text{PA}} c_\ell \sum_{\substack{x=1 \\ x \neq \ell}}^d |\rho_{\ell,x}|}_{\text{constructive interference}} + z_\ell \quad (23) \end{aligned}$$

where  $[\mathbf{R}_\theta]_{\ell,\star}$  denotes the  $\ell$ th row of matrix  $\mathbf{R}_\theta$  in (19), and  $z_\ell$  is the  $\ell$ th element of the noise vector  $\mathbf{z}$ , which is the circularly symmetric additive white Gaussian noise with zero mean and variance  $\sigma^2$ , i.e.,  $z_\ell \sim \mathcal{CN}(0, \sigma^2)$ . From (23), it can be seen that the received signal of the  $\ell$ th user, due to the PA of the co-stream interference, is a factor of only the desired symbol  $c_\ell$  and not the interfering symbols  $c_x$ , as also shown in Fig. 1. Since this interference contributes to the signal power, the effective SINR instead of the conventional form ( $\hat{\eta} = (S/(I + \sigma^2))$ ) can be expressed as  $\hat{\eta} = ((S + I)/\sigma^2)$ , where  $S$  denotes the desired signal's power,  $I$  is the additional signal power due to the constructive interference, and  $\sigma^2$  denotes the noise variance of each user. Therefore, it is basically a case of signal plus noise at the receiver, and consequently, here, we present an SNR calculation, as opposed to SINR. Hence, similar to CI precoding, the output SNR of the  $\ell$ th user based on PA precoding can be shown as

$$\eta_{\text{PA}} = \frac{g_{\text{PA}}^2}{\sigma^2} \left( \sum_{x=1}^d |\rho_{\ell,x}| \right)^2. \quad (24)$$

*Remark 2:* As denoted in (19), PA precoding is only dependent on the amplitudes of the elements of the channel's covariance matrix and the covariance matrix of the transmitted data, which are known at the BS prior to downlink transmission; therefore, by comparing (21) with (2), the major overhead of PA compared with CI is related to computing  $\mathbf{R}_\theta = |\mathbf{R}| \odot \mathbf{Q}$ . However, note that this calculation is going to be done at the BS that has access to sufficient power and computing facilities. Therefore, similar to CI and RCI precoding, the signal processing enhancement of PA precoding is going to be done at BS, and consequently, no overhead is introduced to MTs. This, for example, implies that the MTs do not need to know the relative phases  $\theta_{\ell,x}$  since based on (17) and (18), this phase rotation is translated into amplification of the desired symbols based on the magnitudes of the elements of the covariance matrix of the channel. This is also reflected in (23), indicating that, with PA precoding, the received signal of user  $\ell$ ,  $\ell = 1, \dots, d$  only depends on the intended symbol  $c_\ell$  and does not include the interfering symbols intended for the other  $d - 1$  MTs.

TABLE I  
 STATISTICAL PROPERTIES OF  $|\rho_{\ell,x}|$  (SEE APPENDIX A)

	Type of Random Variable	$\mathbb{E}\{ \rho_{\ell,x} \}$	$\mathbb{E}\{ \rho_{\ell,x} ^2\}$
$\ell \neq x$	Rayleigh	$\sqrt{d\pi}/2$	$d$
$\ell = x$	$\chi$ -squared	$d$	$d(d+1)$

### B. Average Output SNR of PA Precoding

Since no closed-form expression for average output SNR of PA precoding is presented in [9], here, we calculate this value. To compare the power efficiency of PA precoding with that of CI precoding and to find an optimized regularization parameter for our proposed scheme in the following, we should represent the output SNR of PA precoding similar to that of CI precoding in (5). To do this, by considering some simplifying steps and by taking the expectation over  $\eta_{\text{PA}}$  in (24), we can represent the average output SNR of each user as

$$\begin{aligned} \eta'_{\text{PA}} &= \frac{\mathbb{E}\left\{\left(g_{\text{PA}} \sum_{x=1}^d |\rho_{\ell,x}|\right)^2\right\}}{\sigma^2} \\ &= \frac{\mathbb{E}\{g_{\text{PA}}^2\} \mathbb{E}\left\{\left(\sum_{x=1}^d |\rho_{\ell,x}|\right)^2\right\}}{\sigma^2}. \end{aligned} \quad (25)$$

Here, we have assumed that  $g_{\text{PA}}$  is statistically independent of the data and the channel's covariance coefficients. While (22) contradicts this assumption, this is an affordable and common simplification to attain a closed-form approximation of the average output SNR [17]. Moreover, for large  $d$ , this becomes more justifiable, as derived by the law of large numbers. To further proceed, we should derive the statistical properties of random variable  $|\rho_{\ell,x}|$ , which is presented in Table I (for a proof, see Appendix A). Therefore, after some straightforward manipulations, (25) can be expressed as

$$\eta'_{\text{PA}} = \frac{\mathbb{E}\{g_{\text{PA}}^2\}}{\sigma^2} \left[ d(d+1) + d(d-1) \left( 1 + \sqrt{d\pi} + (d-2) \frac{\pi}{4} \right) \right]. \quad (26)$$

To calculate  $\eta'_{\text{PA}}$  in (26), we need to have the value of  $\mathbb{E}\{g_{\text{PA}}^2\}$  using the following theorem.

*Theorem 1:*  $\mathbb{E}\{g_{\text{PA}}^2\}$  can be represented as

$$\mathbb{E}\{g_{\text{PA}}^2\} = \frac{1}{2d^2 \text{Tr}[\mathbf{R}^{-1}]}. \quad (27)$$

*Proof:* See Appendix B. ■

Therefore, (26) can be now written as

$$\eta'_{\text{PA}} = \frac{(d+1) + (d-1) \left( 1 + \sqrt{d\pi} + (d-2) \frac{\pi}{4} \right)}{2d\sigma^2 \text{Tr}[\mathbf{R}^{-1}]}. \quad (28)$$

Analogous to the same procedures of CI precoding in (3)–(6), and with respect to (28) and by considering  $\text{Tr}[\mathbf{R}^{-1}] = \sum_{\ell=1}^d [\mathbf{R}^{-1}]_{\ell,\ell}$ , the average output SNR of PA precoding for the  $\ell$ th user can be represented by

$$\eta'_{\text{PA}} = \frac{\text{snr}_{\text{PA}}}{[\mathbf{R}^{-1}]_{\ell,\ell}}, \quad 1 \leq \ell \leq d \quad (29)$$

where

$$\text{snr}_{\text{PA}} = \frac{(d+1) + (d-1) \left( 1 + \sqrt{d\pi} + (d-2) \frac{\pi}{4} \right)}{2d^2\sigma^2}. \quad (30)$$

Now, the output SNR of PA precoding in (29) is of similar form to that of CI precoding in (5) in the sense that both of these equations have the same denominator. As stated previously, this treatment of output SNR of PA precoding [from (25)–(29)] will help us compare the power efficiency of PA precoding to that of CI precoding, and it also facilitates the selection of the regularization parameter for our proposed scheme in the following.

