Max-Min Fairness Design for MIMO Interference Channels: a Minorization-Maximization Approach

Journal:	*Transactions on Signal Processing*
Manuscript ID	T-SP-22915-2017
Manuscript Type:	Regular Paper
Date Submitted by the Author:	14-Nov-2017
Complete List of Authors:	Naghsh, Mohammad; Isfahan University of Technology, ECE; Uppsala University, Information Technology Masjedy, Maryam Adibi, Arman Stoica, Petre; Uppsala University, Department of Information Technology
EDICS:	125. SPC-MISG MIMO signalling < SPC SIGNAL PROCESSING FOR COMMUNICATIONS, 63. NEG-PHYL Physical layer issues < NEG SIGNAL PROCESSING FOR NETWORKS AND GRAPHS, 75. OPT-NCVX Non-convex methods for SP < OPT OPTIMIZATION METHODS FOR SIGNAL PROCESSING

Max-Min Fairness Design for MIMO Interference Channels: a Minorization-Maximization Approach

Mohammad Mahdi Naghsh*, *Member, IEEE*, Maryam Masjedi, Arman Adibi, and Petre Stoica, *Fellow, IEEE*

***Abstract*—We address the problem of linear precoder (beamformer) design in a multiple-input multiple-output interference channel (MIMO-IC). The aim is to design the transmit covariance matrices in order to achieve max-min utility fairness for all users. The corresponding optimization problem is non-convex and NP-hard in general. We devise an efficient algorithm based on the minorization-maximization (MM) technique to obtain quality solutions to this problem. The proposed method solves a second-order cone convex program (SOCP) at each iteration and converges to a stationary point of the problem under mild conditions. We also extend our algorithm to the case where there are uncertainties in the noise covariance matrices or channel state information (CSI). Simulation results show the effectiveness of the proposed method compared with its main competitor.**

Keywords: Interference channel, Minorization-maximization (MM), Max-min fairness, MIMO, Rate optimization.

TABLE I: Notation

$\|\mathbf{x}\|_n$:	the l_n-norm of the vector $\mathbf{x}$, defined as $\left(\sum_k \|x(k)\|^n\right)^{\frac{1}{n}}$
$\|\mathbf{X}\|_2$:	the spectral norm of the matrix $\mathbf{X}$ i.e. the largest singular value of $\mathbf{X}$
$\mathbf{X}^H$:	the conjugate transpose of matrix $\mathbf{X}$
tr($\mathbf{X}$):	the trace of matrix $\mathbf{X}$
$\lambda_{max}(\mathbf{X})$:	the maximum eigenvalue of hermitian matrix $\mathbf{X}$
$\mathbf{A}\otimes\mathbf{B}$:	the Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{X}\succeq\mathbf{Y}$:	$\mathbf{X}-\mathbf{Y}$ is positive semidefinite
$\mathbf{X}\succ\mathbf{Y}$:	$\mathbf{X}-\mathbf{Y}$ is positive definite
$\mathbf{X}^{\frac{1}{2}}$:	the Hermitian square root of the positive semidefinite matrix $\mathbf{X}$ i.e. $\mathbf{X}=\mathbf{X}^{\frac{1}{2}}(\mathbf{X}^{\frac{1}{2}})^H$
vec($\mathbf{X}$):	the vector obtained by column-wise stacking of $\mathbf{X}$
$\mathbf{I}_n$:	the identity matrix of $\mathbb{C}^{n\times n}$
$\mathbb{R}$:	the set of real numbers
$\mathbb{C}$:	the set of complex numbers
$\Re(x)$:	the real part of x
$\mathbb{R}_+$:	the set of nonnegative real numbers
$\mathbb{S}_N^+$:	the set of positive semidefinite matrices of $\mathbb{C}^{N\times N}$
$\mathbb{S}_N^{++}$:	the set of positive definite matrices of $\mathbb{C}^{N\times N}$

I. Introduction

We consider the linear precoder design problem in a MIMO interference channel in which a set of transmitter-receiver pairs communicate over a shared (time or frequency) resource. The precoder matrices can be designed to improve the network performance from a sum rate or minimum rate (max-min fairness) point of view [1]–[16].

The problem of linear transceiver design under the max-min fairness criterion has been widely studied in the literature [1]–[10]. In [1] and [2], the power control problem under a max-min signal-to-interference-plus-noise ratio (SINR) criterion has been studied and performance bounds for power control algorithms have been obtained. The problem of designing the transmitter precoder that maximizes the minimum rate of users in a multiple-input single-output (MISO) network is also studied in [3]–[6]. The authors of [7] maximized the worst case SINR subject to a power constraint on the design precoder matrices in a MIMO-IC and showed this problem can be solved using standard conic optimization packages. The authors of [17] considered the max-min fairness precoder design in a single-input multiple-output (SIMO) IC and showed that this problem can be solved in polynomial time. In [8], the authors recast the max-min fairness problem in MIMO-IC as the problem of finding the globally optimal transceiver that maximizes the minimum SINR among all users. They showed that when each transmitter (receiver) is equipped with more than one antenna and each receiver (transmitter) is equipped with more than two antennas, the problem is strongly NP-hard. To deal with the problem they proposed two algorithms which decompose the original NP-hard problem into a series of convex subproblems. In [9] and [10], the authors considered the problem of linear precoder design for MIMO-IC under a max-min fairness criterion and showed that when there are at least two antennas at each transmitter and receiver, the problem belongs to a class of NP-hard problems. They proposed an algorithm that computes an approximate solution to the original problem. Note that in the aforementioned works, the precoder matrices are designed for the cases in which the number of symbols in a stream is assumed to be apriori known.

In this paper, we consider the precoder design for multiple-input multiple-output (MIMO) interference channels. We aim to design the transmit covariance matrices (the number of transmitted symbols can be unknown) under a max-min fairness criterion for systems using the conventional linear minimum mean square error (LMMSE) receivers. We propose a computationally efficient algorithm based on the Minorization-maximization (MM)[1] technique to obtain quality solutions to this design problem. The obtained solutions are stationary points of the problem under mild conditions. Compared with [9] and [10], we consider a more general case by designing the precoder covariance matrices, which means that the optimal number of symbols in a stream is also obtained as a by-product. We also extend our algorithm to the case where there are uncertainties in the noise covariance matrices or in the

M. M. Naghsh, M. Masjedi and A. Adibi are with the Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran. P. Stoica is with the Department of Information Technology, Uppsala University, Uppsala, SE 75105, Sweden.

*Please address all the correspondence to M. M. Naghsh, Phone: (+98) 31-33912450; Fax: (+98) 31-33912451; Email: mm_naghsh@cc.iut.ac.ir

[1]Also known as MaMi or MiMa in the literature [16].

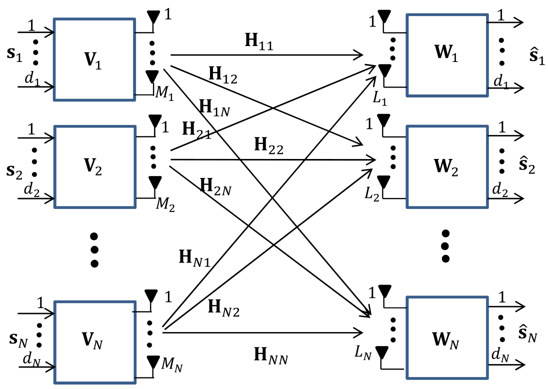


Fig. 1: A generic MIMO-IC.

CSI.

The rest of the paper is organized as follows. The signal and system model along with the associated max-min precoder covariance design problem are described in Section II. The proposed method for designing the precoder covariances and , in particular, the precoder matrices under the max-min fairness criterion is derived in Section III. Precoder design under noise covariance uncertainty and imperfect CSI is considered in Section IV. Numerical results are provided in Section V and, finally, conclusions are drawn in Section VI.

Table I summarizes the notation used throughout this paper.

