Sum-Rate Optimization for the MIMO IC under Imperfect CSI: a Deterministic Equivalent Approach

Mohsen Rezaee[†], Romain Couillet[‡], Maxime Guillaud[†], Gerald Matz[†]
[†]Institute of Telecommunications, Vienna University of Technology, Vienna, Austria
e-mail: {guillaud,mohsen.rezaee,gerald.matz}@tuwien.ac.at
[‡]Supélec, France – e-mail: romain.couillet@supelec.fr

***Abstract*—A new method is proposed to determine precoding matrices that achieve local maxima of the expected sum rate in a multiple input multiple output interference channel (MIMO IC), in the realistic scenario where only partial channel state information (CSI) is available at the transmitters. Relying on a random matrix analysis of the capacity of large dimensional Ricean channels introduced in [1], the expected sum rate of the K-user MIMO IC is approximated by a deterministic equivalent to which an iterative gradient scheme is applied to find local maxima of the approximated sum rate.**

I. Introduction

Multiple input multiple output (MIMO) systems are recognized as an effective means for improving the performance of a wireless communication link as they offer substantial capacity improvements over single-antenna systems without requiring additional power or bandwidth.

In order to optimally exploit the multiplexing gain offered by MIMO systems, spatial precoding is performed by the transmitters, which leads to significant gains particularly in interference channels. When interference is the dominant impairment, the precoding method known as interference alignment has proved capable of achieving the optimal multiplexing gain of interference channels [2]. Many works have then proceeded to analyze the sum rate performance of interference alignment. In particular, in [3] the authors considered sum rate maximization under a total power constraint. In [4], the authors analyzed a family of sum rate maximizing orthonormal precoders (or beamformers) that maintain a low interference level. In [5], a so-called alternating method is proposed to find beamformers under per-transmitter power constraints.

However, these methods design the precoders based on the assumption that perfect CSI is available at the transmitters. In a more realistic scenario where CSI is only partially known, for instance when the channel state available at the transmitters is based on erroneous estimates or on outdated CSI fed back by the receivers, the performance of such methods degrades substantially. While the effect of imperfect CSI has been widely investigated in the single user case, there exist only a few results for the MIMO IC. This case is considered for instance in [3] and [6] under interference alignment constraints.

On the other hand, when only knowledge of the statistics of the channel is available at the transmitters, approximating the expected sum rate will be of particular interest. In single user transmissions, this problem is explored for separately-correlated Rayleigh and uncorrelated Rician MISO channels in [7] and extended to MIMO in [8] and [9]. In [1], a method is proposed to approximate the ergodic mutual information (EMI) for the general correlated Rician fading channel; this method is accurate when the number of antennas is large.

In this paper, the sum rate maximizing precoder design for the MIMO IC is investigated for the case where only a noisy version of the channel is available at the transmitters. We propose to approximate the expected sum rate of the K-user IC by a so-called *deterministic equivalent*, using advanced random matrix tools presented in [1], and we numerically optimize the transmit precoders to determine local maxima of the deterministic equivalent.

The remainder of the paper is organized as follows. In Section II, the system model is described. The deterministic equivalent method is discussed in Section III. The proposed method for approximating the expected sum rate of the K-user IC is presented in Section IV. Simulation results are presented in Section V and conclusions are drawn in Section VI.

Notation: Nonbold letters represent scalar quantities, boldface lowercase and uppercase letters indicate vectors and matrices, respectively. $\mathbf{I}_N$ is the $N \times N$ identity matrix. The trace, conjugate, transpose, Hermitian transpose of a matrix or vector are denoted by $\mathrm{tr}(\cdot), (\cdot)^*, (\cdot)^{\mathrm{T}}, (\cdot)^{\mathrm{H}}$ respectively. The expectation operator is represented by $\mathrm{E}(\cdot)$ and $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semi-definite. Determinant and spectral norm of a matrix are denoted by $|\cdot|$ and $||\cdot||$ respectively and $\mathrm{Bdiag}(\cdot)$ represents a block diagonal matrix with the argument blocks on its diagonal. The complex Gaussian distribution with zero mean and unit variance is denoted by $\mathcal{CN}(0,1)$.

II. System Model

An interference channel is considered in which K transmitters communicate with their respective receivers in a shared medium. Transmitter j and receiver i are equipped with N_j and M_i antennas, respectively. Data symbols are spatially precoded at the transmitters. The number of data streams sent by transmitter i equals d^i. The vector at receiver i reads

$$\mathbf{y}_i = \mathbf{H}_{ii}\sqrt{\lambda_i}\mathbf{V}_i\mathbf{x}_i + \sum_{\substack{1\leq j\leq K\\ j\neq i}} \mathbf{H}_{ij}\sqrt{\lambda_j}\mathbf{V}_j\mathbf{x}_j + \mathbf{n}_i \qquad (1)$$

in which $\mathbf{H}_{ij} \in \mathbb{C}^{M_i\times N_j}$ is the channel matrix between transmitter j and receiver i, $\mathbf{V}_j \in \mathbb{C}^{N_j\times d^j}$ and $\mathbf{x}_j \in \mathbb{C}^{d^j}$ are the precoding matrix and the data vector of transmitter j, respectively. Furthermore, $\mathbf{n}_i \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{M_i})$ is the additive noise at receiver i. Assuming $\mathrm{E}\left[\mathbf{x}_i \mathbf{x}_i^{\mathrm{H}}\right] = \mathbf{I}_{d^i}, \ \ i = 1, \ldots, K$, the covariance matrix of the signal transmitted by user j is given as $\mathbf{Q}_j = \lambda_j \mathbf{V}_j \mathbf{V}_j^{\mathrm{H}}$ in which $\lambda_j = \frac{P_j}{\mathrm{tr}\left(\mathbf{V}_j \mathbf{V}_j^{\mathrm{H}}\right)}$. The transmit power for user j is $\mathrm{tr}\left(\mathbf{Q}_j\right) = P_j$. The transmit signal-to-noise ratio (SNR) for user j is defined as $\frac{\mathrm{tr}(\mathbf{Q}_j)}{\sigma^2} = \frac{P_j}{\sigma^2}$.