### IV. REGULARIZED PHASE ALIGNMENT PRECODING

Earlier, we showed that the PA precoding aims to rotate the phases of the transmitted symbols such that, for each MT, the interference of the remaining  $d-1$  streams add up coherently, and consequently, we can glean more received SNRs for all MTs; however, since it inherently uses CI [see (20)], the PA precoding is still problematic when the channel is ill-conditioned. To overcome this deficiency, we propose to use the concept of RCI precoding by adding a multiple of the identity matrix (i.e.,  $\epsilon \mathbf{I}$ ) to  $\mathbf{R}$  before inverting. Following [8], we seek a regularization parameter being only dependent on  $d$  and noise variance. Since  $\epsilon$  controls the amount of interference introduced to each user, the most important point is how to choose  $\epsilon$  to get the optimum performance since  $\epsilon$  can take on any positive value. In Section II-B, we showed that, for RCI precoding, this amount of  $\epsilon$  is equal to the inverse of (6), which is optimal when  $d$  is large and works well even with small  $d$ , as also discussed in [8]. Since the output SNRs of CI and PA precoding resemble each other [see (5) and (29)], analogous to RCI precoding and by comparing (29) with (4)–(6), it turns out that one good choice of  $\epsilon$  for the proposed RPA precoding can now be obtained via the inverse of (30). In Appendix C, we show that this regularization parameter, i.e.,  $\epsilon = (1/\text{snr}_{\text{PA}})$ , is optimum. Furthermore, in Section VI, we will show that this choice of  $\epsilon$  as a regularization parameter achieves very good performance. In this case, the transmitted signal  $\mathbf{s}$  in (1) is given by

$$\mathbf{s}_{\text{RPA}} = g_{\text{RPA}} \Psi_{\text{RPA}} \mathbf{c} \quad (31)$$

where

$$\Psi_{\text{RPA}} = \mathbf{H}^H \left( \mathbf{R} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{R}_\theta \quad (32)$$

is the precoding matrix, and the scaling factor can be shown as

$$g_{\text{RPA}} = \frac{1}{\sqrt{\text{Tr} \left[ \mathbf{R} \left( \mathbf{R} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{R}_\theta^2 \left( \mathbf{R} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \right]}}. \quad (33)$$

Therefore, the received signal can be now represented as

$$\begin{aligned} \mathbf{y} &= g_{\text{RPA}} \mathbf{H} \mathbf{H}^H \left( \mathbf{R} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{R}_\theta \mathbf{c} + \mathbf{z} \\ &= g_{\text{RPA}} \mathbf{H} \left( \mathbf{H}^H \mathbf{H} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{H}^H \mathbf{R}_\theta \mathbf{c} + \mathbf{z}. \end{aligned} \quad (34)$$

Since  $\mathbf{R}_\theta = |\mathbf{R}| \odot \mathbf{Q}$ , we define

$$\bar{\mathbf{c}} \triangleq \mathbf{R}_\theta \mathbf{c} = \begin{pmatrix} c_1 \sum_{x=1}^d |\rho_{1,x}| \\ \vdots \\ c_d \sum_{x=1}^d |\rho_{d,x}| \end{pmatrix}. \quad (35)$$

Let  $\bar{c}_\ell$  denote the  $\ell$ th element of  $\bar{\mathbf{c}}$  and  $\bar{\mathbf{c}}_\ell$  stand for the subvector obtained by removing  $\bar{c}_\ell$  from  $\bar{\mathbf{c}}$ . Then, the received signal at the  $\ell$ th user can be shown as

$$\begin{aligned} y_\ell &= g_{\text{RPA}} \mathbf{h}_\ell \left( \mathbf{H}^H \mathbf{H} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{H}^H \bar{\mathbf{c}} + z_\ell \\ &= g_{\text{RPA}} \mathbf{h}_\ell \left( \mathbf{H}^H \mathbf{H} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^H \bar{c}_\ell \\ &\quad + \sum_{x \neq \ell}^d g_{\text{RPA}} \mathbf{h}_\ell \left( \mathbf{H}^H \mathbf{H} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{h}_x^H \bar{c}_x + z_\ell \\ &= \underbrace{g_{\text{RPA}} \mathbf{h}_\ell \left( \mathbf{H}^H \mathbf{H} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^H \bar{c}_\ell}_{\text{desired signal}} \\ &\quad + \underbrace{g_{\text{RPA}} \mathbf{h}_\ell \left( \mathbf{H}^H \mathbf{H} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{H}_\ell^H \bar{\mathbf{c}}_\ell}_{\text{interference}} + z_\ell. \end{aligned} \quad (36)$$

Since, based on the matrix inverse lemma, we have

$$\left( \mathbf{H}^H \mathbf{H} + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^H = \frac{\left( \mathbf{H}_\ell^H \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^H}{1 + \mathbf{h}_\ell \left( \mathbf{H}_\ell^H \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^H} \quad (37)$$

by considering i.i.d. input signaling, i.e.,  $\mathbb{E}\{\mathbf{c}\mathbf{c}^H\} = \mathbf{I}$ , the output SINR of the  $\ell$ th user based on RPA precoding is equal to

$$\eta_{\text{RPA}} = \frac{g_{\text{RPA}}^2 G_\ell^2 F_\ell}{g_{\text{RPA}}^2 D_\ell + (1 + G_\ell)^2 \sigma^2} \quad (38)$$

where

$$\begin{aligned} F_\ell &= \left( \sum_{x=1}^d |\rho_{\ell,x}| \right)^2 \\ G_\ell &= \mathbf{h}_\ell \left( \mathbf{H}_\ell^H \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^H \\ D_\ell &= \mathbf{h}_\ell \left( \mathbf{H}_\ell^H \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{H}_\ell^H \boldsymbol{\Upsilon}_\ell \\ &\quad \times \mathbf{H}_\ell \left( \mathbf{H}_\ell^H \mathbf{H}_\ell + \frac{1}{\text{snr}_{\text{PA}}} \mathbf{I} \right)^{-1} \mathbf{h}_\ell^H \end{aligned}$$

such that  $\boldsymbol{\Upsilon}_\ell = \text{diag}\{F_1, \dots, F_{\ell-1}, F_{\ell+1}, \dots, F_d\}$ , and  $\text{diag}\{\cdot\}$  is the diagonal operator.

## V. POWER EFFICIENCY

Here, we investigate the ability of PA and RPA precoding to save the transmit power, which is more appropriate in the sense of green communications. We want to investigate, for a fixed average received SNR by each user, how much power-saving RPA (PA) precoding achieves in comparison with RCI (CI) precoding. If  $P_{\text{PA}}$  and  $P_{\text{CI}}$ , respectively, represent the deployed power for each user by PA and CI precoding (via replacing  $1/\sigma^2$ ), then for the same received SNR for PA and CI precoding, we have

$$\begin{aligned} \eta_{\text{CI}} = \eta'_{\text{PA}} \rightarrow & \frac{P_{\text{CI}}}{d[\mathbf{R}^{-1}]_{\ell,\ell}} = \frac{P_{\text{PA}} \left( (d+1) + (d-1) \left( 1 + \sqrt{d\pi} + (d-2) \frac{\pi}{4} \right) \right)}{2d^2[\mathbf{R}^{-1}]_{\ell,\ell}} \\ \rightarrow \xi \triangleq \frac{P_{\text{PA}}}{P_{\text{CI}}} &= \frac{2d}{(d+1) + (d-1) \left( 1 + \sqrt{d\pi} + (d-2) \frac{\pi}{4} \right)} \end{aligned} \quad (39)$$

which means that, with PA precoding, we can reduce the deployed power by a factor of  $\xi$  to achieve the same average output SNR as CI precoding, which is equivalent to a  $10 \log_{10}(\xi^{-1})$  dB decrease in deployed power for each user. We will show that this analytical result closely matches the simulations.

