Multiplexing More Streams in the MU-MISO Downlink by Interference Exploitation Precoding

Ang Li
School of Electrical and Inf. Eng.
University of Sydney
Sydney, Australia
ang.li2@sydney.edu.au

Christos Masouros
Dept. of Electronic and Electrical Eng.
University College London
London, U.K.
c.masouros@ucl.ac.uk

Yonghui Li, and Branka Vucetic
School of Electrical and Inf. Eng.
University of Sydney
Sydney, Australia
{yonghui.li, branka.vucetic}@sydney.edu.au

Abstract—In this paper, we study the interference exploitation precoding for the scenario where the number of streams simultaneously transmitted by the base station (BS) is larger than that of transmit antennas at the BS, and derive the optimal precoding structure by employing the pseudo inverse. We show that the optimal pre-scaling vector is equal to a linear combination of the right singular vectors that correspond to zero singular values of the coefficient matrix. By formulating the dual problem, the optimal precoding matrix can be expressed as a function of the dual variables in a closed form, and an equivalent quadratic programming (QP) formulation is further derived for computational complexity reduction. Numerical results validate our analysis and demonstrate significant performance improvements for interference exploitation precoding for the considered scenario.

Index Terms—MIMO, symbol-level precoding, constructive interference, optimization, Lagrangian.

I. INTRODUCTION

Multi-antenna wireless communication systems have received extensive research attention due to their significant performance gains over single-antenna systems, where precoding has been widely acknowledged as a promising application [1]. When the channel state information (CSI) is available at the transmitter side, precoding is able to support data transmission to multiple users simultaneously. Well-known precoding approaches include theoretically capacity-achieving dirty paper coding (DPC) [2], non-linear precoding such as Tomlinson-Harashima precoding (THP) [3] and vector perturbation (VP) precoding [4], and low-complexity linear precoding such as zero-forcing (ZF) and regularized ZF (RZF) [5]. Meanwhile, downlink precoding based on optimization has also received increasing research attention in recent years [6]-[8]. Among optimization-based precoding approaches, signal-to-interference-plus-noise ratio (SINR) balancing [7] and power minimization [8] are two most popular designs. These precoding schemes exploit the information of the channel to design the precoding matrices that target at avoiding or limiting the interference.

Compared to the above studies that treat interference as detrimental, recent studies show that interference can also be beneficial and provide further performance improvements on a symbol level [9]. By exploiting the information of the data symbols and their corresponding constellation, the instantaneous interference can be divided into constructive interference (CI) and destructive interference [10]. More specifically, CI is defined as the interference that pushes the received signals away from the detection thresholds, which further improves the detection performance. Based on the above, CI-based precoding for PSK modulations has been proposed in [11], [12] as a modification of ZF precoding. Optimization-based CI precoding has further been proposed in [13] based on symbol scaling and [14]-[16] based on phase rotation, and their extension to multi-level modulations such as QAM is discussed in [17], [18]. More recently, it has been revealed in [19] that there exists an optimal precoding structure for CI precoding. In addition to the performance improvements of CI precoding over traditional precoding approaches, another advantage of CI precoding is its capability of supporting a larger number of streams (single-antenna users) than the number of transmit antennas at the base station (BS) simultaneously, which has only been numerically shown in [14]. Nevertheless, it is still not clear whether the analysis and results in [19] can be extended to this scenario.

Therefore in this paper, we focus on the scenario where the number of streams simultaneously transmitted by the BS is larger than that of transmit antennas at the BS. Based on the Karush-Kuhn-Tucker (KKT) conditions, we derive the optimal precoding structure at the BS and transform the problem into an optimization on the pre-scaling vector. In the derivation, the exact matrix inverse is not applicable due to the rank deficiency, and accordingly we employ the pseudo inverse instead, which introduces an additional constraint to the optimization problem. We further show that the optimal pre-scaling vector is equal to a linear combination of the right singular vectors corresponding to zero singular values of the coefficient matrix. Subsequently, the optimization problem is further transformed into an optimization on the weights for each singular vector, which is finally shown to be a quadratic programming (QP) optimization and can be more efficiently solved than the original second-order cone programming (SOCP) formulation. Based on the QP formulation, we also discuss the condition under which multiplexing more streams than the number of transmit antennas at the BS is feasible with interference exploitation precoding. Numerical results validate our analysis and demonstrate significant performance gains of interference exploitation precoding over traditional precoding methods in the considered scenario.

Notations: a, $\mathbf{a}$, and $\mathbf{A}$ denote scalar, column vector and

matrix, respectively. $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$, and $(\cdot)^+$ denote conjugate, transposition, conjugate transposition, inverse, and pseudo inverse of a matrix, respectively. $diag\,(\cdot)$ is the transformation of a column vector into a diagonal matrix, and $\mathrm{vec}\,(\cdot)$ denotes the vectorization operation. $|\cdot|$ denotes the absolute value or the modulus, and $\|\cdot\|_2$ is the $l2$-norm. $\mathcal{C}^{n\times n}$ and $\mathcal{R}^{n\times n}$ represent an $n \times n$ matrix in the complex and real set, respectively. j is the imaginary unit.

