This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2692222, IEEE Transactions on Communications

A Convex Approach to Near-Optimal Beamforming Designs for Two-User MISO Fading Interference Channels

Sang Won Choi, *Member, IEEE* and Juyeop Kim, *Member, IEEE*

***Abstract*—Based on a convex approach with side-information, we propose transmit beamforming designs for two-user multiple-input single-output fading interference channels. Main contribution of this paper is to provide a novel methodology for solving a non-convex optimization problem efficiently based on the effective side-information. Consequently, the proposed scheme exhibits near-optimal average sum-rate performance under single user detection with Gaussian inputs, which is validated through numerical results. As a by-product, we show that the proposed scheme requires almost no parameter optimization on the average over multiple coding blocks.**

***Index Terms*—Convex optimization, correlated block fading, distributed beamforming design, interference channel, and side-information**

I. Introduction

MULTI-USER interference management is one of significant technical issues for efficient multi-user communications, which is confronted inevitably from resource sharing such as time, frequency, and space. From information-theoretical point of view, understanding two-user Interference Channel (IC) can give us insight on how to manage multi-user interference effectively [1]-[8]. Specifically, it has been shown to be optimal to treat the interference in the following ways:

- Treating interference as signal achieves all rate pairs for ICs with strong interference [1]-[4].
- Treating interference as noise achieves the sum capacity for ICs with noisy interference [6]-[8].

For the ICs where the interference is neither strong nor noisy, the optimal way to treat interference is still unknown. Interestingly, Etkin *et al.* showed that even simplified version of Han and Kobayashi's rate region [3] is within one bit of the capacity region [5].

Basically, following the appropriate principles for interference management is a short cut to achieve efficient multi-user distributed communications. However, the practical distributed communication system bears non-negligible overhead and computation limitations inherently. As for the system overhead, one typical overhead is related with codebook sharing among different cells, which creates higher overhead as the number of shared cells increases. The other is the increase of Channel State Information (CSI) for all channel links. With regard to the computational limitation, mobile stations have difficulty in processing heavy computation. Furthermore, miniaturization of mobile stations limits the number of antennas. However, base stations can cover high computational processing and possibly be equipped with multiple antennas. Motivated by those practical considerations, there have been a few researches on Multiple-Input Single-Output (MISO) ICs [9]-[17].

Under the constraints that each transmitter uses Gaussian codebook, and each receiver performs single user detection, maximizing the weighted sum-rate is generally a non-convex optimization problem. One typical way is to solve the non-convex optimization problem directly using a numerical algorithm, but the convergence to the globally optimal solution is often not guaranteed [18]. Another way is to derive the optimal solution in a closed form, which has been shown with the following interesting facts:

- The non-convex optimization problem can be interpreted as multiple convex optimization problems by fixing interfering or desired signal powers from each transmitter [9], [12]-[14].
- Transmit beamforming scheme is sufficient to achieve all rate pairs with Pareto-optimality [12]-[14].
- Pareto-optimal transmit beamforming vector can be represented as a weight linear sum of Maximum Ratio Combining (MRC) and Zero Forcing (ZF) beamformers [10].

In the framework of optimization problem, the number of parameters dominantly determines computational complexity required for seeking optimal transmit beamformers. Intuitively, the number of parameters to be optimized is 2 because the number of interfering or desired links is 2. Recent results have shown that only one parameter optimization is sufficient to acquire the optimal transmit beamformers [15],[17].

Until now, philosophy behind the transmit beamforming design for the MISO fading IC is to optimize transmit beamforming vectors focussing on each channel instance in the sense of maximizing an arbitrary weighted sum-rate for each MISO IC separately[1]. The motivation of this paper is to seek an appropriate side-information and to design near-optimal

S. W. Choi and J. Kim are with Korea Railroad Research Institute (KRRI), Uiwang-si, Gyeonggi-do, 16105.
E-mail: {swchoi, jykim00}@krri.re.kr, J. Kim is the corresponding author of this paper. The material in this paper was published partly in 2014 International Conference on Information and Communication Technology Convergence (ICTC).

[1]For Multiple-Input Multiple-Output (MIMO) ICs, the beamforming tracking algorithms have been proposed using vector perturbation approaches. However, the algorithms cannot be applied to MISO IC because the interferences have been interpreted as signal [19].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2692222, IEEE Transactions on Communications

transmit beamforming design based on the side-information in the framework of the MISO fading IC.

In this paper, we propose near-optimal transmit beamformers for MISO IC with correlated block fading, i.e., the channel states of adjacent coding blocks are correlated each other. Interestingly, our novel convex approach to the transmit beamforming design achieves both of near-optimal throughput performance and efficient computational complexity. As a by-product, we show that the number of parameters to be optimized can be close to 0 on the average over multiple coding blocks in comparison with the state of the art transmit beamforming scheme [20].

The organization of this paper is as follows: In Section II, we describe our system model with notations. We introduce some preliminaries on a convex approach with problem formulations in Section III. Then, we propose our beamforming schemes along with complexity and performance analysis in Section IV. Numerical results are given in Section V to show the validity of our proposed scheme, and we conclude our paper in Section VI.

II. SYSTEM MODEL AND NOTATIONS

We consider a two-user MISO IC with time-varying block fading. Specifically, the received signals of the k-th receiver for a two-user MISO fading IC are defined as

$$y^{[k]}(n) = \mathbf{h}^{[kk]\dagger}(n)\mathbf{v}^{[k]}(n)s^{[k]}(n) + \sum_{j=1,j\neq k}^{2} \mathbf{h}^{[kj]\dagger}(n)\mathbf{v}^{[j]}(n) + m^{[k]}(n) \quad (1)$$

at the n-th channel use, where n indicates the non-negative integer. In (1), $k \in \{1,\ 2\}$ is the user index, $s^{[k]}(n)$ is the transmitted signal, $\mathbf{v}^{[k]}(n) \in \mathbb{C}^{M\times 1}$ is the transmit beamforming vector, and $\mathbf{h}^{[kj]}(n) \in \mathbb{C}^{M\times 1}$ is the channel vector from the j-th transmitter to the k-th receiver, where the i-th component of $\mathbf{h}^{[kj]*}$, denoted as $\mathbf{h}_i^{[kj]*}$ is the channel coefficient from the i-th antenna at the j-th transmitter to the k-th receiver. Here, M is the number of transmit antennas. For notational convenience, we will omit channel use index n in (1). Then, the received signals in even and odd channel uses indices are represented as

$$y^{[k]} = \mathbf{h}^{[kk]\dagger}\mathbf{v}^{[k]}s^{[k]} + \sum_{j=1,j\neq k}^{K} \mathbf{h}^{[kj]\dagger}\mathbf{v}^{[j]}s^{[k]} + m^{[k]} \quad (2)$$

and

$$z^{[k]} = \mathbf{g}^{[kk]\dagger}\mathbf{w}^{[k]}s^{[k]} + \sum_{j=1,j\neq k}^{K} \mathbf{g}^{[kj]\dagger}\mathbf{w}^{[j]}s^{[k]} + n^{[k]}, \quad (3)$$

respectively[2].

Basically, we have the following assumptions:

- The effective transmit signal vectors $\mathbf{v}^{[k]}s^{[k]}$'s ($\mathbf{w}^{[k]}s^{[k]}$'s) for the k-th transmitter is circularly symmetric complex Gaussian with zero-mean and covariance $\mathbf{S}^{[k]}$'s ($\mathbf{T}^{[k]}$'s).
- Perfect local CSI at each transmitter (CSIT) is available so that the k-th transmitter knows $\mathbf{h}^{[kj]}$'s ($\mathbf{g}^{[kj]}$'s) for all $j \in \{1,\ 2\}$[3].
- The noise $m^{[k]} \in \mathbb{C}$ ($n^{[k]} \in \mathbb{C}$) is circularly symmetric complex Gaussian with zero mean and unit variance.
- The channel is correlated block fading in time, where the channel vector $\mathbf{h}^{[kj]}$'s ($\mathbf{g}^{[kj]}$'s) at the previous (present) coding block are fixed within a block.

For notational convenience, we use the following notations.

- Boldface letters $\mathbf{x}$ and $\mathbf{X}$ represent vector and matrix, respectively.
- $\mathbf{0}_{M\times N}$ and $\mathbf{I}_{M\times M}$ are $M \times N$ zero matrix and $M \times M$ identity matrix, respectively.
- $\|\cdot\|$ represents the norm of a vector [22].
- $(\cdot)^T$ and $(\cdot)^\dagger$ denote transpose and Hermitian, respectively.
- For two symmetric matrices $\mathbf{S}$ and $\mathbf{T}$, $\mathbf{S} \succeq \mathbf{T}$ means that $\mathbf{S}-\mathbf{T}$ is positive semi-definite.
- The operation $\mathrm{tr}(\cdot)$ indicates the trace of a matrix.
- From the Singular Value Decomposition (SVD) of $\mathbf{h}^{[ij]}$ and $\mathbf{g}^{[ij]}$, unitary matrices $\mathbf{U}^{[ij]}$ and $\mathbf{V}^{[ij]}$ are defined as:

$$\mathbf{h}^{[ij]} = \mathbf{U}^{[ij]} \begin{pmatrix} \|\mathbf{h}^{[ij]}\| \\ \mathbf{0}_{(M-1)\times 1} \end{pmatrix} \quad (4)$$

and

$$\mathbf{g}^{[ij]} = \mathbf{V}^{[ij]} \begin{pmatrix} \|\mathbf{g}^{[ij]}\| \\ \mathbf{0}_{(M-1)\times 1} \end{pmatrix} \quad (5)$$

for $i,\ j = 1,\ 2$.

- We write $\theta^{[k]}$'s and $\hat{\theta}^{[k]}$'s as

$$\theta^{[1]} = \cos^{-1}\left(\frac{|\mathbf{h}^{[21]\dagger}\mathbf{h}^{[11]}|}{\|\mathbf{h}^{[11]}\|\cdot\|\mathbf{h}^{[21]}\|}\right), \quad (6)$$

$$\theta^{[2]} = \cos^{-1}\left(\frac{|\mathbf{h}^{[12]\dagger}\mathbf{h}^{[22]}|}{\|\mathbf{h}^{[12]}\|\cdot\|\mathbf{h}^{[22]}\|}\right), \quad (7)$$

$$\hat{\theta}^{[1]} = \cos^{-1}\left(\frac{|\mathbf{g}^{[21]\dagger}\mathbf{g}^{[11]}|}{\|\mathbf{g}^{[11]}\|\cdot\|\mathbf{g}^{[21]}\|}\right), \quad (8)$$

and

$$\hat{\theta}^{[2]} = \cos^{-1}\left(\frac{|\mathbf{g}^{[12]\dagger}\mathbf{g}^{[22]}|}{\|\mathbf{g}^{[12]}\|\cdot\|\mathbf{g}^{[22]}\|}\right), \quad (9)$$

respectively, where $\cos^{-1}(\cdot)$ is the inverse function of $\cos(\cdot)$.