If we define  $P_{\text{RPA}}$  and  $P_{\text{RCI}}$  as the deployed power by RPA and RCI precoding, respectively, by using the numerical simulations in the following, we show that, still,  $P_{\text{RPA}}/P_{\text{RCI}} \approx \xi$ . Unfortunately, due to the complexity of MMSE expressions, it is not possible to prove it mathematically; however, conceptually, we can say that since there is a one-to-one mapping from PA to RPA precoding, which is similar to that of CI to RCI precoding, and all these four precoders are linear; therefore, we can expect that  $P_{\text{RPA}}/P_{\text{RCI}} \approx \xi$ .

One interesting observation from (39) is that the larger the  $d$  value, the more power we can save at the transmit side. For example, in the following, we show that, when  $d = 16$ , we can decrease the transmit power of PA precoding by 9.8 dB (a nearly tenfold reduction in transmit power) to deliver the same average output SNR to each user compared with CI precoding. This tenfold reduction is also there for RPA precoding compared with RCI precoding, which makes the proposed RPA precoding very vital at low input SNR ranges.

## VI. NUMERICAL RESULTS

Here, we provide numerical results to show the superior performance of the proposed RPA precoding compared with the other three precoders.

In simulations and without loss of generality, we assume that each user has one receive antenna, and the total number of users is equal to the number of transmit antennas. We further consider the same fading model as the one discussed in Section II. Moreover, the output SNRs of CI, RCI, PA, and RPA precoding are related to (5), (16), (24), and (38), respectively.

To verify the accuracy of the output SNR analysis and the derived formula in (38), we evaluated the SER based on analytical and simulation results in Fig. 2, and it turned out that the SER curves of these two methods closely match. This confirms the accuracy of the derived formula for the output SINR of RPA.

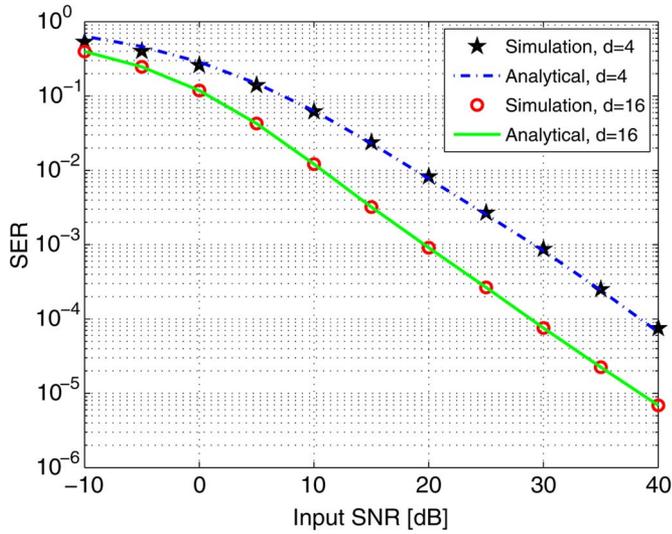


Fig. 2. Comparison of the average SER of the RPA precoding scheme based on analytical and simulation results for  $d = 4$  and  $d = 16$  and QPSK constellation.

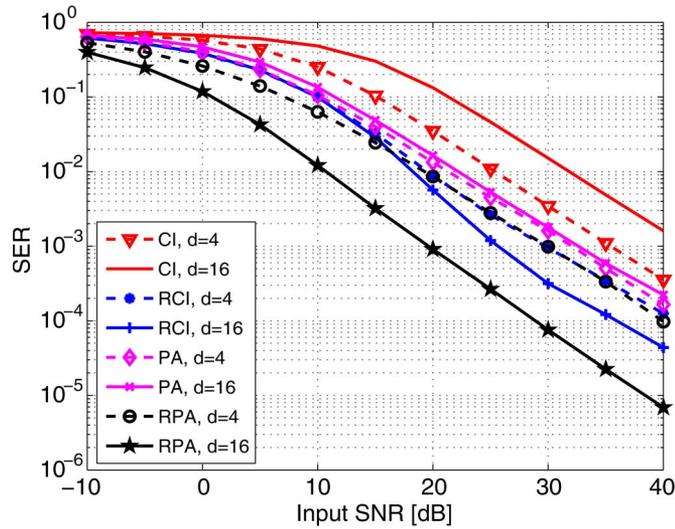


Fig. 3. Comparison of the average SER of CI, RCI, PA, and RPA precoding schemes for  $d = 4$  and  $d = 16$  and QPSK constellation.

Fig. 3 shows the comparison of the average SER based on CI, RCI, PA, and RPA precoding, where, by increasing the number of users from  $d = 4$  to  $d = 16$  (and increasing the number of transmit antennas accordingly), the following behaviors take place.

- 1) *CI Precoding*: Each user experiences inferior SER performance.
- 2) *RCI Precoding*: The SER performance of each user remains constant at low SNRs and improves slightly at high SNRs, as also shown in [8].
- 3) *PA Precoding*: The SER performance of each user remains almost constant for all SNR ranges.
- 4) *RPA Precoding*: Each user experiences remarkably better SER performance for all SNR ranges.

As shown, for  $d = 4$  and to achieve a fixed SER, RPA gives us 2.5-dB gain compared with PA, and for  $d = 16$ , this gain is about 10 dB at low SNRs and 15 dB at high SNRs.

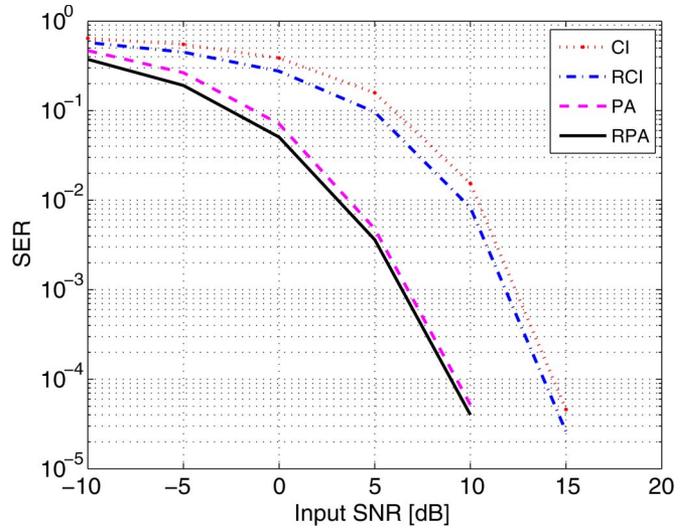


Fig. 4. Comparison of the average SER of CI, RCI, PA, and RPA precoding schemes for  $M = 10$  and  $N = 16$  and QPSK constellation.