II. System Model and Problem Formulation

A. System Model

We focus on a multi-user multiple-input single-ouput (MU-MISO) system in the downlink, where the BS with N_t transmit antennas transmits a total number of K streams simultanesouly, and $K > N_t$. The data symbol vector $\mathbf{s} = [s_1, s_2, \cdots, s_K]^T \in \mathcal{C}^{K\times 1}$ is assumed to be from a normalized PSK constellation, and therefore $s_k s_k^H = 1$ for each stream k [14]. The received signal at user k can be expressed as

$$r_k = \mathbf{h}_k^T \mathbf{W}\mathbf{s} + n_k, \tag{1}$$

where $\mathbf{h}_k \in \mathcal{C}^{N_t\times 1}$ denotes the flat-fading Rayleigh channel between user k and the BS, and perfect CSI is assumed throughout this paper. $\mathbf{W} \in \mathcal{C}^{N_t\times K}$ is the precoding matrix, and n_k is the additive Gaussian nose with zero mean and variance σ^2.

B. Problem Formulation

The optimization of CI precoding can be formulated based on the geometry of the PSK constellation, as shown in Fig. 1 where we employ 8PSK as an example. As discussed in [19], we denote $\vec{OS} = s_k$ and $\vec{OA} = t \cdot s_k$, where $t > 0$ is the objective to be maximized. $\vec{OB}$ is the received signal excluding noise, and based on (1) $\vec{OB}$ is expressed as

$$\vec{OB} = \mathbf{h}_k^T \mathbf{W}\mathbf{s} = \lambda_k s_k, \tag{2}$$

where λ_k is an introduced complex scalar that represents the interference effect on user k's data symbol. Based on the geometry in Fig. 1 and [19], the CI constraint is to locate the noiseless received signal $\vec{OB}$ within the constructive area, i.e, $|\vec{CB}| \leq |\vec{CD}|$, which mathematically leads to

$$\tan\theta_{AB} \leq \tan\theta_t \;\;\Rightarrow\;\; \left(\lambda_k^{\Re} - t\right)\tan\theta_t \geq \left|\lambda_k^{\Im}\right|, \quad (3)$$
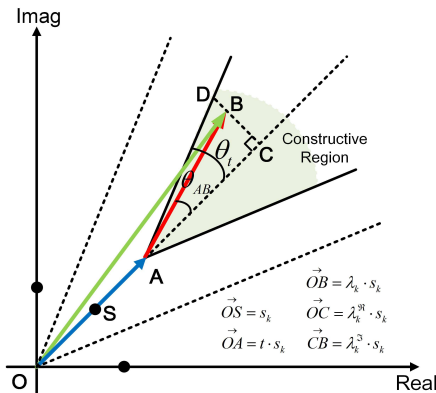

Fig. 1: The phase-rotation metric for 8PSK constellation

where $\lambda_k^{\Re}$ and $\lambda_k^{\Im}$ represent the real and imaginary part of λ_k, respectively. Throughout this paper, we focus on the non-strict phase-rotation CI, while the strict phase-rotation CI can be regarded as a special case by setting each $\lambda_k^{\Im}$ to zero [19]. Based on the constellation, we also obtain $\theta_t = \frac{\pi}{\mathcal{M}}$ for $\mathcal{M}$-PSK modulation. Accordingly, the optimization on CI precoding that maximizes the CI effect subject to the total available transmit power can be constructed as:

$$
\begin{aligned}
\mathcal{P}_1: \;\; &\max_{\mathbf{W}} \; t \\
s.t. \;\; &\mathbf{h}_k^T \mathbf{W}\mathbf{s} = \lambda_k s_k, \; \forall k \in \mathcal{K} \\
&\left(\lambda_k^{\Re} - t\right)\tan\theta_t \geq \left|\lambda_k^{\Im}\right|, \; \forall k \in \mathcal{K} \\
&\|\mathbf{W}\mathbf{s}\|_2^2 \leq p_0
\end{aligned}
\tag{4}
$$

where $\mathcal{K} = \{1, 2, \cdots, K\}$. In $\mathcal{P}_1$, we have enforced a symbol-level power constraint, since the interference exploitation precoding is dependent on the data symbol vector $\mathbf{s}$. $\mathcal{P}_1$ belongs to the second-order cone programming (SOCP) and can be solved via existing convex optimization tools.

III. Interference Exploitation Precoding

In this section, we analyze the interference exploitation precoding problem $\mathcal{P}_1$ based on the KKT conditions. Specifically, our derivations in this section and the numerical results in Section IV show that, by exploiting the information of both the channel and the data symbols, CI precoding is capable of spatially multiplexing more data streams than the number of transmit antennas at the BS simultaneously.

Following [14], [19] and based on the observation that $\mathbf{W}\mathbf{s}$ can be viewed as a single vector in the formulation of $\mathcal{P}_1$, it is safe to assume that each $\mathbf{w}_i s_i$ is identical, which leads to a simpler power constraint in the subsequent analysis, given by

$$\sum_{i=1}^{K} s_i^* \mathbf{w}_i^H \mathbf{w}_i s_i \leq \frac{p_0}{K}. \tag{5}$$