II. System model and Problem formulation

Consider N transmit-receive pairs communicating over a MIMO interference channel as shown in Fig. 1. We assume that the ith transmitter and the jth receiver are equipped with M_i and L_j antennas, respectively. The ith transmitter uses the linear precoder matrix $\mathbf{V}_i \in \mathbb{C}^{M_i \times d_i}$ to convert the symbol stream $\mathbf{s}_i \in \mathbb{C}^{d_i \times 1}$ (consisting of d_i independent data symbols) into the vector $\mathbf{d}_i \in \mathbb{C}^{M_i \times 1}$, i.e.,

$$\mathbf{d}_i = \mathbf{V}_i \mathbf{s}_i \tag{1}$$

and sends it over flat fading channels. The received signal at the ith receiver is given by:

$$\mathbf{y}_i = \underbrace{\mathbf{H}_{ii}\mathbf{d}_i}_{\text{desired signal}} + \underbrace{\sum_{j\neq i}\mathbf{H}_{ji}\mathbf{d}_j + \mathbf{n}_i}_{\text{interference plus noise}} \tag{2}$$

where $\mathbf{H}_{ji} \in \mathbb{C}^{L_i \times M_j}$ denotes the channel matrix between the jth transmitter and the ith receiver. Also, $\mathbf{n}_i \in \mathbb{C}^{L_i \times 1}$ is the circularly symmetric complex Gaussian (CSCG) noise at the ith receiver with zero mean and covariance matrix $\mathbf{\Gamma}_i \in \mathbb{S}^{++}_{L_i}$. The ith receiver uses the linear decoder matrix $\mathbf{W}_i \in \mathbb{C}^{d_i \times L_i}$ to obtain $\hat{\mathbf{s}}_i \in \mathbb{C}^{d_i \times 1}$ which is an estimate of the transmitted vector $\mathbf{s}_i$:

$$\begin{aligned}\hat{\mathbf{s}}_i =& \mathbf{W}_i\mathbf{y}_i \\ =& \mathbf{W}_i\mathbf{H}_{ii}\mathbf{V}_i\mathbf{s}_i + \mathbf{W}_i\sum_{j\neq i}\mathbf{H}_{ji}\mathbf{V}_j\mathbf{s}_j + \mathbf{W}_i\mathbf{n}_i\end{aligned} \tag{3}$$

Assuming the symbol stream $\mathbf{s}_i$ is a Gaussian random vector with zero mean and covariance matrix $\mathbf{I}_{d_i}$, the rate of the ith user is given by [18]:

$$R_i = \log\det\left(\mathbf{I}_{d_i} + \mathbf{W}_i\mathbf{H}_{ii}\mathbf{V}_i\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{W}_i^H\left(\mathbf{W}_i\mathbf{C}_{\bar{i}}\mathbf{W}_i^H\right)^{-1}\right) \tag{4}$$

with $\mathbf{C}_{\bar{i}}$ being the interference plus noise covariance matrix defined as

$$\mathbf{C}_{\bar{i}} = \mathbf{\Gamma}_i + \sum_{j\neq i}\mathbf{H}_{ji}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ji}^H \tag{5}$$

Employing the conventional LMMSE decoder at the receivers means that the ith decoder matrix is given by

$$\mathbf{W}_i^{\text{LMMSE}} = \mathbf{V}_i^H\mathbf{H}_{ii}^H\left(\sum_{j=1}^{N}\mathbf{H}_{ji}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ji}^H + \mathbf{\Gamma}_i\right)^{-1} \tag{6}$$

By substituting (6) into (4), it can be verified that (for completeness we include a proof of (6) and (7) in Appendix A):

$$R_i = \log\det\left(\mathbf{I}_{d_i} + \mathbf{V}_i^H\mathbf{H}_{ii}^H\left[\mathbf{\Gamma}_i + \sum_{j\neq i}\mathbf{H}_{ji}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ji}^H\right]^{-1}\mathbf{H}_{ii}\mathbf{V}_i\right) \tag{7}$$

Remark 1. Interestingly, using the decoder $\mathbf{W}_i' = \mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}$, see (5), leads to the same rate as the LMMSE, see (7). Furthermore, the matrix $\mathbf{W}_i'$ maximizes the rate in (4). To see this, use standard properties of Schur complement to verify that the inequality

$$\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{W}_i^H(\mathbf{W}_i\mathbf{C}_{\bar{i}}\mathbf{W}^H)^{-1}\mathbf{W}_i\mathbf{H}_{ii}\mathbf{V}_i \preceq \mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}\mathbf{H}_{ii}\mathbf{V}_i \tag{8}$$

is equivalent to the positive semi-definitness of the matrix:

$$\mathbf{\Phi}_i = \begin{bmatrix}\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}\mathbf{H}_{ii}\mathbf{V}_i & \mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{W}_i^H \\ \mathbf{W}_i\mathbf{H}_{ii}\mathbf{V}_i & \mathbf{W}_i\mathbf{C}_{\bar{i}}\mathbf{W}^H\end{bmatrix}. \tag{9}$$

Now, observe that the matrix $\mathbf{\Phi}_i$ above indeed is in $\mathbb{S}^{+}_{d_i}$ because it can be decomposed as $\mathbf{\Phi}_i = \mathbf{\Theta}_i\mathbf{\Theta}_i^H$ with

$$\mathbf{\Theta}_i = \begin{bmatrix}\mathbf{V}_i^H\mathbf{H}_{ii}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_i\end{bmatrix}\begin{bmatrix}\mathbf{C}_{\bar{i}}^{-1/2} \\ \mathbf{C}_{\bar{i}}^{1/2}\end{bmatrix} \tag{10}$$

Therefore, (8) holds true. Moreover, it can be verified that by substituting $\mathbf{W}_i = \mathbf{W}_i' = \mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}$ in (8), the left-hand side becomes $\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}\mathbf{H}_{ii}\mathbf{V}_i$ which is equal to the right-hand side. Therefore, $\mathbf{W}_i'$ maximizes the rate in (4). Because the LMMSE decoder in (6) and $\mathbf{W}_i'$ yield the same rate, we conclude that the LMMSE decoder maximizes the rate as well. ■

Using Sylvester's determinant property, i.e. $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$, the rate R_i in (7) can be rewritten as

$$R_i = \log\det\left(\mathbf{I}_{L_i} + \mathbf{H}_{ii}\mathbf{Q}_i\mathbf{H}_{ii}^H\left[\mathbf{\Gamma}_i + \sum_{j\neq i}\mathbf{H}_{ji}\mathbf{Q}_j\mathbf{H}_{ji}^H\right]^{-1}\right) \tag{11}$$

where $\mathbf{Q}_i \triangleq \mathbf{V}_i\mathbf{V}_i^H \in \mathbb{C}^{M_i\times M_i}, i = 1,\ldots,N$, are the precoder covariance matrices. In this paper, the goal is to design the precoder covariance matrices $\{\mathbf{Q}_i\}_{i=1}^N$ to maximize the minimum rate of the users, which can be cast as the following problem:

$$\begin{aligned} \max_{\{\mathbf{Q}_i\}_{i=1}^N} \min_{i=1,2,\ldots,N} \quad & R_i \\ \text{s.t.} \quad & \text{tr}\{\mathbf{Q}_i\} \leq p_i \quad && \forall i = 1,2,\ldots,N \\ & \mathbf{Q}_i \succeq \mathbf{0} \quad && \forall i = 1,2,\ldots,N \end{aligned} \tag{12}$$

where p_i is the power available to the ith transmitter. In the next section, we assume that the noise covariance matrices $\{\mathbf{\Gamma}_i\}_{i=1}^N$ as well as the channel matrices $\{\mathbf{H}_{ij}\}_{i,j=1}^N$ are exactly known. We consider the case of uncertain a priori knowledge in Section IV.

III. The Proposed Method

It can be shown that the design problem in (12) is non-convex and NP-hard in general [10]. In what follows we devise a method based on the minorization-maximization (MM) technique [19] to tackle this problem.

In (12) the constraints are convex but the objective function is non-convex. Therefore we will apply the MM technique to the objective function. For this purpose, we first introduce the following proposition.

Proposition 1. The rate R_i, see (11), can be rewritten as:

$$R_i = \log\det(\mathbf{U}^H \mathbf{B}_i^{-1} \mathbf{U}) \tag{13}$$

where $\mathbf{U}$ and $\mathbf{B}_i$ are defined as,

$$\mathbf{U} \triangleq \begin{bmatrix} \mathbf{I}_{M_i} & \mathbf{0}_{M_i \times L_i} \end{bmatrix}^T \tag{14}$$

and

$$\mathbf{B}_i = \begin{bmatrix} \mathbf{I}_{M_i} & \widetilde{\mathbf{V}}_i^H \mathbf{H}_{ii}^H \\ \mathbf{H}_{ii}\widetilde{\mathbf{V}}_i & \mathbf{\Gamma}_i + \sum_{j=1}^N \mathbf{H}_{ji}\widetilde{\mathbf{V}}_j \widetilde{\mathbf{V}}_j^H \mathbf{H}_{ji}^H \end{bmatrix} \tag{15}$$

with $\widetilde{\mathbf{V}}_i \triangleq \mathbf{Q}_i^{\frac{1}{2}} \in \mathbb{C}^{M_i \times M_i}$.