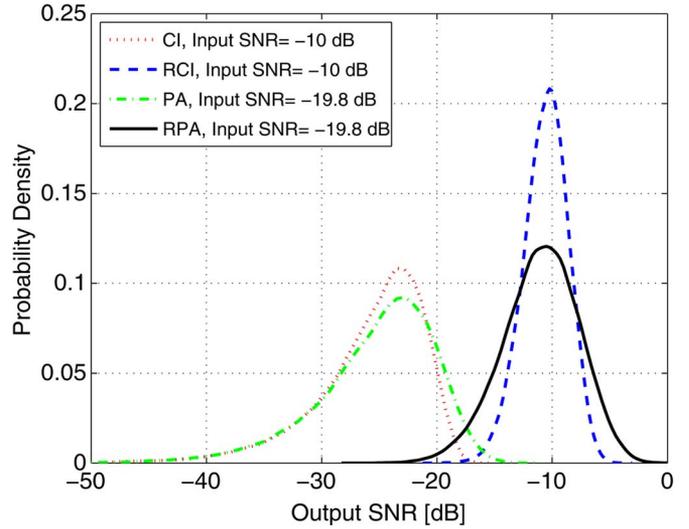


Fig. 5. Probability density of the output SNR of each user based on different precoding schemes and for  $d = 16$ . For CI and RCI precoding schemes, the input SNR is equal to  $-10$  dB, whereas for PA and RPA precoding schemes, the input SNR is equal to  $-19.8$  dB.

Fig. 4 shows the performance of the CI, RCI, PA, and RPA precoding for  $M = 10$ ,  $N = 16$ , and under QPSK signaling. As revealed, the proposed RPA precoding is able to achieve better performance even when the number of receive antennas is less than the number of transmit antennas at the BS. However, when  $N > M$ , the performance of the nonregularized precoders becomes very close to that of the regularized precoders.

In Figs. 5–7, we examine the power efficiency of PA and RPA precoding. Fig. 5 shows the probability density of the output SNR of each user based on different precoding schemes for the case of  $d = 16$ . Based on our discussions in Section V and with respect to (39), we expect that, for  $d = 16$  and for a fixed average output SNR, PA, and RPA precoding, respectively, achieve a 9.8-dB decrease in transmit power compared with CI and RCI precoding for each user. Fig. 5 verifies this behavior. For example, as shown, the *mean* of the output SNR of each user

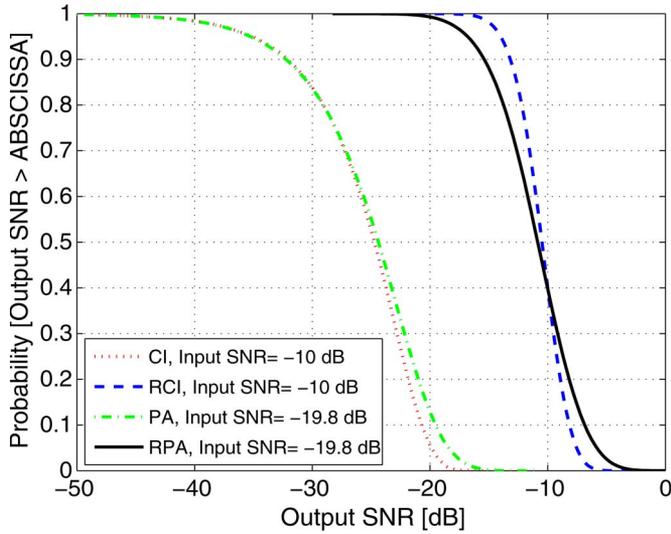


Fig. 6. CCDs of output SNR of each user based on different precoding schemes and for  $d = 16$ . For CI and RCI precoding schemes, the input SNR is equal to  $-10$  dB, whereas for PA and RPA precoding schemes, the input SNR is equal to  $-19.8$  dB.

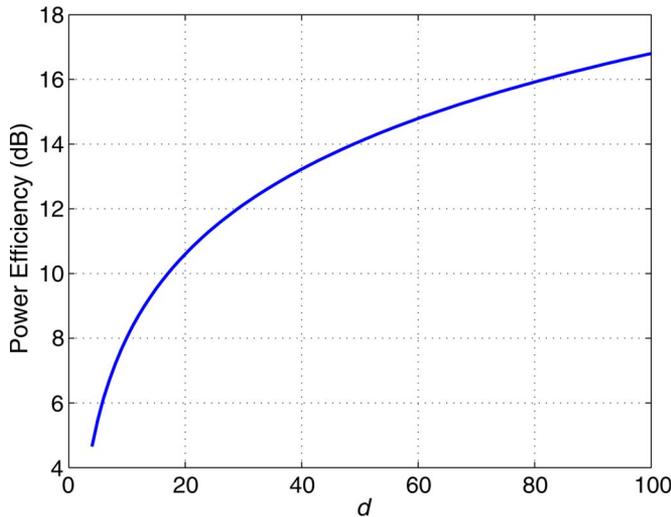


Fig. 7. Average power efficiency of RPA (PA) to RCI (CI) precoding.

based on RPA precoding with input SNR of  $-19.8$  dB is almost the same as that of RCI precoding with input SNR of  $-10$  dB.

In Fig. 6, we compare the complementary cumulative distributions (CCDs) of the output SNR of each user for the case  $d = 16$ . As observed, the CCD of PA precoding with input SNR of  $-19.8$  dB is almost the same as that of CI precoding with input SNR of  $-10$  dB. Moreover, for 40% of channel realizations, the minimum output SNR of each user based on RPA precoding with input SNR of  $-19.8$  dB is the same as that of RCI precoding with input SNR of  $-10$  dB.

Fig. 7 shows the power efficiency of the RPA (PA) to RCI (CI) precoding. As shown, the larger the  $d$  value is, the more power we can save at transmit side. This magnifies the importance of PA and RPA precoding in the context of massive MIMO. For example, with  $d = 100$ , the proposed RPA precoding enables us to save nearly 17 dB (a 50-fold reduction) transmit power compared with RCI precoding for each user, which is significant at low input SNRs.

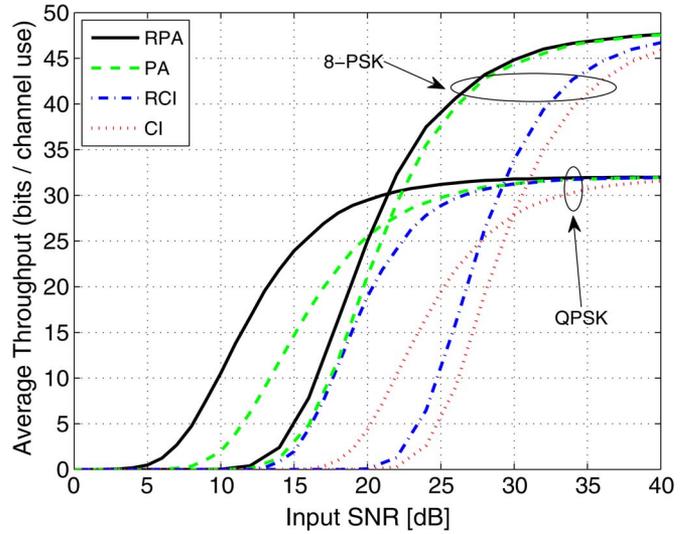


Fig. 8. Average throughput for CI, RCI, PA, and RPA precoding schemes for  $d = 16$  and for QPSK and 8-PSK constellations.