We then express $\mathcal{P}_1$ in a standard minimization form as

$$
\begin{aligned}
\mathcal{P}_2: \;\; &\min_{\mathbf{W}} \; -t \\
s.t. \;\; &\mathbf{h}_k^T \sum_{i=1}^{K} \mathbf{w}_i s_i - \lambda_k s_k = 0, \; \forall k \in \mathcal{K} \\
&\left|\lambda_k^{\Im}\right| - \left(\lambda_k^{\Re} - t\right)\tan\theta_t \leq 0, \; \forall k \in \mathcal{K} \\
&\sum_{i=1}^{K} s_i^* \mathbf{w}_i^H \mathbf{w}_i s_i \leq \frac{p_0}{K}
\end{aligned}
\tag{6}
$$

and the Lagrangian of $\mathcal{P}_1$ is given by [20]

$$
\begin{aligned}
\mathcal{L}\left(\mathbf{W}, t, \delta_k, \tau_k, \delta_0\right) = &-t + \sum_{k=1}^{K} \delta_k \left(\mathbf{h}_k \sum_{i=1}^{K} \mathbf{w}_i s_i - \lambda_k s_k\right) \\
&+ \sum_{k=1}^{K} \tau_k \left[\left|\lambda_k^{\Im}\right| - \left(\lambda_k^{\Re} - t\right)\tan\theta_t\right] \\
&+ \delta_0 \left(\sum_{i=1}^{K} s_i^H \mathbf{w}_i^H \mathbf{w}_i s_i - \frac{p_0}{K}\right),
\end{aligned}
\tag{7}
$$

where δ_k, $\tau_k \geq 0$ and $\delta_0 \geq 0$ are the dual variables corresponding to each constraint of $\mathcal{P}_2$. Based on (6) and the fact that $s_i s_i^H = 1$, $\forall i \in \mathcal{K}$, the KKT conditions can be formulated as [20]

$$\frac{\partial \mathcal{L}}{\partial t} = -1 + \tan \theta_t \sum_{k=1}^{K} \tau_k = 0 \tag{8a}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}_i} = \left(\sum_{k=1}^{K} \delta_k \cdot \mathbf{h}_k \right) s_i + \delta_0 \cdot \mathbf{w}_i^H = \mathbf{0}, \ \forall i \in \mathcal{K} \tag{8b}$$

$$\delta_k \left(\mathbf{h}_k \sum_{i=1}^{K} \mathbf{w}_i s_i - \lambda_k s_k \right) = 0, \ \forall k \in \mathcal{K} \tag{8c}$$

$$\tau_k \left[|\lambda_k^{\Im}| - (\lambda_k^{\Re} - t) \tan \theta_t \right] = 0, \ \forall k \in \mathcal{K} \tag{8d}$$

$$\delta_0 \left(\sum_{i=1}^{K} s_i^H \mathbf{w}_i^H \mathbf{w}_i s_i - \frac{p_0}{K} \right) = 0 \tag{8e}$$

We first obtain that $\delta_0 \neq 0$ based on (8b), and accordingly we can express $\mathbf{w}_i^H$ as

$$\mathbf{w}_i^H = -\left(\sum_{k=1}^{K} \frac{\delta_k}{\delta_0} \cdot \mathbf{h}_k \right) s_i, \ \forall i \in \mathcal{K}. \tag{9}$$

By introducing a new variable $\upsilon_k = -\frac{\delta_k^H}{\delta_0}$, where we note that δ_k can be complex, we can further obtain the expression of $\mathbf{w}_i$, given by

$$\mathbf{w}_i = \left(\sum_{k=1}^{K} \upsilon_k \cdot \mathbf{h}_k^H \right) s_i^H, \ \forall i \in \mathcal{K}, \tag{10}$$

and we can further obtain the expression of each $\mathbf{w}_i s_i$ as

$$\mathbf{w}_i s_i = \sum_{k=1}^{K} \upsilon_k \cdot \mathbf{h}_k^H, \tag{11}$$

which is constant for $\forall i \in \mathcal{K}$ and is consistent with our premise for the power constraint transformation in (5). Based on (10), we are now able to express the precoding matrix $\mathbf{W}$ as

$$\begin{aligned}
\mathbf{W} &= [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_K] \\
&= \left(\sum_{k=1}^{K} \upsilon_k \cdot \mathbf{h}_k^H \right) \cdot \left[s_1^H, s_2^H, \cdots, s_K^H \right] \\
&= \left[\mathbf{h}_1^H, \mathbf{h}_2^H, \cdots, \mathbf{h}_K^H \right] [\upsilon_1, \upsilon_2, \cdots, \upsilon_K]^T \left[s_1^H, s_2^H, \cdots, s_K^H \right] \\
&= \mathbf{H}^H \mathbf{\Upsilon} \mathbf{s}^H,
\end{aligned} \tag{12}$$

Based on (1) and (2), we can express the received signal vector excluding noise $\mathbf{HWs}$ as

$$\mathbf{HWs} = diag(\mathbf{\Lambda}) \mathbf{s}, \tag{13}$$

where $\mathbf{\Lambda} = [\lambda_1, \lambda_2, \cdots, \lambda_K]^T \in \mathcal{C}^{K \times 1}$ denotes the pre-scaling vector. By substituting (12) into (13), we further obtain

$$\begin{aligned}
&\mathbf{HH}^H \mathbf{\Upsilon} \mathbf{s}^H \mathbf{s} = diag(\mathbf{\Lambda}) \mathbf{s} \\
&\Rightarrow \mathbf{\Upsilon} = \frac{1}{K} \cdot (\mathbf{HH}^H)^+ diag(\mathbf{\Lambda}) \mathbf{s},
\end{aligned} \tag{14}$$

where we note that based on the premise that $K > N_t$, the matrix $\mathbf{HH}^H$ is rank-deficient and the exact matrix inverse is

inapplicable. Therefore, pseudo inverse has to be employed in (14), and subsequently we can obtain the optimal precoding structure as a function of the pre-scaling vector, given by

$$\mathbf{W} = \frac{1}{K} \cdot \mathbf{H}^H (\mathbf{HH}^H)^+ diag(\mathbf{\Lambda}) \mathbf{ss}^H. \tag{15}$$