In order to model imperfect CSI at the transmitter, the channel matrix $\mathbf{H}_{ij}$ is assumed estimated as $\bar{\mathbf{H}}_{ij}$ by transmitter j, with

$$\mathbf{H}_{ij} = \bar{\mathbf{H}}_{ij} + \mathbf{E}_{ij},$$

in which $\mathbf{E}_{ij}$ is the estimation error, whose entries are modeled as independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance σ_{ij}^2. Both $\bar{\mathbf{H}}_{ij}$ and σ_{ij}^2 are known to all transmitters. We define the feedback quality for the link between transmitter j and receiver i as

$$\eta_{ij} = \frac{\mathrm{tr}(\bar{\mathbf{H}}_{ij}\bar{\mathbf{H}}_{ij}^{\mathrm{H}})}{\mathrm{tr}(\mathrm{E}\left[\mathbf{E}_{ij}\mathbf{E}_{ij}^{\mathrm{H}}\right])}. \tag{2}$$

The objective function that we wish to study and optimize w.r.t. $\mathbf{Q}_1, \ldots, \mathbf{Q}_K$ is the achievable expected sum rate of the MIMO IC under the assumption that the input signals are Gaussian. Considering the received signal in (1), this reads

$$\begin{aligned} R_{\mathrm{sum}} \;=\; & \mathrm{E}\left[\sum_{i=1}^{K} \log\left|\mathbf{I}_{M_i} + \frac{1}{\sigma^2}\sum_{j=1}^{K}\mathbf{H}_{ij}\mathbf{Q}_j\mathbf{H}_{ij}^{\mathrm{H}}\right|\right. \\ & \left. - \sum_{i=1}^{K}\log\left|\mathbf{I}_{M_i} + \frac{1}{\sigma^2}\sum_{j\neq i}\mathbf{H}_{ij}\mathbf{Q}_j\mathbf{H}_{ij}^{\mathrm{H}}\right|\right] \end{aligned} \tag{3}$$

where the expectation is w.r.t. $\mathbf{E}_{ij}$, $1 \leq i,j \leq K$. $\mathbf{E}_{ij} = \sigma_{ij}\mathbf{W}_{ij}$ and $\mathbf{W}_{ij}$ with i.i.d. $\mathcal{CN}(0,1)$ entries.

In order to determine the sum rate-maximizing precoders, we will use a gradient ascent approach. To simplify the computationally demanding evaluation of the expectations in the expression of R_{sum}, we shall rely on an approximation $\bar{R}_{\mathrm{sum}}$ of R_{sum} provided by tools from random matrix theory, which is asymptotically accurate as the system dimensions N_i, M_j grow large. This approximation, known as *deterministic equivalent*, is introduced in the next section.

III. Deterministic Equivalent

We start with the following Lemma.

Lemma 1: Let $\bar{\mathbf{H}} \in \mathbb{C}^{M\times N}$ and positive semi-definite matrices $\mathbf{Q}, \tilde{\mathbf{C}} \in \mathbb{C}^{N\times N}$ be deterministic matrices and $\sigma^2 > 0$. Define $\mathbf{G}(\delta) = \left(\mathbf{I}_N + \delta\mathbf{Q}^{\frac{1}{2}}\tilde{\mathbf{C}}\mathbf{Q}^{\frac{1}{2}}\right)^{-1}$ and

$$g(\mathbf{Q},\delta,\tilde{\delta}) = \frac{1}{N}\mathrm{tr}\left(\sigma^2(1+\tilde{\delta})\mathbf{I}_M + \bar{\mathbf{H}}\mathbf{Q}^{\frac{1}{2}}\mathbf{G}(\delta)\mathbf{Q}^{\frac{1}{2}}\bar{\mathbf{H}}^{\mathrm{H}}\right)^{-1},$$

$$\tilde{g}(\mathbf{Q},\delta,\tilde{\delta}) = \frac{1}{N}\mathrm{tr}\left(\mathbf{Q}^{\frac{1}{2}}\tilde{\mathbf{C}}\mathbf{Q}^{\frac{1}{2}}\left(\sigma^2(\mathbf{G}(\delta))^{-1} + \frac{\mathbf{Q}^{\frac{1}{2}}\bar{\mathbf{H}}^{\mathrm{H}}\bar{\mathbf{H}}\mathbf{Q}^{\frac{1}{2}}}{1+\tilde{\delta}}\right)^{-1}\right),$$

then the system of equations

$$\left\{\begin{array}{rcl} \delta &=& g(\mathbf{Q},\delta,\tilde{\delta}), \\ \tilde{\delta} &=& \tilde{g}(\mathbf{Q},\delta,\tilde{\delta}), \end{array}\right. \tag{4}$$

has a unique solution $(\delta^\star, \tilde{\delta}^\star) \in (0,\infty)^2$. Moreover, $\delta^\star$ and $\tilde{\delta}^\star$ can be numerically computed via Algorithm 1.

Proof: A sketch of the proof is given in Appendix A. ■

Algorithm 1 Fixed point method

Initialization: $m = 0$ and $\delta^{(0)}, \tilde{\delta}^{(0)} > 0$

Repeat

- $\delta^{(m+1)} = g(\delta^{(m)}, \tilde{\delta}^{(m)})$
- $\tilde{\delta}^{(m+1)} = \tilde{g}(\delta^{(m)}, \tilde{\delta}^{(m)})$
- $m \leftarrow m+1$

until convergence.

