Weighted-Sum-Rate-Maximizing Linear Transceiver Filters for the K-User MIMO Interference Channel

Joonwoo Shin and Jaekyun Moon, *Fellow, IEEE*

Abstract—This letter is concerned with transmit and receive filter optimization for the K-user MIMO interference channel. Specifically, linear transmit and receive filter sets are designed which maximize the weighted sum rate while allowing each transmitter to utilize only the local channel state information. Our approach is based on extending the existing method of minimizing the weighted mean squared error (MSE) for the MIMO broadcast channel to the K-user interference channel at hand. For the case of the individual transmitter power constraint, however, a straightforward generalization of the existing method does not reveal a viable solution. It is in fact shown that there exists no closed-form solution for the transmit filter but simple one-dimensional parameter search yields the desired solution. Compared to the direct filter optimization using gradient-based search, our solution requires considerably less computational complexity and a smaller amount of feedback resources while achieving essentially the same level of weighted sum rate. A modified filter design is also presented which provides desired robustness in the presence of channel uncertainty.

Index Terms—K-user MIMO interference channel, linear transceiver design, sum-rate maximization.

I. Introduction

TO achieve high spectral efficiency, much effort has been focused on improving the achievable rate of multiple-input multiple-output (MIMO) interference channels [1]–[3]. A notable scheme in this area, the interference alignment (IA) technique of [4] confines all undesired interferences from other communication links into a pre-define subspace and achieves a maximum-capacity scaling. However, it is also known that IA can only offer a suboptimal sum rate at finite signal-to-noise ratios (SNRs) [3].

In this letter, we aim at maximizing the sum rate in the K-user MIMO interference channel. We consider two linear transceiver design methods. One is for the sum-power-usage-limit constraint and the other applies to the per-transmit-node power-usage constraint. The former can be viewed as a network-level constraint whereas the latter is more of a device-level constraint. In both designs, to maximize the weighted sum rate (WSR), we pursue minimization of the weighted mean squared error (WMSE). The idea of maximizing the WSR via receiver-side WMSE minimization was originally developed for the multi-user MIMO broadcast channel [5].

Our sum-power-constrained method could be seen as a generalization of the approach of [5] to cover the K-user MIMO interference channel and can be obtained as a direct extension of the method in [5]. However, our individual-power-constrained method is not a direct generalization of the method of [5] due to multiple power constraints. In fact, unlike in the case of the broadcast channel, we show that there is no closed-form solution for the minimum WMSE transmit filter, although a simple one-dimensional search for the power-adjusting parameter leads to the desired solution. Using simulation results and analysis, we verify that both proposed schemes achieve the maximum WSR with lower computational complexity than the gradient-based optimization of the transmit and receive filters [2]. Also, unlike in [2], [4], [6], our schemes require only the local channel state information (CSI) (i.e., each transmitter needs to know only the CSI of the links originating from itself whereas the MIMO interference channel precoder designs in [2], [4], [6] require the CSI for all links). Additionally, we discuss modified transceiver design that provides significant robustness in the presence of inaccurate CSI.

Related ideas for the MIMO interference channel can also be found in [3], [6]–[10]. In [3], [8], the minimum MSE (MMSE) transceiver is designed without considering different weights for the MSEs at multiple receivers. In [6] suboptimal MSE weights are used. In contrast, our weighted MMSE transceiver design relies on a set of MSE weights that provides a direct link between the weighted MMSE (WMMSE) and WSR criteria. The WMMSE-based weighted utility maximization is also considered in [7], but there only a single data stream is assumed between a given user pair. A very similar idea on maximizing WSR via WMSE minimization under the individual power constraint has been discussed in [9]. But, unlike in our approach, the inter-dependency between the transmit-power-adjusting Lagrange multiplier and the precoding matrix has not been considered in [9]. In our individual-power-constrained transceiver design, this inter-dependency is handled by introducing one-dimensional search for the Lagrange multiplier. This means that the method of [9] requires recursive optimization based on exchanges of filter-setting information among all transmitters. Our method does not require recursive filter adjustment and no data exchanges are needed among transmitters.[1] Finally, we present a modified transceiver design method for the imperfect-CSI environment and analyze the computational complexity as well as the required feedback amount in comparison with the gradient

---

[1]The independently conducted and recently published work of [10], which was brought to our attention by an anonymous reviewer, also pursues maximization of the WSR via weighted MSE minimization. The transceivers in [10] become the same as our proposed individual-power-constrained transceivers when each base station serves a single user.
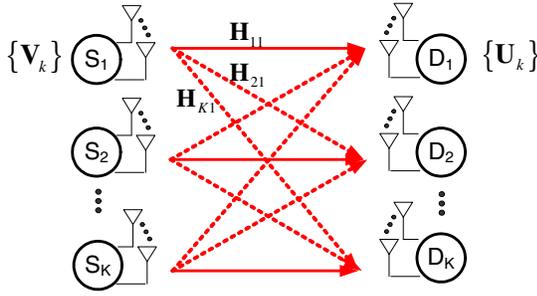
Fig. 1. K-user MIMO interference channel.

descent method [2].

The following notations are used. We employ upper case boldface letters for matrices and lower case boldface for vectors. For any general matrix $\mathbf{X}$, $\mathbf{X}^T$, $\mathbf{X}^*$, $\mathbf{X}^H$, $\text{Tr}(\mathbf{X})$, $\det(\mathbf{X})$, $\text{vec}(\mathbf{X})$, $\text{SVD}(\mathbf{X})$ denote the transpose, the conjugate, the Hermitian transpose, the trace, the determinant, the stack columns, and the singular value decomposition of $\mathbf{X}$, respectively. The symbol $||\cdot||_2$ indicates the 2-norm of a vector. The symbol $\mathbf{I}_n$ denotes an identity matrix of size n.

