Rate Balancing for Multiuser MIMO Systems

Imène Ghamnia[1,2], Dirk Slock[2] and Yi Yuan-Wu[1]

[1]Orange Labs, Châtillon, France,
[2]Communication Systems Department, EURECOM, France
{imene.ghamnia, yi.yuan}@orange.com, dirk.slock@eurecom.fr

Abstract—We investigate and solve the rate balancing problem in the downlink for a multiuser Multiple-Input-Multiple-Output (MIMO) system. In particular, we adopt a transceiver structure to maximize the worst-case rate of the user while satisfying a total transmit power constraint. Most of the existing solutions either perform user Mean Squared Error (MSE) balancing or streamwise rate balancing, which is suboptimal in the MIMO case. The original rate balancing problem in the downlink is complicated due to the coupled structure of the transmit filters. This optimization problem is here solved in an alternating manner by exploiting weighted MSE uplink/downlink duality with proven convergence to a local optimum. Simulation results are provided to validate the proposed algorithm and demonstrate its performance improvement over unweighted MSE balancing.

Index Terms—rate balancing, max-min fairness, MSE duality, tranceiver optimization, multiuser MIMO systems

I. INTRODUCTION

One important criterion in designing wireless networks is ensuring fairness requirements. Fairness is said to be achieved if some performance metric is equally reached by all users of the system, depending on their priority allocations. With respect to applications in communication networks, fairness is closely related to min-max or max-min optimization problems, also referred to as *balancing* problems. Actually, balancing a given metric or a utility function among users implies that the system performances are limited by the weak users. At the optimum, the performance of the latter is brought to be improved [1].

However, most of balancing optimization problems are non-convex and can not be solved directly. Despite that, several works over the litterature have developped optimal solutions. For instance, [2] solved the max-min problem by a sequence of Second Order Cone Programs (SOCP). Also, [3] showed that a semidefinite relaxation is tight for the problem, and the optimal solution can be constructed from the solution to a reformulated semidefinite program. In [4], the authors proposed an algorithm based on fixed-point that alternates between power update and beamformer updates, and the nonlinear Perron-Frobenius theory was applied to prove the convergence of the algorithm.

Another way to solve balancing optimization problems is to convert the problem from the downlink (DL) channel to its equivalent uplink (UL) channel, by exploiting the UL/DL *duality*. Doing so, the transformed problem has better mathematical structure and convexity in the UL, thus, the computational complexity of the original problem can be reduced [5]. The UL/DL duality has been widely used to design optimal transmit and receive filters that ensure fairness requirements
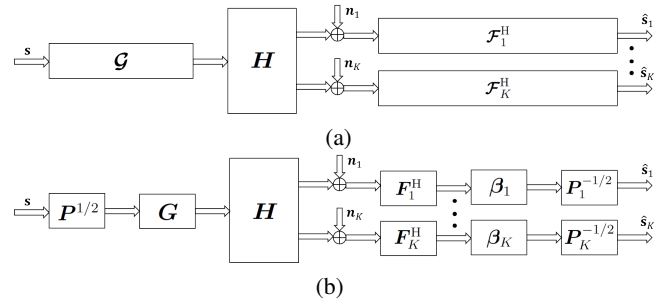

Fig. 1. System model: (a) DL channel, (b) equivalent DL channel.

w.r.t. the Signal-to-Interference-plus-Noise Ratio (SINR), the Mean Square Error (MSE), and the user or stream rate.

With the objective being to equalize all user SINRs, the SINR balancing problem is of particular interest because it is directly related to common performance measures like system capacity and bit error rates. Maximizing the minimum user SINR in the UL can be done straightforwardly since the beamformers can be optimized individually and SINRs are only coupled by the users' transmit powers. In contrast, DL optimization is generally a nontrivial task because the user SINRs depend on all optimization variables and have to be optimized jointly [5]–[10].

Another well-known duality is the stream-wise MSE duality where it has been shown that the same MSE values are achievable in the DL and the UL with the same transmit power constraint. This MSE duality has been exploited to solve various minimum MSE (MMSE) based optimization problems [11]–[13].

In this work, we focus on user rate balancing in a way to maximize the minimum (weighted) rate among all the users in the cell, in order to achieve cell-wide fairness. This balancing problem is studied in [14]. However, the authors do not provide an explicit precoder design. Here we provide a solution via the relation between user rate (summed over its streams) and a weighted sum MSE. But also another ingredient is required: the exploitation of a scale factor that can be freely chosen in the weights for the weighted rate balancing. User-wise rate balancing outperforms user-wise MSE balancing or streamwise rate (or MSE or SINR) balancing when the streams of any MIMO user are quite unbalanced.

II. SYSTEM MODEL

The considered network is a multiuser MIMO DL system, (see Figure 1). We focus on a Base Station (BS) of M transmit antennas serving K users of each N_k antennas, ($k = 1, ..., K$ is the users' index). The channel between the kth user and the

BS is denoted by $\boldsymbol{H}_k^{\mathrm{H}} \in \mathbb{C}^{M \times N_k}$, and $\boldsymbol{H}^{\mathrm{H}} = [\boldsymbol{H}_1^{\mathrm{H}}, ..., \boldsymbol{H}_K^{\mathrm{H}}]$ is the overall channel matrix.