Output: $(\delta^\star, \tilde{\delta}^\star) = (\delta^{(m)}, \tilde{\delta}^{(m)})$.

The deterministic equivalent of interest in this article is provided as follows.

Theorem 1: Let $\mathbf{H} = \bar{\mathbf{H}} + \mathbf{E}$, in which $\bar{\mathbf{H}} \in \mathbb{C}^{M\times N}$ is deterministic with bounded spectral norm, and $\mathbf{E} = \frac{1}{\sqrt{N}}\mathbf{W}\tilde{\mathbf{C}}^{\frac{1}{2}}$ with $\mathbf{W} \in \mathbb{C}^{M\times N}$ having i.i.d. elements from $\mathcal{CN}(0,1)$ and $\tilde{\mathbf{C}}$ diagonal nonnegative with bounded spectral norm. Let also $\mathbf{Q} \in \mathbb{C}^{N\times N}$ be deterministic Hermitian nonnegative with bounded spectral norm. Then, as $M, N \to \infty$ with $M/N \to c > 0$,

$$R - \bar{R} \to 0$$

where $R = \mathrm{E}\left[\log\left|\mathbf{I}_M + \frac{1}{\sigma^2}\mathbf{H}\mathbf{Q}\mathbf{H}^{\mathrm{H}}\right|\right]$,

$$\begin{aligned} \bar{R} \;=\; & \log\left|(1+\tilde{\delta})\mathbf{I}_M + \frac{1}{\sigma^2}\bar{\mathbf{H}}\mathbf{Q}^{\frac{1}{2}}\mathbf{G}(\delta)\mathbf{Q}^{\frac{1}{2}}\bar{\mathbf{H}}^{\mathrm{H}}\right| \\ & - \log\left|\mathbf{G}(\delta)\right| - N\sigma^2\delta\tilde{\delta}, \end{aligned}$$

and δ and $\tilde{\delta}$ are the unique positive solution of (4).

Proof: The proof follows directly from [1, Theorem 2], where only functional uniqueness of δ, $\tilde{\delta}$ (seen as functions of σ^2) was obtained. Lemma 1 completes [1] by adding point-wise uniqueness of δ, $\tilde{\delta}$ for each $\sigma^2 > 0$. ■

$\bar{R}$ is called the deterministic equivalent because it does not involve an expectation. A nice property of $\bar{R}$ is that its partial derivative with respect to δ and $\tilde{\delta}$ vanishes at $(\delta^\star, \tilde{\delta}^\star)$,

$$\left.\frac{\partial\bar{R}}{\partial\delta}\right|_{(\delta^\star,\tilde{\delta}^\star)} = \left.\frac{\partial\bar{R}}{\partial\tilde{\delta}}\right|_{(\delta^\star,\tilde{\delta}^\star)} = 0. \tag{5}$$

In the next section, we introduce our novel gradient ascent technique based on Theorem 1.

IV. Proposed Sum-Rate Optimization Method

Consider again the system model (1). Defining $N = \sum_{j=1}^{K} N_j$, (3) can be rewritten as

$$\begin{aligned} R_{\mathrm{sum}} \;=\; & \mathrm{E}\left[\sum_{i=1}^{K}\log\left|\mathbf{I}_{M_i} + \frac{1}{\bar{\sigma}^2}\mathbf{H}_i\mathbf{Q}\mathbf{H}_i^{\mathrm{H}}\right|\right. \\ & \left. - \sum_{i=1}^{K}\log\left|\mathbf{I}_{M_i} + \frac{1}{\bar{\sigma}^2}\mathbf{H}_i\mathbf{Q}_{-i}\mathbf{H}_i^{\mathrm{H}}\right|\right]. \end{aligned} \tag{6}$$

Here, $\bar{\sigma}^2 = \frac{\sigma^2}{N}$, the equivalent channels of size $M_i \times N$ are defined as $\mathbf{H}_i = \frac{1}{\sqrt{N}}\left[\mathbf{H}_{i1}, \ldots, \mathbf{H}_{iK}\right]$ for $i = 1, \ldots, K$, $\mathbf{Q} =$

$\mathrm{Bdiag}\left(\mathbf{Q}_1,...,\mathbf{Q}_K\right)$ and $\mathbf{Q}_{-i}$ is equal to $\mathbf{Q}$ except that the ith block is replaced by the $N_i \times N_i$ zero matrix. Therefore

$$\mathbf{H}_i = \bar{\mathbf{H}}_i + \mathbf{E}_i, \tag{7}$$

with $\bar{\mathbf{H}}_i = \frac{1}{\sqrt{N}}\left[\bar{\mathbf{H}}_{i1},...,\bar{\mathbf{H}}_{iK}\right]$, $\mathbf{E}_i = \frac{1}{\sqrt{N}}\mathbf{W}_i\tilde{\mathbf{C}}_i^{\frac{1}{2}}$, $\mathbf{W}_i = [\mathbf{W}_{i1},...,\mathbf{W}_{iK}]$ and $\tilde{\mathbf{C}}_i = \mathrm{Bdiag}\left(\sigma_{i1}^2\mathbf{I}_{N_1},...,\sigma_{iK}^2\mathbf{I}_{N_K}\right)$. With these definitions, $\mathbf{W}_i$ has $\mathcal{CN}(0,1)$ elements and $\tilde{\mathbf{C}}_i$ is diagonal nonnegative, like $\mathbf{W}$ and $\tilde{\mathbf{C}}$ in Theorem 1.