II. SYSTEM MODEL

We consider the MIMO interference channel where precoding can only be done over one transmission slot. As shown in Fig. 1, K source nodes simultaneously transmit independent data streams to their desired destination nodes and generate co-channel interference to all other undesired nodes. In this system each source node $\{\mathsf{S}_k\}$ is equipped with M antennas and each destination node $\{\mathsf{D}_k\}$ has N antennas ($k \in \{1 \sim K\}$). The MIMO channels from S_i to D_j are modelled by $\mathbf{H}_{ji} \in \mathcal{C}^{N \times M}$ ($i, j \in \{1 \sim K\}$) whose coefficients are independent and identically distributed (i.i.d) complex Gaussian random variables with $\mathcal{CN}(0, \sigma_h^2)$. We assume that the channel information is only *locally* available, i.e., each node knows only the coefficients for the channel link originating from itself. Note that the precoder designs of [2], [4], [6] are based on the availability of the *global* channel information. Let $\mathbf{s}_k \in \mathcal{C}^{d \times 1}$ denote the symbol vector from S_k with $\mathbb{E}[\mathbf{s}_k \mathbf{s}_k^H] = \mathbf{I}_d$ where d is the number of data streams for D_k, $d \leq M, N$ and the value of d is chosen to meet the feasibility of degree of freedom [11]. Also $\mathbf{V}_k \in \mathcal{C}^{M \times d}$ denotes the precoding matrix for S_k. Then, the $N \times 1$ received signal vector at D_k is represented as

$$\mathbf{y}_k = \mathbf{H}_{kk}\mathbf{V}_k\mathbf{s}_k + \sum_{i \neq k}^{K} \mathbf{H}_{ki}\mathbf{V}_i\mathbf{s}_i + \mathbf{n}_k, \quad (1)$$

where $\mathbf{n}_k$ denotes the i.i.d complex Gaussian noise vector at D_k with $\mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_N)$. Then, D_k combines its received signal with $\mathbf{U}_k \in \mathcal{C}^{d \times N}$ to decode the desired signals:

$$\hat{\mathbf{s}}_k = \mathbf{U}_k\mathbf{y}_k = \mathbf{U}_k\mathbf{H}_{kk}\mathbf{V}_k\mathbf{s}_k + \mathbf{U}_k\sum_{i \neq k}^{K} \mathbf{H}_{ki}\mathbf{V}_i\mathbf{s}_i + \mathbf{U}_k\mathbf{n}_k. \quad (2)$$

Our goal is to find $\{\mathbf{V}_k\}$ and $\{\mathbf{U}_k\}$ that maximize the WSR under the sum-power constraint and also the individual-power constraint. We assume a unit noise variance ($\sigma_n^2 = 1$) without losing generality.

III. WEIGHTED-SUM-RATE MAXIMIZATION

First consider finding $\{\mathbf{V}_k\}$ that maximizes

$$\sum_{k=1}^{K} \mu_k R_k \text{ subject to } \sum_k \text{Tr}(\mathbf{V}_k\mathbf{V}_k^H) \leq P_T \text{ or } \text{Tr}(\mathbf{V}_k\mathbf{V}_k^H)$$
$$\leq P_k \ \forall k \quad (3)$$

where the subscript k points the source node and its intended destination node, μ_k denotes the weight, R_k is the achievable rate, P_T represents the maximum sum power allowed for all transmitters and P_k is the k-th node's maximum transmit power. With Gaussian signaling, the achievable rate takes the well-known form:

$$R_k = \log \left\{ \det\left(\mathbf{I}_N + \boldsymbol{\Phi}_k^{-1}\mathbf{H}_{kk}\mathbf{V}_k\mathbf{V}_k^H\mathbf{H}_{kk}^H\right)\right\}, \quad (4)$$

where $\boldsymbol{\Phi}_k = \mathbf{I}_N + \sum_{i \neq k}^{K} \mathbf{H}_{ki}\mathbf{V}_i\mathbf{V}_i^H\mathbf{H}_{ki}^H$. We attempt to solve this WSR maximization problem by minimizing the weighted receiver MSE, as has been done for the MIMO broadcast channel [5]. This approach was also attempted for the K-user MIMO interference channel in [9] under the individual-power constraint, but our solution is different as elaborated below.

A. Relationship between the achievable rate and the error covariance matrix

To understand the link between the WSR maximization problem and the WMSE minimization problem in the K-user MIMO interference channel, we need to clarify the relationship between the achievable rate and the error covariance matrix. This argument is parallel to one given in [5] for the MIMO broadcast channel. For the MMSE receive filter at D_k, we write

$$\mathbf{U}_k^{(MMSE)} = \arg\min \mathbb{E}||\mathbf{U}_k\mathbf{y}_k - \mathbf{s}_k||_2^2$$
$$= \mathbf{V}_k^H\mathbf{H}_{kk}^H(\sum_{i=1}^{K} \mathbf{H}_{ki}\mathbf{V}_i\mathbf{V}_i^H\mathbf{H}_{ki}^H + \mathbf{I}_N)^{-1}, \quad (5)$$

and the error matrix for D_k is given by

$$\mathbf{E}_k = \mathbb{E}\{(\mathbf{U}_k^{(MMSE)}\mathbf{y}_k - \mathbf{s}_k)(\mathbf{U}_k^{(MMSE)}\mathbf{y}_k - \mathbf{s}_k)^H\}$$
$$= (\mathbf{I}_d + \mathbf{V}_k^H\mathbf{H}_{kk}^H\boldsymbol{\Phi}_k^{-1}\mathbf{H}_{kk}\mathbf{V}_k)^{-1}. \quad (6)$$

Comparing (4) and (6), the relationship between the achievable rate and the error covariance matrix is established as:

$$R_k = \log \{\det(\mathbf{E}_k^{-1})\} \quad (7)$$

which, not surprisingly, is identical to the relationship between the rate and the error covariance matrix for the case of the MIMO broadcast channel [5]. Apparently, though, the error covariance matrix $\mathbf{E}_k$ here is different from that of the broadcast channel due to the presence of multiple sources. Note that this relationship between the achievable rate and the error covariance matrix holds for any $\{\mathbf{V}_k\}$, implying that (7) is true with either transmit power constraint.