We assume zero-mean white Gaussian noise $\boldsymbol{n}_k \in \mathbb{C}^{N_k \times 1}$ with distribution $\mathcal{CN}(0, \sigma_n^2 \boldsymbol{I})$ at the kth user. We assume independent unity-power transmit symbols $\mathbf{s} = [\mathbf{s}_1^{\mathrm{T}} \ldots \mathbf{s}_K^{\mathrm{T}}]^{\mathrm{T}}$, i.e., $\mathbb{E}[\mathbf{ss}^{\mathrm{H}}] = \boldsymbol{I}$, where $\mathbf{s}_k \in \mathbb{C}^{d_k \times 1}$ is the data vector to be transmitted to the kth user, with d_k being the number of streams allowed by user k. The latter are transmitted using the transmit filtering matrix $\boldsymbol{\mathcal{G}} = \boldsymbol{G} \boldsymbol{P}^{1/2} \in \mathbb{C}^{M \times N_d}$, composed of the beamforming matrix $\boldsymbol{G} = [\boldsymbol{G}_1 \ldots \boldsymbol{G}_K] = [\boldsymbol{g}_1 \ldots \boldsymbol{g}_{N_d}]$ with normalized columns $\|\boldsymbol{g}_i\|_2 = 1$ and the diagonal non-negative DL power allocation $\boldsymbol{P}^{1/2} = \mathrm{blkdiag}\{\boldsymbol{P}_1^{1/2}, \ldots, \boldsymbol{P}_K^{1/2}\}$ where $\mathrm{diag}(\boldsymbol{P}_k) \in \mathbb{R}_+^{d_k \times 1}$ contains the transmission powers and $N_d = \sum_{k=1}^K d_k$ is the total number of streams. The total transmit power is limitted, i.e., $\mathrm{tr}(\boldsymbol{P}) \leq P_{\max}$.

Similarly, the receive filtering matrix for each user is defined as $\boldsymbol{\mathcal{F}}_k^{\mathrm{H}} = \boldsymbol{P}_k^{-1/2} \boldsymbol{\beta}_k \boldsymbol{F}_k^{\mathrm{H}} \in \mathbb{C}^{d_k \times N_k}$, composed of beamforming matrix $\boldsymbol{F}_k^{\mathrm{H}} \in \mathbb{C}^{d_k \times N_k}$ and the diagonal matrices $\boldsymbol{\beta}_k$ contain scaling factors which ensure that the columns of $\boldsymbol{F}_k^{\mathrm{H}}$ have unit norm. We define $\boldsymbol{\beta} = \mathrm{blkdiag}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_K\} = \mathrm{diag}\{[\beta_1 \ldots \beta_{N_d}]\}$ and $\boldsymbol{F} = \mathrm{blkdiag}\{\boldsymbol{F}_1, \ldots, \boldsymbol{F}_K\} = [\boldsymbol{f}_1 \ldots \boldsymbol{f}_{N_d}]$ with normalized per-stream receivers, i.e., $\|\boldsymbol{f}_i\|_2 = 1$.

The MSE per stream $\varepsilon_i^{\mathrm{DL}}$ between the decision variable $\hat{s}_i$ and the transmit data symbol s_i is defined as follows

$$\varepsilon_i^{\mathrm{DL}} = \mathbb{E}\{|\hat{s}_i - s_i|^2\} = \beta_i^2/p_i \boldsymbol{f}_i^{\mathrm{H}} \boldsymbol{H} \Big(\sum_{j=1}^{N_d} p_j \boldsymbol{g}_j \boldsymbol{g}_j^{\mathrm{H}}\Big) \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_i$$
$$- 2\beta_i \mathrm{Re}\{\boldsymbol{f}_i^{\mathrm{H}} \boldsymbol{H} \boldsymbol{g}_i\} + \sigma_n^2 \beta_i^2/p_i + 1, \forall i \in \{1, ..., N_d\}. \quad (1)$$

III. PROBLEM FORMULATION

In this work, we aim to solve the weighted user-rate max-min optimization problem under a total transmit power constraint, i.e., the user rate balancing problem expressed as follows

$$\max_{\{\boldsymbol{G},\boldsymbol{P},\boldsymbol{F},\boldsymbol{\beta}\}} \min_k \ r_k/r_k^{\circ}$$
$$\mathrm{s.t.} \ \mathrm{tr}(\boldsymbol{P}) \leq P_{\max} \quad (2)$$

where r_k is the kth user-rate

$$r_k = \ln \det\Big(\boldsymbol{I} + \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_k \boldsymbol{\mathcal{G}}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \big(\sigma_n^2 \boldsymbol{I} + \sum_{j \neq k} \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_j \boldsymbol{\mathcal{G}}_j^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}}\big)^{-1}\Big) \quad (3)$$

and r_k° is the rate scaling factor for user k. However, the problem presented in (2) is complex and can not be solved directly.

Lemma 1. *The rate of user k in (3) can also be represented as*

$$r_k = \max_{\boldsymbol{W}_k, \boldsymbol{\mathcal{F}}_k} \big[\ln \det(\boldsymbol{W}_k) - \mathrm{tr}(\boldsymbol{W}_k \mathbf{E}_k^{\mathrm{DL}}) + d_k\big]. \quad (4)$$

where

$$\mathbf{E}_k^{\mathrm{DL}} = \mathbb{E}\Big[(\hat{\mathbf{s}}_k - \mathbf{s}_k)(\hat{\mathbf{s}}_k - \mathbf{s}_k)^{\mathrm{H}}\Big]$$
$$= (\boldsymbol{I} - \boldsymbol{\mathcal{F}}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_k)(\boldsymbol{I} - \boldsymbol{\mathcal{F}}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_k)^{\mathrm{H}}$$
$$+ \sum_{j \neq k} \boldsymbol{\mathcal{F}}_k^{\mathrm{H}} \boldsymbol{H}_j \boldsymbol{\mathcal{G}}_j \boldsymbol{\mathcal{G}}_j^{\mathrm{H}} \boldsymbol{H}_j^{\mathrm{H}} \boldsymbol{\mathcal{F}}_k + \sigma_n^2 \boldsymbol{\mathcal{F}}_k^{\mathrm{H}} \boldsymbol{\mathcal{F}}_k \quad (5)$$

is the kth-user DL MSE matrix between the decision variable $\hat{\mathbf{s}}_k$ and the transmit signal $\mathbf{s}_k$, and $\boldsymbol{W} = \{\boldsymbol{W}_k\}_{1 \leq k \leq K}$ are auxiliary weight matrix variables with optimal solution $\boldsymbol{W}_k =$
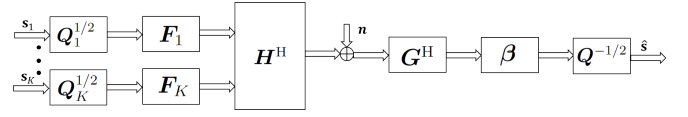

Fig. 2. Dual UL channel.