Defining R_i^+ and R_i^- as

$$R_i^+ = \mathrm{E}\left[\log\left|\mathbf{I}_{M_i} + \frac{1}{\bar{\sigma}^2}\mathbf{H}_i\mathbf{Q}\mathbf{H}_i^{\mathrm{H}}\right|\right],$$
$$R_i^- = \mathrm{E}\left[\log\left|\mathbf{I}_{M_i} + \frac{1}{\bar{\sigma}^2}\mathbf{H}_i\mathbf{Q}_{-i}\mathbf{H}_i^{\mathrm{H}}\right|\right],$$

where the expectation is over $\mathbf{E}_i$, the expected sum rate is

$$R_{\mathrm{sum}} = \sum_{i=1}^{K}\left(R_i^+ - R_i^-\right).$$

Using Theorem 1, we approximate R_i^+ and R_i^- by deterministic equivalents. In particular, we have

$$R_i^+ - \bar{R}_i^+ \to 0$$

as $N_j, M_j \to \infty$ with $N_j/M_j \to c_j > 0$ for all j, where

$$\begin{aligned}\bar{R}_i^+ &= f_i(\mathbf{Q},\delta_i^+,\tilde{\delta}_i^+)\\ &= \log\left|(1+\tilde{\delta}_i^+)\mathbf{I}_{M_i} + \frac{1}{\bar{\sigma}^2}\bar{\mathbf{H}}_i\mathbf{Q}^{\frac{1}{2}}\mathbf{G}_i(\delta_i^+)\mathbf{Q}^{\frac{1}{2}}\bar{\mathbf{H}}_i^{\mathrm{H}}\right|\\ &\quad - \log\left|\mathbf{G}_i(\delta_i^+)\right| - N\bar{\sigma}^2\delta_i^+\tilde{\delta}_i^+\end{aligned} \tag{8}$$

in which $\mathbf{G}_i(x) = (\mathbf{I}_N + x\mathbf{Q}^{\frac{1}{2}}\tilde{\mathbf{C}}_i\mathbf{Q}^{\frac{1}{2}})^{-1}$, and δ_i^+ and $\tilde{\delta}_i^+$ are the unique positive solutions to

$$\begin{aligned}\delta_i^+ &= g_i(\mathbf{Q},\delta_i^+,\tilde{\delta}_i^+)\\ &= \frac{1}{N}\mathrm{tr}\left(\bar{\sigma}^2(1+\tilde{\delta}_i^+)\mathbf{I}_{M_i} + \bar{\mathbf{H}}_i\mathbf{Q}^{\frac{1}{2}}\mathbf{G}_i(\delta_i^+)\mathbf{Q}^{\frac{1}{2}}\bar{\mathbf{H}}_i^{\mathrm{H}}\right)^{-1},\end{aligned} \tag{9}$$

$$\begin{aligned}\tilde{\delta}_i^+ &= \tilde{g}_i(\mathbf{Q},\delta_i^+,\tilde{\delta}_i^+)\\ &= \frac{1}{N}\mathrm{tr}\left(\mathbf{Q}^{\frac{1}{2}}\tilde{\mathbf{C}}_i\mathbf{Q}^{\frac{1}{2}}\left(\bar{\sigma}^2(\mathbf{G}_i(\delta_i^+))^{-1} + \frac{\mathbf{Q}^{\frac{1}{2}}\bar{\mathbf{H}}_i^{\mathrm{H}}\bar{\mathbf{H}}_i\mathbf{Q}^{\frac{1}{2}}}{1+\tilde{\delta}_i^+}\right)^{-1}\right).\end{aligned} \tag{10}$$

Similarly we define $\bar{R}_i^- = f_i(\mathbf{Q}_{-i},\delta_i^-,\tilde{\delta}_i^-)$ and $(\delta_i^-,\tilde{\delta}_i^-)$ the unique nonnegative solution to

$$\begin{cases}\delta_i^- = g_i(\mathbf{Q}_{-i},\delta_i^-,\tilde{\delta}_i^-),\\ \tilde{\delta}_i^- = \tilde{g}_i(\mathbf{Q}_{-i},\delta_i^-,\tilde{\delta}_i^-).\end{cases} \tag{11}$$

Defining $\bar{R}_{\mathrm{sum}} = \sum_{i=1}^{K}\left(\bar{R}_i^+ - \bar{R}_i^-\right)$, with K finite, we have from Theorem 1 that, for all $\mathbf{Q}_i$ with bounded spectral norm,

$$\sup_{\substack{\mathbf{Q},\\ \|\mathbf{Q}\|\ \mathrm{bounded}}} R_{\mathrm{sum}}(\mathbf{Q}) - \sup_{\substack{\mathbf{Q},\\ \|\mathbf{Q}\|\ \mathrm{bounded}}} \bar{R}_{\mathrm{sum}}(\mathbf{Q}) \to 0, \tag{12}$$

as N_i, M_i grow large. Thus, optimizing R_{sum} over $\mathbf{Q}$, for any family of bounded precoders, is equivalent, in the asymptotic regime, to optimizing $\bar{R}_{\mathrm{sum}}$ over $\mathbf{Q}$. $\bar{R}_{\mathrm{sum}}$ is deterministic and its evaluation does not require heavy Monte Carlo simulations. Hence we propose to use a gradient ascent method to determine a local maximum as summarized in Algorithm 2. The gradient ascent algorithm consists in starting from an initial precoder $\mathbf{Q} = \mathbf{Q}^0$, with which $\delta_i^+, \delta_i^-, \tilde{\delta}_i^+, \tilde{\delta}_i^-$ are evaluated using Algorithm 1. This allows for an evaluation of the gradient at the initial step, from which a new precoder $\mathbf{Q}^1$ is derived, and the algorithm unfolds similarly as in the initial step until convergence.