Based on the fact that $\delta_0 \neq 0$, we obtain that the power constraint is strictly active. Similar to the analysis for the case of $K \leq N_t$ in [19], by substituting the expression of $\mathbf{W}$ into the power constraint $\|\mathbf{Ws}\|_2^2 = p_0$, one can similarly transform the power constraint on $\mathbf{W}$ into a power constraint on the pre-scaling vector $\mathbf{\Lambda}$, given by

$$\begin{aligned}
&\|\mathbf{Ws}\|_2^2 = p_0 \\
&\Rightarrow \mathbf{s}^H \mathbf{W}^H \mathbf{Ws} = p_0 \\
&\Rightarrow \mathbf{s}^H diag(\mathbf{\Lambda}^H) (\mathbf{HH}^H)^+ diag(\mathbf{\Lambda}) \mathbf{s} = p_0 \\
&\Rightarrow \mathbf{\Lambda}^H \underbrace{diag(\mathbf{s}^H) (\mathbf{HH}^H)^+ diag(\mathbf{s})}_{\mathbf{P}} \mathbf{\Lambda} = p_0.
\end{aligned} \tag{16}$$

Then, one can follow a similar approach in [19] to obtain an equivalent QP formulation.

However, in the case of $K > N_t$ considered in this paper, we note that following the above procedure and using (16) will only lead to an erroneous solution for CI precoding, which is due to the fact that the inclusion of the pseudo inverse does not guarantee equality to the original constraint. To be more specific, let's first consider the conventional case of $K \leq N_t$, where the optimal precoding structure is given by [19]

$$\mathbf{W} = \frac{1}{K} \cdot \mathbf{H}^H (\mathbf{HH}^H)^{-1} diag(\mathbf{\Lambda}) \mathbf{ss}^H. \tag{17}$$

In this case, by substituting the expression of $\mathbf{W}$ in (17) into (13), we obtain

$$\begin{aligned}
&\mathbf{HWs} = diag(\mathbf{\Lambda}) \mathbf{s} \\
&\Rightarrow \mathbf{H} \left[\frac{1}{K} \cdot \mathbf{H}^H (\mathbf{HH}^H)^{-1} diag(\mathbf{\Lambda}) \mathbf{ss}^H \right] \mathbf{s} = diag(\mathbf{\Lambda}) \mathbf{s} \\
&\Rightarrow diag(\mathbf{\Lambda}) \mathbf{s} = diag(\mathbf{\Lambda}) \mathbf{s},
\end{aligned} \tag{18}$$

which is always true. This in fact means that the pre-scaling constraint in (13) is already included in the power constraint implicitly, for the case of $K \leq N_t$. On the contrary, in the considered scenario of $K > N_t$ in this paper where the pseudo inverse is included, the above equality will not hold, and simply following a similar approach to the case of $K \leq N_t$ in [19] will lead to invalid and erroneous solutions. Therefore, an additional constraint is required to guarantee that the inclusion of pseudo inverse still meets the pre-scaling requirement, given

by

$$\mathbf{HWs} = diag\left(\boldsymbol{\Lambda}\right)\mathbf{s}$$
$$\Rightarrow \mathbf{H}\left[\frac{1}{K}\cdot\mathbf{H}^H\left(\mathbf{HH}^H\right)^+diag\left(\boldsymbol{\Lambda}\right)\mathbf{ss}^H\right]\mathbf{s} = diag\left(\boldsymbol{\Lambda}\right)\mathbf{s}$$
$$\Rightarrow \mathbf{HH}^H\left(\mathbf{HH}^H\right)^+diag\left(\boldsymbol{\Lambda}\right)\mathbf{s} = diag\left(\boldsymbol{\Lambda}\right)\mathbf{s}$$
$$\Rightarrow \left[\mathbf{HH}^H\left(\mathbf{HH}^H\right)^+ - \mathbf{I}\right]diag\left(\boldsymbol{\Lambda}\right)\mathbf{s} = \mathbf{0}$$
$$\Rightarrow \underbrace{\left[\mathbf{HH}^H\left(\mathbf{HH}^H\right)^+ - \mathbf{I}\right]diag\left(\mathbf{s}\right)}_{\mathbf{T}}\boldsymbol{\Lambda} = \mathbf{0}$$
(19)

Based on (19), first we observe that $\boldsymbol{\Lambda} = \mathbf{0}$ is obviously not a valid solution to the original CI precoding. Accordingly, this additional constraint is equivalent to finding non-zero solutions to the linear equation set $\mathbf{T}\boldsymbol{\Lambda} = \mathbf{0}$. Noting that both $\mathbf{T}$ and $\boldsymbol{\Lambda}$ are complex, we first transform them into their real equivalence, given by

$$\mathbf{T_E} = \begin{bmatrix} \Re\left(\mathbf{T}\right) & -\Im\left(\mathbf{T}\right) \\ \Im\left(\mathbf{T}\right) & \Re\left(\mathbf{T}\right) \end{bmatrix}, \ \boldsymbol{\Lambda_E} = \begin{bmatrix} \Re\left(\boldsymbol{\Lambda}\right) \\ \Im\left(\boldsymbol{\Lambda}\right) \end{bmatrix}, \quad (20)$$

and we further express the singular value decomposition (SVD) of $\mathbf{T_E}$ as