$(\mathbf{E}_k^{\mathrm{DL}})^{-1}$ *and* $\boldsymbol{\mathcal{F}}_k = (\sigma_n^2 \boldsymbol{I} + \sum_{j=1}^K \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_j \boldsymbol{\mathcal{G}}_j^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}})^{-1} \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_k$, [15].

Now considering both (2) and (4), and introducing $t = \min_k r_k/r_k^{\circ}$, we have $\forall k$

$$r_k/(t\, r_k^{\circ}) \geq 1 \text{ or } r_k/r_k^{\circ} \geq t$$
$$\overset{(a)}{\iff} \ln \det(\boldsymbol{W}_k) + d_k - \mathrm{tr}(\boldsymbol{W}_k \mathbf{E}_k^{\mathrm{DL}}) \geq tr_k^{\circ} \quad (6)$$
$$\iff \frac{\mathrm{tr}(\boldsymbol{W}_k \mathbf{E}_k^{\mathrm{DL}})}{\ln \det(\boldsymbol{W}_k) + d_k - tr_k^{\circ}} \overset{(b)}{=} \frac{\epsilon_{w,k}^{\mathrm{DL}}}{\xi_k} \leq 1$$

where (a) follows from (4) (with optimal $\boldsymbol{W}_k$) and (b) from $\epsilon_{w,k}^{\mathrm{DL}} = \mathrm{tr}(\boldsymbol{W}_k \mathbf{E}_k^{\mathrm{DL}})$, the matrix-weighted MSE (WMSE), and $\xi_k = \ln \det(\boldsymbol{W}_k) + d_k - r_k^{\triangle}$ the WMSE requirement, with $r_k^{\triangle} = tr_k^{\circ}$ the individual rate target, i.e. $r_k \geq r_k^{\triangle}$. What we exploit here is a scale factor t that can be chosen freely in the rate weights r_k^{o} in (2), to transform the rate weights r_k° into target rates $r_k^{\triangle} = tr_k^{\circ}$, which at the same time allows to interpret the WMSE weights ξ_k as target WMSE values.

Doing so, the initial rate balancing optimization problem (2) can be transformed into a matrix-weighted MSE balancing problem expressed as follows

$$\min_{\{\boldsymbol{G},\boldsymbol{P},\boldsymbol{F},\boldsymbol{\beta}\}} \max_k \ \epsilon_{w,k}^{\mathrm{DL}}/\xi_k$$
$$\mathrm{s.t.} \ \mathrm{tr}(\boldsymbol{P}) \leq P_{\max}, \quad (7)$$

which needs to be complemented with an outer loop in which $\boldsymbol{W}_k = (\mathbf{E}_k^{\mathrm{DL}})^{-1}$, $t = \min_k r_k/r_k^{\circ}$, $r_k^{\triangle} = tr_k^{\circ}$ and $\xi_k = d_k + r_k - r_k^{\triangle}$ get updated.

The problem in (7) is still difficult to be handled directly. In the next sections, we solve the problem via UL and DL MSE duality. To this aim, we model an equivalent UL-DL channel plus transceivers pair by separating the filters into two parts: a matrix with unity-norm columns and a scaling matrix [16]. Then, the UL and DL are proved to share the same MSE by switching the role of the normalized filters in the UL and DL. Doing so, an algorithmic solution can be derived for the optimization problem (7).

IV. DUAL UL CHANNEL

In the equivalent UL model represented in Figure 2, we switch between the role of the normalized transmit and receive filters. In fact, $\boldsymbol{F}_k \boldsymbol{Q}_k^{1/2}$ is the kth transmit filter and $\boldsymbol{Q}^{-1/2} \boldsymbol{\beta} \boldsymbol{G}^{\mathrm{H}}$ is a multiuser receive filter, where $\boldsymbol{Q} = \mathrm{blkdiag}\{\boldsymbol{Q}_1, ..., \boldsymbol{Q}_K\}$ with $\mathrm{diag}(\boldsymbol{Q}_k) \in \mathbb{R}_+^{d_k \times 1}$ being the UL power allocation.

Although the quantities $\boldsymbol{H}, \boldsymbol{G}, \boldsymbol{F}$ and $\boldsymbol{\beta}$ are the same, the UL power allocation $\boldsymbol{q} = [q_1 \ldots q_{N_d}]^{\mathrm{T}} = \mathrm{diag}(\boldsymbol{Q})$ may differ from the DL allocation $\boldsymbol{p} = [p_1 \ldots p_{N_d}]^{\mathrm{T}} = \mathrm{diag}(\boldsymbol{P})$, both verifying the same sum power constraint $\|\boldsymbol{p}\|_1 = \|\boldsymbol{q}\|_1 \leq P_{\max}$.

The corresponding UL per stream MSE $\varepsilon_i^{\mathrm{UL}}$ is given by

$$\varepsilon_i^{\mathrm{UL}} = \beta_i^2/q_i \boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \big(\sum_{j=1}^{N_d} q_j \boldsymbol{f}_j \boldsymbol{f}_j^{\mathrm{H}}\big) \boldsymbol{H} \boldsymbol{g}_i - 2\beta_i \mathrm{Re}\{\boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_i\}$$
$$+ \sigma_n^2 \beta_i^2/q_i + 1, \forall i. \tag{8}$$

V. MSE DUALITY

With the equivalent DL channel and its dual UL, it has been shown that the same per stream MSE values are achieved in both links, i.e., $\boldsymbol{\varepsilon}^{\mathrm{UL/DL}} = \mathrm{diag}\{[\varepsilon_1^{\mathrm{UL/DL}} \ldots \varepsilon_{N_d}^{\mathrm{UL/DL}}]\} = \mathrm{diag}\{[\varepsilon_1 \ldots \varepsilon_{N_d}]\} = \boldsymbol{\varepsilon}$ [16].