Proof: See Appendix B. ■

By using (13), the problem in (12) can be rewritten as follows

$$\begin{aligned} \max_{\{\widetilde{\mathbf{V}}_i\}_{i=1}^N,} \min_{i=1,\cdots,N} \; & \log\det(\mathbf{U}^H \mathbf{B}_i^{-1} \mathbf{U}) \\ \text{s.t.} \; & \text{tr}\{\widetilde{\mathbf{V}}_i \widetilde{\mathbf{V}}_i^H\} \leq p_i, \; \forall i = 1,2,\ldots,N \end{aligned} \tag{16}$$

The following lemma (see, e.g., [20]) lays the ground for applying MM to (16).

Lemma 1. The function $f(\mathbf{X}) = \log\det(\mathbf{U}^H \mathbf{X}^{-1} \mathbf{U})$: $\mathbb{S}_N^{++} \to \mathbb{R}_+$ is convex for any full column rank matrix $\mathbf{U}$. ■

Using Lemma 1 and noting that $\mathbf{B}_i \succ \mathbf{0}$, $\forall i = 1,2,\ldots,N$ (see Appendix C), the objective function in problem (16) can be minorized at a given $\overline{\mathbf{B}}_i$ as follows

$$\begin{aligned} \log\det(\mathbf{U}^H \mathbf{B}_i^{-1} \mathbf{U}) \geq\; & \log\det(\mathbf{U}^H \overline{\mathbf{B}}_i^{-1} \mathbf{U}) \\ & - \text{tr}\{\mathbf{F}_i(\mathbf{B}_i - \overline{\mathbf{B}}_i)\} \end{aligned} \tag{17}$$

where $\mathbf{F}_i$ is given by (see Appendix D):

$$\mathbf{F}_i = \overline{\mathbf{B}}_i^{-1} \mathbf{U}(\mathbf{U}^H \overline{\mathbf{B}}_i^{-1} \mathbf{U})^{-1} \mathbf{U}^H \overline{\mathbf{B}}_i^{-1}. \tag{18}$$

Note that $\overline{\mathbf{B}}_i$ can be chosen as the value of $\mathbf{B}_i$ at the $(\kappa-1)$th iteration. Consequently, let

$$\begin{aligned} g_i^{(\kappa)}(\widetilde{\mathbf{V}}_1, \cdots, \widetilde{\mathbf{V}}_N) \triangleq\; & \log\det(\mathbf{U}^H (\mathbf{B}_i^{(\kappa-1)})^{-1} \mathbf{U}) \\ & - \text{tr}\{\mathbf{F}_i(\mathbf{B}_i - \mathbf{B}_i^{(\kappa-1)})\} \end{aligned} \tag{19}$$

(we omit the dependence of $\mathbf{F}_i$ on the iteration number to simplify the notation). Then it follows from (17) that the objective function in (16) can be minorized at the κth iteration by:

$$\min_{i=1,\cdots,N} \log\det(\mathbf{U}^H \mathbf{B}_i^{-1} \mathbf{U}) \geq \min_{i=1,\cdots,N} g_i^{(\kappa)}(\widetilde{\mathbf{V}}_1, \cdots, \widetilde{\mathbf{V}}_N) \tag{20}$$

The MM technique that makes use of (20), consists of iteratively solving the following problem (for $\kappa = 1,2,...$):

$$\begin{aligned} \max_{\{\widetilde{\mathbf{V}}_i\}_{i=1}^N} \min_{i=1,\cdots,N} \; & g_i^{(\kappa)}(\widetilde{\mathbf{V}}_1, \cdots, \widetilde{\mathbf{V}}_N) \\ \text{s.t.} \; & \text{tr}\{\widetilde{\mathbf{V}}_i \widetilde{\mathbf{V}}_i^H\} \leq p_i, \; \forall i = 1,2,\ldots,N \end{aligned} \tag{21}$$

Next, we rewrite (21) using an auxiliary variable t:

$$\begin{aligned} \max_{\{\widetilde{\mathbf{V}}_i\}_{i=1}^N, t} \quad & t \\ \text{s.t.} \quad & g_i^{(\kappa)}(\widetilde{\mathbf{V}}_1, \cdots, \widetilde{\mathbf{V}}_N) \geq t, \quad && \forall i = 1,2,\ldots,N \\ & \text{tr}\{\widetilde{\mathbf{V}}_i \widetilde{\mathbf{V}}_i^H\} \leq p_i, \quad && \forall i = 1,2,\ldots,N \end{aligned} \tag{22}$$

To derive an explicit expression for the constraints $g_i^{(\kappa)}(\widetilde{\mathbf{V}}_1, \cdots, \widetilde{\mathbf{V}}_N) \geq t$ in terms of the design variables $\{\widetilde{\mathbf{V}}_i\}_{i=1}^N$, let

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_{11_{M_i \times M_i}} & \mathbf{F}_{12_{M_i \times L_i}} \\ \mathbf{F}_{21_{L_i \times M_i}} & \mathbf{F}_{22_{L_i \times L_i}} \end{pmatrix}. \tag{23}$$

Then, combining (23) and (15), it can be verified that:

$$\begin{aligned} \text{tr}\{\mathbf{F}_i \mathbf{B}_i\} =\; & 2\Re\{\text{tr}\{(\mathbf{F}_i)_{12} \mathbf{H}_{ii} \widetilde{\mathbf{V}}_i\}\} + \text{tr}\{(\mathbf{F}_i)_{11}\} \\ & + \text{tr}\{(\mathbf{F}_i)_{22} \mathbf{\Gamma}_i\} \\ & + \text{tr}\{(\mathbf{F}_i)_{22} \sum_{j=1}^N \mathbf{H}_{ji} \widetilde{\mathbf{V}}_i \widetilde{\mathbf{V}}_i^H \mathbf{H}_{ji}^H\}. \end{aligned} \tag{24}$$

Defining $\mathbf{x}_i \triangleq \text{vec}(\widetilde{\mathbf{V}}_i)$ and using properties of the vectorization operator, viz. $\text{tr}(\mathbf{AB}) = (\text{vec}(\mathbf{A}^T))^T \text{vec}(\mathbf{B})$ and $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$ (for any arbitrary matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$), we can rewrite the first and the last terms in (24) as:

$$\text{tr}\{(\mathbf{F}_i)_{12} \mathbf{H}_{ii} \widetilde{\mathbf{V}}_i\} = \mathbf{b}_i^H \mathbf{x}_i \tag{25}$$

$$\text{tr}\{(\mathbf{F}_i)_{22} \sum_{j=1}^N \mathbf{H}_{ji} \widetilde{\mathbf{V}}_i \widetilde{\mathbf{V}}_i^H \mathbf{H}_{ji}^H\} = \sum_{j=1}^N \mathbf{x}_j^H \mathbf{G}_{ji} \mathbf{x}_j \tag{26}$$

where $\mathbf{b}_i \triangleq \text{vec}(\mathbf{H}_{ii}^H (\mathbf{F}_i)_{12}^H)$ and $\mathbf{G}_{ji} \triangleq \mathbf{I}_M \otimes (\mathbf{H}_{ji}^H (\mathbf{F}_i)_{22} \mathbf{H}_{ji})$. Note that according to the Kronecker product properties, $\mathbf{G}_{ji} \succeq \mathbf{0}$, because $\mathbf{H}_{ji}^H (\mathbf{F}_i)_{22} \mathbf{H}_{ji}$ is positive semidefinite.