$$\mathbf{T_E} = \mathbf{S}\boldsymbol{\Sigma}\hat{\mathbf{D}}^H, \quad (21)$$

where $\hat{\mathbf{D}} = \left[\hat{\mathbf{d}}_1, \hat{\mathbf{d}}_2, \cdots, \hat{\mathbf{d}}_{2K}\right]$ is the right singular matrix that consists of right singular vectors. Based on the linear algebra theory [21], the non-zero solution $\boldsymbol{\Lambda_E}$ is therefore in the null space of $\mathbf{T_E}$, which can be expressed as a linear combination of the right singular vectors that correspond to zero singular values, given by

$$\boldsymbol{\Lambda_E} = \sum_{n=1}^{2K-\mathrm{rank}\{\mathbf{T_E}\}} \beta_n\cdot\hat{\mathbf{d}}_{\mathrm{rank}\{\mathbf{T_E}\}+n} = \mathbf{D}\boldsymbol{\beta}, \quad (22)$$

where each β_n is real and $\boldsymbol{\beta} = \left[\beta_1, \beta_2, \cdots, \beta_{2K-\mathrm{rank}\{\mathbf{T_E}\}}\right]^T$. $\mathbf{D}$ consists of right singular vectors corresponding to zero singular values, given by

$$\mathbf{D} = \left[\hat{\mathbf{d}}_{\mathrm{rank}\{\mathbf{T_E}\}+1}, \hat{\mathbf{d}}_{\mathrm{rank}\{\mathbf{T_E}\}+2}, \cdots, \hat{\mathbf{d}}_{2K}\right]$$
$$= \left[\mathbf{d}_1^T, \mathbf{d}_2^T, \cdots, \mathbf{d}_{2K}^T\right]^T, \quad (23)$$

where each $\mathbf{d}_k^T$ represents the k-th row of $\mathbf{D}$. Subsequently, we expand the left-hand side of (16) into its real equivalence, given by

$$\boldsymbol{\Lambda_E}^T\underbrace{\begin{bmatrix} \Re\left(\mathbf{P}\right) & -\Im\left(\mathbf{P}\right) \\ \Im\left(\mathbf{P}\right) & \Re\left(\mathbf{P}\right) \end{bmatrix}}_{\mathbf{P_E}}\boldsymbol{\Lambda_E} = p_0 \Rightarrow \boldsymbol{\beta}^T\underbrace{\mathbf{D}^T\mathbf{P_E}\mathbf{D}}_{\mathbf{Q_E}}\boldsymbol{\beta} = p_0,$$
(24)

which is the valid power constraint for the case of $K > N_t$ considered in this paper, and we further note that $\mathbf{Q_E}$ is symmetric.

Based on the above analysis, we can now formulate an equivalent optimization on the weight vector $\boldsymbol{\beta}$, given by

$$\mathcal{P}_3: \ \min_{\boldsymbol{\beta}} \ -t$$
$$s.t. \quad \boldsymbol{\beta}^T\mathbf{Q_E}\boldsymbol{\beta} - p_0 = 0$$
$$\frac{\mathbf{d}_{k+K}^T\boldsymbol{\beta}}{\tan\theta_t} + t - \mathbf{d}_k^T\boldsymbol{\beta} \le 0, \ \forall k\in\mathcal{K} \quad (25)$$
$$-\frac{\mathbf{d}_{k+K}^T\boldsymbol{\beta}}{\tan\theta_t} + t - \mathbf{d}_k^T\boldsymbol{\beta} \le 0, \ \forall k\in\mathcal{K}$$

where we have transformed the CI constraint with the absolute value into two separate constraints. The Lagrangian of $\mathcal{P}_3$ is constructed as

$$\mathcal{L}\left(\boldsymbol{\beta}, t, \alpha_0, \mu_k, \nu_k\right) = -t + \alpha_0\left(\boldsymbol{\beta}^T\mathbf{Q_E}\boldsymbol{\beta} - p_0\right)$$
$$+ \sum_{k=1}^{K}\mu_k\left(\frac{\mathbf{d}_{k+K}^T\boldsymbol{\beta}}{\tan\theta_t} + t - \mathbf{d}_k^T\boldsymbol{\beta}\right)$$
$$+ \sum_{k=1}^{K}\nu_k\left(-\frac{\mathbf{d}_{k+K}^T\boldsymbol{\beta}}{\tan\theta_t} + t - \mathbf{d}_k^T\boldsymbol{\beta}\right),$$
(26)

where α_0, $\mu_k \ge 0$ and $\nu_k \ge 0$ are the corresponding dual variables. By defining

$$\mathbf{u} = \left[\mu_1, \cdots, \mu_K, \nu_1, \cdots, \nu_K\right]^T, \ \mathbf{S} = \begin{bmatrix} \mathbf{I} & -\frac{1}{\tan\theta_t}\cdot\mathbf{I} \\ \mathbf{I} & \frac{1}{\tan\theta_t}\cdot\mathbf{I} \end{bmatrix},$$
(27)

where $\mathbf{u}\in\mathcal{R}^{2K\times 1}$ and $\mathbf{S}\in\mathcal{R}^{2K\times 2K}$. The Lagrangian of $\mathcal{P}_3$ can be further simplified into

$$\mathcal{L}\left(\boldsymbol{\beta}, t, \alpha_0, \mathbf{u}\right) = \left(\mathbf{1}^T\mathbf{u} - 1\right)t + \alpha_0\cdot\boldsymbol{\beta}^T\mathbf{Q_E}\boldsymbol{\beta} - \mathbf{u}^T\mathbf{SD}\boldsymbol{\beta} - \alpha_0 p_0,$$
(28)