B. MSE weight design

Now consider finding $\{\mathbf{V}_k\}$ that solves the following WMMSE problem:

$$\min \sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k \mathbf{E}_k) \tag{8}$$

$$\text{subject to} \sum_{k} \text{Tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq P_T \text{ or } \text{Tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq P_k \ \forall k$$

where $\mathbf{W}_k \in \mathcal{C}^{d \times d}$ represents the MSE weight. Again following the argument of [5], the MSE weights can be chosen so that both WSR and WMMSE problems have a common solution. For this, set up the Lagrangians for (3) and (8) as in the top of the next page, where θ selects the desired power constraint ('$\theta = 1$' for the sum power constraint and '$\theta = 0$' for the individual power constraint), λ and $\{\lambda_k\}$ denote the Lagrange multipliers for the two transmit power constraints. Next, equate their gradients obtained via the matrix derivative formulas: $d\{\ln(\det(\mathbf{X}))\} = \text{Tr}\{\mathbf{X}^{-1}d(\mathbf{X})\}$, $d\{\text{Tr}(\mathbf{X})\} = \text{Tr}\{d(\mathbf{X})\}$, $\text{vec}\{d(\mathbf{X})\} = d\{\text{vec}(\mathbf{X})\}$, $\text{Tr}(\mathbf{X}^T \mathbf{Y}) = \text{vec}(\mathbf{X})^T \text{vec}(\mathbf{Y})$. Subsequently, the resulting MSE weight can be found as

$$\mathbf{W}_k = \frac{\mu_k}{\ln(2)} \mathbf{E}_k^{-1}. \tag{9}$$

Note that the choice of the MSE weights $\{\mathbf{W}_k\}$ is irrelevant to the transmit power constraint, which makes sense as $\{\mathbf{W}_k\}$ are receiver-side design parameters.

C. Sum-power-constrained precoder design

We are now ready to find the transmit precoding matrix that minimizes the WMSE under the sum-power constraint, i.e., find $\{\mathbf{V}_k\}$ that minimizes

$$\sum_{k=1}^{K} \mathbb{E}[\text{Tr}\{\mathbf{W}_k(\mathbf{s}_k - \beta^{-1}\hat{\mathbf{s}}_k)(\mathbf{s}_k - \beta^{-1}\hat{\mathbf{s}}_k)^H\}] \tag{10}$$

$$\text{subject to} \sum_{k} \text{Tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq P_T$$

where $\{\mathbf{W}_k\}$ is set according to (9) and β is a scaling parameter. With matrix derivative formulas, the WMMSE transmit filter that satisfies (10) can be shown to be

$$\mathbf{V}_k = \beta \mathbf{V}_k', \tag{11}$$

where $\mathbf{V}_k' = \left(\mathbf{\Psi}_k + \frac{\sum_i \text{Tr}(\mathbf{W}_i \mathbf{U}_i \mathbf{U}_i^H)}{P_T}\mathbf{I}_M\right)^{-1} \mathbf{H}_{kk}^H \mathbf{U}_k^H \mathbf{W}_k$, $\mathbf{\Psi}_k = \sum_{i=1}^{K} \mathbf{H}_{ik}^H \mathbf{U}_i^H \mathbf{W}_i \mathbf{U}_i \mathbf{H}_{ik}$, and $\beta = \sqrt{\frac{P_T}{\sum_k \text{Tr}(\mathbf{V}_k' \mathbf{V}_k'^H)}}$.

This result is a rather straightforward generalization of the WMMSE precoder in the broadcast channel. It can indeed be seen that setting $\mathbf{H}_{ki} = \mathbf{H}_{kk}$ for all i, our solutions (5), (9), and (11) reduce to the respective receive filter, MSE weight and transmit filter solutions obtained for the multi-user MIMO broadcast channels through WMSE minimization [5].

D. Individual-power-constrained transceiver design

Now let us consider the individual-power-constrained network. We proceed to find the transmit filter that minimizes the weighted MSE:

$$\sum_{k=1}^{K} \mathbb{E}[\text{Tr}\{\mathbf{W}_k(\mathbf{s}_k - \hat{\mathbf{s}}_k)(\mathbf{s}_k - \hat{\mathbf{s}}_k)^H\}] \tag{12}$$

$$\text{subject to} \text{Tr}(\mathbf{V}_k \mathbf{V}_k^H) \leq P_k \quad \forall k.$$

Again equating the gradients of the Lagrangians corresponding to the WMMSE and WSR maximization procedures and using the matrix derivative formulas, the WMMSE transmit filter at S_k is found as:

$$\mathbf{V}_k = \left(\mathbf{\Psi}_k + \lambda_k \mathbf{I}_M\right)^{-1} \mathbf{H}_{kk}^H \mathbf{U}_k^H \mathbf{W}_k \tag{13}$$

where λ_k is set to satisfy the transmit power constraint at S_k and again $\{\mathbf{W}_k\}$ are as given in (9). Unlike the sum-power-constrained WMMSE precoders of (11), for which the power control parameters are found in closed form, here we resort to a numerical method to find λ_k, due to the inter-dependency between $\mathbf{V}_k$ and λ_k in (13). Fortunately, based on the following lemma, λ_k can be found with simple one-dimensional (1-D) numerical search.

Lemma 1: The per-node transmit power, $\text{Tr}(\mathbf{V}_k(\lambda_k)\mathbf{V}_k(\lambda_k)^H)$, is a monotonically decreasing function of λ_k.

Proof: Let $\text{SVD}(\mathbf{\Psi}_k) = \mathbf{Q}_k \mathbf{\Sigma}_k \mathbf{Q}_k^H$. Then, the transmit power at S_k is given by

$$\begin{aligned}
&\text{Tr}\{(\mathbf{V}_k(\lambda_k)\mathbf{V}_k(\lambda_k)^H\} \\
&=\text{Tr}\{(\mathbf{Q}_k \mathbf{\Sigma}_k \mathbf{Q}_k^H + \lambda_k \mathbf{I_M})^{-2}\mathbf{H}_{kk}^H \mathbf{U}_k^H \mathbf{W}_k \mathbf{U}_k \mathbf{H}_{kk}\} \\
&=\text{Tr}\{(\mathbf{\Sigma}_k + \lambda_k \mathbf{I_M})^{-2}\mathbf{Q}_k^H \mathbf{H}_{kk}^H \mathbf{U}_k^H \mathbf{W}_k \mathbf{U}_k \mathbf{H}_{kk}\mathbf{Q}_k\} \\
&=\sum_{i=1}^{K} \frac{[\mathbf{\Pi}_k]_{i,i}}{(\sigma_{k,i} + \lambda_k)^2},
\end{aligned} \tag{14}$$

where $\mathbf{\Pi}_k = \mathbf{Q}_k^H \mathbf{H}_{kk}^H \mathbf{U}_k^H \mathbf{W}_k \mathbf{U}_k \mathbf{H}_{kk}\mathbf{Q}_k$, $[\mathbf{\Pi}_k]_{i,i}$ is the ii-th element of $\mathbf{\Pi}_k$, and $\sigma_{k,i}$ is the i-th element of $\mathbf{\Sigma}_k$. Because $\lambda_k \geq 0$, $\text{Tr}\{\mathbf{V}_k(\lambda_k)\mathbf{V}_k(\lambda_k)^H\}$ is monotonically decreasing with λ_k. ∎