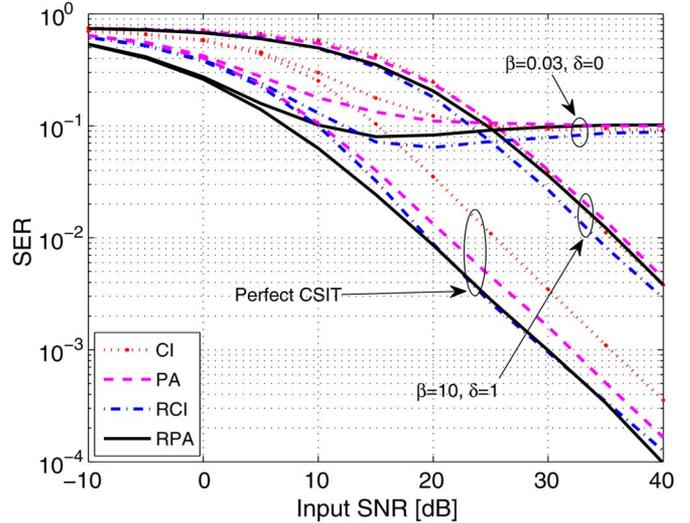


Fig. 9. Average SER performance for CI, RCI, PA, and RPA precoding schemes for  $d = 4$  and QPSK constellation. The CSIT error can depend on SNR ( $\beta = 10, \delta = 1$ ) or be independent of SNR ( $\beta = 0.03, \delta = 0$ ).

The throughput benefits of different linear precoding techniques are examined in Fig. 8 for the case of  $d = 16$  and for both QPSK and 8-PSK constellations. In the results depicted, the throughput is expressed as  $(1 - \text{blkerr}) d \log_2 \mathcal{M}$  bits per channel use (bits/cu), where blkerr is the block error rate (here, we considered each block consists of 128 symbols),  $\mathcal{M} = 4$  for QPSK, and  $\mathcal{M} = 8$  for 8-PSK constellations. As revealed, the proposed RPA precoding achieves better throughput compared with the other three precoders. For example, at input SNR of 7.5 dB and for QPSK modulation, while CI, RCI, and PA precoders give no throughput, RPA precoding attains 5 bits/cu.

So far, we have assumed that perfect CSIT is available at the BS. However, since it is not practically easy to obtain perfect CSIT, we examine the performance (or sensitivity) of the proposed method to channel estimation errors. Fig. 9 shows the average SER of different precoding schemes for QPSK constellation. The SER results of perfect CSIT are also repeated

for comparison. Following [18], we model the imperfect CSIT as

$$\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{E} \quad (40)$$

where the CSIT accuracy is characterized by the error matrix  $\mathbf{E}$ , which is assumed independent of the actual channel matrix  $\mathbf{H}$ .  $\mathbf{E}$  can be further considered complex Gaussian with i.i.d. elements, i.e.,

$$\text{vec}(\mathbf{E}) \sim \mathcal{NC}(\mathbf{0}, \tau \mathbf{I}) \text{ with } \tau \triangleq \beta \left(\frac{1}{\sigma^2}\right)^{-\delta}, \quad \beta > 0; \quad \delta \geq 0 \quad (41)$$

where  $1/\sigma^2$  is the input SNR. With this model, the error variance  $\tau$  can depend on the SNR ( $\delta \neq 0$ ) or be independent of SNR ( $\delta = 0$ ).  $\tau$  can be further assumed a parameter that captures the quality of the channel estimation, which is possible to be known *a priori*, depending on the channel dynamics and channel estimation schemes [19]. Fig. 9 shows that, for both SNR-dependent and SNR-independent error models, the performance trend of RPA and PA precoding follows that of RCI and CI precoding, which further implies that RPA and PA are as sensitive as RCI and CI to channel imperfections.

## VII. CONCLUSION

We have considered linear precoders in multiantenna downlink communications. We reformulated the PA precoding that aims to rotate the phases of transmit symbols such that they cause constructive interference. Unlike CI precoding where we null out the interference completely, there is no need to remove the interference by using PA precoding. Because of this and by considering fixed transmit power, PA precoding gives more output SNR to each user compared with CI precoding. However, the PA precoding is still problematic when the channel is ill-conditioned. Therefore, in this paper, we have proposed an enhanced version of PA precoding (named RPA), and we showed that it achieves better SER and throughput than CI, RCI, and PA precoding, particularly when the number of users becomes larger. We also showed that PA and RPA precoding enable us to decrease the deployed power at transmit side to achieve the same average output SNR for each user, compared with CI and RCI precoding, respectively. This transmit-power reduction is more significant when there is a large number of transmit antennas. We also illustrated that, even with imperfect CSIT, the performance trend of the proposed RPA precoding follows that of conventional precoders.

### APPENDIX A

#### STATISTICAL PROPERTIES OF $|\rho_{\ell, x}|$

By expanding the complex multiplications of matrix  $\mathbf{R} = \mathbf{H}\mathbf{H}^H$  for the case  $\ell \neq x$ , we have

$$|\rho_{\ell, x}| = \left( \left[ \sum_{n=1}^d (h_{\ell, n}^r h_{x, n}^r + h_{\ell, n}^i h_{x, n}^i) \right]^2 + \left[ \sum_{n=1}^d (h_{\ell, n}^i h_{x, n}^r - h_{\ell, n}^r h_{x, n}^i) \right]^2 \right)^{\frac{1}{2}} \quad (42)$$

where the notations  $h_{\ell, n}^r = \Re(h_{\ell, n})$ ,  $h_{\ell, n}^i = \Im(h_{\ell, n})$  are used for convenience, and  $h_{\ell, n}$  is used to denote the generic channel coefficient of the  $n$ th transmit antenna to the  $\ell$ th MT. Since we assumed that  $h_{\ell, n}^r, h_{\ell, n}^i, h_{x, n}^r, h_{x, n}^i \in \mathcal{CN}(0, (1/2))$ , we have

$$\mathbb{E}\{h_{\ell, n}^r h_{x, n}^r\} = 0 \text{ and } \text{var}\{h_{\ell, n}^r h_{x, n}^r\} = \frac{1}{4}. \quad (43)$$

The same applies to all combinations of real and imaginary coefficients that appear in (42). Therefore

$$\begin{aligned} \mathbb{E}\{h_{\ell, n}^r h_{x, n}^r + h_{\ell, n}^i h_{x, n}^i\} &= 0 \\ \text{var}\{h_{\ell, n}^r h_{x, n}^r + h_{\ell, n}^i h_{x, n}^i\} &= \frac{1}{2} \end{aligned} \quad (44)$$

$$\begin{aligned} \mathbb{E}\left\{\sum_{n=1}^d (h_{\ell, n}^r h_{x, n}^r + h_{\ell, n}^i h_{x, n}^i)\right\} &= 0 \\ \text{var}\left\{\sum_{n=1}^d (h_{\ell, n}^r h_{x, n}^r + h_{\ell, n}^i h_{x, n}^i)\right\} &= \frac{d}{2} \triangleq \mu. \end{aligned} \quad (45)$$

Due to the symmetry of the real and imaginary parts of the channel taps, the values of (45) also apply to the second term on the right-hand side of (42). Consequently,  $|\rho_{\ell, x}|$  is a Rayleigh variable with  $\mathbb{E}\{|\rho_{\ell, x}|\} = \sqrt{2\mu}\Gamma(3/2) = (\sqrt{d\pi}/2)$  and  $\mathbb{E}\{|\rho_{\ell, x}|^2\} = 2\mu\Gamma(2) = d$  [20], where  $\Gamma(\cdot)$  is the gamma function such that  $\Gamma(1) = 1$ ,  $\Gamma(1/2) = \sqrt{\pi}$ , and  $\Gamma(1+x) = x\Gamma(x)$ .