Algorithm 2 Iterative optimization

Initialization: $m = 0$ and $\mathbf{Q}^0$ arbitrary.
Repeat
- Compute $\delta_i^+, \delta_i^-, \tilde{\delta}_i^+, \tilde{\delta}_i^-$, $i = 1,...,K$, using Algorithm 1
- Evaluate the gradient $\nabla\bar{R}_{\mathrm{sum}}$ w.r.t. $\mathbf{V} = [\mathbf{V}_1,...,\mathbf{V}_K]^{\mathrm{T}}$
- Let $\mathbf{V} = \mathbf{V} + \beta\nabla\bar{R}_{\mathrm{sum}}$ (for some step-size β)
- Let $\mathbf{Q}^{m+1} = \mathrm{Bdiag}\left(P_1\frac{\mathbf{V}_1\mathbf{V}_1^{\mathrm{H}}}{\mathrm{tr}(\mathbf{V}_1\mathbf{V}_1^{\mathrm{H}})},...,P_K\frac{\mathbf{V}_K\mathbf{V}_K^{\mathrm{H}}}{\mathrm{tr}(\mathbf{V}_K\mathbf{V}_K^{\mathrm{H}})}\right)$
- $m \leftarrow m+1$

until convergence.

The remainder of this section is dedicated to deriving explicit expression for the gradients required in Algorithm 2.

We know from (5) that the partial derivative of $\bar{R}_i^+$ w.r.t. $\delta_i^+, \tilde{\delta}_i^+$ for $i = 1,...,K$ is equal to zero (the same applies to $\bar{R}_i^-$ w.r.t. $\delta_i^-, \tilde{\delta}_i^-$). This fact simplifies the calculation of the gradient when using the differentiation chain rule.

In order to differentiate $\bar{R}_{\mathrm{sum}}$, we rewrite $\bar{R}_i^+$ as

$$\begin{aligned}\bar{R}_i^+ &= \log\left|(1+\tilde{\delta}_i^+)\mathbf{I}_{M_i} + \frac{1}{\bar{\sigma}^2}\sum_{j=1}^{K}\bar{\mathbf{H}}_{ij}\mathbf{Q}_j^{\frac{1}{2}}\mathbf{Z}_{ij}(\delta_i^+)\mathbf{Q}_j^{\frac{1}{2}}\bar{\mathbf{H}}_{ij}^{\mathrm{H}}\right|\\ &\quad - \sum_{j=1}^{K}\log\left|\mathbf{Z}_{ij}(\delta_i^+)\right| - N\bar{\sigma}^2\delta_i^+\tilde{\delta}_i^+\end{aligned} \tag{13}$$

in which $\mathbf{Z}_{ij}(x) = (\mathbf{I}_{N_j} + x\sigma_{ij}^2\mathbf{Q}_j)^{-1}$, and

$$\delta_i^+ = \frac{1}{N}\mathrm{tr}\left(\bar{\sigma}^2(1+\tilde{\delta}_i^+)\mathbf{I}_{M_i} + \sum_{j=1}^{K}\bar{\mathbf{H}}_{ij}\mathbf{Q}_j^{\frac{1}{2}}\mathbf{Z}_{ij}(\delta_i^+)\mathbf{Q}_j^{\frac{1}{2}}\bar{\mathbf{H}}_{ij}^{\mathrm{H}}\right)^{-1}.$$

Expanding $\tilde{\delta}_i^+$ in (10) using standard matrix inversion formulas, and using the fact that $\mathbf{Z}_{ij}(x)$ and $\mathbf{Q}_j^{\frac{1}{2}}$ commute, we obtain

$$\begin{aligned}\tilde{\delta}_i^+ =& \frac{1}{N}\sum_{j=1}^{K}\mathrm{tr}\left[\frac{\mathbf{Q}_j\mathbf{Z}_{ij}(\delta_i^+)}{\bar{\sigma}^2}\right.\\ &\left.\cdot\left(\mathbf{I}_{N_j} - \mathbf{Q}_j^{\frac{1}{2}}\bar{\mathbf{H}}_{ij}^{\mathrm{H}}(\mathbf{F}_i(\delta_i^+))^{-1}\bar{\mathbf{H}}_{ij}\mathbf{Q}_j^{\frac{1}{2}}\frac{\mathbf{Z}_{ij}(\delta_i^+)}{\bar{\sigma}^2}\right)\right]\end{aligned} \tag{14}$$

where

$$\mathbf{F}_i(\delta_i^+) = (1+\tilde{\delta}_i^+)\mathbf{I}_{M_i} + \frac{1}{\bar{\sigma}^2}\sum_{j=1}^{K}\bar{\mathbf{H}}_{ij}\mathbf{Z}_{ij}(\delta_i^+)\mathbf{Q}_j\bar{\mathbf{H}}_{ij}^{\mathrm{H}}. \tag{15}$$

Using those notations, (13) becomes

$$\bar{R}_i^+ = \log\left|\mathbf{F}_i(\delta_i^+)\right| - \sum_{j=1}^{K}\log\left|\mathbf{Z}_{ij}(\delta_i^+)\right| - N\bar{\sigma}^2\delta_i^+\tilde{\delta}_i^+. \quad (16)$$

Defining $d_n\mathbf{X} = \frac{\partial \mathbf{X}}{\partial \mathbf{V}_n^*}d\mathbf{V}_n^*$, the differential of $\bar{R}_i^+$ w.r.t. $\mathbf{V}_n$ is

$$\begin{aligned} d_n\bar{R}_i^+ &= \mathrm{tr}\left((\mathbf{F}_i(\delta_i^+))^{-1}d_n\mathbf{F}_i(\delta_i^+) - (\mathbf{Z}_{in}(\delta_i^+))^{-1}d_n(\mathbf{Z}_{in}(\delta_i^+))\right) \\ &= \mathrm{tr}\left(\mathbf{\Omega}_{in}^+ d_n\mathbf{Q}_n\right) \end{aligned} \quad (17)$$

where the last equality is detailed in Appendix B, and where

$$\mathbf{\Omega}_{in}^+ = \mathbf{Z}_{in}(\delta_i^+)\left(\frac{1}{\bar{\sigma}^2}\bar{\mathbf{H}}_{in}^{\mathrm{H}}(\mathbf{F}_i(\delta_i^+))^{-1}\bar{\mathbf{H}}_{in}\mathbf{Z}_{in}(\delta_i^+) + \delta_i^+\sigma_{in}^2\mathbf{I}_{N_n}\right). \quad (18)$$