based on which the KKT conditions for $\mathcal{P}_3$ are given by

$$\frac{\partial\mathcal{L}}{\partial t} = \mathbf{1}^T\mathbf{u} - 1 = 0 \quad (29a)$$
$$\frac{\partial\mathcal{L}}{\partial\boldsymbol{\beta}} = 2\alpha_0\cdot\mathbf{Q_E}\boldsymbol{\beta} - \mathbf{D}^T\mathbf{S}^T\mathbf{u} = \mathbf{0} \quad (29b)$$
$$\boldsymbol{\beta}^T\mathbf{Q_E}\boldsymbol{\beta} - p_0 = 0 \quad (29c)$$
$$\mu_k\left(\frac{\mathbf{d}_{k+K}^T\boldsymbol{\beta}}{\tan\theta_t} + t - \mathbf{d}_k^T\boldsymbol{\beta} \le 0\right) = 0, \ \forall k\in\mathcal{K} \quad (29d)$$
$$\nu_k\left(-\frac{\mathbf{d}_{k+K}^T\boldsymbol{\beta}}{\tan\theta_t} + t - \mathbf{d}_k^T\boldsymbol{\beta} \le 0\right) = 0, \ \forall k\in\mathcal{K} \quad (29e)$$

Based on (29b) we obtain $\alpha_0 \ne 0$ and the expression of $\boldsymbol{\beta}$ as a function of $\mathbf{u}$, given by

$$\boldsymbol{\beta} = \frac{1}{2\alpha_0}\cdot\mathbf{Q_E}^{-1}\mathbf{D}^T\mathbf{S}^T\mathbf{u}. \quad (30)$$

By substituting the expression of $\boldsymbol{\beta}$ into (24), we can further obtain the expression of α_0, given by

$$\left(\frac{1}{2\alpha_0}\cdot\mathbf{Q_E}^{-1}\mathbf{D}^T\mathbf{S}^T\mathbf{u}\right)^T\mathbf{Q_E}\left(\frac{1}{2\alpha_0}\cdot\mathbf{Q_E}^{-1}\mathbf{D}^T\mathbf{S}^T\mathbf{u}\right) = p_0$$
$$\Rightarrow \alpha_0 = \sqrt{\frac{\mathbf{u}^T\mathbf{SD}\mathbf{Q_E}^{-1}\mathbf{D}^T\mathbf{S}^T\mathbf{u}}{4p_0}}$$
(31)

$$\mathbf{W} = \frac{1}{K} \cdot \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^+ diag \left\{ \sqrt{\frac{p_0}{\mathbf{u}^T \mathbf{S} \mathbf{D} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u}}} \cdot \mathbf{U} \mathbf{D} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u} \right\} \mathbf{s}\mathbf{s}^H \tag{34}$$

For $\mathcal{P}_3$, it is easy to verify that the Slater's condition is satisfied, and therefore we consider the dual problem of $\mathcal{P}_3$:

$$\begin{aligned}
\mathcal{D} &= \max_{\alpha_0, \mathbf{u}} \min_{\boldsymbol{\beta}, t} \mathcal{L}(\beta, t, \alpha_0, \mathbf{u}) \\
&= \max_{\alpha_0, \mathbf{u}} \alpha_0 \left(\frac{1}{2\alpha_0} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u} \right)^T \mathbf{Q}_{\mathbf{E}} \left(\frac{1}{2\alpha_0} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u} \right) \\
&\quad - \mathbf{u}^T \mathbf{S} \mathbf{D} \left(\frac{1}{2\alpha_0} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u} \right) - \alpha_0 p_0 \\
&= \max_{\alpha_0, \mathbf{u}} \frac{1}{4\alpha_0} \cdot \mathbf{u}^T \mathbf{S} \mathbf{D} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u} - \alpha_0 p_0 \\
&= \max_{\mathbf{u}} -\frac{\mathbf{u}^T \mathbf{S} \mathbf{D} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u}}{4 \sqrt{\frac{\mathbf{u}^T \mathbf{S} \mathbf{D} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u}}{4 p_0}}} - p_0 \sqrt{\frac{\mathbf{u}^T \mathbf{S} \mathbf{D} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u}}{4 p_0}} \\
&= \max_{\mathbf{u}} -\sqrt{p_0 \cdot \mathbf{u}^T \mathbf{S} \mathbf{D} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \mathbf{u}}.
\end{aligned} \tag{32}$$

Based on the monotonicity of the square root function, the maximization on the dual in (32) is equivalent to the following optimization:

$$\begin{aligned}
\mathcal{P}_4: \quad & \min_{\mathbf{u}} \mathbf{u}^T \left(\mathbf{S} \mathbf{D} \mathbf{Q}_{\mathbf{E}}^{-1} \mathbf{D}^T \mathbf{S}^T \right) \mathbf{u} \\
s.t. \quad & \mathbf{1}^T \mathbf{u} - 1 = 0 \\
& u_k \geq 0, \ \forall k \in \hat{\mathcal{K}}
\end{aligned} \tag{33}$$

where the first constraint is from (29a), $\hat{\mathcal{K}} = \{1, 2, \cdots, 2K\}$, and u_k is the k-th entry in $\mathbf{u}$.