- For $k = 1,\ 2$, we define matrices $\tilde{\mathbf{S}}^{[k]}$ and $\tilde{\mathbf{T}}^{[k]}$ as

$$\tilde{\mathbf{S}}^{[k]} = \mathbf{U}^{[kk]\dagger}\mathbf{S}^{[k]}\mathbf{U}^{[kk]} \quad (10)$$

and

$$\tilde{\mathbf{T}}^{[k]} = \mathbf{V}^{[kk]\dagger}\mathbf{T}^{[k]}\mathbf{V}^{[kk]}, \quad (11)$$

where $\tilde{\mathbf{S}}_{11}^{[k]}$ ($\tilde{\mathbf{T}}_{11}^{[k]}$) is the $(1,\ 1)$-th component of $\tilde{\mathbf{S}}^{[k]}$ ($\tilde{\mathbf{T}}^{[k]}$).

[2]Here, we allow abuse of notations for classifying received signals with even and odd channel uses with omitting the channel use index n.

[3]We assume perfect local CSIT to show validity of the proposed beamforming scheme because imperfect local CSI affects throughput performance so that the performance enhancement of the proposed scheme might be not clarified. In a practical sense, perfect local CSIT can be achievable approximately by allocating separate reference signal [21] for channel estimation at each receiver with CSI feedback. In addition, the proposed scheme considers distributed beamforming scheme in a sense of using local CSIT, which reduces the amount of CSI feedback overhead.

III. Problem formulations and preliminaries on a convex approach

In this paper, we assume that each receiver implements single user detection, i.e., the interference is treated as noise. Essentially, weighted sum-rate maximization problem under the assumption is formulated as

$$\begin{aligned}&\text{Maximize} \quad \sum_{k=1}^{2} \mu^{[k]} R^{[k]}\\ &\text{Subject to} \quad \mathrm{tr}(\mathbf{S}^{[k]}) \leq P^{[k]},\ \mathbf{S}^{[k]} \succeq 0,\\ &R^{[k]} = \log_2\left(1+\frac{\mathbf{h}^{[kk]\dagger}\mathbf{S}^{[k]}\mathbf{h}^{[kk]}}{1+\sum_{j=1,j\neq k}^{2}\mathbf{h}^{[kj]\dagger}\mathbf{S}^{[j]}\mathbf{h}^{[kj]}}\right) \end{aligned}\tag{12}$$

given a predetermined weight $\mu^{[k]}$ for the k-th transmitter with $k \in \{1,\ 2\}$. The original non-convex optimization problem (12) has been shown to be transformed into multiple convex optimization problems by fixing the interfering (desired) signal power while maximizing (minimizing) the desired (interfering) signal power [12]-[14]. Here, the transformation is feasible without loss of optimality. Specifically, the problem (12) is equivalent to the following optimization problem.

$$\begin{aligned}&\text{Maximize} \quad \sum_{k=1}^{2} \mu^{[k]} R^{[k]}\\ &\text{Subject to} \quad \mathrm{tr}(\mathbf{S}^{[k]}) \leq P^{[k]},\ \mathbf{S}^{[k]} \succeq 0,\\ &R^{[k]} = \log_2\left(1+\frac{\mathbf{h}^{[kk]\dagger}\mathbf{S}^{[k]}\mathbf{h}^{[kk]}}{1+\sum_{j=1,j\neq k}^{2}\mathbf{h}^{[kj]\dagger}\mathbf{S}^{[j]}\mathbf{h}^{[kj]}}\right),\\ &\mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]} = \left(i^{[1]}\right)^2, \mathbf{h}^{[12]\dagger}\mathbf{S}^{[2]}\mathbf{h}^{[12]} = \left(i^{[2]}\right)^2, \end{aligned}\tag{13}$$

where $\left(i^{[k]}\right)^2$ indicates interfering signal power from the k-th transmitter with the covariance matrix $\mathbf{S}^{[k]}$. Implicitly, the precedent search for optimal interfering or desired signal power is required for solving transformed multiple convex optimization problems as an alternative of solving non-convex optimization problem. Consequently, the form of subsequent possible convex optimization problem for solving (12) can be characterized as one of the followings:

$$\begin{aligned}&\text{Maximize} && \mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]}\\ &\text{Subject to} && \mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]} = \left(i^{[1]}\right)^2,\\ &&& \mathrm{tr}(\mathbf{S}^{[1]}) \leq P^{[1]},\ \mathbf{S}^{[1]} \succeq 0, \end{aligned}\tag{14}$$

$$\begin{aligned}&\text{Maximize} && \mathbf{h}^{[22]\dagger}\mathbf{S}^{[2]}\mathbf{h}^{[22]}\\ &\text{Subject to} && \mathbf{h}^{[12]\dagger}\mathbf{S}^{[2]}\mathbf{h}^{[12]} = \left(i^{[2]}\right)^2,\\ &&& \mathrm{tr}(\mathbf{S}^{[2]}) \leq P^{[2]},\ \mathbf{S}^{[2]} \succeq 0, \end{aligned}\tag{15}$$

$$\begin{aligned}&\text{Minimize} && \mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]}\\ &\text{Subject to} && \mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]} = \left(d^{[1]}\right)^2,\\ &&& \mathrm{tr}(\mathbf{S}^{[1]}) \leq P^{[1]},\ \mathbf{S}^{[1]} \succeq 0, \end{aligned}\tag{16}$$

and

$$\begin{aligned}&\text{Minimize} && \mathbf{h}^{[12]\dagger}\mathbf{S}^{[2]}\mathbf{h}^{[12]}\\ &\text{Subject to} && \mathbf{h}^{[22]\dagger}\mathbf{S}^{[2]}\mathbf{h}^{[22]} = \left(d^{[2]}\right)^2,\\ &&& \mathrm{tr}(\mathbf{S}^{[2]}) \leq P^{[2]},\ \mathbf{S}^{[2]} \succeq 0, \end{aligned}\tag{17}$$

where $z^{[k]} \geq 0$ and $d^{[k]} \geq 0$ for $k = 1,\ 2$. The original non-convex optimization problem (12) can be solved by considering two sets of convex optimization problems. Specifically, one set corresponds to (14) or (16), and the other set becomes (15) or (17). Here, key concept is that maximizing desired signal power given constant interfering signal power is equivalent to minimizing interfering signal power given constant desired signal power for achieving optimal signal to interference and noise ratio (SINR).

Furthermore, by relaxing a constraint of interfering or desired signal power, (14) and (16) can be written as

$$\begin{aligned}&\text{Maximize} && \mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]}\\ &\text{Subject to} && \mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]} \leq \left(i^{[1]}\right)^2,\\ &&& \mathrm{tr}(\mathbf{S}^{[1]}) \leq P^{[1]},\ \mathbf{S}^{[1]} \succeq 0 \end{aligned}\tag{18}$$

and

$$\begin{aligned}&\text{Minimize} && \mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]}\\ &\text{Subject to} && \mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]} \geq \left(d^{[1]}\right)^2,\\ &&& \mathrm{tr}(\mathbf{S}^{[1]}) \leq P^{[1]},\ \mathbf{S}^{[1]} \succeq 0. \end{aligned}\tag{19}$$

For any combinations of two sets with the framework of convex optimization problems, the 2 parameters are to be optimized for Pareto-optimality in two-user MISO ICs. However, if we use the representation of transmit beamformers with a linear combination of MRC and ZF transmit beamformers, the number of parameters can be further reduced to 1 [17], which is feasible by using the following facts:

- When we relax the constraints from equality to inequality, the Karush-Kuhn-Tucker (KKT) conditions are sufficient conditions for global optimality.
- By using parameterizations in [10], a parameter determining Pareto-optimal transmit beamformer for one transmitter determines the other transmitter's parameter with Pareto-optimality.

A. Convex approach 1. Maximizing desired signal power with an interfering signal power constraint

In the following, we consider the optimal covariance matrix for (14). For that, we begin with the following notations.

- We define $\tilde{\mathbf{h}}^{[11]}$, $\tilde{\mathbf{h}}^{[22]}$, $\tilde{\mathbf{h}}^{[21]}$, and $\tilde{\mathbf{h}}^{[12]}$ as

$$\tilde{\mathbf{h}}^{[11]} = \mathbf{U}^{[21]\dagger}\mathbf{h}^{[11]} = \begin{pmatrix}\tilde{\mathbf{h}}_1^{[11]}\\ \alpha^{[1]}\end{pmatrix},\tag{20}$$

$$\tilde{\mathbf{h}}^{[22]} = \mathbf{U}^{[12]\dagger}\mathbf{h}^{[22]} = \begin{pmatrix}\tilde{\mathbf{h}}_1^{[22]}\\ \alpha^{[2]}\end{pmatrix},\tag{21}$$

$$\tilde{\mathbf{h}}^{[21]} = \mathbf{U}^{[11]\dagger}\mathbf{h}^{[21]} = \begin{pmatrix}\tilde{\mathbf{h}}_1^{[21]}\\ \eta^{[1]}\end{pmatrix},\tag{22}$$

and

$$\tilde{\mathbf{h}}^{[12]} = \mathbf{U}^{[22]\dagger}\mathbf{h}^{[12]} = \begin{pmatrix}\tilde{\mathbf{h}}_1^{[12]}\\ \eta^{[2]}\end{pmatrix},\tag{23}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2692222, IEEE Transactions on Communications

respectively, where $\mathbf{U}^{[kj]}$'s are unitary matrices, $\tilde{\mathbf{h}}_1^{[11]}$ ($\tilde{\mathbf{h}}_1^{[21]}$) is the first component of a column vector $\tilde{\mathbf{h}}^{[11]}$ ($\tilde{\mathbf{h}}^{[21]}$), and $\alpha^{[k]}$ ($\eta^{[k]}$) is a $(M-1)\times 1$ vector for $k=1,\ 2$.

- We write the SVD of $\eta^{[k]}$ as

$$\eta^{[k]} = \mathbf{U}_{\eta^{[k]}} \begin{pmatrix} \|\eta^{[k]}\| \\ \mathbf{0}_{(M-1)\times 1} \end{pmatrix}. \tag{24}$$

- We denote $\tilde{\mathbf{\Delta}}^{[k]}$ and $\tilde{\delta}^{[k]}$ as

$$\tilde{\mathbf{\Delta}}^{[k]} = \mathbf{U}_{\eta^{[k]}}^{\dagger} \mathbf{\Delta}^{[k]} \mathbf{U}_{\eta^{[k]}} \tag{25}$$

and

$$\tilde{\delta}^{[k]} = \mathbf{U}_{\eta^{[k]}}^{\dagger} \delta^{[k]}, \tag{26}$$

respectively, where $\tilde{\mathbf{\Delta}}_{11}^{[k]}$ is the $(1,\ 1)$-th component of $\tilde{\mathbf{\Delta}}^{[k]}$, and $\tilde{\delta}_1^{[k]}$ is the first component of a column vector $\tilde{\delta}^{[k]}$.