Considering the normalization factors λ_j, note that the power constraint is always satisfied. In other words the optimization finds the precoding matrices $\mathbf{V}_j$ such that the normalized version of the corresponding covariance matrices $\mathbf{Q}_j = \lambda_j\mathbf{V}_j\mathbf{V}_j^{\mathrm{H}}$ will maximize the expected sum rate. It can be shown that $d_n\lambda_n = -\frac{\lambda_n^2}{P_n}\mathrm{tr}\left(\mathbf{V}_n d_n\mathbf{V}_n^{\mathrm{H}}\right)$, therefore the differential of $\mathbf{Q}_n$ can be evaluated as

$$\begin{aligned} d_n\mathbf{Q}_n &= d_n\left(\lambda_n\mathbf{V}_n\mathbf{V}_n^{\mathrm{H}}\right) \\ &= -\frac{\lambda_n^2}{P_n}\mathrm{tr}\left(\mathbf{V}_n d_n\mathbf{V}_n^{\mathrm{H}}\right)\mathbf{V}_n\mathbf{V}_n^{\mathrm{H}} + \lambda_n\mathbf{V}_n d_n\mathbf{V}_n^{\mathrm{H}}. \end{aligned} \quad (19)$$

Inserting (19) into (17) yields

$$\begin{aligned} d_n\bar{R}_i^+ &= -\frac{\lambda_n^2}{P_n}\mathrm{tr}\left(\mathbf{\Omega}_{in}^+\mathbf{V}_n\mathbf{V}_n^{\mathrm{H}}\right)\mathrm{tr}\left(\mathbf{V}_n d_n\mathbf{V}_n^{\mathrm{H}}\right) \\ &\quad + \lambda_n\mathrm{tr}\left(\mathbf{\Omega}_{in}^+\mathbf{V}_n d_n\mathbf{V}_n^{\mathrm{H}}\right) \\ &= \lambda_n\mathrm{tr}\left(\left[\mathbf{\Omega}_{in}^+\mathbf{V}_n - \frac{\mathrm{tr}\left(\mathbf{\Omega}_{in}^+\mathbf{V}_n\mathbf{V}_n^{\mathrm{H}}\right)}{\mathrm{tr}\left(\mathbf{V}_n\mathbf{V}_n^{\mathrm{H}}\right)}\mathbf{V}_n\right]^{\mathrm{T}} d_n\mathbf{V}_n^*\right). \end{aligned} \quad (20)$$

Therefore the gradient with respect to $\mathbf{V}_n$ is given by

$$\nabla_n\bar{R}_i^+ = \mathbf{\Omega}_{in}^+\mathbf{V}_n - \frac{\mathrm{tr}\left(\mathbf{\Omega}_{in}^+\mathbf{V}_n\mathbf{V}_n^{\mathrm{H}}\right)}{\mathrm{tr}\left(\mathbf{V}_n\mathbf{V}_n^{\mathrm{H}}\right)}\mathbf{V}_n. \quad (21)$$

The same procedure holds for $\bar{R}_i^-$ with $\mathbf{Q}_{-i}$ instead of $\mathbf{Q}$, and therefore we have $\nabla_n\bar{R}_{\text{sum}} = \sum_{i=1}^{K}(\nabla_n\bar{R}_i^+ - \nabla_n\bar{R}_i^-)$. Defining $\mathbf{V} = \left[\mathbf{V}_1, ..., \mathbf{V}_K\right]^{\mathrm{T}}$, the gradient with respect to $\mathbf{V}$ is $\nabla\bar{R}_{\text{sum}} = \left[\nabla_1\bar{R}_{\text{sum}}, ..., \nabla_K\bar{R}_{\text{sum}}\right]$.

V. Simulation Results

In this section, the performance of the proposed scheme is evaluated through numerical simulations. The performance metric is the expected sum rate (3) evaluated through Monte-Carlo simulations employing the covariance matrices designed with Algorithm 2. The whole process is repeated and averaged for many realizations of $\bar{\mathbf{H}}_{ij}$, with entries $\mathcal{CN}(0,1)$.

Figure 1 shows the expected sum rate versus transmit SNR for a three-user IC with four antennas per node and two data streams for each transmitter, using different precoder optimization schemes. We assume $\sigma_{ij} = 0.5$ and $\eta_{ij} = 3$ for

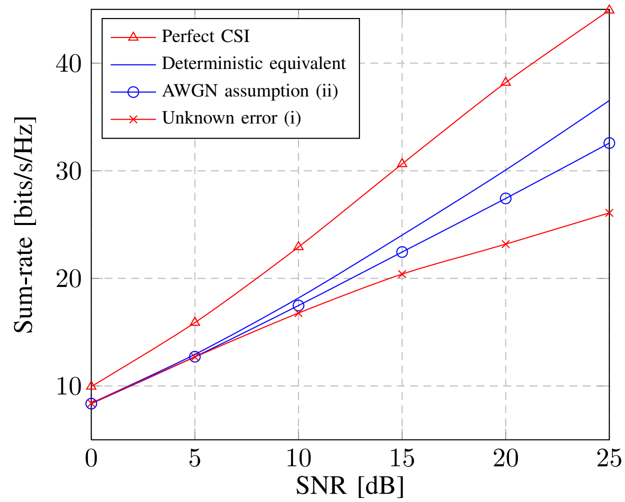

Fig. 1. Performance comparison of precoder optimization methods, for 3-user MIMO IC, $N_j = 4$, $M_i = 4$, $\eta_{ij} = 3$.