Based on our above transformations and analysis, we have shown that the original CI precoding problem $\mathcal{P}_1$ can be equivalently solved by a simplified optimization $\mathcal{P}_4$. Moreover, based on the expression of $\boldsymbol{\Lambda}_{\mathbf{E}}$ in (22), $\boldsymbol{\beta}$ in (30), and α_0 in (31), we can express the optimal precoding matrix $\mathbf{W}$ as a function of the dual vector $\mathbf{u}$ in a closed form, which is shown in (34) at the top of this page, where $\mathbf{U} = \begin{bmatrix} \mathbf{I} & j \cdot \mathbf{I} \end{bmatrix}$ transforms the expanded pre-scaling vector $\boldsymbol{\Lambda}_{\mathbf{E}}$ into its complex equivalence $\boldsymbol{\Lambda}$.

Compared to the original CI precoding optimization in $\mathcal{P}_1$ which is a SOCP optimization, it is observed that $\mathcal{P}_4$ is a standard QP optimization over a simplex. It has been shown in the literature that such a QP formulation can be more efficiently solved than the SOCP formulation using the simplex or interior-point methods [22], [23], and the iterative closed-form algorithm proposed in [19] can also be directly employed to solve $\mathcal{P}_4$ with reduced computational costs.

A. Condition for Spatially Multiplexing $K > N_t$ Streams

Based on the above, we can also obtain the expression of t^* when the optimality of $\mathcal{P}_3$ is achieved, given by

$$t^* = \min_k \left\{ d_k^T \boldsymbol{\beta} - \frac{d_{k+K}^T \boldsymbol{\beta}}{\tan \theta_t}, \ d_k^T \boldsymbol{\beta} + \frac{d_{k+K}^T \boldsymbol{\beta}}{\tan \theta_t} \right\}, \ \forall k \in \hat{\mathcal{K}}. \tag{35}$$

If $t^* > 0$, we obtain a valid pre-scaling vector $\boldsymbol{\Lambda}$ and a corresponding valid precoding matrix $\mathbf{W}$. Otherwise if $t^* \leq 0$, it means that the data symbols will be scaled and rotated to other three quarters of the constellation, which only leads to erroneous detection. Accordingly, whether the obtained

$$\min_k \left\{ d_k^T \boldsymbol{\beta} - \frac{d_{k+K}^T \boldsymbol{\beta}}{\tan \theta_t}, \ d_k^T \boldsymbol{\beta} + \frac{d_{k+K}^T \boldsymbol{\beta}}{\tan \theta_t} \right\} > 0 \tag{36}$$

is the condition under which multiplexing $K > N_t$ streams is feasible.

IV. NUMERICAL RESULTS

Numerical results are presented based on Monte Carlo simulations in this section. We assume the total transmit power $p_0 = 1$, and define the transmit SNR as $\rho = 1/\sigma^2$. Since ZF precoding and SINR balancing precoding are not feasible for the scenario of $K > N_t$, we compare our QP-based CI precoding with closed-form RZF precoding and traditional SOCP-based CI precoding. Both QPSK and 8PSK modulations are considered in the simulations.

Before we present the bit error rate (BER) performance, we first depict the feasibility probability with respect to the number of streams K in Fig. 2, where the number of transmit antennas N_t varies from $N_t = 4$ to $N_t = 12$. Generally, for a specific feasibility target, we observe that a larger number of transmit antennas at the BS can support more streams than that of transmit antennas, i.e., a larger N_t leads to a larger $(K - N_t)$. Specifically when $N_t = 12$ for QPSK, CI precoding is able to support $K = 18$ streams simultaneously with a feasibility probability higher than 95%. For the following BER results in Fig. 3 and 4, RZF precoding is employed instead when CI precoding is not feasible in the simulations.

The BER results of CI precoding are presented in Fig. 3 with respect to the transmit SNR ρ, where we consider two scenarios $K = 9$, $N_t = 8$ and $K = 10$, $N_t = 8$ for both QPSK and 8PSK modulation. When $K > N_t$, traditional ZF
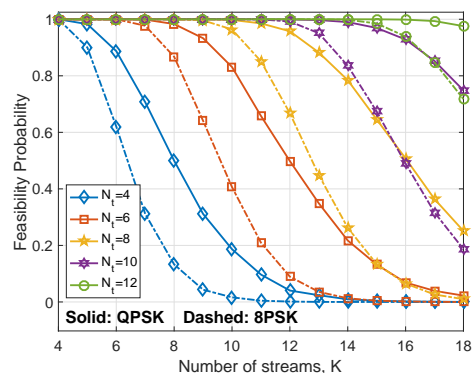

Fig. 2: Feasibility v.s. number of streams K

of interference exploitation precoding over traditional RZF precoding.