Note that the proper set of MSE weights for the K-user MIMO interference channel has already been derived in [9] in the process of establishing a connection between the WMMSE problem and the WSR maximization problem. In [9], though, the transmitter S_k is expressed as a function of itself as well as transmitters at the other nodes, i.e. $\mathbf{V}_k = f(\{\mathbf{V}_1, \cdots, \mathbf{V}_K\})$. The consequence of this formulation is that the transmitter solution in [9] cannot be found without recursive calculation and additional filter-setting information exchanges among all transmit nodes. In contrast, our transmit filter design is based on a clear recognition of the inter-dependency between λ_k and $\mathbf{V}_k$, and as a result the proposed transmit filter (13) can be found through a simple 1-D numerical search with no additional information ($\{\mathbf{V}_l\}$ ($l \neq k$)) exchanges needed among the transmit nodes.

E. Iterative algorithm to maximize the weighted-sum-rate

In the previous sections, we found the MSE weights and then subsequently WMMSE receive and transmit filters with

$$\mathcal{L}_{WSR} = -\sum_{k=1}^{K} \mu_k R_k + \theta\lambda(\sum_{k=1}^{K} \text{Tr}(\mathbf{V}_k\mathbf{V}_k^H) - P_T) + (1-\theta)\Big(\sum_{k=1}^{K} \lambda_k(\text{Tr}(\mathbf{V}_k\mathbf{V}_k^H) - P_k)\Big)$$

$$\mathcal{L}_{WMSE} = \sum_{k=1}^{K} \text{Tr}(\mathbf{W}_k\mathbf{E}_k) + \theta\lambda(\sum_{k=1}^{K} \text{Tr}(\mathbf{V}_k\mathbf{V}_k^H) - P_T) + (1-\theta)\Big(\sum_{k=1}^{K} \lambda_k(\text{Tr}(\mathbf{V}_k\mathbf{V}_k^H) - P_k)\Big)$$

both the sum-power constraint and the individual-power constraint. Each of three sets of parameters - MSE weights, transmit filters and receive filters - is derived assuming the other sets are given. In practice, to find optimum WSR solutions, the inter-dependencies between the parameters are handled with the following iterative or alternating optimization algorithm.

Algorithm 1 Obtaining the optimal WSR transceivers via the WMMSE criterion

---

Initialize $l = 0$ and $\{\mathbf{V}_k^{(0)}\}$, calculate $R_{sum}^{(0)}$.
repeat
  $l := l + 1$
  Step 1: Calculate $\mathbf{U}_k^{(l)}|\{\mathbf{V}_i^{(l-1)}\}$ for all k using (5).
  Step 2: Calculate $\mathbf{W}_k^{(l)}|\{\mathbf{V}_i^{(l-1)}\}$ for all k using (9).
  Step 3: Calculate $\mathbf{V}_k^{(l)}|\{\mathbf{U}_i^{(l)}\}$, $\{\mathbf{W}_i^{(l)}\}$ for all k using (11) for the sum power constrained case or (13) for the individual power constrained case.
until $|R_{sum}^{(l)} - R_{sum}^{(l-1)}| < \epsilon$, where ϵ is some arbitrarily small value and $R_{sum} = \sum_k \mu_k R_k$.

The algorithm is common to both the sum-power-constrained design and the individual-power-constrained design. This algorithm is provably convergent to a local optimum; this can be shown by proving monotonic convergence of an equivalent optimization problem based on expanding the WSR maximization problem of (3) to add the MMSE weights and receive filters as optimization variables, as has been done for the MIMO broadcast channel in [5]. We note, however, that this algorithm does not guarantee the global optimal solution, since the WMMSE minimization (8) is not *jointly* convex over all input variables. To reasonably approach the optimal solution one must resort to repeated runs of the algorithm using different initial settings, or, for computationally efficient initialization, choose $\{\mathbf{V}_k^{(0)}\}$ in Step 1 from the right singular matrices of $\{\mathbf{H}_{kk}\}$ or from random matrices generated according to the normal distribution with zero mean and unit variance [8].

IV. ROBUST TRANSCEIVER DESIGN FOR IMPERFECT CHANNEL INFORMATION

In practical scenarios, mismatch between the true channel $\{\mathbf{H}_{ij}\}$ and the estimated channel (denoted by $\{\tilde{\mathbf{H}}_{ij}\}$) is inevitable because of the channel estimation errors [12]. In this section, we design robust transceivers for mitigating the performance degradation caused by channel mismatch. We assume that $\{\tilde{\mathbf{H}}_{ij}\}$ is related to $\{\mathbf{H}_{ij}\}$ by $\tilde{\mathbf{H}}_{ij} = \mathbf{H}_{ij} + \boldsymbol{\Delta}_{ij}$ where the elements of $\boldsymbol{\Delta}_{ij}$ are independent and identically distributed (i.i.d.) complex Gaussian random variables with

variance σ_Δ^2 [12]. Then, the received signal can be rewritten as

$$\tilde{\mathbf{s}}_k = \tilde{\mathbf{U}}_k(\tilde{\mathbf{H}}_{kk} - \boldsymbol{\Delta}_{kk})\tilde{\mathbf{V}}_k\mathbf{s}_k + \tilde{\mathbf{U}}_k\sum_{i \neq k}^{K}(\tilde{\mathbf{H}}_{ki} - \boldsymbol{\Delta}_{ki})\tilde{\mathbf{V}}_i\mathbf{s}_i + \tilde{\mathbf{U}}_k\mathbf{n}_k \tag{15}$$

where $\{\tilde{\mathbf{V}}_k\}$ and $\{\tilde{\mathbf{U}}_k\}$ are computed from $\{\tilde{\mathbf{H}}_{ij}\}$ with no knowledge of the presence of $\{\boldsymbol{\Delta}_{ij}\}$. We try to mitigate the effect of channel mismatch by minimizing the appropriate metrics averaged over $\boldsymbol{\Delta}_{ij}$'s.