For the case of  $\ell = x$ , we have

$$|\rho_{\ell, \ell}| = \sum_{n=1}^d |h_{\ell, n}|^2 = \sum_{n=1}^d [(h_{\ell, n}^r)^2 + (h_{\ell, n}^i)^2]. \quad (46)$$

Since  $h_{\ell, n}^r, h_{\ell, n}^i \in \mathcal{CN}(0, (1/2))$ ,  $|\rho_{\ell, \ell}|$  is a  $\chi$ -square random variable with  $2d$  degrees of freedom, i.e.,  $|\rho_{\ell, \ell}| \sim \chi_{2d}^2$ , and  $\mathbb{E}\{|\rho_{\ell, \ell}|\} = 2d\varepsilon = d$ , and  $\text{var}\{|\rho_{\ell, \ell}|\} = 4d\varepsilon^2 = d$  [20], where  $\varepsilon \triangleq \text{var}\{h_{\ell, n}^r\} = (1/2)$ , therefore,  $\mathbb{E}\{|\rho_{\ell, \ell}|^2\} = [\mathbb{E}\{|\rho_{\ell, \ell}|\}]^2 + \text{var}\{|\rho_{\ell, \ell}|\} = d(d+1)$ .

### APPENDIX B

#### $\mathbb{E}\{g_{\text{PA}}^2\}$

Based on (22), we have

$$g_{\text{PA}}^2 = \frac{1}{\text{Tr}[\mathbf{R}_{\theta}^2 \mathbf{R}^{-1}]} \quad (47)$$

Therefore, to calculate  $\mathbb{E}\{g_{\text{PA}}^2\}$ , we should find the value of  $\mathbb{E}\{\text{Tr}[\mathbf{R}_{\theta}^2 \mathbf{R}^{-1}]\}^{-1}$ . Since both  $\mathbf{R}_{\theta}^2$  and  $\mathbf{R}^{-1}$  are Hermitian matrices, by using eigendecomposition, we have

$$\mathbf{R}_{\theta}^2 = \mathbf{U}_{\theta} \mathbf{\Lambda}_{\theta} \mathbf{U}_{\theta}^H \quad (48)$$

$$\mathbf{R}^{-1} = \mathbf{U}_r \mathbf{\Lambda}_r \mathbf{U}_r^H \quad (49)$$

where  $\mathbf{U}_{\theta}$  and  $\mathbf{U}_r$  are unitary matrices containing the eigenvectors of  $\mathbf{R}_{\theta}^2$  and  $\mathbf{R}^{-1}$ , respectively, and  $\mathbf{\Lambda}_{\theta}$  and  $\mathbf{\Lambda}_r$  are diagonal matrices consisting of eigenvalues of  $\mathbf{R}_{\theta}^2$  and  $\mathbf{R}^{-1}$ , respectively. Note that since we assumed i.i.d. input signaling,  $\mathbf{R}_{\theta}^2$  and

$\mathbf{U}_\theta$  are random matrices. Therefore,  $\mathbb{E}\{\text{Tr}[\mathbf{R}_\theta^2 \mathbf{R}^{-1}]\}$  can be written as

$$\begin{aligned} \mathbb{E}\{\text{Tr}[\mathbf{R}_\theta^2 \mathbf{R}^{-1}]\} &= \mathbb{E}\{\text{Tr}[\mathbf{U}_\theta \mathbf{\Lambda}_\theta \mathbf{U}_\theta^H \mathbf{U}_r \mathbf{\Lambda}_r \mathbf{U}_r^H]\} \\ &= \mathbb{E}\{\text{Tr}[\mathbf{U}_r^H \mathbf{U}_\theta \mathbf{\Lambda}_\theta \mathbf{U}_\theta^H \mathbf{U}_r \mathbf{\Lambda}_r]\}. \end{aligned} \quad (50)$$

Now, if we define  $\mathbf{U} \triangleq \mathbf{U}_r^H \mathbf{U}_\theta$ , (50) can be shown by

$$\begin{aligned} \mathbb{E}\{\text{Tr}[\mathbf{R}_\theta^2 \mathbf{R}^{-1}]\} &= \mathbb{E}\{\text{Tr}[\mathbf{U} \mathbf{\Lambda}_\theta \mathbf{U}^H \mathbf{\Lambda}_r]\} \\ &= \mathbb{E}\left\{\sum_{i=1}^d \sum_{k=1}^d \lambda_{r,i} \lambda_{\theta,k} |u_{i,k}|^2\right\} \end{aligned} \quad (51)$$

where  $\lambda_{r,i}$  and  $\lambda_{\theta,k}$  are the  $i$ th and the  $k$ th diagonal elements of matrices  $\mathbf{\Lambda}_r$  and  $\mathbf{\Lambda}_\theta$ , respectively, and  $u_{i,k}$  denotes the  $(i, k)$ th element of unitary matrix  $\mathbf{U}$ .

To continue, we consider the following theorem [21].

*Theorem 2:* If the Hermitian unitary invariant<sup>2</sup> random matrix  $\mathbf{W}$  can be eigendecomposed as  $\mathbf{W} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H$ , then the unitary matrix  $\mathbf{U}$ , which is a Haar<sup>3</sup> matrix, is independent of the diagonal matrix  $\mathbf{\Lambda}$ .

Since both  $\mathbf{U}_\theta$  and  $\mathbf{U}_r$  are random matrices, so is  $\mathbf{U}$ ; hence,  $\mathbb{E}\{|u_{i,k}|^2\} = (1/d)$  [21], and based on *Theorem 2*, we can write (51) as

$$\begin{aligned} \mathbb{E}\{\text{Tr}[\mathbf{R}_\theta^2 \mathbf{R}^{-1}]\} &= \sum_{i=1}^d \sum_{k=1}^d \lambda_{r,i} \mathbb{E}\{\lambda_{\theta,k}\} \mathbb{E}\{|u_{i,k}|^2\} \\ &= \frac{1}{d} \mathbb{E}\{\text{Tr}[\mathbf{R}_\theta^2]\} \text{Tr}[\mathbf{R}^{-1}]. \end{aligned} \quad (52)$$

Note that, since  $\mathbf{R}^{-1}$  is a square matrix,  $\mathbb{E}\{\text{Tr}[\mathbf{R}^{-1}]\}$  is not defined [21]. To further proceed, we consider the following lemma.

*Lemma 1:*  $\text{Tr}[\mathbf{R}_\theta^2] = \text{Tr}[\mathbf{R}^2]$ .