Theorem 1: [9], [*Theorem* 1, 14] For a two-user MISO fading IC, the single user detection rate region is

$$\bigcup_{\phi^{[k]}\in\left[0,\ \frac{\pi}{2}-\theta^{[k]}\right],\ k=1,\ 2} \left\{ \begin{array}{l} R^{[1]} \le \log_2\left(1+\frac{P^{[1]}\|\mathbf{h}^{[11]}\|^2 \sin^2(\theta^{[1]}+\phi^{[1]})}{1+P^{[2]}\|\mathbf{h}^{[12]}\|^2\sin^2(\phi^{[2]})}\right), \\ R^{[2]} \le \log_2\left(1+\frac{P^{[2]}\|\mathbf{h}^{[22]}\|^2 \sin^2(\theta^{[2]}+\phi^{[2]})}{1+P^{[1]}\|\mathbf{h}^{[21]}\|^2\sin^2(\phi^{[1]})}\right) \end{array} \right\}. \tag{27}$$

By introducing two variables $\phi^{[1]}\in\left[0,\ \frac{\pi}{2}\right]$ and $\phi^{[2]}\in\left[0,\ \frac{\pi}{2}\right]$ we can write

$$\left(i^{[1]}\right)^2 = P^{[1]}\|\mathbf{h}^{[21]}\|^2 \sin^2\phi^{[1]}, \tag{28}$$

$$\left(i^{[2]}\right)^2 = P^{[2]}\|\mathbf{h}^{[12]}\|^2 \sin^2\phi^{[2]}, \tag{29}$$

$$\mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]} = P^{[1]}\|\mathbf{h}^{[11]}\|^2 \sin^2\left(\theta^{[1]}+\phi^{[1]}\right), \tag{30}$$

and

$$\mathbf{h}^{[22]\dagger}\mathbf{S}^{[2]}\mathbf{h}^{[22]} = P^{[2]}\|\mathbf{h}^{[22]}\|^2 \sin^2\left(\theta^{[2]}+\phi^{[2]}\right). \tag{31}$$

Remark 1: In (30) and (31), it is observed that given $\theta^{[k]}$'s and $\phi^{[k]}\in\left[\pi/2-\theta^{[k]},\ \pi/2\right]$ for $k=1,\ 2$, desired signal power $P^{[1]}\|\mathbf{h}^{[11]}\|^2\sin^2\left(\theta^{[1]}+\phi^{[1]}\right)$ decreases as the interfering signal power $P^{[1]}\|\mathbf{h}^{[21]}\|^2\sin^2\phi^{[1]}$ increases. Thus, in order to maximize the weighted sum-rate, it does not lose optimality to reduce the range of $\phi^{[1]}$'s, which results in $\phi^{[1]}\in\left[0,\ \pi/2-\theta^{[1]}\right]$ in (27).

Remark 2: Theorem 1 provides an alternative way to find the maximum of a weighted sum-rate. Instead of exhaustively search through all covariance matrices for the original non-convex optimization problem (12), we only need to find the $\phi^{[1]}$ and $\phi^{[2]}$ in (27) that maximize $\mu^{[1]}R^{[1]}+\mu^{[2]}R^{[2]}$.

*Lemma 1 (*Lemma 2, *14):* Suppose that the optimization problem (14) is feasible, i.e., $0\le\left(i^{[1]}\right)^2\le\|\mathbf{h}^{[21]}\|^2P^{[1]}$. Then, the covariance matrix $\mathbf{S}^{[1]}$ given the fixed interfering signal power $i^{[1]}$ is as follows:

- When $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly independent, $\mathbf{S}^{[1]}$ is

$$\mathbf{S}^{[1]} = \mathbf{v}^{[1]}\mathbf{v}^{[1]\dagger}P^{[1]}, \tag{32}$$

and the maximum value of the objective function is given by (33), where

$$\mathbf{v}^{[1]} = \frac{\mathbf{U}^{[21]}}{\sqrt{P^{[1]}}}\begin{pmatrix} \frac{i^{[1]}}{\|\mathbf{h}^{[21]}\|} \\ \gamma^{[1]}\sqrt{P^{[1]}-\frac{\left(i^{[1]}\right)^2}{\|\mathbf{h}^{[21]}\|^2}}\cdot\frac{\alpha^{[1]}}{\|\alpha^{[1]}\|} \end{pmatrix} \tag{34}$$

with $\alpha^{[1]}$ from (20) and

$$\gamma^{[1]} = \begin{cases} \frac{\tilde{\mathbf{h}}_1^{[11]\dagger}}{|\tilde{\mathbf{h}}_1^{[11]}|}, & \text{if } \tilde{\mathbf{h}}_1^{[11]}\neq 0, \\ 1, & \text{otherwise.} \end{cases} \tag{35}$$

- When $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly dependent and $\|\mathbf{h}^{[21]}\|\neq 0$,

$$\mathbf{S}^{[1]} = \frac{\mathbf{h}^{[21]}\mathbf{h}^{[21]\dagger}}{\|\mathbf{h}^{[21]}\|^2}\cdot\left(i^{[1]}\right)^2 \tag{36}$$

and

$$\mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]} = \frac{|\mathbf{h}^{[11]\dagger}\mathbf{h}^{[21]}|^2}{\|\mathbf{h}^{[21]}\|^2}\left(i^{[1]}\right)^2. \tag{37}$$

- When $\|\mathbf{h}^{[21]}\|=0$,

$$\mathbf{S}^{[1]} = \frac{\mathbf{h}^{[11]}\mathbf{h}^{[11]\dagger}}{\|\mathbf{h}^{[11]}\|^2}P^{[1]} \tag{38}$$

and

$$\mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]} = P^{[1]}\|\mathbf{h}^{[11]}\|^2. \tag{39}$$

B. Convex approach 2. Minimizing interfering signal power with a desired signal power constraint

Lemma 2: Suppose that the optimization problem (16) is feasible, i.e., $0\le\left(d^{[1]}\right)^2\le\|\mathbf{h}^{[11]}\|^2P^{[1]}$. Defining $|\tilde{\delta}_1^{[1]}|$ as (40), the covariance matrix $\mathbf{S}^{[1]}$ is as follows:

- When $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly independent, $\mathbf{S}^{[1]}$ is

$$\mathbf{S}^{[1]} = \mathbf{v}^{[1]}\mathbf{v}^{[1]\dagger}P^{[1]} \tag{41}$$

and the minimum value of the objective function is

$$\frac{\|\eta^{[1]}\|^2\|\mathbf{h}^{[11]}\|^2}{\left(d^{[1]}\right)^2}\left\{|\tilde{\delta}_1^{[1]}|-\frac{|\tilde{\mathbf{h}}_1^{[21]}|\left(d^{[1]}\right)^2}{\|\eta^{[1]}\|\|\mathbf{h}^{[11]}\|^2}\right\}^2, \tag{42}$$

where

$$\mathbf{v}^{[1]} = \frac{\mathbf{U}^{[11]}}{\sqrt{P^{[1]}}}\begin{pmatrix} \frac{d^{[1]}}{\|\mathbf{h}^{[11]}\|} \\ \rho^{[1]}\frac{\|\mathbf{h}^{[11]}\||\tilde{\delta}_1^{[1]}|}{d^{[1]}}\cdot\frac{\eta^{[1]}}{\|\eta^{[1]}\|} \end{pmatrix} \tag{43}$$

and

$$\rho^{[1]} = \begin{cases} \frac{-\tilde{\mathbf{h}}_1^{[21]\dagger}}{|\tilde{\mathbf{h}}_1^{[21]}|}, & \text{if } \tilde{\mathbf{h}}^{[21]}\neq 0, \\ 1, & \text{otherwise.} \end{cases} \tag{44}$$

- When $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly dependent and $\|\mathbf{h}^{[11]}\|\neq 0$,

$$\mathbf{S}^{[1]} = \frac{\mathbf{h}^{[11]}\mathbf{h}^{[11]\dagger}}{\|\mathbf{h}^{[11]}\|^2}\cdot\left(d^{[1]}\right)^2 \tag{45}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2692222, IEEE Transactions on Communications

$$\mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]} = \left(\sqrt{\left(\|\mathbf{h}^{[11]}\|^2 - \frac{|\mathbf{h}^{[21]\dagger}\mathbf{h}^{[11]}|^2}{\|\mathbf{h}^{[21]}\|^2}\right)\left(P^{[1]} - \frac{\left(i^{[1]}\right)^2}{\|\mathbf{h}^{[21]}\|^2}\right)} + \frac{i^{[1]}|\mathbf{h}^{[21]\dagger}\mathbf{h}^{[11]}|}{\|\mathbf{h}^{[21]}\|^2}\right)^2. \tag{33}$$

$$|\tilde{\delta}_1^{[1]}| = \begin{cases} \frac{|\tilde{\mathbf{h}}_1^{[21]}|\left(d^{[1]}\right)^2}{\|\eta^{[1]}\|\|\mathbf{h}^{[11]}\|^2}, & \text{if } \frac{|\tilde{\mathbf{h}}_1^{[21]}|\left(d^{[1]}\right)^2}{\|\eta^{[1]}\|\|\mathbf{h}^{[11]}\|^2} \le \frac{\left(d^{[1]}\right)}{\|\mathbf{h}^{[11]}\|}\sqrt{\left(P^{[1]} - \frac{(d^{[1]})^2}{\|\mathbf{h}^{[11]}\|^2}\right)} \\ \frac{d^{[1]}}{\|\mathbf{h}^{[11]}\|}\sqrt{\left(P^{[1]} - \frac{(d^{[1]})^2}{\|\mathbf{h}^{[11]}\|^2}\right)}, & \text{otherwise.} \end{cases} \tag{40}$$

and

$$\mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]} = \frac{|\mathbf{h}^{[21]\dagger}\mathbf{h}^{[11]}|^2}{\|\mathbf{h}^{[11]}\|^2}\left(d^{[1]}\right)^2. \tag{46}$$

- When $\|\mathbf{h}^{[11]}\| = 0$,

$$\mathbf{S}^{[1]} = \mathbf{0}_{M\times M} \tag{47}$$

and

$$\mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]} = 0. \tag{48}$$

Proof 1: The proof is given in Appendix A.

Theorem 2: For the two-user MISO IC, all the rate pairs are represented as (49). under single user detection, where

$$\theta^{[1]} = \cos^{-1}\left(\frac{|\mathbf{h}^{[21]\dagger}\mathbf{h}^{[11]}|}{\|\mathbf{h}^{[11]}\|\cdot\|\mathbf{h}^{[21]}\|}\right) \tag{50}$$

and

$$\theta^{[2]} = \cos^{-1}\left(\frac{|\mathbf{h}^{[12]\dagger}\mathbf{h}^{[22]}|}{\|\mathbf{h}^{[12]}\|\cdot\|\mathbf{h}^{[22]}\|}\right). \tag{51}$$

Proof 2: The proof is in Appendix B.