Finally, the problem in (22) that is solved at the κth iteration of MM can be rewritten as the following optimization:

$$\max_{t,\{\mathbf{x}_i\}_{i=1}^N} \; t \tag{27}$$

$$\text{s.t.} \quad C_i^{(\kappa-1)} + 2\Re\{(\mathbf{b}_i^{(\kappa-1)})^H\mathbf{x}_i\} + \sum_{j=1}^N \mathbf{x}_j^H\mathbf{G}_{ji}^{(\kappa-1)}\mathbf{x}_j \leq -t \quad \forall i = 1,2,\ldots,N$$

$$\|\mathbf{x}_i\|_2^2 \leq p_i \quad \forall i = 1,2,\ldots,N$$

where C_i is the following real-valued constant:

$$C_i^{(\kappa-1)} = -\log\det(\mathbf{U}^H(\mathbf{B}_i^{(\kappa-1)})^{-1}\mathbf{U}) - \mathrm{tr}\{\mathbf{F}_i^{(\kappa-1)}\mathbf{B}_i^{(\kappa-1)}\} + \mathrm{tr}\{(\mathbf{F}_i^{(\kappa-1)})_{11}\} + \mathrm{tr}\{(\mathbf{F}_i^{(\kappa-1)})_{22}\mathbf{\Gamma}_i\}. \tag{28}$$

Note that (27) is a convex problem with a linear objective and quadratic constrains. Hence it can be expressed as a second-order cone program (SOCP). The proposed algorithm, which is based on iteratively solving (27), is summarized in Table II. In the first step, we initialize the algorithm with i.i.d. CSCG random variables after making them feasible by normalization i.e. $\|\mathbf{x}_i\|_2^2 \leq p_i$. In the second step, we use efficient methods such as interior point algorithms to solve the problem in (27) [21]. After updating the parameters in step 3, the stop criterion is checked and steps 1 to 3 are repeated until this criterion is satisfied.

Remark 2. (Convergence): The proposed algorithm can be shown to be locally convergent. To this end, observe that for the minimum rate at the κth iteration, we have that:

$$\min_i \log\det(\mathbf{U}^H(\mathbf{B}_i^{(\kappa-1)})^{-1}\mathbf{U}) = \min_i g_i^{(\kappa)}(\widetilde{\mathbf{V}}_1^{(\kappa-1)},\cdots,\widetilde{\mathbf{V}}_N^{(\kappa-1)}) \tag{29}$$

$$\leq \min_i g_i^{\kappa}(\widetilde{\mathbf{V}}_1^{(\kappa)},\cdots,\widetilde{\mathbf{V}}_N^{(\kappa)}) \leq \min_i \log\det(\mathbf{U}^H(\mathbf{B}_i^{(\kappa)})^{-1}\mathbf{U})$$

The first inequality in (29) holds due to the maximization step at the κth iteration and the second one is satisfied due the definition of the minorizer, see (20). Combining (29) and the fact that the objective function is upper bounded, it follows that the sequence of objective values converges. ∎

Remark 3. (Calculating $\mathbf{V}_i$ from $\mathbf{Q}_i$): At the convergence of the proposed method, the optimized transmit covariances $\{\mathbf{Q}_i = \widetilde{\mathbf{V}}_i\widetilde{\mathbf{V}}_i^H\} \in \mathbb{C}^{M_i\times M_i}$ are obtained. Next, the precoder matrices $\{\mathbf{V}_i\} \in \mathbb{C}^{M_i\times d_i}$ are obtained as square roots of the $\{\mathbf{Q}_i\}$: $\mathbf{V}_i\mathbf{V}_i^H = \mathbf{Q}_i$. Note that the so-obtained precoder matrix $\mathbf{V}_i$ is not unique but this has no effect on the rate. Indeed the rate R_i in (7) is a many-to-one function of $\mathbf{V}_i$ as $\mathbf{V}_i$ and $\mathbf{V}_i\mathbf{A}$ lead to the same R_i for any matrix $\mathbf{A}$ satisfying $\mathbf{A}\mathbf{A}^H = \mathbf{I}$. Also note that whenever $\mathbf{Q}_i$ is (nearly) singular, one can perform a thresholding operation on its eigenvalues and reduce the number of columns of $\mathbf{V}_i$ accordingly. Finally observe that the optimized stream lengths $\{d_i\}_{i=1}^N$ are given once we have $\{\mathbf{V}_i\} \in \mathbb{C}^{M_i\times d_i}$. ∎

Remark 4. (Precoder design for given $\{d_i\}_{i=1}^N$): As explained above, by designing the precoder covariance matrices $\{\mathbf{Q}_i\}_{i=1}^N$, we simultaneously design the precoder matrices $\{\mathbf{V}_i\}_{i=1}^N$ and the number of their columns $\{d_i\}_{i=1}^N$, i.e. the lengths of the symbol streams. In some cases, the length of symbols streams $\{d_i\}_{i=1}^N$ are given and the precoder matrices $\{\mathbf{V}_i\}_{i=1}^N$ should be designed directly. To deal with this case, we can modify the proposed method simply by replacing $\widetilde{\mathbf{V}}_i \in \mathbb{C}^{M_i\times M_i}$ in (16) with $\mathbf{V}_i \in \mathbb{C}^{M_i\times d_i}$. ∎

TABLE II: The proposed method for the max-min rate design of the transmit covariance matrices in MIMO-IC.

Step 1: Initialize $\{\mathbf{x}_i\}_{i=1}^N$ with complex random vectors in $\mathbb{C}^{M_i^2\times 1}$ such that they satisfy $\|\mathbf{x}_i\|_2^2 \leq p_i$.
Step 2: Solve the (convex) SOCP problem in (27).
Step 3: Update $\mathbf{b}_i$, $\mathbf{G}_{ji}$, and C_i according to equations (25), (26), and (28), respectively.
Step 4: Repeat steps 1 and 2 until a pre-defined stop criterion is satisfied, e.g $\lvert t^{(\kappa)} - t^{(\kappa-1)}\rvert \leq \epsilon$, for a given $\epsilon > 0$.

IV. Precoder Design in the Presence of a Priori Knowledge Uncertainty

In practice there always exist uncertainties in the noise covariance and the channel state information. In this section we will consider these uncertainties in the design problem.

We first consider the effect of imperfect CSI due to channel estimation errors. Using the conventional LMMSE estimator, the channels can be modeled as [22]:

$$\mathbf{H}_{ji} = \hat{\mathbf{H}}_{ji} + \mathbf{Z}_{ji} \tag{30}$$

where $\hat{\mathbf{H}}_{ji}$ is the estimate of the true channel $\mathbf{H}_{ji}$ and $\mathbf{Z}_{ji}$ is the channel estimation error which is assumed to be uncorrelated with $\hat{\mathbf{H}}_{ji}$. Assuming the entries of $\mathbf{H}_{ji}$ are i.i.d random variables (RVs) with variances σ_{ji}^2, the entries of $\hat{\mathbf{H}}_{ji}$ and $\mathbf{Z}_{ji}$ will be i.i.d RVs with variances $\rho_{ji}^2\sigma_{ji}^2$ and $(1-\rho_{ji}^2)\sigma_{ji}^2$, respectively. The parameter $\rho_{ji} \in [0,1]$ quantifies the estimation accuracy, in particular if $\rho_{ji} = 1$, $\hat{\mathbf{H}}_{ji} = \mathbf{H}_{ji}$ and CSI is perfect.

Substituting (30) in (2), we obtain:

$$\mathbf{y}_i = \underbrace{\hat{\mathbf{H}}_{ii}\mathbf{V}'_i\mathbf{s}_i}_{\text{desired signal}} + \underbrace{\sum_{j\neq i}\hat{\mathbf{H}}_{ji}\mathbf{V}'_j\mathbf{s}_j + \sum_{j=1}^N \mathbf{Z}_{ji}\mathbf{V}'_j\mathbf{s}_j + \mathbf{n}_i}_{\text{interference plus estimation error and noise}} \tag{31}$$

with $\mathbf{V}'_i$ being the precoder matrix of the ith transmitter designed under imperfect CSI. It can be proved that:

$$\mathbb{E}\left\{\mathbf{Z}_{ji}\mathbf{V}'_j\mathbf{V}'^H_j\mathbf{Z}_{ji}^H\right\} = (1-\rho_{ji}^2)\sigma_{ji}^2\,\mathrm{tr}\{\mathbf{V}'_j\mathbf{V}'^H_j\}\mathbf{I}_{L_i} \tag{32}$$

Therefore, the LMMSE decoder will be:

$$\hat{\mathbf{W}}_i^{\text{LMMSE}} = \mathbf{V}'^H_i\hat{\mathbf{H}}_{ii}^H\left(\sum_{j=1}^N \hat{\mathbf{H}}_{ji}\mathbf{V}'_j\mathbf{V}'^H_j\hat{\mathbf{H}}_{ji}^H + \sum_{j=1}^N (1-\rho_{ji}^2)\sigma_{ji}^2\,\mathrm{tr}\{\mathbf{V}'_j\mathbf{V}'^H_j\}\mathbf{I}_{L_i} + \mathbf{\Gamma}_i\right)^{-1} \tag{33}$$

Let $\mathbf{Q'}_j \triangleq \mathbf{V'}_j \mathbf{V'}_j^H$ be the precoder covariance matrices in the imperfect CSI case. The rate $\hat{R}_i$ of the ith user for this case becomes:

$$\hat{R}_i = \log\det\left(\mathbf{I}_{L_i} + \hat{\mathbf{H}}_{ii}\mathbf{Q'}_i\hat{\mathbf{H}}_{ii}^H\left[\mathbf{\Gamma}_i + \sum_{j\neq i}\hat{\mathbf{H}}_{ji}\mathbf{Q'}_j\hat{\mathbf{H}}_{ji}^H + \sum_{j=1}^{N}(1-\rho_{ji}^2)\sigma_{ji}^2 \,\mathrm{tr}\{\mathbf{Q}'_j\}\mathbf{I}_{L_j}\right]^{-1}\right) \tag{34}$$

Next, we also consider the uncertainty of the noise covariance matrices, which can be modeled as [20]:

$$\|\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i\|_2 \leq \zeta_i, \quad \forall i = 1, \cdots, N \tag{35}$$

where $\hat{\mathbf{\Gamma}}_i$s are known positive definite matrices (initial guesses of the covariance matrices) and ζ_is are positive scalars that determine the size of the uncertainty regions.