1) Modified MSE weight: Following the design procedure in previous sections, a modified version of the MMSE receiver filter is found as $\tilde{\mathbf{U}}_k = \tilde{\mathbf{V}}_k^H\tilde{\mathbf{H}}_{kk}^H(\sum_{i=1}^{K}\tilde{\mathbf{H}}_{ki}\tilde{\mathbf{V}}_i\tilde{\mathbf{V}}_i^H\tilde{\mathbf{H}}_{ki}^H + \sum_{i=1}^{K}\sigma_\Delta^2\text{Tr}(\boldsymbol{\Lambda}_{\tilde{\mathbf{V}}_i})\mathbf{I}_N + \mathbf{I}_N)^{-1}$, where $\text{SVD}(\tilde{\mathbf{V}}_i\tilde{\mathbf{V}}_i^H) = \mathbf{Q}_i\boldsymbol{\Lambda}_{\tilde{\mathbf{V}}_i}\mathbf{Q}_i^H$. The modified MSE weights that force the optimum solutions of the WSR maximization and WMMSE problems to be identical are derived as $\tilde{\mathbf{W}}_k = \frac{\mu_k}{\ln(2)}\tilde{\mathbf{E}}_k^{-1}$, where $\tilde{\mathbf{E}}_k = (\mathbf{I}_d + \tilde{\mathbf{V}}_k^H\tilde{\mathbf{H}}_{kk}^H\tilde{\boldsymbol{\Phi}}_k^{-1}\tilde{\mathbf{H}}_{kk}\tilde{\mathbf{V}}_k)^{-1}$ and $\tilde{\boldsymbol{\Phi}}_k = \mathbf{I}_N + \sum_{i \neq k}^{K}\tilde{\mathbf{H}}_{ki}\tilde{\mathbf{V}}_i\tilde{\mathbf{V}}_i^H\tilde{\mathbf{H}}_{ki}^H + \sum_{i=1}^{K}\sigma_\Delta^2\text{Tr}(\boldsymbol{\Lambda}_{\tilde{\mathbf{V}}_i})\mathbf{I}_N$.

2) Robust transceiver design with the sum power constraint: The modified transmit filters are derived based on the following optimization problem:

$$\min\sum_{k=1}^{K}\mathbb{E}[\text{Tr}\{\tilde{\mathbf{W}}_k(\mathbf{s}_k - \tilde{\beta}^{-1}\tilde{\mathbf{s}}_k)(\mathbf{s}_k - \tilde{\beta}^{-1}\tilde{\mathbf{s}}_k)^H\}] \tag{16}$$

$$\text{subject to }\sum_k\text{Tr}(\tilde{\mathbf{V}}_k\tilde{\mathbf{V}}_k^H) \leq P_T.$$

Utilizing matrix derivative formulas, the resultant modified-WMMSE transmit filters are obtained as

$$\tilde{\mathbf{V}}_k = \tilde{\beta}\tilde{\mathbf{V}}_k', \tag{17}$$

where $\tilde{\mathbf{V}}_k' = \Big(\tilde{\boldsymbol{\Psi}}_k + \frac{\sum_{i=1}^{K}\text{Tr}(\tilde{\mathbf{W}}_i\tilde{\mathbf{U}}_i\tilde{\mathbf{U}}_i^H)}{P_T}\mathbf{I}_M + \sum_{i=1}^{K}\sigma_\Delta^2\text{Tr}(\boldsymbol{\Lambda}_{\tilde{\mathbf{U}}_i})\mathbf{I}_M\Big)^{-1}\tilde{\mathbf{H}}_{kk}^H\tilde{\mathbf{U}}_k^H\tilde{\mathbf{W}}_k$, $\tilde{\beta} = \sqrt{\frac{P_T}{\sum_k\text{Tr}(\tilde{\mathbf{V}}_k'\tilde{\mathbf{V}}_k'^H)}}$, $\tilde{\boldsymbol{\Psi}}_k = \sum_{i=1}^{K}\tilde{\mathbf{H}}_{ik}^H\tilde{\mathbf{U}}_i^H\tilde{\mathbf{W}}_i\tilde{\mathbf{U}}_i\tilde{\mathbf{H}}_{ik}$, and $\text{SVD}(\tilde{\mathbf{U}}_i^H\tilde{\mathbf{W}}_i\tilde{\mathbf{U}}_i) = \mathbf{Q}_i'\boldsymbol{\Lambda}_{\tilde{\mathbf{U}}_i}\mathbf{Q}_i'^H$.

3) Robust transceiver design with the individual power constraint: The optimization problem to derive the modified precoder is

$$\min\sum_{k=1}^{K}\mathbb{E}[\text{Tr}\{\tilde{\mathbf{W}}_k(\mathbf{s}_k - \tilde{\mathbf{s}}_k)(\mathbf{s}_k - \tilde{\mathbf{s}}_k)^H\}] \tag{18}$$

$$\text{subject to }\text{Tr}(\tilde{\mathbf{V}}_k\tilde{\mathbf{V}}_k^H) \leq P_k \quad \forall k.$$

With the matrix derivative formulas, the modified-WMMSE transmit precoder at S_k with the individual power constraint
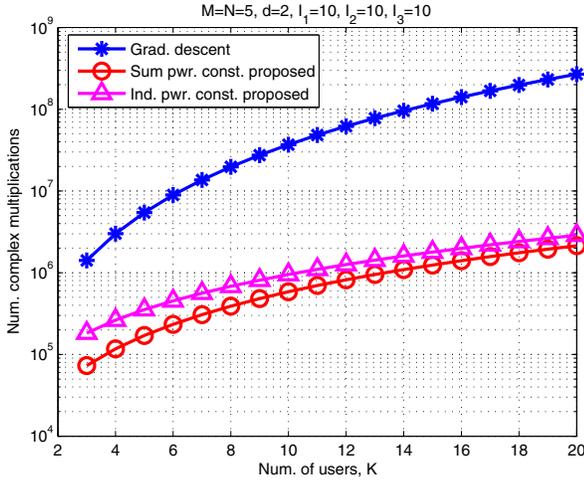
Fig. 2. Complexity comparison between the gradient-descent-based method and the proposed methods.