Remark 3: In *Theorem* 2, when we use the following replace variables $\tilde{\psi}^{[k]} = \psi^{[k]} - \theta^{[k]} \in \left[0,\ \pi/2 - \theta^{[k]}\right]$ for $k = 1,\ 2$, then (49) is equivalent to (27). Furthermore, from (74) and (75), we get the mathematical relation (52) between interfering and desired signal powers. where $0 \le \left(d^{[1]}\right)^2 \le \|\mathbf{h}^{[11]}\|^2 P^{[1]}$. As shown in Fig. 1, we have

- when $0 \le \left(d^{[1]}\right)^2 < \|\mathbf{h}^{[11]}\|^2 P^{[1]}\sin^2\left(\theta^{[1]}\right)$, $\mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]}$ is a nonincreasing function of $\left(d^{[1]}\right)^2$.
- when $\|\mathbf{h}^{[11]}\|^2 P^{[1]}\sin^2\left(\theta^{[1]}\right) \le \left(d^{[1]}\right)^2 \le \|\mathbf{h}^{[11]}\|^2 P^{[1]}$, $\mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]}$ is a nondecreasing function of $\left(d^{[1]}\right)^2$.

IV. Proposed scheme

In the two-user MISO IC where each coding block experiences correlated block fading in time domain, one naive way is to optimize beamforming vectors using *Lemma* 1 and *Theorem* 1 at each coding block in order to maximize an arbitrary weighted sum-rate.

In the naive way, optimized interfering or signal power levels in the previous coding block are approximately same as that in the present coding block if the channel correlation level is sufficiently high between adjacent coding blocks. Hence, we can approach to the following convex optimization problem by exploiting the side-information, i.e., interfering or desired signal power from each transmitter derived from the previous coding block.

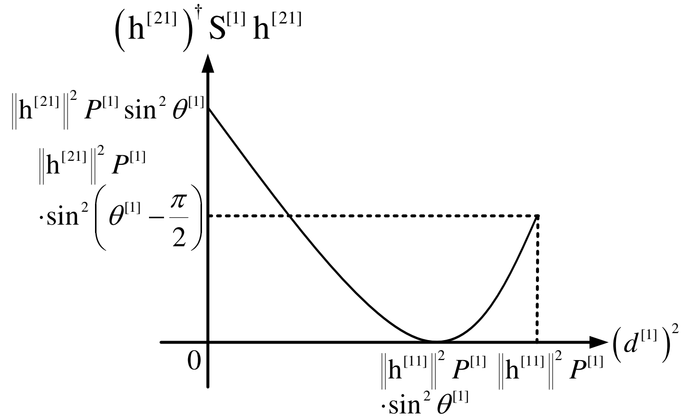

Fig. 1. Optimization problem (16)

A. Transmit beamforming design based on the convex approach 1

We begin with the following *Lemma* to get effective side information so that near-optimal transmit beamforming vector is derived assuming that the convex approach 1 is used.

Lemma 3: When all channel vectors $\mathbf{h}^{[kj]}$'s are drawn i.i.d. from a continuous distribution with $0 \le \left(i^{[1]}\right)^2 \le \|\mathbf{h}^{[21]}\|^2 P^{[1]}$, the optimal solution $\mathbf{v}_*^{[1]}$'s for the optimization problems (18) becomes

$$\mathbf{v}_*^{[1]} = \frac{\mathbf{U}^{[21]}}{\sqrt{P^{[1]}}}\begin{bmatrix} \frac{\left(i_*^{[1]}\right)}{\|\mathbf{h}^{[21]}\|} \\ \frac{\tilde{\mathbf{h}}_1^{[11]}}{|\tilde{\mathbf{h}}_1^{[11]}|}\sqrt{P^{[1]} - \frac{\left(i_*^{[1]}\right)^2}{\|\mathbf{h}^{[21]}\|^2}}\cdot\frac{\alpha^{[1]}}{\|\alpha^{[1]}\|} \end{bmatrix} \tag{53}$$

with $\alpha^{[1]}$ from (20), where $\left(i_*^{[1]}\right)^2$ is computed from (28) with optimal angles $\phi_*^{[1]}$ and $\phi_*^{[2]}$ which maximize the weighted sum-rate $\mu^{[1]}R^{[1]} + \mu^{[2]}R^{[2]}$ in (27), and is given by (60).

Proof 3: Based on the following two facts along with *Lemma* 1, *Lemma* 4 is easily proved.

- Since all channel vectors are drawn i.i.d. from a continuous distribution, any two channel vectors are linearly independent with probability 1.
- After some manipulations, we get (55) with $0 \le \left(i^{[1]}\right)^2 \le \|\mathbf{h}^{[21]}\|^2 P^{[1]}$, which leads to the mathematical relation (55) between the interfering and desired signal powers from (28) and (30). As depicted in Fig. 2, we have
  - when $0 \le \left(i^{[1]}\right)^2 < \|\mathbf{h}^{[21]}\|^2 P^{[1]}\sin^2\left(\frac{\pi}{2} - \theta^{[1]}\right)$, $\mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]}$ is a nondecreasing function of $\left(i^{[1]}\right)^2$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2692222, IEEE Transactions on Communications

$$\bigcup_{\psi^{[k]}\in\left[\theta^{[k]},\ \frac{\pi}{2}\right],\ k=1,\ 2}\left\{\begin{array}{c} R^{[1]}\le\log_2\left(1+\frac{P^{[1]}\|\mathbf{h}^{[11]}\|^2\sin^2(\psi^{[1]})}{1+P^{[2]}\|\mathbf{h}^{[12]}\|^2\sin^2(\theta^{[2]}-\psi^{[2]})}\right), \\ R^{[2]}\le\log_2\left(1+\frac{P^{[2]}\|\mathbf{h}^{[22]}\|^2\sin^2(\psi^{[2]})}{1+P^{[1]}\|\mathbf{h}^{[21]}\|^2\sin^2(\theta^{[1]}-\psi^{[1]})}\right)\end{array}\right\}. \tag{49}$$

$$\mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]}=\|\mathbf{h}^{[21]}\|^2P^{[1]}\left\{\sin^2\theta^{[1]}+\frac{\cos\left(2\theta^{[1]}\right)}{\|\mathbf{h}^{[11]}\|^2\sqrt{P^{[1]}}}\left(d^{[1]}\right)^2-\frac{\sin\left(2\theta^{[1]}\right)}{\|\mathbf{h}^{[11]}\|\sqrt{P^{[1]}}}d^{[1]}\sqrt{1-\frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2P^{[1]}}}\right\}. \tag{52}$$

$$\left(i_*^{[1]}\right)^2=\begin{cases}\left(i_*^{[1]}\right)^2, & 0\le\left(i_*^{[1]}\right)^2\le\|\mathbf{h}^{[21]}\|^2P^{[1]}\sin^2\left(\frac{\pi}{2}-\theta^{[1]}\right)\\ \|\mathbf{h}^{[21]}\|^2P^{[1]}\sin^2\left(\frac{\pi}{2}-\theta^{[1]}\right), & \text{else.}\end{cases} \tag{54}$$

$$\begin{aligned}&\mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]}\\&=\|\mathbf{h}^{[11]}\|^2P^{[1]}\left\{\sin^2\theta^{[1]}+\frac{\cos\left(2\theta^{[1]}\right)}{\|\mathbf{h}^{[21]}\|^2\sqrt{P^{[1]}}}\left(i^{[1]}\right)^2+\frac{\sin\left(2\theta^{[1]}\right)}{\|\mathbf{h}^{[21]}\|\sqrt{P^{[1]}}}i^{[1]}\sqrt{1-\frac{\left(i^{[1]}\right)^2}{\|\mathbf{h}^{[21]}\|^2P^{[1]}}}\right\}.\end{aligned} \tag{55}$$

  - when $\|\mathbf{h}^{[21]}\|^2P^{[1]}\sin^2\left(\frac{\pi}{2}-\theta^{[1]}\right)\le\left(i^{[1]}\right)^2\le\|\mathbf{h}^{[21]}\|^2P^{[1]}$, $\mathbf{h}^{[11]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[11]}$ is a nonincreasing function of $\left(i^{[1]}\right)^2$.

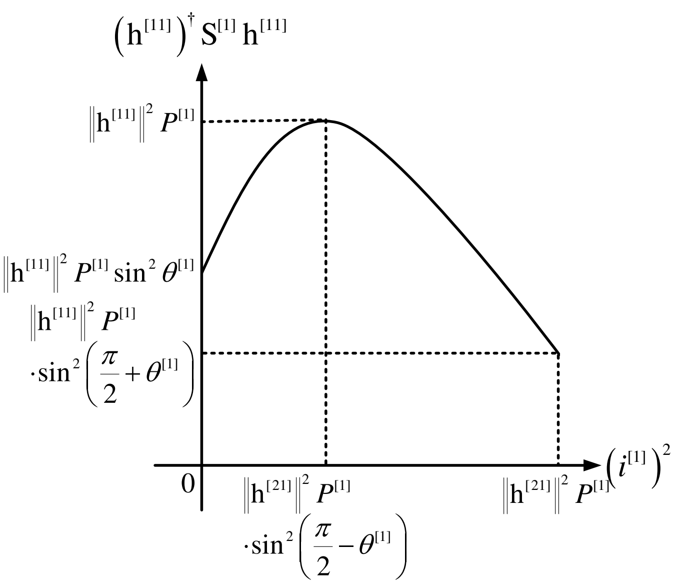

Fig. 2. Optimization problem (14)

Based on the effective side-information, i.e., interfering signal power, the following principle for the transmit beamforming design can be considered.

Each transmit beamformer is derived by restricting the present interfering signal power no greater than that of the previous coding block while maximizing the desired signal power at the present coding block.