The UL and DL power allocation, obtained by solving the MSE expressions as in (8) for UL w.r.t. the powers, are given by

$$\boldsymbol{q} = \sigma_n^2 (\boldsymbol{\varepsilon} - \boldsymbol{D} - \boldsymbol{\beta}^2 \boldsymbol{\Psi})^{-1} \boldsymbol{\beta}^2 \mathbf{1}_{N_d} \tag{9}$$

and

$$\boldsymbol{p} = \sigma_n^2 (\boldsymbol{\varepsilon} - \boldsymbol{D} - \boldsymbol{\beta}^2 \boldsymbol{\Psi}^T)^{-1} \boldsymbol{\beta}^2 \mathbf{1}_{N_d} \tag{10}$$

respectively, where the diagonal matrix $\boldsymbol{D}$ is defined as

$$[\boldsymbol{D}]_{ii} = \beta_i^2 \boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_i \boldsymbol{f}_i^{\mathrm{H}} \boldsymbol{H} \boldsymbol{g}_i - 2\beta_i \mathrm{Re}\{\boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_i\} + 1$$

and

$$[\boldsymbol{\Psi}]_{ij} = \begin{cases} \boldsymbol{g}_i^{\mathrm{H}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{f}_j \boldsymbol{f}_j^{\mathrm{H}} \boldsymbol{H} \boldsymbol{g}_i, & i \neq j \\ 0, & i = j. \end{cases}$$

In fact, the MSE duality allows to optimize the transceiver design by switching between the virtual UL and actual DL channels. The optimal receive filtering matrices in both UL and DL are MMSE filters and given by

$$\boldsymbol{G}_k \beta_k \boldsymbol{Q}_k^{-1/2} = \big(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{F} \boldsymbol{Q} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{H} + \sigma_n^2 \boldsymbol{I}\big)^{-1} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{F}_k \boldsymbol{Q}_k^{1/2} \tag{11}$$

and

$$\boldsymbol{F}_k \beta_k \boldsymbol{P}_k^{-1/2} = \big(\boldsymbol{H}_k \boldsymbol{G} \boldsymbol{P} \boldsymbol{G}^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} + \sigma_n^2 \boldsymbol{I}\big)^{-1} \boldsymbol{H}_k \boldsymbol{G}_k \boldsymbol{P}_k^{1/2}. \tag{12}$$

VI. THE MATRIX WEIGHTED USER-MSE OPTIMIZATION

In this section, the problem (7) with respect to the matrix weighted user-MSE is studied. First, we start by the UL power allocation strategies. Then, the joint optimization will follow given the MSE duality. In fact, the MSE duality opens up a way to obtain optimal MMSE receiver designs in (11) and (12). The DL matrix weighted user-MSE optimization problems can be solved by optimizing the weighted MSE values of the dual UL system. The latter can be formulated as

$$\min_{\{\boldsymbol{G},\boldsymbol{F},\boldsymbol{W}\}} \max_k \epsilon_{w,k}^{\mathrm{UL}}/\xi_k$$
$$\text{s.t. } \mathrm{tr}(\boldsymbol{Q}) \leq P_{\max} \tag{13}$$

where $\epsilon_{w,k}^{\mathrm{UL}} = \mathrm{tr}(\boldsymbol{W}_k \boldsymbol{E}_k^{\mathrm{UL}})$, and

$$\boldsymbol{E}_k^{\mathrm{UL}} = (\boldsymbol{I} - \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{F}_k \boldsymbol{Q}_k^{1/2})$$
$$\times (\boldsymbol{I} - \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{F}_k \boldsymbol{Q}_k^{1/2})^{\mathrm{H}}$$
$$+ \sum_{j \neq k} \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_j^{\mathrm{H}} \boldsymbol{F}_j \boldsymbol{Q}_j \boldsymbol{F}_j^{\mathrm{H}} \boldsymbol{H}_j \boldsymbol{G}_k \beta_k \boldsymbol{Q}_k^{-1/2}$$
$$+ \sigma_n^2 \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k \beta_k \boldsymbol{Q}_k^{-1/2}. \tag{14}$$

Then, based on the equivalent UL/DL channel pair, we derive a general framework for joint DL MSE design. First, in the *UL channel*, we find the globally optimal powers $\boldsymbol{Q}$ according to the optimization problem under consideration; then, we update the UL receivers as MMSE filters (11) and we compute the associated per stream MSE values $\varepsilon_i^{\mathrm{UL}}$, $\forall i$. Second, in the *DL*

channel, we find the DL power allocation $\boldsymbol{P}$ which achieves the same UL MSE values; and we update the DL receivers as MMSE filters (12). Finally, we update $\boldsymbol{W}_k$.

The matrix weighted per user MSE can be expressed as follows

$$\epsilon_{w,k}^{\mathrm{UL}} = \mathrm{tr}(\boldsymbol{W}_k \boldsymbol{E}_k^{\mathrm{UL}}) \tag{15}$$
$$= \mathrm{tr}(\boldsymbol{W}_k) + \mathrm{tr}(\boldsymbol{W}_k \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{F}_k \boldsymbol{Q}_k \boldsymbol{F}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{G}_k \beta_k \boldsymbol{Q}_k^{-1/2})$$
$$- 2\mathrm{Re}\big\{\mathrm{tr}(\boldsymbol{Q}_k^{1/2} \boldsymbol{W}_k \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{F}_k)\big\}$$
$$+ \sigma_n^2 \mathrm{tr}(\boldsymbol{W}_k \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k \beta_k \boldsymbol{Q}_k^{-1/2})$$
$$+ \sum_{j \neq k} \mathrm{tr}(\boldsymbol{W}_k \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_j^{\mathrm{H}} \boldsymbol{F}_j \boldsymbol{Q}_j \boldsymbol{F}_j^{\mathrm{H}} \boldsymbol{H}_j \boldsymbol{G}_k \beta_k \boldsymbol{Q}_k^{-1/2}), \forall k.$$