$i, j = 1, ..., K$. We compare the performance of the precoders provided by Algorithm 2 to two alternative approaches: (i) precoders are designed to maximize the sum rate under the assumption that $\bar{\mathbf{H}}_{ij}$ is the true channel (i.e. the transmitter assumes $\sigma_{ij}^2 = 0$), (ii) the signal resulting from the channel estimation error is modeled as an additive white Gaussian noise term. In that case, the last two terms in

$$\begin{aligned} \mathbf{y}_i = \bar{\mathbf{H}}_{ii}\sqrt{\lambda_i}\mathbf{V}_i\mathbf{x}_i + \sum_{j=1,j\neq i}^{K}\bar{\mathbf{H}}_{ij}\sqrt{\lambda_j}\mathbf{V}_j\mathbf{x}_j + \mathbf{n}_i \\ + \mathbf{E}_{ii}\sqrt{\lambda_i}\mathbf{V}_i\mathbf{x}_i + \sum_{j=1,j\neq i}^{K}\mathbf{E}_{ij}\sqrt{\lambda_j}\mathbf{V}_j\mathbf{x}_j \end{aligned}$$

are considered as noise and therefore the covariance matrix of the equivalent noise vector is $\sigma_i^2\mathbf{I}_{M_i}$ with $\sigma_i^2 = M_i\sum_{j=1}^{K}P_jN_j\sigma_{ij}^2 + \sigma^2$. In this case the precoders are designed as in (i) except for the different noise variance. We also provide as a reference the performance of precoders optimized with perfect CSI.

The results clearly show that our deterministic equivalent approach is superior to (i) and (ii). This suggests that, even though the system dimensions are small in this example, and therefore we operate far from the asymptotic regime $N_i, M_j \to \infty$, the approximation through deterministic equivalents outperforms the classical simplifying assumptions (i) and (ii).

VI. Conclusion

Precoder design for the MIMO IC was investigated when imperfect channel state information is available at the transmitters. The expected sum rate of the MIMO IC was approximated by extending the method presented in [1] to the interference channel and an iterative method was proposed to evaluate precoders that locally maximize the sum rate approximation.

Simulation results suggest that this method provides superior precoder design than classical alternative approximations.

Appendix A
Proof of Lemma 1

From the fact that, for $\mathbf{A}, \mathbf{B} \succeq 0$, $\frac{1}{N}\mathrm{tr}\left(\mathbf{A}(\mathbf{I}_N + \mathbf{B})^{-1}\right) \leq \|\mathbf{A}\|$, we have $g(\mathbf{Q}, \delta, \tilde{\delta}) \leq (M/N)\sigma^{-2}(1+\tilde{\delta})^{-1}$ and $\tilde{g}(\mathbf{Q}, \delta, \tilde{\delta}) \leq \sigma^{-2}\|\mathbf{Q}\tilde{\mathbf{C}}\| \triangleq U$. Define now $\hat{\delta} = \frac{U}{\tilde{\delta}} - 1$ which is one-to-one with $\tilde{\delta}$ and therefore we can equivalently discuss about the convergence and uniqueness of the following system of equations,

$$\begin{cases} \delta &= h(\mathbf{Q}, \delta, \hat{\delta}) = g(\mathbf{Q}, \delta, \frac{U}{1+\hat{\delta}}), \\ \hat{\delta} &= \hat{h}(\mathbf{Q}, \delta, \hat{\delta}) = \frac{U}{\tilde{g}(\mathbf{Q}, \delta, \frac{U}{1+\hat{\delta}})} - 1. \end{cases} \tag{22}$$

To prove uniqueness and convergence of the fixed-point algorithm, we use a result on *standard interference functions*.

Definition 1 ([10]): A function $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), ..., h_n(\mathbf{x}))$ in which $h_i : R_+^n \to R_+$, $i = 1, ..., K$ is a standard interference function if the following assumptions hold for all $i \in \{1, ..., n\}$:

I. *Positivity*: $h_i(\mathbf{x}) > 0$ for all $\mathbf{x} \geq 0$
II. *Monotonicity*: if $\mathbf{x}' \geq \mathbf{x}$, then $h_i(\mathbf{x}') \geq h_i(\mathbf{x})$
III. *Scalability*: for $\alpha > 1$, $h_i(\alpha\mathbf{x}) < \alpha h_i(\mathbf{x})$.

Also a vector $\mathbf{x}$ is said to be a feasible point if $\mathbf{h}(\mathbf{x}) < \mathbf{x}$ where the inequality is element-wise.

Theorem 2 ([10]): If $\mathbf{h}$ is a standard interference function, and there exists a feasible $\mathbf{x}$, then the fixed-point equation $\mathbf{h}(\mathbf{x}) = \mathbf{x}$ has a unique solution $\mathbf{x}^*$, given as the limit $\mathbf{x}^* = \lim_{t\to\infty} \mathbf{x}^t$, where, for all $t \geq 0$,

$$\mathbf{x}^{t+1} = \mathbf{h}(\mathbf{x}^t)$$

and $\mathbf{x}^0 > 0$ is arbitrary.

In our setting, we define $\mathbf{x} = (\delta, \hat{\delta})$ and $\mathbf{h}(\mathbf{x}) = (h(\delta, \hat{\delta}), \hat{h}(\delta, \hat{\delta}))$. Then, from Theorem 2, we only need to prove that $\mathbf{h}$ is a standard interference function that admits a feasible point. This implies the convergence of Algorithm 1 (considering the one-to-one map between $\tilde{\delta}$ and $\hat{\delta}$ at each iteration).