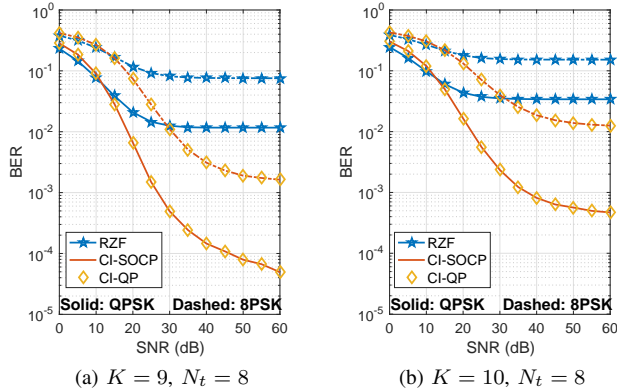(a) $K = 9$, $N_t = 8$ (b) $K = 10$, $N_t = 8$

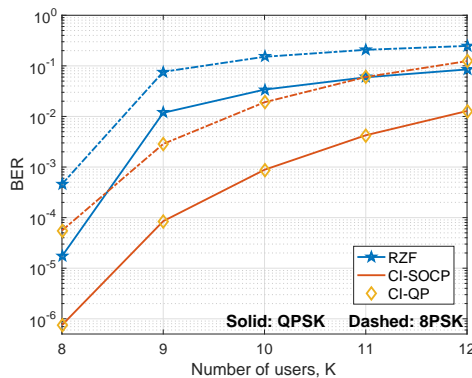Fig. 3: BER v.s. transmit SNR, $N_t = 8$

Fig. 4: BER v.s. number of streams K, $N_t = 8$, SNR=40dB

precoding and SINR balancing precoding are inapplicable, and therefore we compare with RZF precoding only. In Fig. 3, compared to RZF precoding where an error floor is observed, CI precoding achieves a significant performance gain in the medium-to-high SNR regime, which validates the superiority of CI precoding over traditional RZF precoding.

In Fig. 4, we show the BER results of CI precoding with an increasing number of users K for both QPSK and 8PSK modulation, where $N_t = 8$ and $\rho = 40$dB. For both modulations considered in Fig. 4, we observe a significant gain of CI precoding over traditional RZF precoding, when $K > N_t$. The performance gains become less significant when K increases, which is due to a lower feasibility probability for CI precoding, as observed in Fig. 2.

V. CONCLUSION

In this paper, the interference exploitation precoding for the scenario where the BS serves a larger number of users than that of the transmit antennas is studied. By analyzing the optimization problem with KKT conditions and by formulating the dual problem, we obtain the closed-form optimal precoding matrix as a function of the dual vector, as well a QP optimization that efficiently obtains the optimal dual vector. Numerical results validate the optimality of the closed-form precoding matrix, and reveal significant performance gains

REFERENCES

[1] L. Zheng and D. N. C. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[2] M. Costa, "Writing on Dirty Paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[3] L. Sun and M. Lei, "Quantized CSI-based Tomlinson-Harashima Precoding in Multiuser MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1118–1126, Mar. 2013.

[4] 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.

[5] 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. 195–202, Jan. 2005.

[6] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear Precoding via Conic Optimization for Fixed MIMO Receivers," *IEEE Trans. Sig. Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[7] M. Schubert and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem with Individual SINR Constraints," *IEEE Trans. Veh. Tech.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[8] M. Bengtsson and B. Ottersten, "Optimal and Suboptimal Transmit Beamforming," *Handbook of Antennas in Wireless Communications*, Jan. 2001.

[9] 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.

[10] G. Zheng, I. Krikidis, C. Masouros, S. Timotheou, D. A. Toumpakaris, and Z. Ding, "Rethinking the Role of Interference in Wireless Networks," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 152–158, Nov. 2014.

[11] C. Masouros and E. Alsusa, "Dynamic Linear Precoding for the Exploitation of Known Interference in MIMO Broadcast Systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1396–1401, Mar. 2009.

[12] C. Masouros, "Correlation Rotation Linear Precoding for MIMO Broadcast Communicaitons," *IEEE Trans. Sig. Process.*, vol. 59, no. 1, pp. 252–262, Jan. 2011.

[13] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector Perturbation based on Symbol Scaling for Limited Feedback MISO Downlinks," *IEEE Trans. Sig. Process.*, vol. 62, no. 3, pp. 562–571, Feb. 2014.

[14] C. Masouros and G. Zheng, "Exploiting Known Interference as Green Signal Power for Downlink Beamforming Optimization," *IEEE Trans. Sig. Process.*, vol. 63, no. 14, pp. 3628–3640, July 2015.

[15] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive Multiuser Interference in Symbol Level Precoding for the MISO Downlink Channel," *IEEE Trans. Sig. Process.*, vol. 63, no. 9, pp. 2239–2252, May 2015.

[16] ——, "Energy-Efficient Symbol-Level Precoding in Multiuser MISO based on Relaxed Detection Region," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3755–3767, May 2016.

[17] ——, "Symbol-Level Multiuser MISO Precoding for Multi-Level Adaptive Modulation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5511–5524, Aug. 2017.

[18] M. Alodeh, D. Spano, A. Kalantari, C. Tsinos, D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Symbol-Level and Multicast Precoding for Multiuser Multiantenna Downlink: A Survey, Classification and Challenges," *arXiv preprint*, available online: https://arxiv.org/abs/1703.03617, 2018.

[19] A. Li and C. Masouros, "Interference Exploitation Precoding Made Practical: Optimal Closed-Form Solutions for PSK Modulations," *IEEE Trans. Wireless Commun.*, p. DOI:10.1109/TWC.2018.2869382, 2018.

[20] L. Vandenberghe and S. Boyd, *Convex Optimization*. Cambridge University Press, 2004.

[21] D. C. Lay, S. R. Lay, and J. J. McDonald, *Linear Algebra and Its Applications*, 5th ed. Pearson, 2015.

[22] G. Cornuejols and R. Tutuncu, *Optimization Methods in Finance*. Cambridge University Press, Dec. 2006.

[23] F. Alizadeh and D. Goldfarb, "Second-Order Cone Programming," *Mathematical Programming*, vol. 95, no. 1, pp. 3–51, 2003.