*Proof:* To continue with the proof, we only need to show that the  $\ell$ th diagonal elements of  $\mathbf{R}_\theta^2$  and  $\mathbf{R}^2$  are the same. Since matrix  $\mathbf{R}_\theta$  is Hermitian, we have  $\mathbf{R}_\theta^2 = \mathbf{R}_\theta \mathbf{R}_\theta^H$ . Hence, we can write

$$\alpha_{\ell,\ell} = \sum_{i=1}^d r_{\ell,i} r_{\ell,i}^H \quad (53)$$

where  $\alpha_{\ell,\ell}$  is the  $\ell$ th diagonal element of  $\mathbf{R}_\theta^2$ , and  $r_{\ell,i}$  is the  $(\ell, i)$ th element of  $\mathbf{R}_\theta$ . Based on (18),  $r_{\ell,i} = |\rho_{\ell,i}| c_{\ell,i}^H$ , where  $c_{\ell}$  is the  $\ell$ th element of the data vector  $\mathbf{c}$ . Therefore, we can rewrite (53) as

$$\begin{aligned} \alpha_{\ell,\ell} &= \sum_{i=1}^d |\rho_{\ell,i}| c_{\ell,i}^H (|\rho_{\ell,i}| c_{\ell,i}^H)^H = \sum_{i=1}^d |\rho_{\ell,i}| c_{\ell,i}^H c_{\ell,i}^H |\rho_{\ell,i}|^H \\ &= \sum_{i=1}^d |\rho_{\ell,i}|^2 = \sum_{i=1}^d \rho_{\ell,i} \rho_{\ell,i}^H. \end{aligned} \quad (54)$$

<sup>2</sup>A Hermitian random matrix  $\mathbf{W}$  is called *unitary invariant* if the joint distribution of its entries equals that of  $\mathbf{V} \mathbf{W} \mathbf{V}^H$  for any unitary matrix  $\mathbf{V}$  independent of  $\mathbf{W}$ .

<sup>3</sup>A  $d \times d$  random matrix  $\mathbf{U}$  is a *Haar* matrix (also called isotropic in the multiantenna literature) if it is uniformly distributed on the set of  $d \times d$  unitary matrices.

Therefore, from (54), we can deduce that the  $\ell$ th diagonal element of  $\mathbf{R}_\theta^2$ , i.e.,  $\alpha_{\ell,\ell}$ , is equal to that of  $\mathbf{R}^2 = \mathbf{R} \mathbf{R}^H$ , which is denoted by  $\sum_{i=1}^d \rho_{\ell,i} \rho_{\ell,i}^H$ . ■

Therefore, we have

$$\mathbb{E}\{g_{\text{RPA}}^2\} = \frac{d}{\mathbb{E}\{\text{Tr}[\mathbf{R}^2]\} \text{Tr}[\mathbf{R}^{-1}]} \quad (55)$$

and since  $\mathbb{E}\{\text{Tr}[\mathbf{R}^2]\} = 2d^3$  [21],  $\mathbb{E}\{g_{\text{RPA}}^2\}$  can be shown as

$$\mathbb{E}\{g_{\text{RPA}}^2\} = \frac{1}{2d^2 \text{Tr}[\mathbf{R}^{-1}]}. \quad (56)$$

## APPENDIX C

### ON THE OPTIMALITY OF THE REGULARIZATION PARAMETER $1/\text{SNR}_{\text{RPA}}$

We show the regularization parameter used in Section IV for the proposed RPA precoding is optimum.

By considering (32), we redefine the RPA precoder as

$$\mathbf{\Psi}_{\text{RPA}} = \mathbf{\Phi} \mathbf{R}_\theta. \quad (57)$$

$\mathbf{\Phi}$  can be then found by using the following MMSE criterion:

$$\arg \min_{\mathbf{\Phi}} \mathbb{E}\{\|\mathbf{H} \mathbf{\Phi} \bar{\mathbf{c}} + f \mathbf{z} - \bar{\mathbf{c}}\|_2^2\} \quad (58)$$

where  $\bar{\mathbf{c}}$  is defined in (35), and

$$f = \frac{1}{g_{\text{RPA}}} = \sqrt{\text{Tr}[\mathbf{\Psi}_{\text{RPA}}^H \mathbf{\Psi}_{\text{RPA}}]} = \sqrt{\text{Tr}[\mathbf{R}_\theta^H \mathbf{\Phi}^H \mathbf{\Phi} \mathbf{R}_\theta]} \quad (59)$$

where  $g_{\text{RPA}}$  is the scaling factor of the RPA precoding. The inclusion of  $f$  in (58) is due to the fact that, in all precoding schemes (e.g., CI and RCI) the power of noise is affected by the precoding matrix, and consequently, this effect can be reflected through a multiplicative factor such as  $f$ . This can be perceived with respect to the fact that, at the transmit side, the transmitted signals are scaled by  $g_{\text{RPA}}$  to meet the power constraints; consequently, at the receive side, the received signals should be scaled back by  $1/g_{\text{RPA}}$ , which further appears as a multiplicative factor for the noise vector.

To further proceed, we consider the following two lemmas.

*Lemma 2:*  $\mathbb{E}\{\bar{\mathbf{c}} \bar{\mathbf{c}}^H\} = \omega \mathbf{I}$ , where

$$\omega = d \left[ (d+1) + (d-1) \left( 1 + \sqrt{d\pi} + (d-2) \frac{\pi}{4} \right) \right]. \quad (60)$$

*Proof:* With respect to the fact that  $\mathbb{E}\{\mathbf{c} \mathbf{c}^H\} = \mathbf{I}$ , we have

$$\begin{aligned} \mathbb{E}\{\bar{\mathbf{c}} \bar{\mathbf{c}}^H\} &= \text{diag} \left( \mathbb{E} \left\{ \left( \sum_{x=1}^d |\rho_{1,x}| \right)^2 \right\} \right. \\ &\quad \left. \dots, \mathbb{E} \left\{ \left( \sum_{x=1}^d |\rho_{d,x}| \right)^2 \right\} \right). \end{aligned} \quad (61)$$

By considering the statistical properties of  $|\rho_{\ell,x}|$  presented in Table I and after some straightforward manipulations, the claim follows. ■

*Lemma 3:*  $\mathbb{E}\{\mathbf{R}_\theta^2\} = \nu\mathbf{I}$ , where

$$\nu = 2d^2. \quad (62)$$

*Proof:* Based on (35) and by considering the fact that  $\mathbb{E}\{\mathbf{c}\mathbf{c}^H\} = \mathbf{I}$ , we have

$$\mathbb{E}\{\mathbf{R}_\theta^2\} = \text{diag}\left(\mathbb{E}\{[\mathbf{R}^2]_{1,1}\}, \dots, \mathbb{E}\{[\mathbf{R}^2]_{d,d}\}\right). \quad (63)$$

Since  $\mathbf{R}$  is a Hermitian matrix, we have

$$[\mathbf{R}^2]_{\ell,\ell} = \sum_{i=1}^d |\rho_{\ell,i}|^2. \quad (64)$$

Thus, based on Table I, we have  $\mathbb{E}\{[\mathbf{R}^2]_{\ell,\ell}\} = 2d^2$ , and the claim follows. ■

Therefore, the objective function in (58) can be shown as

$$J = \mathbb{E}\left\{\text{Tr}\left[(\mathbf{H}\Phi\bar{\mathbf{c}} + f\mathbf{z} - \bar{\mathbf{c}})(\mathbf{H}\Phi\bar{\mathbf{c}} + f\mathbf{z} - \bar{\mathbf{c}})^H\right]\right\}. \quad (65)$$

With respect to the fact that the data and noise are independent of each other and both are also independent of channel matrix  $\mathbf{H}$ , the objective function in (65) can be rewritten as

$$J = \mathbb{E}\left\{\text{Tr}\left[\mathbf{H}\Phi\bar{\mathbf{c}}\bar{\mathbf{c}}^H\Phi^H\mathbf{H}^H + f^2\mathbf{z}\mathbf{z}^H + \bar{\mathbf{c}}\bar{\mathbf{c}}^H - \mathbf{H}\Phi\bar{\mathbf{c}}\bar{\mathbf{c}}^H - \bar{\mathbf{c}}\bar{\mathbf{c}}^H\Phi^H\mathbf{H}^H\right]\right\}. \quad (66)$$

Therefore, by considering *Lemma 2* and *Lemma 3*, the objective function can be shown as

$$J = \omega\text{Tr}\left[\Phi^H\mathbf{H}^H\mathbf{H}\Phi\right] + d\nu\sigma^2\text{Tr}\left[\Phi^H\Phi\right] - \omega\text{Tr}\left[\mathbf{H}\Phi\right] - \omega\text{Tr}\left[\Phi^H\mathbf{H}^H\right] + d\omega. \quad (67)$$

To further proceed, we consider the following assumptions [22].