We define $\boldsymbol{Q}_k = \tilde{q}_k \bar{\boldsymbol{Q}}_k$ where $\mathrm{tr}(\bar{\boldsymbol{Q}}_k) = 1$ and $\tilde{q}_k$ is the individual power of the kth user. Then, the transmit covariance matrix $\boldsymbol{R}_k = \boldsymbol{F}_k \boldsymbol{Q}_k \boldsymbol{F}_k^{\mathrm{H}}$ can be written as $\boldsymbol{R}_k = \tilde{q}_k \bar{\boldsymbol{R}}_k$ with $\mathrm{tr}(\bar{\boldsymbol{R}}_k) = 1$. Thus, the matrix weighted MSE $\epsilon_{w,k}$ becomes a function of $\tilde{\boldsymbol{q}} = [\tilde{q}_1, ..., \tilde{q}_K]^T$

$$\epsilon_{w,k}^{\mathrm{UL}} = a_k + \tilde{q}_k^{-1} \sum_{j \neq k} \tilde{q}_j b_{kj} + \tilde{q}_k^{-1} c_k \sigma_n^2, \forall k \tag{16}$$

where

$$a_k = \mathrm{tr}(\boldsymbol{W}_k) + \mathrm{tr}(\boldsymbol{W}_k \bar{\boldsymbol{Q}}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \bar{\boldsymbol{R}}_k \boldsymbol{H}_k \boldsymbol{G}_k \beta_k \bar{\boldsymbol{Q}}_k^{-1/2})$$
$$- 2\mathrm{Re}\big\{\mathrm{tr}(\boldsymbol{Q}_k^{1/2} \boldsymbol{W}_k \boldsymbol{Q}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{F}_k)\big\},$$
$$b_{kj} = \mathrm{tr}(\boldsymbol{W}_k \bar{\boldsymbol{Q}}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{H}_j^{\mathrm{H}} \bar{\boldsymbol{R}}_j \boldsymbol{H}_j \boldsymbol{G}_k \beta_k \bar{\boldsymbol{Q}}_k^{-1/2})$$

and $c_k = \mathrm{tr}(\boldsymbol{W}_k \bar{\boldsymbol{Q}}_k^{-1/2} \beta_k \boldsymbol{G}_k^{\mathrm{H}} \boldsymbol{G}_k \beta_k \bar{\boldsymbol{Q}}_k^{-1/2})$.

Actually, problem (13) always has a global minimizer $\tilde{\boldsymbol{q}}^{\mathrm{opt}}$ characterized by the following equations:

$$\Delta^{\mathrm{UL}} = \frac{\epsilon_{w,k}^{\mathrm{UL}}(\boldsymbol{q}^{\mathrm{opt}})}{\xi_k}, \ \forall k, \tag{17}$$

$$\|\boldsymbol{q}^{\mathrm{opt}}\|_1 = P_{\max} \tag{18}$$

where Δ^{UL} is the minimum balanced matrix-weighted user MSE.

We aim to form an eigensystem by combining (17) and (18). For that, we rewrite (16) as

$$\boldsymbol{\epsilon}_w^{\mathrm{UL}} \tilde{\boldsymbol{q}} = \boldsymbol{A} \tilde{\boldsymbol{q}} + \sigma_n^2 \boldsymbol{C} \mathbf{1}_K \tag{19}$$

where $\boldsymbol{\epsilon}_w^{\mathrm{UL}} = \mathrm{diag}\{\epsilon_{w,1}^{\mathrm{UL}}, \ldots, \epsilon_{w,K}^{\mathrm{UL}}\}$, $\boldsymbol{C} = \mathrm{diag}\{c_1, \ldots, c_K\}$ and

$$[\boldsymbol{A}]_{kj} = \begin{cases} b_{kj}, & k \neq j \\ a_k, & k = j. \end{cases}$$

Now, we define $\boldsymbol{\xi} = \mathrm{diag}\{[\xi_1 \ldots \xi_K]\}$ and multiply both sides by $\boldsymbol{\xi}^{-1}$ to have

$$\boldsymbol{\xi}^{-1} \boldsymbol{\epsilon}_w^{\mathrm{UL}} \tilde{\boldsymbol{q}} = \boldsymbol{\xi}^{-1} \boldsymbol{A} \tilde{\boldsymbol{q}} + \sigma_n^2 \boldsymbol{\xi}^{-1} \boldsymbol{C} \mathbf{1}_K. \tag{20}$$

From (17), we have $\boldsymbol{\xi}^{-1} \boldsymbol{\epsilon}_w^{\mathrm{UL}}(\tilde{\boldsymbol{q}}^{\mathrm{opt}}) = \Delta^{\mathrm{UL}} \boldsymbol{I}$. Thus, (20) becomes

$$\Delta^{\mathrm{UL}} \tilde{\boldsymbol{q}} = \boldsymbol{\xi}^{-1} \boldsymbol{A} \tilde{\boldsymbol{q}} + \sigma_n^2 \boldsymbol{\xi}^{-1} \boldsymbol{C} \mathbf{1}_K. \tag{21}$$

From (18), we can reparameterize $\tilde{\boldsymbol{q}} = \frac{P_{\max}}{\mathbf{1}_K^T \boldsymbol{q}'} \boldsymbol{q}'$ where $\boldsymbol{q}'$ is unconstrained. This allows to rewrite (21) as [7]

$$\boldsymbol{\Lambda} \tilde{\boldsymbol{q}}' = \Delta^{\mathrm{UL}} \tilde{\boldsymbol{q}}', \ \boldsymbol{\Lambda} = \boldsymbol{\xi}^{-1} \boldsymbol{A} + \frac{\sigma_n^2}{P_{\max}} \boldsymbol{\xi}^{-1} \boldsymbol{C} \mathbf{1}_K \mathbf{1}_K^T \tag{22}$$