Fig. 3. The amount of feedback information at each source node to design precoder.

is written as

$$\tilde{\mathbf{V}}_k = \left(\tilde{\boldsymbol{\Psi}}_k + \tilde{\lambda}_k \mathbf{I}_M + \sum_{i=1}^{K} \sigma_\Delta^2 \text{Tr}(\boldsymbol{\Lambda}_{\tilde{\mathbf{U}}_i}) \mathbf{I}_M \right)^{-1} \tilde{\mathbf{H}}_{kk}^H \tilde{\mathbf{U}}_k^H \tilde{\mathbf{W}}_k \tag{19}$$

where the power control parameter $\tilde{\lambda}_k$ is also found by numerical 1-D search.

Note that, for the above derivations, we have assumed that the value of the channel error variance σ_Δ^2 is perfectly known. In the practical systems, the channel error variance can be estimated through an appropriate statistical approach [13]. Below, we also present numerical performance results corresponding to the cases where the error variance is not perfectly known.

V. DISCUSSION: COMPUTATIONAL COMPLEXITY AND CHANNEL STATE INFORMATION

In this section, we analyze computational complexity and required feedback resources. For comparison, we also analyze those of the gradient descent method of [2].

A. Computational complexity

We consider the number of complex multiplications as a complexity measure. As summarized in the Table I, the number of complex multiplications is proportional to the number of iterations. The proposed method with the sum-power constraint which has a single iteration loop is computationally the most efficient. Whereas both the proposed method with the individual-power constraint and the gradient descent method require double iteration loops, i.e., the outer loop for updating the sum rate and the inner loop for adjusting the Lagrange multiplier (in the case of the proposed method) or for updating the step size (in the case of the gradient-based method). Calculating the gradient and adjusting the step size require more computational resources. According to simulation, when SNR = 10 dB which is in the mid SNR regime, $K = 4$, $M = N = 5$, and $d = 2$, the minimum average numbers
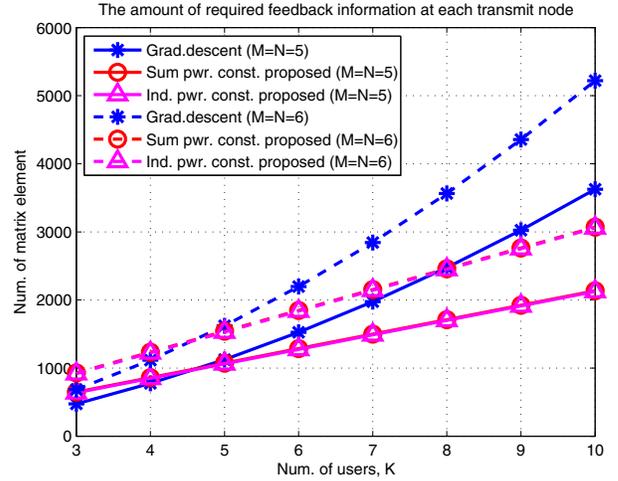
of iteration for the convergence of sum rate, updating the step size of gradient method and 1-D search with bisection method are 10, 10 and 10, respectively. In accordance with these, $I_1 = 10$, $I_2 = 10$ and $I_3 = 10$ are chosen. The symbols c_N^1, c_{NM}^2 and c_N^3 denote the computational complexity of a matrix inversion of $N \times N$ matrix, a singular value decomposition of $N \times M$ matrix, and a Cholesky factorization of $N \times N$ matrix, respectively. The corresponding values for those variables are $\frac{2}{3}N^3$, $7NM^2 + 4M^3$, and $\frac{1}{3}N^3$, respectively [14]. Fig. 2 shows comparison when $M = N = 5$ and $d = 2$ [2]. As expected, for the same WSR values the proposed method with the sum-power constraint has the least complexity while the gradient descent algorithm is the most computationally complex.

B. The amount of required feedback information

To find the optimized transmit precoders, each transmit node requires feedback information. As illustrated in Table II, feedback information is composed of CSI and coefficients for filter updating. For a given transmission slot, CSI feedback is required once, but the filter coefficients are updated several times due to the iterative optimization algorithm. Although the proposed method requires a larger amount of feedback information for the iteratively updated coefficients such as MSE weights $\{\mathbf{W}_k\}$ and receive filter coefficients $\{\mathbf{U}_k\}$ than the gradient descent method does, the amount of CSI feedback for the proposed method is smaller than for the gradient descent method. This is because, unlike the global CSI requirement of the gradient-based method, the proposed methods need only local CSI. From Table II, we observe that as the network size grows (i.e., K increases) the required feedback resources for local CSI and coefficient updating increase linearly, but those for global CSI increases quadratically. Fig. 3 clearly shows that with $I_1 = 10$ the proposed methods are advantageous in terms of required feedback resources, especially for larger K. Note that, for the transmit power adjustment, the sum-power-constrained method additionally requires iterative update of

---

[2]To see the effect of the number of K, we fixed $M = N = 5$, even though the degree of freedom (DoF) is not achievable when $K \geq 5$

TABLE I
COMPUTATIONAL COMPLEXITY COMPARISON

(a) Description of each stage for gradient descent method

STAGE		Index
Initialization		a.1
Outer loop	Calculating gradient	a.2
	Inner loop: calculating step size	a.3
	Calculating sum rate	a.4
Calculating optimal precoders and decoders		a.5

(b) Description of each stage for proposed methods

STAGE		Index
Initialization		b.1
Loop	Calculating the variance of noise and interference	b.2
	Calculating the receive filter	b.3
	Calculating the error covariance matrix	b.4
	Calculating the MSE weights	b.5
	Calculating the transmit filter (1-D search is needed for individual power constraint)	b.6-1 (for sum power constraint) b.6-2 (for individual power constraint)
	Calculating sum rate	b.7