We can robustify the design method with respect to a priori knowledge uncertainty by considering the following reformulation of the optimization problem:

$$\begin{aligned}
\max_{\{\mathbf{Q'}_i\}_{i=1}^N} \; &\min_{i=1,\dots,N} \; \min_{\{\mathbf{\Gamma}_i\}_{i=1}^N} \quad \hat{R}_i \\
\text{s.t.} \quad & \mathrm{tr}\{\mathbf{Q}'_i\} \leq p_i, \qquad \forall i = 1, 2, \dots, N \\
& \|\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i\|_2 \leq \zeta_i, \qquad \forall i = 1, 2, \cdots, N \\
& \mathbf{\Gamma}_i \succeq \mathbf{0}, \mathbf{Q}'_i \succeq \mathbf{0} \qquad \forall i = 1, 2, \cdots, N
\end{aligned} \tag{36}$$

where $\hat{R}_i$ is as given in (34). In what follows we present a theorem which shows that the problem in (36) can be dealt with via a modified version of the method proposed in Section III.

Theorem 1. Let R'_i be defined as:

$$R'_i = \log\det\left(\mathbf{I}_{L_i} + \hat{\mathbf{H}}_{ii}\mathbf{Q'}_i\hat{\mathbf{H}}_{ii}^H\left[\mathbf{\Gamma}'_i + \sum_{j\neq i}\hat{\mathbf{H}}_{ji}\mathbf{Q'}_j\hat{\mathbf{H}}_{ji}^H + \sum_{j=1}^{N}(1-\rho_{ji}^2)\sigma_{ji}^2 \,\mathrm{tr}\{\mathbf{Q}'_j\}\mathbf{I}_{L_j}\right]^{-1}\right) \tag{37}$$

where $\mathbf{\Gamma}'_i = \hat{\mathbf{\Gamma}}_i + \zeta_i \mathbf{I}_{L_i}$. The problem

$$\begin{aligned}
\max_{\{\mathbf{Q}'_i\}_{i=1}^N} \; &\min_{i=1,2,\dots,N} \quad R'_i \\
\text{s.t.} \quad & \mathrm{tr}\{\mathbf{Q}'_i\} \leq p_i, \quad \forall i = 1, 2, \dots, N \\
& \mathbf{Q}'_i \succeq \mathbf{0} \qquad \forall i = 1, 2, \cdots, N
\end{aligned} \tag{38}$$

is equivalent to the problem in (36) in the sense that these two problems share the same solution $\{\mathbf{Q}'_i\}_{i=1}^N$.

Proof: Noting that the inner problem of (36) is separable w.r.t i, we consider it for a fixed i:

$$\begin{aligned}
\min_{\mathbf{\Gamma}_i \succeq \mathbf{0}} \quad & \hat{R}_i \\
\text{s.t.} \quad & \|\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i\|_2 \leq \zeta_i
\end{aligned} \tag{39}$$

Note that $\|\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i\|_2 = \sqrt{\lambda_{max}\left((\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i)^H(\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i)\right)}$ and the matrix $\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i$ is Hermitian; therefore, the constraint $\|\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i\|_2 \leq \zeta_i$ is equivalent to $\max_m |\lambda_m(\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i)| \leq \zeta_i$ with $\lambda_m(\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i)$ being the mth eigenvalue of the matrix $\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i$. Therefore, we have that

$$\lambda_m(\mathbf{\Gamma}_i - \hat{\mathbf{\Gamma}}_i) \in [-\zeta_i, \zeta_i], \forall m = 1, 2, ..., L_i \tag{40}$$

Consequently, it can be verified that the constraint in (39) is equivalent to

$$\hat{\mathbf{\Gamma}}_i - \zeta_i \mathbf{I}_{L_i} \preceq \mathbf{\Gamma}_i \preceq \hat{\mathbf{\Gamma}}_i + \zeta_i \mathbf{I}_{L_i} \tag{41}$$

and therefore, that the problem (39) is equivalent to the following optimization:

$$\begin{aligned}
\min_{\mathbf{\Gamma}_i \succeq \mathbf{0}} \quad & \hat{R}_i \\
\text{s.t.} \quad & \hat{\mathbf{\Gamma}}_i - \zeta_i \mathbf{I}_{L_i} \preceq \mathbf{\Gamma}_i \preceq \hat{\mathbf{\Gamma}}_i + \zeta_i \mathbf{I}_{L_i}
\end{aligned} \tag{42}$$

Note that $\mathbf{H}_{ji}\mathbf{Q}_j\mathbf{H}_{ji}^H \succeq \mathbf{0}$ and also that $(1-\rho_{ji}^2)\sigma_{ji}^2\,\mathrm{tr}\{\mathbf{Q}'_j\}\mathbf{I}_{L_j} \succeq \mathbf{0}, \forall i, j$. Consequently, using (41) we have that:

$$\left[\mathbf{\Gamma}_i + \sum_{j\neq i}\mathbf{H}_{ji}\mathbf{Q}_j\mathbf{H}_{ji}^H + \sum_{j=1}^{N}(1-\rho_{ji}^2)\sigma_{ji}^2\,\mathrm{tr}\{\mathbf{Q}'_j\}\mathbf{I}_{L_j}\right]^{-1} \succeq \left[\hat{\mathbf{\Gamma}}_i + \zeta_i\mathbf{I}_{L_i} + \sum_{j\neq i}\mathbf{H}_{ji}\mathbf{Q}_j\mathbf{H}_{ji}^H + \sum_{j=1}^{N}(1-\rho_{ji}^2)\sigma_{ji}^2\,\mathrm{tr}\{\mathbf{Q}'_j\}\mathbf{I}_{L_j}\right]^{-1} \tag{43}$$

The stated result follows from (43). ■

Corollary 1. The robust design problem in (36) can be solved using the proposed algorithm (see Table II) after replacing $\mathbf{\Gamma}_i$ with $\mathbf{\Gamma}'_i + \sum_{j=1}^{N}(1-\rho_{ji}^2)\sigma_{ji}^2\,\mathrm{tr}\{\mathbf{Q}'_j\}\mathbf{I}_{L_j}$, and after modifying $\mathbf{F}_i$, $\mathbf{G}_{ji}$ and C_i accordingly. ■

V. Numerical Results

In this section, we present several numerical examples to illustrate the performance of the proposed method. In all cases, unless otherwise stated, we assume that $N = 3$, $M_i = L_i \triangleq M = 4$, and SNR$\triangleq \frac{L_i p_i}{\mathrm{tr}\{\mathbf{\Gamma}_i\}}$=15dB, $\forall i = 1, 2, \cdots, N$. The receiver noise vectors are assumed to be white with unit variances, i.e., $\mathbf{\Gamma}_i = \mathbf{I}_{L_i}$, and the elements of channel matrices are i.i.d. CSCG random variables with zero mean and unit variances.

To investigate the convergence behaviour of the proposed algorithm, in Fig. 2 we plot the minimum rate achieved at various iterations for different values of the SNR. It can be observed that the minimum rate, i.e. the value of the objective function, increases at each iteration in agreement with the results in Section III. As expected, the higher the SNR, the larger the minimum rate.

To show the importance of considering the max-min fairness criterion for the precoder design, in Fig. 3 we compare the performance of the proposed method with that of the method for sum-rate maximization presented in [15], for $N = 10$.