1. initialize: $\boldsymbol{F}_k^{\mathrm{H}(0,0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})$, $\bar{\boldsymbol{Q}}^{(0,0)} = \frac{P_{\max}}{N_d}\boldsymbol{I}$, $m = n = 0$ and $n_{\max}, m_{\max}$ and fix $r_k^{\circ(0)}$
2. compute UL receive filter $\boldsymbol{G}^{(0,0)}$ and $\boldsymbol{\beta}^{(0,0)}$ with (11)
3. set $\boldsymbol{W}_k^{(0)} = \boldsymbol{I}$ and $\xi_k^{(0)} = d_k$
4. find optimal user power allocation $\tilde{\boldsymbol{q}}^{(0,0)}$ by solving (22) and compute $\boldsymbol{Q}_k^{(0,0)} = \tilde{q}_k^{(0,0)}\bar{\boldsymbol{Q}}_k^{(0,0)}$
5. **repeat**

    5.1 **repeat**
        $n \leftarrow n + 1$
        UL channel:
- update $\boldsymbol{G}^{(n,m-1)}$ and $\boldsymbol{\beta}^{(tmp,tmp)}$ with (11)
- compute the MSE values $\boldsymbol{\varepsilon}^{\mathrm{UL},(n)}$ with (8)
        DL channel:
- compute $\boldsymbol{P}^{(n,m-1)}$ with (10)
- update $\boldsymbol{F}^{(n,m-1)}$ and $\boldsymbol{\beta}^{(tmp,tmp)}$ with (12)
- compute the MSE values $\boldsymbol{\varepsilon}^{\mathrm{DL}(n)}$ with (1)
        UL channel:
- compute $\boldsymbol{Q}^{(tmp,tmp)}$ with (9) and $\bar{\boldsymbol{Q}}_k^{(n,m-1)} = \boldsymbol{Q}_k^{(tmp,tmp)}/\mathrm{tr}(\boldsymbol{Q}_k^{(tmp,tmp)})$
- find optimal user power allocation $\tilde{\boldsymbol{q}}^{(n,m-1)}$ by solving (22) and compute $\boldsymbol{Q}_k^{(n,m-1)} = \tilde{q}_k^{(n,m-1)}\bar{\boldsymbol{Q}}_k^{(n,m-1)}$
    5.2 **until** required accuracy is reached or $n \geq n_{\max}$
    5.3 $m \leftarrow m + 1$
    5.4 update $\boldsymbol{W}_k^{(m)} = (\mathbf{E}_k^{\mathrm{UL}(m)})^{-1}$, $r_k^{(m)} = \ln\det(\boldsymbol{W}_k^{(m)})$, $t = \min_k \frac{r_k^{(m)}}{r_k^{\circ(m-1)}}$, $r_k^{\circ(m)} = t\,r_k^{\circ(m-1)}$, and $\xi_k^{(m)} = d_k + r_k^{(m)} - r_k^{\circ(m)}$
    5.5 do $n \leftarrow 0$ and set $(.)^{(n_{\max},m-1)} \rightarrow (.)^{(0,m)}$ in order to re-enter the inner loop
6. **until** required accuracy is reached or $m \geq m_{\max}$

It can be observed that Δ^{UL} is an eigenvalue of the non-negative extended coupling matrix $\boldsymbol{\Lambda}$. However, not all eigenvalues represent physically meaningful values. In particular, $\tilde{\boldsymbol{q}}^{\mathrm{opt}} > 0$ and $\Delta^{\mathrm{UL}} > 0$ must be fulfilled.

It is known that for any non-negative irreducible real matrix $\boldsymbol{X}$ with spectral radius $\rho(\boldsymbol{X})$, there exists a unique vector $\boldsymbol{q} > 0$ and $\lambda_{\max}(\boldsymbol{X}) = \rho(\boldsymbol{X})$ such that $\boldsymbol{X}\boldsymbol{q} = \lambda_{\max}(\boldsymbol{X})\boldsymbol{q}$. The uniqueness of $\lambda_{\max}(\boldsymbol{\Lambda})$ also follows from immediately from the function $\Delta^{\mathrm{UL}}(P_{\max})$ being strictly monotonically decreasing in $P_{\max}$. This rules out the existence of two different balanced levels with the same sum power. Hence, the balanced level is given by

$$\Delta^{\mathrm{UL,\ opt}} = \lambda_{\max}(\boldsymbol{\Lambda}). \tag{23}$$

Therefore, the optimal power allocation $\tilde{\boldsymbol{q}}'$ is the principal eigenvector of the matrix $\boldsymbol{\Lambda}$ in (22). As noted in [5], we have in fact

$$\lambda_{\max}(\boldsymbol{\Lambda}) = \min_{\tilde{\boldsymbol{p}}}\max_{\tilde{\boldsymbol{q}}} \frac{\tilde{\boldsymbol{p}}^H \boldsymbol{\Lambda}\tilde{\boldsymbol{q}}}{\tilde{\boldsymbol{p}}^H\tilde{\boldsymbol{q}}} = \max_{\tilde{\boldsymbol{p}}}\min_{\tilde{\boldsymbol{q}}} \frac{\tilde{\boldsymbol{p}}^H \boldsymbol{\Lambda}\tilde{\boldsymbol{q}}}{\tilde{\boldsymbol{p}}^H\tilde{\boldsymbol{q}}} \tag{24}$$

where in [5] $\tilde{\boldsymbol{p}}$ was said to have no particular meaning but actually can be shown to relate to the DL powers. So, the proposed algorithm provides in the inner loop an alternating optimization of (24) w.r.t. $\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{q}}$, $\boldsymbol{F}$, $\boldsymbol{G}$ [5], [16]. If we take for $\tilde{\boldsymbol{p}}$ the K standard basis vectors, then we get

$$\lambda_{\max}(\boldsymbol{\Lambda}) = \min_{\tilde{\boldsymbol{q}}}\max_k \frac{(\boldsymbol{\Lambda}\tilde{\boldsymbol{q}})_k}{\tilde{\boldsymbol{q}}_k} \tag{25}$$

which from (17), (20), (22) can be seen to be exactly the WMSE balancing problem we want to solve.