Suppose that given any pair of $\mu^{[1]}\ge0$ and $\mu^{[2]}\ge0$, optimized $\left(i_*^{[1]}\right)^2$ and $\left(i_*^{[2]}\right)^2$ lead to maximizing the weighted sum-rate $\mu^{[1]}R^{[1]}+\mu^{[2]}R^{[2]}$ (27) at the previous coding block. Then, the two present interfering signal powers are close to $\left(i_*^{[1]}\right)^2$ and $\left(i_*^{[2]}\right)^2$ at the previous coding block, respectively. Thus, we can formulate the following optimization problem to derive a near-optimal transmit beamforming vector.

$$\begin{array}{ll}\text{Maximize} & \mathbf{g}^{[11]\dagger}\mathbf{T}^{[1]}\mathbf{g}^{[11]}\\ \text{Subject to} & \mathbf{g}^{[21]\dagger}\mathbf{T}^{[1]}\mathbf{g}^{[21]}\le\left(i_*^{[1]}\right)^2,\\ & \mathrm{tr}(\mathbf{T}^{[1]})\le P^{[1]},\ \mathbf{T}^{[1]}\succeq0.\end{array} \tag{56}$$

For (56), the near-optimal solution is $\mathbf{T}^{[1]}=\hat{\mathbf{v}}^{[1]}\hat{\mathbf{v}}^{[1]\dagger}P^{[1]}$, where

$$\begin{aligned}\hat{\mathbf{v}}^{[1]}&=\frac{\mathbf{v}^{[1]}+\Delta_{\mathbf{v}^{[1]}}}{\|\mathbf{v}^{[1]}+\Delta_{\mathbf{v}^{[1]}}\|}\\&=\frac{\mathbf{V}^{[21]}}{\sqrt{P^{[1]}}}\left[\begin{array}{c}\frac{\left(j^{[1]}\right)}{\|\mathbf{h}^{[21]}+\Delta_{\mathbf{h}^{[21]}}\|}\\ \frac{\tilde{\mathbf{g}}^{[11]\dagger}}{|\tilde{\mathbf{g}}^{[11]}|}\sqrt{P^{[1]}-\frac{\left(j^{[1]}\right)^2}{\|\mathbf{h}^{[21]}+\Delta_{\mathbf{h}^{[21]}}\|^2}}\cdot\frac{\beta^{[1]}}{\|\beta^{[1]}\|}\end{array}\right],\end{aligned} \tag{57}$$

and the present interfering signal power $\left(j^{[1]}\right)^2$ is given by (58).

This comes from (53) and facts that

- when $\left(i_*^{[1]}\right)^2\le P^{[1]}\|\mathbf{g}^{[21]}\|^2\sin^2\left(\frac{\pi}{2}-\hat{\theta}^{[1]}\right)$, $\mathbf{g}^{[11]}\mathbf{T}^{[1]}\mathbf{g}^{[11]}$ is a nondecreasing function of $\left(j^{[1]}\right)^2$.
- when $\left(i_*^{[1]}\right)^2\ge P^{[1]}\|\mathbf{g}^{[21]}\|^2\sin^2\left(\frac{\pi}{2}-\hat{\theta}^{[1]}\right)$, $\mathbf{g}^{[11]}\mathbf{T}^{[1]}\mathbf{g}^{[11]}$ is a nonincreasing function of $\left(j^{[1]}\right)^2$.

Fig. 3 describes the convex behavior of desired signal power according to the interfering signal power.

B. Transmit beamforming design based on the convex approach 2

Lemma 4: When all channel vectors $\mathbf{h}^{[kj]}$'s are drawn i.i.d. from a continuous distribution with $0\le\left(d^{[1]}\right)^2\le$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2692222, IEEE Transactions on Communications

$$\left(j^{[1]}\right)^2 = \begin{cases} \left(i_*^{[1]}\right)^2, & \text{if } \left(i_*^{[1]}\right)^2 \le P^{[1]}\|\mathbf{g}^{[21]}\|^2 \sin^2\left(\frac{\pi}{2} - \hat{\theta}^{[1]}\right), \\ P^{[1]}\|\mathbf{g}^{[21]}\|^2 \sin^2\left(\frac{\pi}{2} - \hat{\theta}^{[1]}\right), & \text{otherwise.} \end{cases} \quad (58)$$

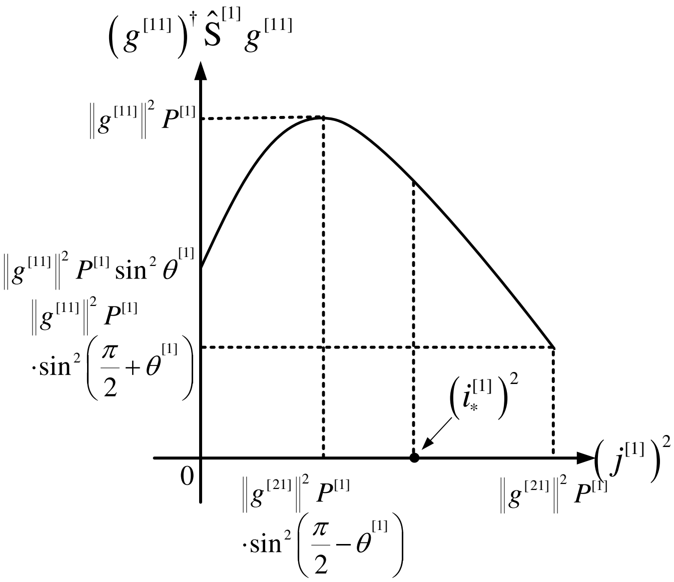

Fig. 3. Desired signal power characteristics with respect to an interfering signal power

$\|\mathbf{h}^{[11]}\|^2 P^{[1]}$, the optimal solution $\mathbf{v}_*^{[1]}$'s for the optimization problems (19) becomes

$$\mathbf{v}^{[1]} = \frac{\mathbf{U}^{[11]}}{\sqrt{P^{[1]}}} \begin{pmatrix} \frac{d^{[1]}}{\|\mathbf{h}^{[11]}\|} \\ \rho^{[1]} \frac{\|\mathbf{h}^{[11]}\| |\tilde{\delta}_1^{[1]}|}{d^{[1]}} \cdot \frac{\eta^{[1]}}{\|\eta^{[1]}\|} \end{pmatrix}, \quad (59)$$

where $\rho^{[1]}$ from (44) and $\eta^{[1]}$ from (22), and $\left(d_*^{[1]}\right)^2$ is computed from (74) with optimal angles $\psi_*^{[1]}$ and $\psi_*^{[2]}$ which maximize the weighted sum-rate $\mu^{[1]}R^{[1]} + \mu^{[2]}R^{[2]}$ in (49), and is given by (60).

Proof 4: Based on *Lemma* 1 with *Remark* 3, *Lemma* 4 is easily proved and we omitted the proof. 4.

Fig. 2 describes the convexity of desired signal power behavior according to interfering signal power. Now, we can consider the following transmit beamforming design principle as an alternative.

Each transmitter updates beamforming vectors by making the desired signal power greater than or equal to that at the previous coding block while minimizing the interfering signal power at the present coding block.

Assuming that $\left(d_*^{[1]}\right)^2$ and $\left(d_*^{[2]}\right)^2$ are optimized desired signal powers at the previous coding block, we get the near-optimal transmit beamforming vector using the following optimization problem:

$$\begin{aligned} &\text{Minimize} && \mathbf{g}^{[12]\dagger}\mathbf{T}^{[2]}\mathbf{g}^{[12]} \\ &\text{Subject to} && \mathbf{g}^{[11]\dagger}\mathbf{T}^{[1]}\mathbf{g}^{[11]} \ge \min\left\{\left(d_*^{[1]}\right)^2,\ \|\mathbf{g}^{[11]}\|^2 P^{[1]}\right\}, \\ &&& \mathrm{tr}(\mathbf{T}^{[1]}) \le P^{[1]},\ \mathbf{T}^{[1]} \succeq 0. \end{aligned} \quad (61)$$

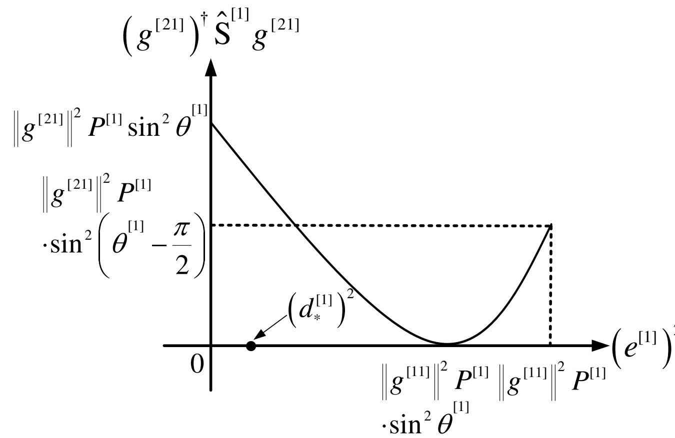

Fig. 4. Interfering signal power characteristics with respect to a desired signal power

The near-optimal solution for (61) is $\mathbf{T}^{[1]} = \hat{\mathbf{v}}^{[1]}\hat{\mathbf{v}}^{[1]\dagger}P^{[1]}$, where

$$\hat{\mathbf{v}}^{[1]} = \frac{\mathbf{V}^{[11]}}{\sqrt{P^{[1]}}} \begin{pmatrix} \frac{e^{[1]}}{\|\mathbf{g}^{[11]}\|} \\ \frac{-\mathbf{g}_1^{[21]\dagger}}{|\mathbf{g}_1^{[21]}|} \frac{\|\mathbf{g}^{[11]}\| |\tilde{\zeta}_1^{[1]}|}{e^{[1]}} \cdot \frac{\omega^{[1]}}{\|\omega^{[1]}\|} \end{pmatrix} \quad (62)$$

and $|\tilde{\zeta}_1^{[1]}|$ is defined as (63) with (64). This comes from (43) and facts that

- when $0 \le \left(e^{[1]}\right)^2 \le \|\mathbf{g}^{[11]}\|^2 P^{[1]} \sin^2\left(\hat{\theta}^{[1]}\right)$, $\mathbf{g}^{[21]\dagger}\mathbf{T}^{[1]}\mathbf{g}^{[21]}$ is a nonincreasing function of $\left(e^{[1]}\right)^2$.
- when $\left(e^{[1]}\right)^2 > \|\mathbf{g}^{[11]}\|^2 P^{[1]} \sin^2\left(\hat{\theta}^{[1]}\right)$, $\mathbf{g}^{[21]\dagger}\mathbf{T}^{[1]}\mathbf{g}^{[21]}$ is a nondecreasing function of $\left(e^{[1]}\right)^2$.

Fig. 4 describes the concave behavior of interfering signal power according to the desired signal power.

C. Complexity and performance analysis

The proposed transmit beamforming designs can be summarized in Table I, which will be analyzed further in the perspective of complexity and throughput performance with overhead. In Table I, b is the coding block index.

1) Convex approach 1 *Vs. Convex approach* 2*:* In the time-varying MISO fading IC, adjacent channel coding blocks have correlation to an extent each other. To get an insight on the two methods, let's assume that the two corresponding channel vectors in (2) and (3) are asymptotically same for adjacent channel coding blocks. In that case, effective interfering power in (56) and effective desired signal power (61) act as equivalent side-information in a sense that the optimal transmit beamforming vectors in the previous coding block is still optimal for the present coding block in an asymptotic sense.