A feasible point can always be found since the functions h and $\hat{h}$ are bounded for every positive $(\delta, \hat{\delta})$. It is also easy to show that the positivity assumption always holds. For monotonicity, we have to show that $\delta' \geq \delta$ and $\hat{\delta}' \geq \hat{\delta}$ results in $h(\delta', \hat{\delta}') \geq h(\delta, \hat{\delta})$ and $\hat{h}(\delta', \hat{\delta}') \geq \hat{h}(\delta, \hat{\delta})$. For $\delta' \geq \delta$, we get $\mathrm{tr}\left(\delta'\mathbf{Q}^{\frac{1}{2}}\tilde{\mathbf{C}}\mathbf{Q}^{\frac{1}{2}}\right) \geq \mathrm{tr}\left(\delta\mathbf{Q}^{\frac{1}{2}}\tilde{\mathbf{C}}\mathbf{Q}^{\frac{1}{2}}\right)$. Therefore using

$$\mathrm{tr}(\mathbf{A}^{-1} - \mathbf{B}^{-1}) \geq 0 \;\Leftrightarrow\; \mathbf{B} \succeq \mathbf{A} \tag{23}$$

it is clear that

$$\left(\mathbf{I}_N + \delta\mathbf{Q}^{\frac{1}{2}}\tilde{\mathbf{C}}\mathbf{Q}^{\frac{1}{2}}\right)^{-1} \succeq \left(\mathbf{I}_N + \delta'\mathbf{Q}^{\frac{1}{2}}\tilde{\mathbf{C}}\mathbf{Q}^{\frac{1}{2}}\right)^{-1}. \tag{24}$$

Also from $\hat{\delta}' \geq \hat{\delta}$, we have

$$\sigma^2\left(1 + \frac{U}{1+\hat{\delta}}\right)\mathbf{I}_M \succeq \sigma^2\left(1 + \frac{U}{1+\hat{\delta}'}\right)\mathbf{I}_M. \tag{25}$$

Multiplying both sides of (24) by $\bar{\mathbf{H}}\mathbf{Q}^{\frac{1}{2}}$ and its Hermitian from left and right respectively, adding the result to (25), and using the equivalence in (23) results in $h(\delta', \hat{\delta}') \geq h(\delta, \hat{\delta})$ in which from (22), $h(\delta, \hat{\delta})$ is written as

$$h(\delta, \hat{\delta}) = \frac{1}{N}\mathrm{tr}\left(\sigma^2\left(1 + \frac{U}{1+\hat{\delta}}\right)\mathbf{I}_M + \bar{\mathbf{H}}\mathbf{Q}^{\frac{1}{2}}\mathbf{G}(\delta)\mathbf{Q}^{\frac{1}{2}}\bar{\mathbf{H}}^{\mathrm{H}}\right)^{-1}. \tag{26}$$

Using the same line of arguments, monotonicity of $\hat{h}$ and scalability of both h and $\hat{h}$ are then proved.

Appendix B
Calculating the Differentials

Using $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$,

$$\begin{aligned} d_n\bar{R}_i^+ = &\frac{1}{\bar{\sigma}^2}\mathrm{tr}\bigg(\bar{\mathbf{H}}_{in}^{\mathrm{H}}(\mathbf{F}_i(\delta_i^+))^{-1}\bar{\mathbf{H}}_{in}d_n\left(\mathbf{Z}_{in}(\delta_i^+)\mathbf{Q}_n\right)\bigg) \\ &- \mathrm{tr}\left(\left(\mathbf{Z}_{in}(\delta_i^+)\right)^{-1}d_n(\mathbf{Z}_{in}(\delta_i^+))\right). \end{aligned} \tag{27}$$

Using the facts that $d(\mathbf{XY}) = d(\mathbf{X})\mathbf{Y} + \mathbf{X}d(\mathbf{Y})$, $d(\mathbf{X}^{-1}) = -\mathbf{X}^{-1}d(\mathbf{X})\mathbf{X}^{-1}$ and $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, with some manipulations, (27) can be rewritten as (17).

Acknowledgment

This work was supported by the FP7 project HIATUS (grant 265578) of the European Commission and by the Austrian Science Fund (FWF) through grant NFN SISE (S106).

References

[1] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim, "On the capacity achieving covariance matrix for Rician MIMO channels: An asymptotic approach," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1048–1069, Mar. 2010.

[2] V. Cadambe and S. Jafar, "Interference alignment and the degrees of freedom of the K user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[3] S. H. Park, H. Park, Y. D. Kim, and I. Lee, "Regularized interference alignment based on weighted sum-MSE criterion for MIMO interference channels," in *Proc. IEEE Int. Conf. Commun.*, Cape Town, South Africa, May 2010.

[4] I. Santamaria, O. Gonzales, R. W. Heath, Jr., and S. W. Peters, "Maximum sum-rate interference alignment algorithms for MIMO channels," in *Proc. IEEE Global Telecommunications Conference (Globecom)*, Miami, Florida USA, Dec. 2010.

[5] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. M. Slock, "On the MIMO interference channel," in *Proc. Information Theory and Applications Workshop*, Jan. 2010.

[6] R. Tresch and M. Guillaud, "Cellular interference alignment with imperfect channel knowledge," in *Proc. IEEE International Conference on Communications (ICC)*, Dresden, Germany, Jun. 2004.

[7] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.

[8] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1175, Jul. 2004.

[9] S. Venkatesan, S. H. Simon, and R. A. Valenzuela, "Capacity of a gaussian MIMO channel with nonzero mean," in *Proc. IEEE Vehicular Technology Conference*, Orlando, Florida USA, Oct. 2003.

[10] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.