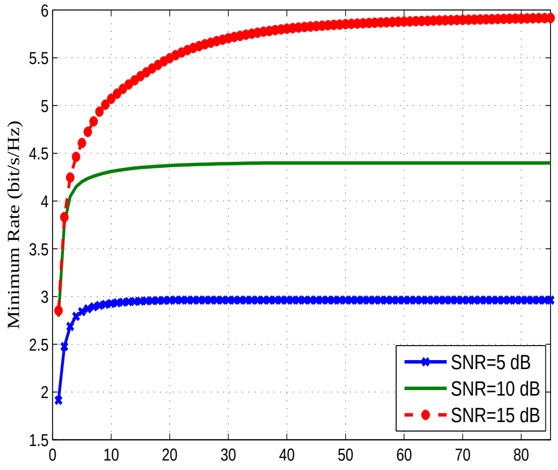


Fig. 2: The min rate versus number of iterations for the proposed method.

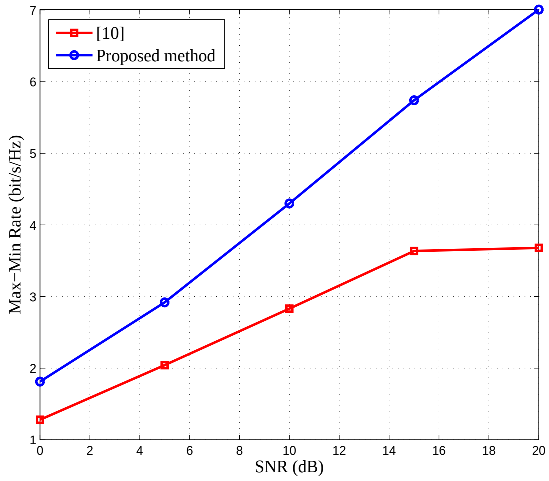


Fig. 4: The max-min rate achieved by the proposed algorithm and the method in [10], versus SNR.

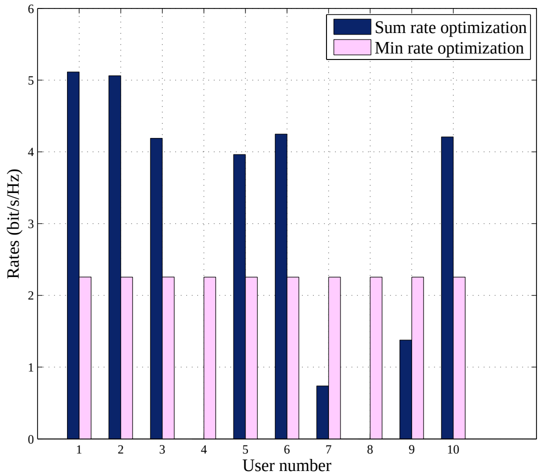


Fig. 3: Comparison between min-rate maximization and sum-rate maximization for $N = 10$. The rate of users 4 and 8 are too small to be visible.

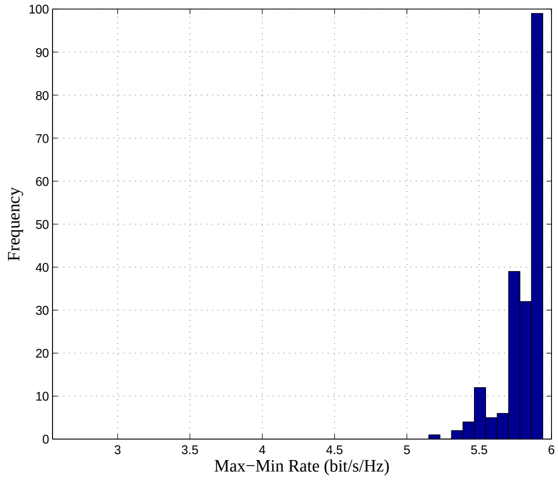
As expected, when the sum-rate is maximized, there exist users with very low rates (e.g, users 4 and 8) and users with high rates (e.g. users 1 and 2). Thus, the major benefit of our proposed method in comparison with sum rate maximization methods like that in [15] is that by using our method the rate will be distributed more fairly among the users.

Next, we compare the proposed method with the method in [10] that has been suggested for min rate optimization. Fig. 4 shows the max-min rates, averaged over 30 random channel realizations, for the proposed method and the one in [10]. In this example, we set $d_i = 2, \forall i$ for both methods (see Remark 3). We observe that the rate obtained by the proposed algorithm is significantly higher than the one obtained by the method in [10], which shows that the method introduced in this paper can provide higher quality solutions to the design problem, (12), than its competitor. We also see that the rate achieved by the proposed algorithm improves as SNR increases, while for the method of [10] the rate exhibits a saturation beyond an SNR of about 15dB. This behaviour is also seen in the simulation results reported in [10].

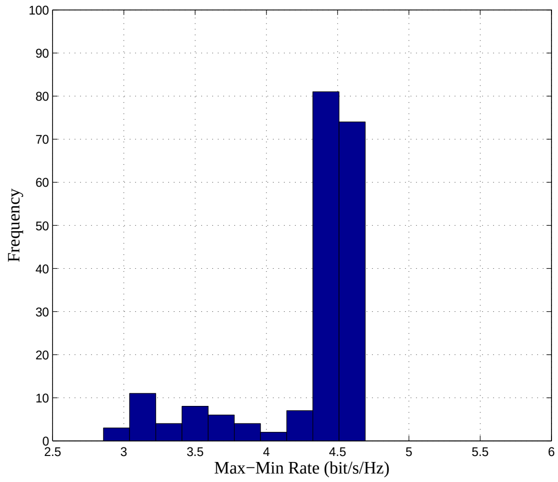
As stated earlier, the considered optimization problem is NP-hard and, as a result, any solution depends on the employed initial point. To investigate the dependency of the proposed method on the employed initial points, in Fig. 5.a we plot the histogram of the max-min rates corresponding to 200 randomly chosen initial points. The histogram for the algorithm in [10] is also depicted in Fig. 5.b. The rates achieved by the proposed method are in the interval $[5.11 - 5.98]$ with a variance of about 0.02, while those achieved by the method in [10] are in the interval $[2.77 - 4.78]$ with a variance of about 0.2. Consequently, in this example, the proposed method achieves higher rates and its performance depends on the initial points only mildly.

Unlike the method in [10] that directly designs the precoder matrices $\{\mathbf{V}_i\}_{i=1}^N$ (given $\{d_i\}_{i=1}^N$), the proposed algorithm designs the precoder covariance matrices $\{\mathbf{Q}_i\}_{i=1}^N$ (the precoder matrices $\{\mathbf{V}_i\}_{i=1}^N$ can be obtained as a by-product of the proposed method, see Remark 2). Therefore, by using the proposed method, the optimum precoder matrices $\{\mathbf{V}_i\}_{i=1}^N$ as well as the optimum number of their columns $\{d_i\}_{i=1}^N$ (i.e. the lengths of symbol streams) will be determined. To show the importance of this design aspect, in Fig. 6 we plot the max-min rate achieved by designing $\{\mathbf{Q}_i\}_{i=1}^N$ and, respectively, by designing $\{\mathbf{V}_i\}_{i=1}^N$ for certain values of $\{d_i\}_{i=1}^N = d, \forall i$, versus the number of antennas. As expected, the rates achieved by designing $\{\mathbf{Q}_i\}_{i=1}^N$ are higher than (or equal to) those obtained by designing $\{\mathbf{V}_i\}_{i=1}^N$ with fixed $\{d_i\}_{i=1}^N$. This can be explained by the fact that the optimal values for $\{d_i\}_{i=1}^N$ are also determined in the design of $\{\mathbf{Q}_i\}_{i=1}^N$ (see Remark 2).

Next we study the effect of channel estimation errors and noise covariance uncertainty on the performance of the method proposed in Section IV. To this end, we set $\rho_{ji} = \rho$ as well



(a)



(b)

Fig. 5: Histogram of max-min rates achieved using (a) the proposed method and (b) the algorithm in [10] for 200 randomly chosen initial points.