- 1)  $\Phi$  and  $\Phi^H$  are treated as independent variables.
- 2)  $(\partial\text{Tr}[\mathbf{A}\Phi]/\partial\Phi) = (\partial\text{Tr}[\Phi\mathbf{A}]/\partial\Phi) = \mathbf{A}$ .

Now, with respect to the aforementioned assumptions, the sought precoder  $\Phi$  can be found by differentiating  $J$  with respect to  $\Phi$  and setting it equal to zero, which yields

$$\begin{aligned} \frac{\partial J}{\partial\Phi} &= \omega\Phi^H\mathbf{H}^H\mathbf{H} + d\nu\sigma^2\Phi^H - \omega\mathbf{H} = \mathbf{0} \\ \Rightarrow \Phi &= \mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H + \frac{d\nu\sigma^2}{\omega}\mathbf{I}\right)^{-1}. \end{aligned} \quad (68)$$

With respect to (57), the RPA precoder is defined as  $\Psi_{\text{RPA}} = \Phi\mathbf{R}_\theta$ . Now, as revealed in (68), the regularization parameter, i.e., the multiplicative factor of the identity matrix, is  $(d\nu\sigma^2/\omega)$ , which is exactly equal to  $1/\text{snr}_{\text{PA}}$ , where  $\text{snr}_{\text{PA}}$  has been already defined in (30).

Note that since the Hessian matrix of the MSE objective function is positive definite, the expression in (68) is a global

minimizer for the considered MMSE optimization problem, which implies on the optimality of the derived regularization parameter.

## REFERENCES

- [1] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [2] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, Mar. 2005.
- [3] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Computationally efficient vector perturbation precoding using thresholded optimization," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1880–1890, May 2013.
- [4] L. Sanguinetti and M. Morelli, "Non-linear pre-coding for multiple-antenna multi-user downlink transmissions with different QoS requirements," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 852–856, Mar. 2007.
- [5] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Interference optimization for transmit power reduction in Tomlinson–Harashima precoded MIMO downlinks," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2470–2481, May 2012.
- [6] T. Haustein, C. von Helmolt, E. Jorswieck, V. Jungnickel, and V. Pohl, "Performance of MIMO systems with channel inversion," in *Proc. IEEE VTC*, May 2002, vol. 1, pp. 35–39.
- [7] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
- [8] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 192–202, Jan. 2005.
- [9] C. Masouros, "Correlation rotation linear precoding for MIMO broadcast communications," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 252–262, Jan. 2011.
- [10] C. Masouros and T. Ratnarajah, "Interference as a source of green signal power in cognitive relay assisted co-existing MIMO wireless transmissions," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 525–536, Feb. 2012.
- [11] F. A. Khan, C. Masouros, and T. Ratnarajah, "Interference-driven linear precoding in multiuser MISO downlink cognitive radio network," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2531–2543, Jul. 2012.
- [12] C. Masouros, T. Ratnarajah, M. Sellathurai, C. B. Papadias, and A. K. Shukla, "Known interference in the cellular downlink: A performance limiting factor or a source of green signal power?" *IEEE Commun. Mag.*, vol. 51, no. 10, pp. 162–171, Oct. 2013.
- [13] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Large-scale MIMO transmitters in fixed physical spaces: The effect of transmit correlation and mutual coupling," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2794–2804, Jul. 2013.
- [14] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.
- [15] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: A large system analysis," in *Proc. IEEE GLOBECOM*, 2008, pp. 1–4.
- [16] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.
- [17] X. Shao, J. Yuan, and Y. Shao, "Error performance analysis of linear zero forcing and MMSE precoders for MIMO broadcast channels," *IET Commun.*, vol. 1, no. 5, pp. 1067–1074, Oct. 2007.
- [18] S. M. Razavi and T. Ratnarajah, "Performance analysis of interference alignment under CSI mismatch," *IEEE Trans. Veh. Technol.*
- [19] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.
- [20] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York, NY, USA: McGraw-Hill, 2007.
- [21] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.
- [22] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

**S. Morteza Razavi** received the Ph.D. degree in electrical engineering from The University of Edinburgh, Edinburgh, U.K., in 2014.

His research interests are in the areas of wireless communications and signal processing, with an emphasis on multiuser MIMO systems.

**Tharmalingam Ratnarajah** (A'96–M'05–SM'05) has held various positions with the Queen's University Belfast, Belfast, U.K., since 1993; the University of Ottawa, Ottawa, ON, Canada; Nortel Networks, Ottawa; McMaster University, Hamilton, ON; and Imperial College, London, U.K. He is currently a Reader of signal processing and communications with the Institute for Digital Communications, The University of Edinburgh, Edinburgh, U.K. He is the author of over 200 publications in these areas and holds four U.S. patents. He is currently the Coordinator of the FP7 Project HARP (3.2M€) in the area of highly distributed multiple-input multiple-output (MIMO) and FP7 Project ADEL (2.6M€) in the area of licensed shared access. Previously, he was the Coordinator of FP7 Future and Emerging Technologies Project CROWN (2.3M€) in the area of cognitive radio networks and Project HIATUS (2.7M€) in the area of interference alignment. His research interests include random matrix theory, information theoretic aspects of MIMO channels and ad hoc networks, wireless communications, signal processing for communications, statistical and array signal processing, biomedical signal processing, and quantum information theory.

Dr. Ratnarajah is a member of the American Mathematical Society and Information Theory Society and a Fellow of the Higher Education Academy, U.K.

**Christos Masouros** (M'06–SM'14) received the Diploma in electrical and computer engineering from the University of Patras, Patras, Greece, in 2004 and the M.Sc. degree and the Ph.D. degree in electrical and electronic engineering from the University of Manchester, Manchester, U.K., in 2006 and 2009, respectively.

He was a Research Associate with the University of Manchester and a Research Fellow with Queen's University Belfast, Belfast, U.K. He is currently a Lecturer with the Department of Electrical and Electronic Engineering, University College London, London, U.K. His research interests include wireless communications and signal processing, with particular focus on green communications, cognitive radio, interference mitigation techniques for multiple-input–multiple-output, code-division multiple access, and multicarrier communications.

Dr. Masouros received a Royal Academy of Engineering Research Fellowship for 2011–2016.