However, in case that the channel vectors for the adjacent coding blocks are independent each other, we lose the equivalence of the two side-information. This is because (61) has

$$\left(d_*^{[1]}\right)^2 = \begin{cases} \left(d_*^{[1]}\right)^2, & 0 \le \left(d_*^{[1]}\right)^2 \le \|\mathbf{h}^{[11]}\|^2 P^{[1]} \sin^2\left(\theta^{[1]}\right) \\ \|\mathbf{h}^{[11]}\|^2 P^{[1]} \sin^2\left(\theta^{[1]}\right), & \text{else.} \end{cases} \tag{60}$$

$$|\tilde{\zeta}_1^{[1]}| = \begin{cases} \frac{|\tilde{\mathbf{g}}_1^{[21]}|\left(e^{[1]}\right)^2}{\|\omega^{[1]}\|\|\mathbf{g}^{[11]}\|^2}, & \text{if } \frac{|\tilde{\mathbf{g}}_1^{[21]}|\left(e^{[1]}\right)^2}{\|\omega^{[1]}\|\|\mathbf{g}^{[11]}\|^2} \le \frac{\left(e^{[1]}\right)}{\|\mathbf{g}^{[11]}\|}\sqrt{\left(P^{[1]} - \frac{(e^{[1]})^2}{\|\mathbf{g}^{[11]}\|^2}\right)} \\ \frac{e^{[1]}}{\|\mathbf{g}^{[11]}\|}\sqrt{\left(P^{[1]} - \frac{(e^{[1]})^2}{\|\mathbf{g}^{[11]}\|^2}\right)}, & \text{otherwise.} \end{cases} \tag{63}$$

$$\left(e^{[1]}\right)^2 = \begin{cases} \|\mathbf{g}^{[11]}\|^2 P^{[1]} \sin^2 \hat{\theta}^{[1]}, & \text{if } \left(d_*^{[1]}\right)^2 \le \|\mathbf{g}^{[11]}\|^2 P^{[1]} \sin^2 \hat{\theta}^{[1]}, \\ \min\left\{\left(d_*^{[1]}\right)^2, \ \|\mathbf{g}^{[11]}\|^2 P^{[1]}\right\}, & \text{otherwise.} \end{cases} \tag{64}$$

TABLE I
PROPOSED BEAMFORMING UPDATE ALGORITHM IN PSEUDO CODE.

Proposed scheme using the convex Approach 1	Proposed scheme using the convex Approach 2
for $b = B(n-1)+1 : Bn$ if $(b == B(n-1)+1)$ Use the transmit beamformer $v_*^{[k]}$'s from *Theorem* 1 Store $i_*^{[k]}$ at the k-th transmitter else Calculate $j^{[k]}$ from at (58) the k-th transmitter Use the transmit beamformer $v^{[k]}$'s from (57) end end	for $b = B(n-1)+1 : Bn$ if $(b == B(n-1)+1)$ Use the transmit beamformer $v_*^{[k]}$'s from *Theorem* 2 Store $d_*^{[k]}$ at the k-th transmitter else Calculate $e^{[k]}$ from at (64) the k-th transmitter Use the transmit beamformer $v^{[k]}$'s from (62) end end

additional constraint from $\min$ function in (61). Thus, transmit beamforming design based on the convex approach 2 has more likely to be deviated from the near-optimal value of the side-information value in comparison with transmit beamforming design based on the convex approach 1.

In summary, the latter one can lead to worse performance degradation for a weighted sum-rate than the former one except for the highly correlated channel environment.

2) Computational complexity: For two-user MISO ICs, the state of the art transmit beamforming scheme with Pareto-optimality requires 1 real parameter optimization, which is constant with respect to the number of transmit antennas. When we consider multiple coding blocks, the state of the art transmit beamformer still requires 1 on the average.

On the contrary, the proposed scheme can be applied more efficiently than the state of the art beamforming scheme[4]. Specifically, in the first coding block, Pareto-optimal transmit beamformers are derived to be used. From the second coding blocks to the B-th coding blocks, transmit beamformers are acquired based on the side-information, i.e., interfering or desired signal power in the first coding block. Finally, $1/B$ real parameters are required on the average, which is close to 0 as B increases. Table II depicts the required average number of optimizing parameters of beamforming schemes including the state of the art beamforming scheme.

From the numerical results in the next section, we show feasibility that the proposed scheme can guarantee average sum-rate performance without significant performance degradation even under lowly correlated block fading environments. This means that B can be sufficiently large so that the computational complexity in the sense of the number of optimizing parameters can be negligible with comparable throughput performance with the optimal scheme.

3) Overhead of CSI and side-information: In the perspective of overhead, the proposed scheme requires essentially same amount of CSI in comparison with the state of the art beamforming scheme.

For acquiring an optimal weighted sum-rate, the state of the art scheme should search transmit beamforming vectors exhaustively as shown in *Theorem* 1 and *Theorem* 2. Instead, the proposed scheme reduces the computational complexity significantly since the exhaustive search is only needed every B coding blocks. The proposed scheme, however, requires a overhead for storing effective interfering and desired signal power at the transmitter, which is required every B coding blocks. Thus, the overhead of proposed scheme is almost negligible for storing the side-information when B is large.

There is trade-off between the computation complexity and throughput performance according to the value of B. Specifically, when B is small, the computational complexity is not significantly reduced, but the throughput behavior is expected

[4]In the system model, B was assumed to be 2. However, it can be directly extended to B with $B > 2$. In this case, optimal transmit beamforming vectors should be searched exhaustively every B coding blocks, which will be used as side-information for deriving near-optimal transmit beamforming vectors. Note that B should be determined by considering channel coherence time in the time-varying fading channel.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2692222, IEEE Transactions on Communications

TABLE II
COMPLEXITY COMPARISON OF THE PROPOSED BEAMFORMER WITH CONVENTIONAL ONES

Computational complexity when $K=2$ (Average number of parameters to be optimized over multiple coding blocks)			
[10]	[14], [16]	[17]	Proposed scheme
2 complex variables	2 real variables	1 real variables	$\frac{1}{B}$ real variables

to be close to the optimal throughput behavior especially for highly correlated channel environment. When B is large, it is expected that the computational complexity will be more reduced at the cost of a degraded throughput performance.

V. Numerical results

A. Simulation environment

In this section, the two proposed schemes are compared with optimal scheme using the metric of average sum-rate, which is justified from the facts that average power constraint across all users are same and fading channel characteristics for all channel links are isotropic.

In the numerical simulation, we assume that the correlation between adjacent coding blocks is modeled using the first order AR model[5]. Specifically, from the autoregressive (AR) model [23] of order 1, the channel vectors $\mathbf{h}^{[kj]}$'s and $\mathbf{g}^{[kj]}$'s for adjacent two coding blocks satisfy

$$\begin{aligned}\mathbf{g}^{[kj]} &= \mathbf{h}^{[kj]} + \Delta_{\mathbf{h}^{[kj]}} \\ &= \rho\mathbf{h}^{[kj]} + \sqrt{1-\rho^2}\,\mathbf{i}^{[kj]}, \end{aligned} \quad (65)$$

where ρ represents the correlation coefficient between $\mathbf{h}_i^{[kj]}$ and $\mathbf{g}_i^{[kj]}$ with $i = 1, 2, \cdots, M$. Here, all components of channel vectors $\mathbf{h}^{[kj]}$'s are independent and identically distributed (i.i.d.). Each element of $\mathbf{i}^{[kj]}$'s is i.i.d. circularly symmetric complex Gaussian with zero mean and unit variance.

We assume that only adjacent coding blocks experience correlated fading in time domain. Specifically, we generate channel vectors $\mathbf{g}^{[kj]}$ randomly from (65), where $\mathbf{h}^{[kj]}$ is circularly complex Gaussian distributed with zero mean and covariance $\mathbf{I}_{M\times M}$. We consider the average sum-rates of optimal scheme (MAX-SINR) (*Theorem* 1), proposed schemes, and MAX-SINR with no transmit beamformer update. Here, MAX-SINR is the beamforming scheme maximizing average sum-rate as a function of SINR for each user at each coding block, and proposed scheme with convex approach l is denoted as 'Proposed scheme l' with $l = 1, 2$. Note that optimal scheme essentially corresponds to maximize SINR at each coding block. In our numerical evaluations, we assume that B is equal to 2. We use (56) for proposed scheme to see the maximized effect of side-information.

[5]The proposed scheme can be applied to any correlated channel model. We use the AR model to see the performance behavior of the proposed scheme effectively according to the channel correlation, which is determined by one parameter (ρ) in the framework of the AR model.

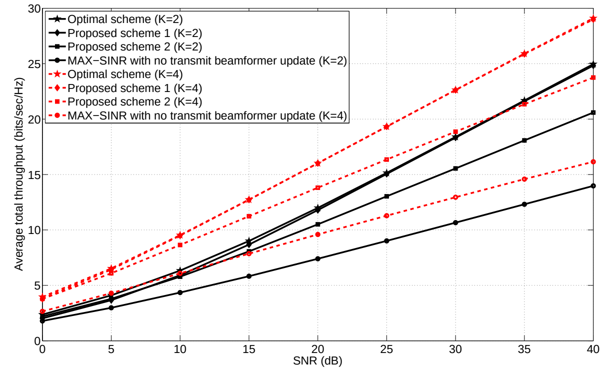

Fig. 5. Average sum-rate with respect to SNR with $\rho = 0.3$

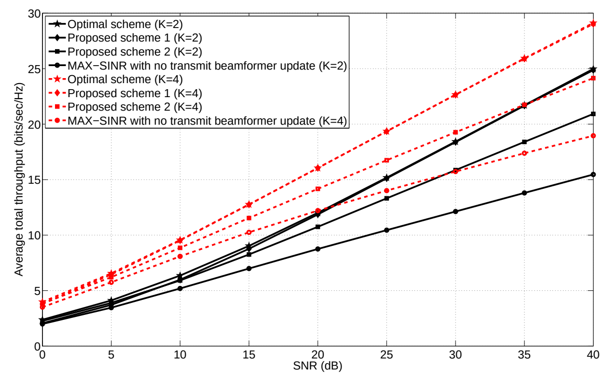

Fig. 6. Average sum-rate with respect to SNR with $\rho = 0.9$

Figs. 5 and 6 illustrates average sum-rates according to Signal to Noise Ratio (SNR) given the number (M) of transmit antennas. In Figs. 7 and 8, average sum-rate behaviors with respect to the correlation coefficient ρ for some given M and SNR. We investigate the average sum-rate behaviors according to the following arguments.

- The effect of the number of transmit antennas
- The effect of the correlation coefficient ρ in (65)
- The performance difference between the two proposed schemes

B. The effect of the number of transmit antennas

In an asymptotic performance sense, optimal multiplexing gain [24] for two-user MISO IC is always 2 under the condition of $M \geq 2$ [25]. As depicted in Figs. 5 and 6, optimal schemes with $M = 2$ and $M = 4$ are observed to have the same average sum-rate behavior in the sense of slope with respect to SNR. Interestingly, the proposed scheme is observed to have comparable multiplexing gain, which shows that the proposed scheme can be effectively used at high SNR regime.