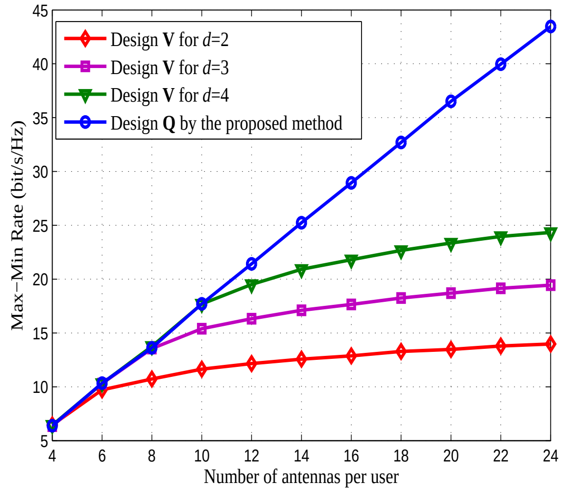


Fig. 6: The max-min rate achieved by the proposed algorithm for designing $\{\mathbf{Q}_i\}_{i=1}^N$ and, respectively, designing $\{\mathbf{V}_i\}_{i=1}^N$ for different values of d, versus the number of antennas per user (M).

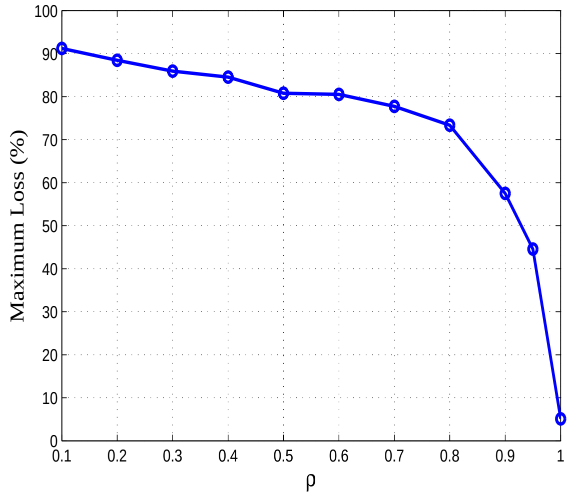


Fig. 7: The Loss $\mathcal{L}$ (in percentage) versus the CSI error parameter ρ ($\zeta = 0.25$).

as $\zeta_i = \zeta,\ \forall i, j = 1, \cdots, N$ and define the loss parameter

$$\mathcal{L}(\rho, \zeta) = 1 - \frac{R_{nr}(\rho, \zeta)}{R_r(\rho, \zeta)} \tag{44}$$

where $R_{nr}(\rho, \zeta)$ and $R_r(\rho, \zeta)$ denote the max-min rates achieved by the non-robust and the robust methods, respectively, for uncertainty parameters (ρ, ζ). Note that the loss parameter $\mathcal{L}(\rho, \zeta)$ quantifies the performance degradation caused by employing the non-robust method instead of the robust one. Note also that $\mathcal{L}$ depends on the realizations of the channel matrices as well as noise covariances. In Fig. 7, we plot the maximum value of $\mathcal{L}(\rho, \zeta)$ versus ρ for 100 realizations of channel matrices. In this example, we set $\mathbf{\Gamma}_i = \hat{\mathbf{\Gamma}}_i + \zeta \mathbf{I}, \forall i$ with $\zeta = 0.25$. It can be seen that even for large values of ρ (i.e., relatively low channel estimation errors), employing the robust method provides a significantly larger max-min rates. As expected, the loss decreases as the estimation quality improves, i.e., as ρ increases. Note that the loss is non-zero even for the case of $\rho = 1$ in which CSI is perfect. This is due to the uncertainty in the noise covariances. Finally, note that in this example we have numerically observed that the performance loss is more sensitive to CSI uncertainty than the noise covariance uncertainty.

VI. Conclusion

In this paper, we considered a MIMO interference channel network with conventional LMMSE decoder matrices at the receivers for which we designed the transmit covariance matrices under the max-min fairness criterion. The problem is non-convex and NP-hard in the number of users. We proposed an efficient algorithm based on the MM optimization technique that computes a locally optimum solution to this design problem. We showed that the proposed algorithm is convergent. We also considered uncertainties in the noise covariances and the CSI, and extended our algorithm to design precoder covariance matrices in these cases. Numerical results were included to illustrate the effectiveness of the proposed method in various scenarios.

Appendix A
Proof of (6) and (7)

We begin by proving the expression of the LMMSE in (6). Assuming $\mathbb{E}\{\mathbf{y}_i\} = \mathbf{0}$ and $\mathbb{E}\{\mathbf{s}_i\} = \mathbf{0}$, the LMMSE estimator of $\mathbf{s}_i$ for given $\mathbf{y}_i$ has the following expression [22]:

$$\hat{\mathbf{s}}_i = \underbrace{\mathbf{C}_{\mathbf{s}_i\mathbf{y}_i}\mathbf{C}_{\mathbf{y}_i}^{-1}}_{\triangleq \mathbf{W}_i}\mathbf{y}_i \tag{45}$$

where $\mathbf{C}_{\mathbf{s}_i\mathbf{y}_i}$ is the cross-covariance matrix between $\mathbf{s}_i$ and $\mathbf{y}_i$ and $\mathbf{C}_{\mathbf{y}_i}$ is the auto-covariance matrix of $\mathbf{y}_i$. Using (2) and noting that $\mathbb{E}\{\mathbf{s}_i\mathbf{s}_i^H\} = \mathbf{I}_{d_i}$ and $\mathbb{E}\{\mathbf{s}_i\mathbf{s}_j^H\} = \mathbf{0}, i \neq j$, we have:

$$\begin{aligned} \mathbf{C}_{\mathbf{s}_i\mathbf{y}_i} = \mathbb{E}\{\mathbf{s}_i\mathbf{y}_i^H\} &= \mathbf{V}_i^H\mathbf{H}_{ii}^H \\ \mathbf{C}_{\mathbf{y}_i} = \mathbb{E}\{\mathbf{y}_i\mathbf{y}_i^H\} &= \sum_{j=1}^{N}\mathbf{H}_{ji}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ji}^H + \mathbf{\Gamma}_i \end{aligned} \tag{46}$$

The expression of the LMMSE in (6) is obtained by substituting (46) in (45).

Next, we show that substituting (6) in (4) yields the expression for rate R_i in (7). To this end, we rewrite (6) by using the matrix inversion identity $(\mathbf{A}+\mathbf{BCD})^{-1}\mathbf{BC} = \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B})^{-1}$ as follows:

$$\mathbf{W}_i^{\text{LMMSE}} = \underbrace{\left(\mathbf{I}_{d_i} + \mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}\mathbf{H}_{ii}\mathbf{V}_i\right)^{-1}}_{\triangleq \mathbf{C}_e}\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1} \tag{47}$$

Let $\mathbf{\Omega}_i \triangleq \mathbf{W}_i\mathbf{H}_{ii}\mathbf{V}_i\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{W}_i^H\left(\mathbf{W}_i\mathbf{C}_{\bar{i}}\mathbf{W}_i^H\right)^{-1}$, then

$$\begin{aligned} \mathbf{\Omega}_i =& \mathbf{C}_e\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}\mathbf{H}_{ii}\mathbf{V}_i\mathbf{C}_e^{-1}\times \\ & \mathbf{C}_e\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}\mathbf{H}_{ii}\mathbf{V}_i\mathbf{C}_e\left(\mathbf{C}_e\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}\mathbf{H}_{ii}\mathbf{V}_i\mathbf{C}_e\right)^{-1} \\ =& \mathbf{C}_e\mathbf{V}_i^H\mathbf{H}_{ii}^H\mathbf{C}_{\bar{i}}^{-1}\mathbf{H}_{ii}\mathbf{V}_i\mathbf{C}_e^{-1} \end{aligned} \tag{48}$$

Finally, it is readily verified that by substituting (48) in (4) and using Sylvester determinant property, (7) is obtained.

Appendix B
Proof of Proposition 1

First, note that $\mathbf{Q}_i$ can be decomposed as $\mathbf{Q}_i = \widetilde{\mathbf{V}}_i\widetilde{\mathbf{V}}_i^H$. Let $\mathbf{B}_{i,11}$ denote the left upper block of $\mathbf{B}_i^{-1}$. By using the blockwise matrix inversion lemma (see, e.g., [23]), we have that:

$$\mathbf{B}_{i,11}=\left(\mathbf{I}_{M_i} - \widetilde{\mathbf{V}}_i^H\mathbf{H}_{ii}^H\left[\mathbf{\Gamma}_i + \sum_{j=1}^{N}\mathbf{H}_{ji}\widetilde{\mathbf{V}}_j\widetilde{\mathbf{V}}_j^H\mathbf{H}_{ji}^H\right]^{-1}\mathbf{H}_{ii}\widetilde{\mathbf{V}}_i\right)^{-1} \tag{49}$$

Then, by using Woodbury matrix identity, (49) can be rewritten as:

$$\mathbf{B}_{i,11}=\mathbf{I}_{M_i}+\widetilde{\mathbf{V}}_i^H\mathbf{H}_{ii}^H\left[\mathbf{\Gamma}_i + \sum_{j\neq i}\mathbf{H}_{ji}\widetilde{\mathbf{V}}_j\widetilde{\mathbf{V}}_j^H\mathbf{H}_{ji}^H\right]^{-1}\mathbf{H}_{ii}\widetilde{\mathbf{V}}_i \tag{50}$$

Finally, substituting (50) in (13) and using Sylvester determinant property, (11) is obtained.