VII. ALGORITHMIC SOLUTION AND SIMULATIONS

A. Algorithm

The proposed optimization framework is summarized hereafter in Table I. Superscripts $(.)^{(n)}$ and $(.)^{(tmp)}$ denote the nth iteration and a temporary value, respectively. This algorithm is based on a double loop. The inner loop solves the WMSE balancing problem in (7) whereas the outer loop iteratively transforms the WMSE balancing problem into the original rate balancing problem in (2).

B. Proof of Convergence

In case the rate weights r_k° would not satisfy $r_k \geq r_k^\circ$, this issue will be rectified by the scale factor t after one iteration (of the outer loop). It can be shown that $t = \min_k \frac{r_k^{(m)}}{r_k^{\circ(m-1)}} \geq 1$. By contradiction, if this was not the case, it can be shown to lead to $\frac{\mathrm{tr}(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)})}{\xi_k^{(m-1)}} > 1$, $\forall k$ and hence $\Delta^{(m)} > 1$. But we have

$$\Delta^{(m)} = \frac{\mathrm{tr}(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)})}{\xi_k^{(m-1)}}, \forall k, = \max_k \frac{\mathrm{tr}(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)})}{\xi_k^{(m-1)}}$$
$$\overset{(a)}{<} \max_k \frac{\mathrm{tr}(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m-1)})}{\xi_k^{(m-1)}} = \max_k \frac{d_k}{\xi_k^{(m-1)}} \overset{(b)}{<} 1. \tag{26}$$

Let $\mathbf{E} = \{\mathbf{E}_k, k = 1, ..., K\}$ and $f^{(m)}(\mathbf{E}) = \max_k \frac{\mathrm{tr}(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k)}{\xi_k^{(m-1)}}$. Then (a) is due to the fact that the algorithm in fact performs alternating minimization of $f^{(m)}(\mathbf{E})$ w.r.t. $\boldsymbol{G}$, $\boldsymbol{F}$, $\tilde{\boldsymbol{q}}$ and hence will lead to $f^{(m)}(\mathbf{E}^{(m)}) < f^{(m)}(\mathbf{E}^{(m-1)})$. On the other hand, (b) is due to $\xi_k^{(m-1)} = d_k + r_k^{(m-1)} - r_k^{\circ(m-1)} > d_k$, for $m \geq 3$.

Hence, $t \geq 1$. Of course, during the convergence $t > 1$. The increasing rate targets $r_k^{\circ(m)}$ constantly catch up with the increasing rates $r_k^{(m)}$. Now, the rates are upper bounded by the single user MIMO rates (using all power), and hence the rates will converge and the sequence t will converge to 1. That means that for at least one user k, $r_k^{(\infty)} = r_k^{\circ(\infty)}$. The question is whether this will be the case for all users, as is required for rate balancing. Now, the WMSE balancing leads at every outer iteration m to $\frac{\mathrm{tr}(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)})}{\xi_k^{(m-1)}} = \Delta^{(m)}, \forall k$. At convergence, this becomes $\frac{d_k}{\xi_k^{(\infty)}} = \Delta^{(\infty)}$ where $\xi_k^{(\infty)} = d_k + r_k^{(\infty)} - r_k^{\circ(\infty)}$. Hence, if we have convergence because for one user k_∞ we arrive at $r_{k_\infty}^{(\infty)} = r_{k_\infty}^{\circ(\infty)}$, then this implies $\Delta^{(\infty)} = 1$ which implies $r_k^{(\infty)} = r_k^{\circ(\infty)}, \forall k$. Hence, the rates will be maximized and balanced.

Remark 1. *In fact, the algorithm also converges with* $n_{\max} = 1$, *i.e., with only a single loop.*

C. Simulation results

In this section, we numerically illustrate the performance of the proposed algorithm. The simulations are obtained under a channel modeled as follows : $\boldsymbol{H}_k^{\mathrm{H}} = \boldsymbol{B}_k\boldsymbol{\mathcal{U}}_k\boldsymbol{A}_k$ where $\boldsymbol{B}_k, \boldsymbol{A}_k$ are of dimensions $(M \times N_k)$ and $(N_k \times N_k)$ respectively, and have i.i.d. elements distributed as $\mathcal{CN}(0,1)$; $\boldsymbol{\mathcal{U}}_k = \mu\boldsymbol{U}_k$, with the normalization parameter $\mu = (\mathrm{trace}(\boldsymbol{U}_k))^{-1/2}$ and $\boldsymbol{U}_k = \mathrm{diag}\{1, \alpha, \alpha^2, \ldots, \alpha^{N_k-1}\}$ ($\alpha \in \mathbb{R}$ being a scalar parameter). This model allows to control the rank profile of the MIMO
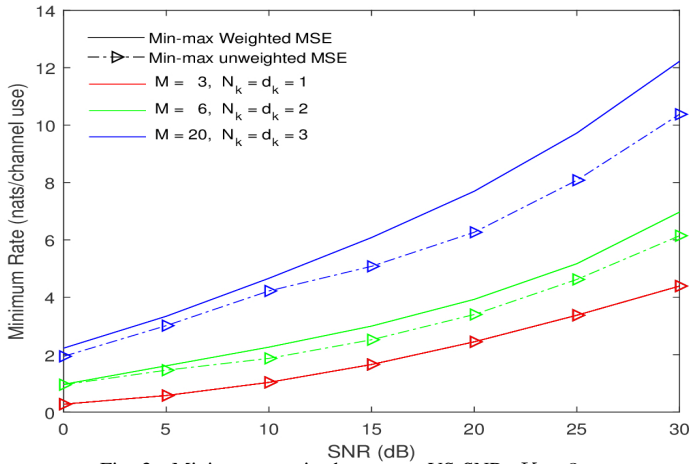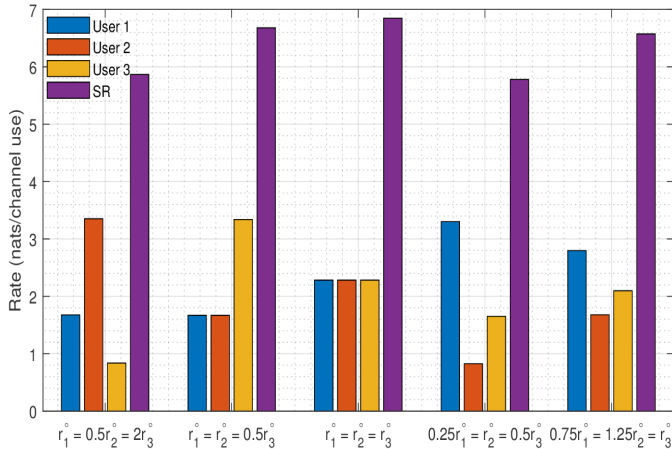
Fig. 3. Minimum rate in the system VS SNR: $K = 3$.