On the other hand, it observed that finite SNR performance of optimal scheme and proposed scheme are affected by the number of transmit antennas. As M increases with $M \geq 2$, finite SNR gain is expected to be increased. The proposed

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2692222, IEEE Transactions on Communications

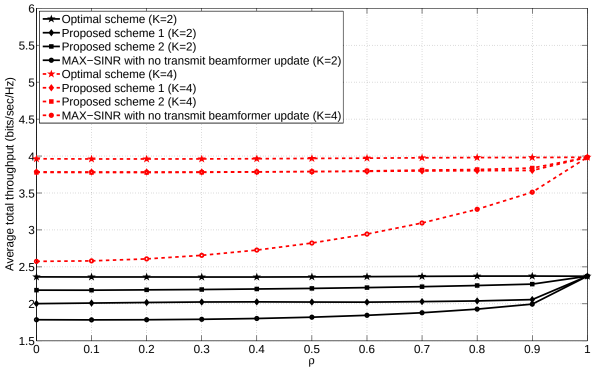

Fig. 7. Average sum-rate with respect to the channel correlation level with SNR=0dB

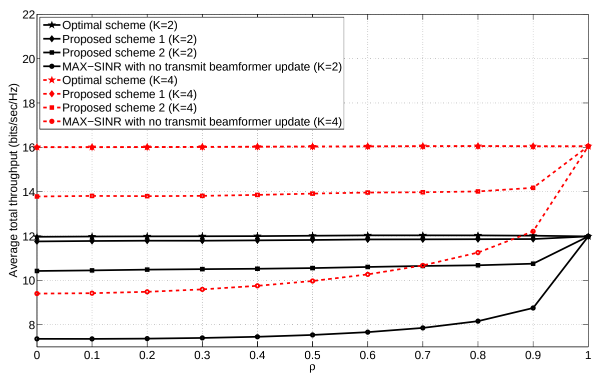

Fig. 8. Average sum-rate with respect to the channel correlation level with SNR=20dB

scheme exhibits near-optimal finite SNR performance in addition to have optimal multiplexing gain. When optimal transmit beamformer is not updated at each coding block, the performance degradation is not negligible, and the performance gap is expected to be increased as SNR increases.

From Figs. 5 and 6, the proposed scheme has comparable throughput behavior with the optimal scheme not only low SNR regime but also high SNR regime[6].

C. The effect of the correlation coefficient

The proposed scheme essentially uses transmit beamformers assuming that channel correlation level is sufficiently high. Thus, the proposed scheme is expected to show near-optimal sum-rate performance at high correlation level regimes, which is confirmed from Figs. 7 and 8. Furthermore, the average sum-rate behavior of the proposed scheme is seen to be robust to the correlation level counter-intuitively. From the observations of proposed scheme's robustness, the proposed scheme can be used even when the optimal transmit beamformer update period of B is strictly greater than 2 without significant performance degradation. On average, the number of required real parameters to be optimized is $1/B$ real parameters, which can be asymptotically 0 since B can be sufficiently extended with comparable performance with the optimal scheme.

[6] In a practical sense, main operating range of SNR is around 5$\sim$10dB. In this paper, we show throughput performance at the two extreme cases, i.e., low and high SNRs. The throughput behavior of mid SNR range (around 5$\sim$10dB) is expected to be between that of low SNR range (around 0dB) and that of high SNR range (around 20dB), which is omitted.

D. The performance comparison between the two proposed schemes

Essentially, the two proposed schemes have an equivalent transmit beamformers when two adjacent coding blocks' channel instances are exactly same. The equivalence is observed to be almost maintained when the correlation level is sufficiently high as shown in Figs. 7 and 8.

On the other hand, when the channel correlation level between the two adjacent channel instances is low or medium, the average sum-rate of the Proposed scheme 2 is more degraded in comparison with the Proposed scheme 1, which is confirmed from Figs. 7 and 8. Furthermore, the performance degradation was expected to be worse as the correlation level becomes low. However, it is observed that the throughput performance of the Proposed scheme 2 at high SNR is almost uniformly degraded according to the correlation level except for the range of $0.9{\sim}1$ unexpectedly.

VI. Conclusion

In an information-theoretical perspective, the total amount of information for two random variables is decreased when the correlation between the two variables is increased [26]. In other words, in the correlated environment, there is a possibility of using a side-information effectively. In this paper, we found the side-information in the framework of convex optimization problems for two-user MISO ICs. Based on the side-information such as interfering or desired signal power derived from the optimal solutions, subsequent transmit beamformers were shown to be derived in a closed form with no further parameter optimization. Consequently, the proposed scheme reduced computational complexity on the average in comparison with the state of the art optimal beamforming scheme. From numerical results, we observed that the proposed scheme exhibited near-optimal average sum-rate performance in highly correlated fading channel environment. Furthermore, our proposed scheme were observed to be robust to the channel correlation level, which gives us feasibility that the proposed scheme can be still used even when B is strictly greater than 2. In a practical sense, the computational complexity can be asymptotically negligible without significant performance degradation, which was justified from near-optimal throughput performance in lowly correlated channel environment.

Acknowledgement

This work was supported by the ICT R&D program of MSIP/IITP. [R7117160130001002, International standardization for railway mobile communications and smart transportation ICT convergence services based on V2X]

APPENDIX

A. Proof of Lemma 2

When $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly dependent, we can easily show (45)-(48), and it is sufficient to prove when $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly independent.

Using (10) and (22), the optimization problem (16) reduces to

$$\begin{aligned} \text{Minimize } & |\tilde{\mathbf{h}}_1^{[21]}|^2\tilde{\mathbf{S}}_{11}^{[1]} + \tilde{\mathbf{h}}_1^{[21]}\eta^{[1]\dagger}\delta^{[1]} + \tilde{\mathbf{h}}_1^{[21]\dagger}\delta^{[1]\dagger}\eta^{[1]} \\ & + \eta^{[1]\dagger}\mathbf{\Delta}^{[1]}\eta^{[1]} \\ \text{Subject to } & \|\mathbf{h}^{[11]}\|^2\tilde{\mathbf{S}}_{11}^{[1]} = \left(d^{[1]}\right)^2, \\ & \mathrm{tr}(\mathbf{\Delta}^{[1]}) \le P^{[1]} - \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}, \\ & \tilde{\mathbf{S}}_{11}^{[1]}\mathbf{\Delta}^{[1]} \succeq \delta^{[1]}\delta^{[1]\dagger}, \end{aligned} \tag{66}$$

where

$$\tilde{\mathbf{S}}_{11}^{[1]}\mathbf{\Delta}^{[1]} \succeq \delta^{[1]}\delta^{[1]\dagger} \tag{67}$$

comes from $\tilde{\mathbf{S}}^{[1]} \succeq 0$. Note that in (66), $\|\mathbf{h}^{[11]}\| \neq 0$ since $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly independent, and from the constraint $\tilde{\mathbf{S}}_{11}^{[1]}\mathbf{\Delta}^{[1]} \succeq \delta^{[1]}\delta^{[1]\dagger}$ with (25) and (26), we get the following relation.

$$\frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\mathbf{\Delta}_{11}^{[1]} \ge |\tilde{\delta}_1^{[1]}|^2. \tag{68}$$

Now, the objective function in (66) is further derived as (69). which leads to the following solutions.

- When $\frac{|\tilde{\mathbf{h}}_1^{[21]}|\left(d^{[1]}\right)^2}{\|\eta^{[1]}\|\|\mathbf{h}^{[11]}\|^2} \le \sqrt{\frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\left(P^{[1]} - \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\right)}$, we have (70) and $\mathbf{h}^{[21]\dagger}\mathbf{S}_*^{[1]}\mathbf{h}^{[21]} = 0$.
- When $\frac{|\tilde{\mathbf{h}}_1^{[21]}|\left(d^{[1]}\right)^2}{\|\eta^{[1]}\|\|\mathbf{h}^{[11]}\|^2} > \sqrt{\frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\left(P^{[1]} - \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\right)}$, we have (71) and (72).

Note that we have (40) using the range of $|\tilde{\delta}_1^{[1]}|^2$

$$0 \le |\tilde{\delta}_1^{[1]}|^2 \le \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\tilde{\mathbf{\Delta}}_{11}^{[1]} \le \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\left(P^{[1]} - \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\right). \tag{73}$$

B. Proof of Theorem 2

We first consider the case in which $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly independent. The proof for the case in which $\mathbf{h}^{[22]}$ and $\mathbf{h}^{[12]}$ are linearly independent can be derived similarly.

- When $\frac{|\tilde{\mathbf{h}}_1^{[21]}|\left(d^{[1]}\right)^2}{\|\eta\|\|\mathbf{h}^{[11]}\|^2} > \sqrt{\frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\left(P^{[1]} - \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\right)}$, we let
$$\left(d^{[1]}\right)^2 = \|\mathbf{h}^{[11]}\|^2 P^{[1]}\sin^2\psi^{[1]},\ 0 \le \psi^{[1]} \le \pi/2. \tag{74}$$
If $0 \le \left(d^{[1]}\right)^2 \le \|\mathbf{h}^{[11]}\|^2 P^{[1]}$, then we have
$$\mathbf{h}^{[21]\dagger}\mathbf{S}_*^{[1]}\mathbf{h}^{[21]} = \|\mathbf{h}^{[21]}\|^2 P^{[1]}\sin^2\left(\theta^{[1]} - \psi^{[1]}\right). \tag{75}$$
Thus, all rate pairs under single user detection are given by (76). Now, when $\psi^{[k]} \in \left[0,\ \theta^{[k]}\right]$ for $k = 1,\ 2$, $\left(d^{[1]}\right)^2$ and $\mathbf{h}^{[21]\dagger}\mathbf{S}^{[1]}\mathbf{h}^{[21]}$ are monotonically increasing and decreasing functions, respectively. Thus, (76) is reduced to (49) without loss of optimality.
- When $\frac{|\tilde{\mathbf{h}}_1^{[21]}|\left(d^{[1]}\right)^2}{\|\eta\|\|\mathbf{h}^{[11]}\|^2} \le \sqrt{\frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\left(P^{[1]} - \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\right)}$, using (50) and (74), $\frac{|\tilde{\mathbf{h}}_1^{[21]}|\left(d^{[1]}\right)^2}{\|\eta\|\|\mathbf{h}^{[11]}\|^2} \le \sqrt{\frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\left(P^{[1]} - \frac{\left(d^{[1]}\right)^2}{\|\mathbf{h}^{[11]}\|^2}\right)}$ is equivalent to $\psi^{[1]} \le \theta^{[1]}$ from $\tan\psi^{[1]} \le \tan\theta^{[1]}$. Here, note that only $\psi^{[1]} = \theta^{[1]}$ is valid in (76). Similarly, only $\psi^{[2]} = \theta^{[2]}$ when $\psi^{[2]} \le \theta^{[2]}$ is also valid, which means that (76) is proved to be all the rate pairs under single user detection at each receiver.