(c) Number of complex multiplication at each stage

Index	Number of complex multiplication
a.1	$K(M^2d+1) + K(K-1)(1+2MNd+N^2d) + K(2+2MNd+N^2d+N^3+c_N^1)$
a.2	$I_1\Big\{ K(2K-1)(1+2MNd+N^2d)$ $+K(2K-1)(9+2c_N^1+2MN^2+2M^2N+2M^2d+Md^2)\Big\}$
a.3	$I_1\Big\{ KI_2(I_2+1)/2 + KI_2\{2K(K-1)(1+2MNd+N^2d)$ $+2K(2+2MNd+N^2d+N^3+c_N^1)+K(M^2d+1)+2+Md^2\}\Big\}$
a.4	$I_1\{K(M^2d+1)+K(K-1)(1+2MNd+N^2d)+K(2+2MNd+N^2d+N^3+c_N^1)\}$
a.5	$K(1+2Md+2M^2d+c_{Md}^2)+K(K-1)(2MNd+N^2d)$ $+K(2MNd+2N^2d+4Nd^2+Md^2+d^3+c_N^1+c_d^3+c_d^1+c_{dd}^2)$
b.1	$K(M^2d+1) + K(K-1)(1+2MNd+N^2d) + K(2+2MNd+N^2d+N^3+c_N^1)$
b.2	$I_1K(K-1)(1+2MNd+N^2d)$
b.3	$I_1K(3MNd+2N^2d+c_N^1)$
b.4	$I_1K(2MNd+N^2d+Nd^2+c_N^1+c_d^1)$
b.5	$I_1Kc_d^1$
b.6-1	$I_1\Big\{ K(K-1)(2NMd+Md^2+M^2d)+K(Nd^2+d^3)$ $+K(3MNd+2Md^2+M^2d+1+c_M^1)+K(M^2d+Md)\Big\}$
b.6-2	$I_1\Big\{ K(K-1)(2NMd+Md^2+M^2d)+I_3KM^2d$ $+(I_3+1)K(3MNd+2Md^2+M^2d+1+c_M^1)\Big\}$
b.7	$I_1\{K(M^2d+1)+K(K-1)(1+2MNd+N^2d)+K(2+2MNd+N^2d+N^3+c_N^1)\}$

TABLE II
SUMMARY OF REQUIRED FEEDBACK INFORMATION AT THE k-TH TRANSMIT NODE, S_k, $i,j = 1 \sim K$

	Grad. descent method		Prop. method	
	Global CSI	Updating coefficients	Local CSI	Updating coefficients
Feedback information	$\{\mathbf{H}_{ij}\}$	$\{\mathbf{V}_i\}$, $(i \neq k)$	$\{\mathbf{H}_{ik}\}$	$\{\mathbf{U}_i\}$, $\{\mathbf{W}_i\}$ (Ind. pwr.)
				$\{\mathbf{U}_i\}$, , $\{\mathbf{W}_i\}$, $\sum_{i\neq k}\mathrm{Tr}\{\mathbf{V}_i\mathbf{V}_i'^H\}$ (Sum pwr.)
Matrix size	MNK^2	$Md(K-1)$	MNK	$(Md+d^2)K$ (Ind. pwr.)
				$(Md+d^2)K+1$ (Sum. pwr.)
Feedback resource amount	$MNK^2 + Md(K-1)I_1$		$MNK + (Md+d^2)KI_1$ (Ind. pwr.)	
			$MNK + ((Md+d^2)K+1)I_1$ (Sum. pwr.)	

the scalar parameter $\mathrm{Tr}\{\sum_{i\neq k}\mathbf{V}_i'\mathbf{V}_i'^H\}$, but the size of this parameter is negligible compared to other matrix parameters.

VI. NUMERICAL RESULTS

In this section, we provide the numerical results related to the WSR performances. The SNR for the sum-power-constrained network, SNR $= \frac{P_T\sigma_h^2}{K\sigma_n^2}$, and that for the individual power constrained network, SNR$_k = \frac{P_k\sigma_h^2}{\sigma_n^2}$, $\forall k$, are derived assuming $P_T = K$, $P_k = 1 \forall k$ and $\sigma_n^2 = 1$, i.e.,

SNR $=$ SNR$_k = \sigma_h^2$. The results are averaged over 1000 independent trials. Fig. 4 shows the average WSR performance of the proposed methods for $M = N = 5$ (when $K = 4$), $M = N = 6$ (when $K = 5$), and $d = 2$. For fairness, all schemes are initialized with the right singular matrices of the intended channels. For the sum-power constraint, we set the weights to be $\mu_1 = 2$ and $\mu_k = 0.25$ ($k \neq 1$), which were chosen rather arbitrarily except that μ_1 is made considerably larger than μ_k to bring out the performance advantage of
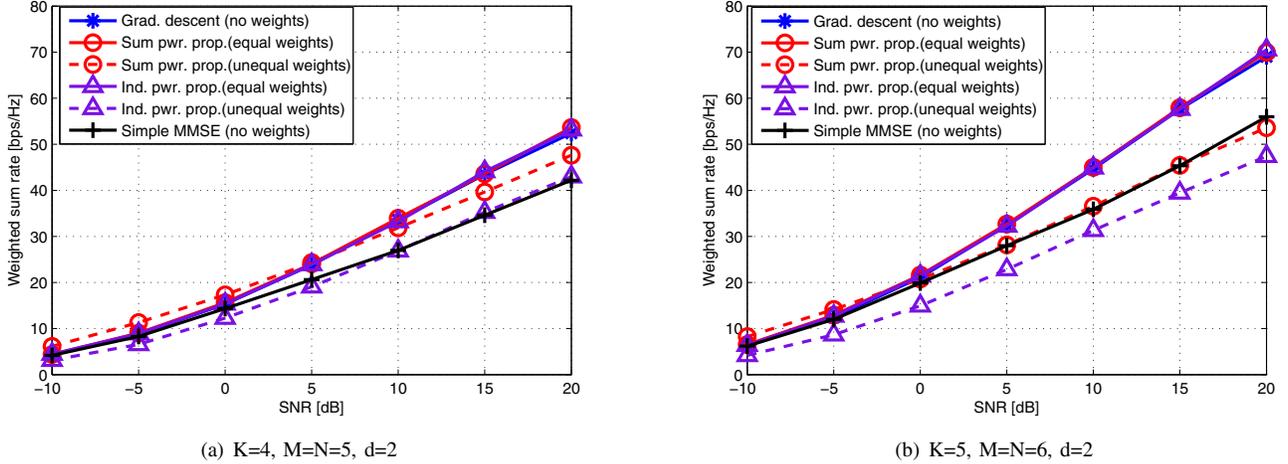
(a) K=4, M=N=5, d=2