Appendix C
Proof that $\mathbf{B}_i \succ \mathbf{0}$

The matrix $\mathbf{B}_i$ is defined in (15). First it is obvious that $\mathbf{I}_{M_i} \succ \mathbf{0}$. Thus, it suffices to prove that the Schur complement of $\mathbf{I}_{M_i}$ in $\mathbf{B}_i$ is positive definite, i.e. [21] [23]:

$$\begin{aligned} \mathbf{S}_i \triangleq& \mathbf{\Gamma}_i + \sum_{j=1}^{N}\mathbf{H}_{ji}\mathbf{Q}_j\mathbf{H}_{ji}^H - \mathbf{H}_{ii}\mathbf{Q}_i\mathbf{H}_{ii}^H \\ =& \mathbf{\Gamma}_i + \sum_{\substack{j=1\\ j\neq i}}^{N}\mathbf{H}_{ji}\mathbf{Q}_j\mathbf{H}_{ji}^H \succ \mathbf{0} \end{aligned} \tag{51}$$

The matrices $\mathbf{H}_{ji}\mathbf{Q}_j\mathbf{H}_{ji}^H, \forall i,j$ are obviously positive semidefinite. Therefore, $\mathbf{S}_i$ is positive definite because it is the sum of a positive definite matrix ($\mathbf{\Gamma}_i$) and a number of positive semidefinite matrices, and as a result $\mathbf{B}_i \succ 0$.

Appendix D
Proof of Eq. (17)

We begin the proof by presenting the following theorem from [24].

Theorem 2. Let $\mathbf{X} \in \mathbb{S}_N^+$ and define

$$\nabla_{\mathbf{X}}f(\mathbf{X}) = \begin{bmatrix} \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{11}} & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{12}} & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{13}} & \cdots & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{1N}} \\ \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{21}} & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{22}} & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{23}} & \cdots & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{2N}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{N1}} & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{N2}} & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{N3}} & \cdots & \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}_{NN}} \end{bmatrix}$$

for a differentiable function $f(\mathbf{X}) : \mathbb{S}_N^+ \to \mathbb{R}$. Then, the following inequality holds for any convex (differentiable) function $f(\mathbf{X})$:

$$f(\mathbf{Y}) \geq f(\mathbf{X}) + \mathrm{tr}\{(\nabla_{\mathbf{X}}f(\mathbf{X}))^H(\mathbf{Y}-\mathbf{X})\}, \quad \forall \mathbf{X},\mathbf{Y} \succeq \mathbf{0} \tag{52}$$

■

To make use of the above lemma, we note the following differentiation formulas (which hold for $\mathbf{X} \in \mathbb{S}_N^{++}$ and $\mathbf{A}$, $\mathbf{B}$ of proper dimensions):

$$\nabla_{\mathbf{X}}\left(\det(\mathbf{A}\mathbf{X}^{-1}\mathbf{B})\right) = -\det(\mathbf{A}\mathbf{X}^{-1}\mathbf{B})\mathbf{X}^{-1}\mathbf{A}^H(\mathbf{B}^H\mathbf{X}^{-1}\mathbf{A}^H)^{-1}\mathbf{B}^H \tag{53}$$

$$\nabla_{\mathbf{X}}(\log(f(\mathbf{X})) = \frac{\nabla_{\mathbf{X}}(f(\mathbf{X}))}{f(\mathbf{X})} \tag{54}$$

To compute $\nabla_{\mathbf{X}}$ for the function $\log\det(\mathbf{U}^H\mathbf{X}^{-1}\mathbf{U})$ in Lemma 1, define $\psi(\mathbf{X}) \triangleq \det(\mathbf{U}^H\mathbf{X}^{-1}\mathbf{U})$ and note that using

(53)-(54) we obtain:

$$\nabla_{\mathbf{X}}(\log(\psi(\mathbf{X}))) = \left(\frac{1}{\psi(\mathbf{X})}\right)\nabla_{\mathbf{X}}(\psi(\mathbf{X}))$$
$$= -(\det(\mathbf{U}^H\mathbf{X}^{-1}\mathbf{U}))^{-1}\nabla_{\mathbf{X}}(\det(\mathbf{U}^H\mathbf{X}^{-1}\mathbf{U}))$$
$$= -\mathbf{X}^{-1}\mathbf{U}(\mathbf{U}^H\mathbf{X}^{-1}\mathbf{U})^{-1}\mathbf{U}^H\mathbf{X}^{-1}. \quad (55)$$

The proof of (17) is concluded using the expression for $\nabla_{\mathbf{X}} \log\det(\mathbf{U}^H\mathbf{X}^{-1}\mathbf{U})$ above in the inequality (52).

REFERENCES

[1] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 1, pp. 57–62, 1992.

[2] ——, "Distributed cochannel interference control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 3, pp. 305–311, Aug 1992.

[3] H. Boche and M. Schubert, "Resource allocation in multiantenna systems-achieving max-min fairness by optimizing a sum of inverse SIR," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 1990–1997, June 2006.

[4] Y. F. Liu, Y. H. Dai, and Z. Q. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1142–1157, March 2011.

[5] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, Jan 2004.

[6] W. Yang and G. Xu, "Optimal downlink power assignment for smart antenna systems," in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, 1998.*, vol. 6, 1998, pp. 3337–3340 vol.6.

[7] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 161–176, Jan 2006.

[8] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Max-min fairness linear transceiver design for a multi-user MIMO interference channel," in *2011 IEEE International Conference on Communications (ICC)*. IEEE, 2011, pp. 1–5.

[9] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "Linear transceiver design for a MIMO interfering broadcast channel achieving max-min fairness," in *2011 Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2011, pp. 1309–1313.

[10] ——, "Linear transceiver design for a MIMO interfering broadcast channel achieving max–min fairness," *Signal Processing*, vol. 93, no. 12, pp. 3327–3340, 2013.

[11] P. Aquilina, A. C. Cirik, and T. Ratnarajah, "Weighted sum rate maximization in full-duplex multi-user multi-cell MIMO networks," *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1590–1608, 2017.

[12] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2205–2217, 2015.

[13] X. Wang, Y. Wang, and S. Ma, "Upper bound on uplink sum rate for multi-cell massive MU-MIMO systems with ZF receivers," *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 250–253, 2017.

[14] C. Zhang, Y. Huang, Y. Jing, S. Jin, and L. Yang, "Sum-rate analysis for massive MIMO downlink with joint statistical beamforming and user scheduling," *IEEE Transactions on Wireless Communications*, vol. 16, no. 4, pp. 2181–2194, 2017.

[15] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, "Linear transceiver design for interference alignment: Complexity and computation," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2896–2910, 2012.

[16] M. M. Naghsh, M. Soltanalian, P. Stoica, M. Masjedi, and B. Ottersten, "Efficient sum-rate maximization for medium-scale MIMO AF-relay networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6400–6411, 2016.

[17] Y.-F. Liu, M. Hong, and Y.-H. Dai, "Max-min fairness linear transceiver design problem for a multi-user SIMO interference channel is polynomial time solvable," *IEEE Signal Processing Letters*, vol. 20, no. 1, pp. 27–30, 2013.

[18] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. Slock, "On the MIMO interference channel," in *IEEE Information Theory and Applications Workshop (ITA), 2010*, 2010, pp. 1–9.

[19] P. Stoica and Y. Selen, "Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: a refresher," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 112–114, 2004.

[20] M. M. Naghsh, M. Modarres-Hashemi, M. Alaee, and E. Alian, "An information theoretic approach to robust constrained code design for MIMO radars," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3647–3661, 2017.

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

[22] S. M. Kay, *Fundamentals of statistical signal processing*. Prentice Hall, 1993.

[23] P. Stoica, R. L. Moses *et al.*, *Spectral analysis of signals*. Prentice Hall, 2005, vol. 452.

[24] C. Zalinescu, *Convex Analysis in General Vector Spaces*. World Scientific, 2002.