Second, when $\mathbf{h}^{[11]}$ and $\mathbf{h}^{[21]}$ are linearly dependent, (46) and (48) correspond to $\left(\theta^{[1]} = 0,\ \psi^{[1]} = \pi/2\right)$ and $\left(\theta^{[1]},\ \psi^{[1]} = 0\right)$ in (75), respectively. Similarly for the case where $\mathbf{h}^{[22]}$ and $\mathbf{h}^{[12]}$ are linearly dependent, which completes the proof.

REFERENCES

[1] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory,* vol. 21, no. 5, pp. 569-570, Sep. 1975.

[2] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory,* vol. 27, no. 6, pp. 786-788, Nov. 1981.

[3] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory,* vol. IT-27, no. 1, pp. 49-60, Jan. 1981.

[4] M. H. M. Costa and A. A. El Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. Inf. Theory,* vol. 33, no. 5, pp. 710-711, Sep. 1987.

[5] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory,* vol. 54, no. 12, pp. 5534-5562, Dec. 2008.

[6] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inf. Theory,* vol. 55, no. 2, pp. 620-643, Feb. 2009.

[7] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory,* vol. 55, no. 2, pp. 689-699, Feb. 2009.

[8] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory,* vol. 55, no. 7, pp. 3032-3050, July 2009.

[9] X. Shang and B. Chen, "Achievable rate region for downlink beamforming in the presence of interference," in *Proc. 41st Ailomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2007.

[10] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Tr. Signal Process.*, vol. 56, no. 10, pp. 5292-5296, Oct. 2008.

[11] E. A. Jorswieck and E. G. Larsson, "Monotonic optimization framework for the two-user MISO interference channel," *IEEE Tr. Comm.*, vol. 58, no. 7, pp. 2159-2168, July 2010.

[12] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Tr. Signal Process.*, vol. 58, no. 10, pp. 5450-5458, Oct. 2010.

[13] R. Mochaourab and E. Jorswieck, "Optimal beamforming in interference networks with perfect local channel information," *IEEE Tr. Signal Process.*, vol. 59. no. 3, pp. 1128-1141, Mar. 2011.

[14] X. Shang, B. Chen, and H. V. Poor, "Multi-user MISO interference channel with single-user detection: optimality of beamforming and the achievable rate region," *IEEE Tr. Inf. Theory,* vol. 57, no. 7, pp. 4255-4273, July 2011.

$$
\begin{aligned}
&|\tilde{\mathbf{h}}_1^{[21]}|^2\tilde{\mathbf{S}}_{11}^{[1]} + \tilde{\mathbf{h}}_1^{[21]}\eta^{[1]\dagger}\delta^{[1]} + \tilde{\mathbf{h}}_1^{[21]\dagger}\delta^{[1]\dagger}\eta^{[1]} + \eta^{[1]\dagger}\mathbf{\Delta}^{[1]}\eta^{[1]} \\
&= |\tilde{\mathbf{h}}_1^{[21]}|^2\tilde{\mathbf{S}}_{11}^{[1]} + \tilde{\mathbf{h}}_1^{[21]}||\eta^{[1]}||\tilde{\delta}_1^{[1]} + \tilde{\mathbf{h}}_1^{[21]\dagger}||\eta^{[1]}||\tilde{\delta}_1^{[1]\dagger} + \eta^{[1]\dagger}\mathbf{\Delta}^{[1]}\eta^{[1]} \\
&= |\tilde{\mathbf{h}}_1^{[21]}|^2\tilde{\mathbf{S}}_{11}^{[1]} + \tilde{\mathbf{h}}_1^{[21]}||\eta^{[1]}||\tilde{\mathbf{\Delta}}_1^{[1]} + \tilde{\mathbf{h}}_1^{[21]\dagger}||\eta^{[1]}||\tilde{\mathbf{\Delta}}_1^{[1]\dagger} + \begin{pmatrix} ||\eta^{[1]}|| & 0 \end{pmatrix}\tilde{\mathbf{\Delta}}^{[1]}\begin{pmatrix} ||\eta^{[1]}|| \\ \mathbf{0}_{(M-1)\times 1} \end{pmatrix} \\
&\geq |\tilde{\mathbf{h}}_1^{[21]}|^2\frac{\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2} + \tilde{\mathbf{h}}_1^{[21]}||\eta^{[1]}||\tilde{\delta}_1^{[1]} + \tilde{\mathbf{h}}_1^{[21]\dagger}||\eta^{[1]}||\tilde{\delta}_1^{[1]\dagger} + ||\eta^{[1]}||^2\frac{||\mathbf{h}^{[11]}||^2}{\left(d^{[1]}\right)^2}|\tilde{\delta}_1^{[1]}|^2 \\
&\geq |\tilde{\mathbf{h}}_1^{[21]}|^2\frac{\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2} - 2||\eta^{[1]}||||\tilde{\mathbf{h}}_1^{[21]}|||\tilde{\delta}_1^{[1]}| + ||\eta^{[1]}||^2\frac{||\mathbf{h}^{[11]}||^2}{\left(d^{[1]}\right)^2}|\tilde{\delta}_1^{[1]}|^2 \\
&= \frac{||\eta^{[1]}||^2||\mathbf{h}^{[11]}||^2}{\left(d^{[1]}\right)^2}\left\{|\tilde{\delta}_1^{[1]}| - \frac{|\tilde{\mathbf{h}}_1^{[21]}\left(d^{[1]}\right)^2}{||\eta^{[1]}||||\mathbf{h}^{[11]}||^2}\right\}^2. \qquad (69)
\end{aligned}
$$

$$
\mathbf{S}_*^{[1]} = \mathbf{U}^{[11]}\begin{pmatrix} \frac{\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2} & -\frac{\tilde{\mathbf{h}}_1^{[21]}\left(d^{[1]}\right)^2}{||\eta^{[1]}||^2||\mathbf{h}^{[11]}||^2}\eta^{[1]\dagger} \\ -\frac{\tilde{\mathbf{h}}_1^{[21]\dagger}\left(d^{[1]}\right)^2}{||\eta^{[1]}||^2||\mathbf{h}^{[11]}||^2}\eta^{[1]} & \frac{|\tilde{\mathbf{h}}_1^{[21]}|^2\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2}\frac{\eta^{[1]}\eta^{[1]\dagger}}{||\eta^{[1]}||^4} \end{pmatrix}\mathbf{U}^{[11]\dagger}. \qquad (70)
$$

$$
\mathbf{S}_*^{[1]} = \mathbf{U}^{[11]}\begin{pmatrix} \frac{\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2} & -\frac{\tilde{\mathbf{h}}_1^{[21]}d^{[1]}}{|\tilde{\mathbf{h}}_1^{[21]}|||\mathbf{h}^{[11]}||}\sqrt{P^{[1]} - \frac{\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2}}\frac{\eta^{[1]\dagger}}{||\eta^{[1]}||} \\ -\frac{\tilde{\mathbf{h}}_1^{[21]\dagger}d^{[1]}}{|\tilde{\mathbf{h}}_1^{[21]}|||\mathbf{h}^{[11]}||}\sqrt{P^{[1]} - \frac{\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2}}\frac{\eta^{[1]}}{||\eta^{[1]}||} & \left(P^{[1]} - \frac{\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2}\right)\frac{\eta^{[1]}\eta^{[1]\dagger}}{||\eta^{[1]}||^2} \end{pmatrix}\mathbf{U}^{[11]\dagger}. \qquad (71)
$$

$$
\mathbf{h}^{[21]\dagger}\mathbf{S}_*^{[1]}\mathbf{h}^{[21]} = \left(\sqrt{\left(||\mathbf{h}^{[21]}||^2 - \frac{|\mathbf{h}^{[11]\dagger}\mathbf{h}^{[21]}|^2}{||\mathbf{h}^{[11]}||^2}\right)\left(P^{[1]} - \frac{\left(d^{[1]}\right)^2}{||\mathbf{h}^{[11]}||^2}\right)} - \frac{d^{[1]}|\mathbf{h}^{[11]\dagger}\mathbf{h}^{[21]}|}{||\mathbf{h}^{[11]}||^2}\right)^2. \qquad (72)
$$

$$
\bigcup_{\psi^{[k]}\in\left[0,\ \frac{\pi}{2}\right],\ k=1,\ 2}\left\{\begin{array}{l} R^{[1]} \leq \log_2\left(1 + \frac{P^{[1]}||\mathbf{h}^{[11]}||^2\sin^2(\psi^{[1]})}{1+P^{[2]}||\mathbf{h}^{[12]}||^2\sin^2(\theta^{[2]}-\psi^{[2]})}\right), \\ R^{[2]} \leq \log_2\left(1 + \frac{P^{[2]}||\mathbf{h}^{[22]}||^2\sin^2(\psi^{[2]})}{1+P^{[1]}||\mathbf{h}^{[21]}||^2\sin^2(\theta^{[1]}-\psi^{[1]})}\right) \end{array}\right\}. \qquad (76)
$$

[15] R. Mochaourab and E. Jorswieck, “Exchange economy in two-user multiple-input single-output interference channels, *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 151-164, Apr. 2012.

[16] E. Björnson, M. Bengtsson, and B. Otterson, “Pareto characterization of the multicell MIMO performance region with simple receivers,” *IEEE Tr. Signal Process.*, vol. 60, no. 8, pp. 4464-4469, Aug. 2012.

[17] J. Lindblom, E. Karipidis, and E. G. Larsson, “Efficient computation of Pareto optimal beamforming vectors for the MISO interference channel with successive interference cancellation,” *IEEE Tr. Signal Process.*, vol. 61, no. 19, pp. 4782-4795, Oct. 2013.

[18] S. Boyd and L. Vandenberghe, *Convex optimization,* New York: Cambridge Univ. Press, 2004.

[19] H. Yu, Y. Sung, H. Kim, and Y. H. Lee, “Beam tracking for interference alignment in slowly fading MIMO interference channels: A perturbations approach under a linear framework,” *IEEE Tr. Signal Process.*, vol. 60. no. 4, pp. 1910-1926, Apr. 2012.

[20] E. Björnson, M. Bengtsson, and B. Ottersten, “Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure,” *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142-148, July 2014.

[21] 3GPP TR 36.363 v12.0.0, “Technical Specification Group Radio Access Network; Study on Cell-specific Reference Signals (CRS) interference mitigation for homogenous deployments of LTE,” 2014.

[22] R. A. Horn and C. R. Johnson, *Matrix analysis,* New York: Cambridge Univ. Press, 1988.

[23] M. H. Hayes, *Statistical digital signal processing and modeling,* John Wiley & Sons, 1996.

[24] L. Zheng and D. N. C. Tse, “Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels,” *IEEE Tr. Inf. Theory,* vol. 49. no. 5, pp. 1073-1096, May 2003.

[25] S. A. Jafar and M. Fakhereddin, “Degrees of freedom for the MIMO interference channel,” *IEEE Tr. Inf. Theory,* vol. 53, No. 7, pp. 2637-2642, July 2007.

[26] T. M. Cover and J. A. Thomas, *Elements of information theory,* John Wiley & Sons, 1991.