Fig. 4. Rate distribution among users: $K = 3$, SNR= 10 dB, $M = 6, N_k = d_k = 2$.

channels. For all simulations, we fix $\alpha = 0.3$ and take 1000 channel realisations and $n_{\max} = 20$. The algorithm converges after 4-5 (or 13-15) iterations of m at SNR = $\frac{P_{\max}}{\sigma_n^2}$ =10dB (or 30dB).

Figure 3 plots the minimum achieved per user rate using *i)* our max-min user rate approach with equal user priorities and *ii)* the user MSE balancing approach [16], as a function of the Signal to Noise Ratio (SNR). We observe that our approach outperforms significantly the unweighted MSE balancing optimization, and the gap gets larger with more streams. Note that we observe the same behavior with the classical i.i.d. channel $\boldsymbol{H}_k^{\mathrm{H}} = \boldsymbol{B}_k$, but with a smaller gap (e.g., for 15dB, $\frac{\min_k r_k (\text{weighted-MSE})}{\min_k r_k (\text{unweighted-MSE})}$ = 1.05 instead of 1.18 with $M = 6, N_k = d_k = 2$ in Figure 3).

In Figure 4, we illustrate how rate is distributed among users according to their priorities represented by the rate targets r_k°. We can see that, using the min-max weighted MSE approach, the rate is equally distributed between the users with equal user priorities, i.e., $r_k^\circ = r_1^\circ \ \forall k$, whereas with different user priorities, the rate differs from one user to another accordingly. Furthermore, the Sum Rate (SR) reaches its maximum when user priorities are equal, as the channel statistics are identical for each user.

VIII. CONCLUSIONS

In this work, we addressed the multiple streams per user case (MIMO links) for which we considered user rate balancing, not stream rate balancing. Actually, we optimized the rate distribution over the streams of a user, within the rate balancing of the users. In this regard, we proposed an iterative algorithm to balance the rate between the users in a MIMO system. The latter was derived by transforming the max-min rate optimization problem into a min-max weighted MSE optimization problem to enable MSE duality. We also provided numerical comparisons between the proposed weighted rate balancing approach and unweighted MSE balancing.

REFERENCES

[1] L. Zheng, D. W. H. Cai, and C. W. Tan, "Max-min Fairness Rate Control in Wireless Networks: Optimality and Algorithms by Perron-Frobenius Theory," *IEEE Trans. Mobile Computing*, Jan 2018.

[2] 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.

[3] M. Bengtsson and B. Ottersten, "Optimal and suboptimal Transmit Beamforming," *Handbk Antennas in Wireless Communications*, 2001.

[4] D. Cai, T. Quek, and C. W. Tan, "A Unified Analysis of Max-min Weighted SINR for MIMO Downlink System," *IEEE Trans. on Signal Processing*, Aug 2011.

[5] M. Schubert and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem with Individual SINR Constraints," *IEEE Trans. on Vehicular Tech.*, Jun. 2004.

[6] G. Montalbano and D. Slock, "Matched Filter Bound Optimization For Multiuser Downlink Transmit Beamforming," *IEEE Int'l Conf. Universal Personal Communications (ICUPC)*, Oct. 1998.

[7] F. Negro, M. Cardone, I. Ghauri, and D. Slock, "SINR Balancing and Beamforming for the MISO Interference Channel," *IEEE Int'l Symp. on Personal, Indoor and Mobile Radio Comm's (PIMRC)*, Sep. 2011.

[8] W. Yu and T. Lan, "Transmitter Optimization for the Multi-antenna Downlink with Per-antenna Power Constraints," *IEEE Trans. Signal Processing*, June 2007.

[9] L. Zhang, R. Zhang, Y. C. Liang, Y. Xin, and H. V. Poor, "On Gaussian MIMO BC-MAC Duality with Multiple Transmit Covariance Constraints," *IEEE Trans. Inform. Theory*, vol. 58, no. 4, Apr 2012.

[10] K. Cumanan, L. Musavian, S. Lambotharan, and A. B. Gershman, "SINR Balancing Technique for Downlink Beamforming in Cognitive Radio Networks," *IEEE Signal Process. Lett.*, Feb 2010.

[11] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: Duality and sum-MSE Minimization," *IEEE Trans. Signal Process*, Nov. 2007.

[12] ——, "Capacity Balancing for Multiuser MIMO Systems," *Proc. IEEE ICASSP*, Apr. 2007.

[13] R. Hunger, M. Joham, and W. Utschick, "On the MSE-Duality of the Broadcast Channel and the Multiple Access Channel," *IEEE Trans. Signal Processing*, Feb. 2009.

[14] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "Linear Transceiver Design for a MIMO Interfering Broadcast Channel Achieving Max-min Fairness," *Proc. Asilomar Conf. Signals, Systems and Computers*, Nov. 2011.

[15] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An Iteratively Weighted MMSE Approach to Distributed Sum-Utility Maximization for a MIMO Interfering Broadcast Channel," *IEEE Trans. Signal Processing*, Sept. 2011.

[16] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: MMSE Balancing," *IEEE Trans. Signal Processing*, vol. 56, no. 8, pp. 606–619, Aug 2008.