(b) K=5, M=N=6, d=2

Fig. 4. The weighted sum rate performance (equal weight $\mu_k = 1, \forall k$ and unequal weights $\mu_1 = 2, \mu_{k'} = 0.25, k' \neq 1$).
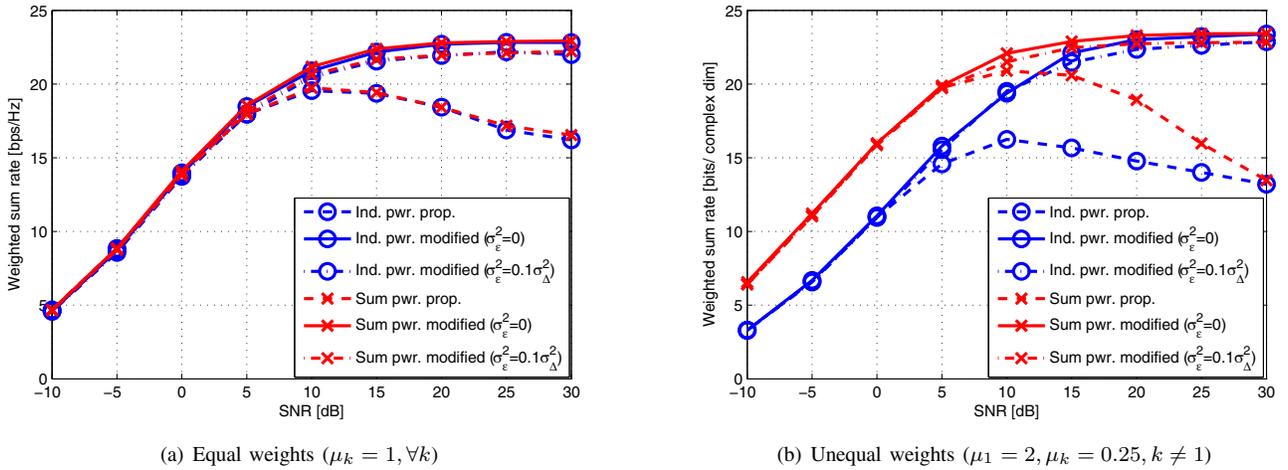

(a) Equal weights ($\mu_k = 1, \forall k$)

(b) Unequal weights ($\mu_1 = 2, \mu_k = 0.25, k \neq 1$)

Fig. 5. The weighted sum rate performances with imperfect channel state information ($\sigma_\Delta^2 = 0.1 \times \sigma_h^2, K = 4$).

the sum-power constraint. The performance of the sum power constraint method should be better than that of the individual power constraint method because the former, which is less stringent, is able to allocate more power to the higher weighted transmitter to maximize the WSR. When the weights are equal, $\mu_k = 1 \ \forall k$, the performance of both proposed schemes and that of the conventional gradient descent method are nearly identical. Note that, as explained in section IV, the proposed methods achieve these performances with less computational complexity and a smaller amount of feedback resources than the gradient descent method. Compared to the performance of the MMSE transceiver without the MSE weights [3], [8] (curves labelled "Simple MMSE"), the advantage of designed MSE weights is clearly shown as SNR grows. Fig. 5 demonstrates the effectiveness of the robust design with either transmit power constraint in presence of channel uncertainty when $K = 4$ and $\sigma_\Delta^2 = 0.1\sigma_h^2$. As SNR grows, the amount of leakage interference due to CSI imperfection also increases. This is why the performance is saturated in the high SNR regime in Fig. 5. To reflect a potential error in estimating σ_Δ^2, we model the channel error variance as $\sigma_\Delta^2 + \sigma_\epsilon^2$, where σ_Δ^2

is the actual channel error variance and σ_ϵ^2 indicates over-estimation. As shown in Fig. 5, at SNR = 15 dB at most 3 % sum rate losses are shown when $\sigma_\epsilon^2 = 0.1\sigma_\Delta^2$. Although not shown, same results were observed for under-estimating the channel estimation error variance.

VII. CONCLUSIONS

In this letter, we have studied a linear transceiver design method for the K-user MIMO interference channel. To maximize the weighted sum rate with less computational complexity and a smaller amount of feedback resources, the proposed transceivers are designed in the weighted MMSE sense with suitably chosen MSE weights. Also, the proposed transceiver design considers both the sum-power-usage constraint and the individual-power constraint. Through numerical simulation, we have demonstrated that the weighed-sum-rate performances of the proposed schemes approach that of the existing gradient descent method. The proposed methods have clear advantage in terms of processing requirements as well as feedback resources over the gradient-based technique. Also,

modified versions of proposed schemes have been provided for compensating channel mismatch.

REFERENCES

[1] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," ArXiv pre-print cs.IT/1011.3816. Available: http://arxiv.org/abs/0803.3816.

[2] H. Sung, S. H. Park, K. J. Lee, and I. Lee, "Linear precoder designs for K-user interference channels," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 291–301, Jan. 2010.

[3] S. W. Peters and R. W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 60, pp. 206–218, Jan. 2011.

[4] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3425–3441, Aug. 2008.

[5] S. S. Christensen, R. Agarwal, E. Carvalho, and J. M. Cioffi, "Weighted sum rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4792–4799, Dec. 2008.

[6] 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. 2010 IEEE Int. Conf. Commun.*.

[7] D. A. Schmidt, S. Changxin, R. Berry, M. L. Honig, and W. Utschick, "Minimum mean squared error interference alignment," in *2009 Asilomar Conf. Signals, Syst. Comput.*.

[8] H. Shen, B. Li, M. Tao, and X. Wang, "MSE-based transceiver designs for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3480–3489, Nov. 2010.

[9] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. M. Slock, "Weighted sum rate maximization in the MIMO interference channel," in *Proc. 2010 IEEE Conf. Personal Indoor Mobile Radio Commun.*.

[10] Q. Shi, M. Razaviyayn, Z. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, pp. 4331–4340, Sep. 2011.

[11] G. Bresler, D. Cartwright, and D. Tse, "Settling the feasibility of interference alignment for the MIMO interference channel: the symmetric square case," ArXiv pre-print cs.IT/1104.0888. Available: http://arxiv.org/abs/1104.0888.

[12] R. Tresch and M. .Guillaud, "Cellular interference alignment with imperfect channel knowledge," in *Proc. 2009 IEEE Int. Conf. Commun.*.

[13] M. B. Shenouda and T. N. Davidson, "Tomlinson-Harashima precoding for broadcast channels with uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 1380–1389, Sep. 2007.

[14] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Johns Hopkins University Press, 1996.