Multi-User MIMO Max-Min User Rate via Weighted MSE Balancing

Imène Ghamnia[1,2], Dirk Slock[1] and Yi Yuan-Wu[2]
[1]Communication Systems Department, EURECOM, France
[2]Orange Labs, Châtillon, France,
{imene.ghamnia, yi.yuan}@orange.com, dirk.slock@eurecom.fr

***Abstract*—This paper investigates the problem of user rate balancing with power constraints for the downlink transmission of multiuser Multiple-Input-Multiple-Output (MIMO) system. However, this max-min rate problem is not jointly convex. Therefore, we transform the original problem into matrix-weighted user Mean Squared Error (MSE) balancing. In particular, we enhance user fairness by solving the min-max matrix-weighted MSE optimization problem by power allocation and choice of transceiver filters in a way to maximize the minimum downlink rate of the users. Formulating the balancing operation as constraints, we introduce the corresponding Lagrangian for which we reach the saddle point by alternating optimization. Various aspects of Perron Frobenius theory are exploited in the process. Simulation results are provided to validate the proposed algorithm and demonstrate its performance improvement over e.g. unweighted MSE balancing.**

***Index Terms*—rate balancing, max-min fairness, Lagrange duality, transceiver optimization, multi-user MIMO**

I. Introduction

One important criterion in designing wireless networks is ensuring fairness requirements. In fact, 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 nonconvex 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 channel to its equivalent uplink channel, by exploiting the uplink-downlink *duality*. Doing so, the transformed problem has better mathematical structure and convexity in the uplink, thus, the computational complexity of the original problem can be reduced [5]. The uplink-downlink duality has been widely used to design optimal transmit and receive filters that ensure faireness requirements w.r.t the Signal-to-Interference-plus-Noise Ratio (SINR), the Mean Square Error (MSE), and the user rate.

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 uplink can be done straightforwardly since the beamformers can be optimized individually and SINRs are only coupled by the users' transmit powers. In contrast, downlink 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 downlink and the uplink 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 per user (weighted) rate in the network. This balancing problem is studied in [14] without providing an explicit precoder design. As in [15], we provide here 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 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/SINR) balancing when the streams of any MIMO user are quite unbalanced. In [15] the problem is transformed into weighted MSE balancing using non-diagonal weight matrices. Here we solve the user rate balancing problem using diagonal weight matrices by diagonalizing the user signal error covariance matrices, which allows to link the per stream and per user power allocation problems. Also, the optimization of the matrix-weighted MSE balancing problem is held in the downlink (DL) to stick to the problem definition, unlike in [15] where the optimization is related to the equivalent uplink (UL) problem. However, we exploit Lagrangian duality to transform WMSE balancing into a weighted sum MSE optimization problem. This also leads to limited use of UL/DL duality with transmit filters

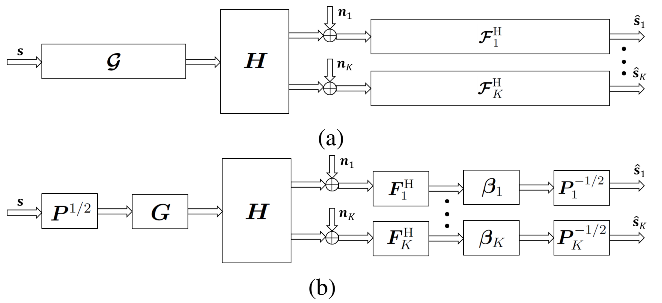

Fig. 1. System model: (a) downlink channel, (b) equivalent downlink channel.

appearing as MMSE receivers in a dual UL, and Lagrange multipliers for the balancing constraints becoming uplink user powers. Actually, UL and DL user powers will be left and right Perron Frobenius eigenvectors of a matrix of weighted interference plus noise to signal power ratios..

II. System Model

The considered network is a multiuser MIMO downlink 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., $\mathbb{E}\big[\mathbf{s}\mathbf{s}^{\mathrm{H}}\big] = \boldsymbol{I}$, 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 downlink 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., $\mathrm{tr}\big(\boldsymbol{P}\big) \leq P_{\max}$.

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

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

III. Rate - Weighted MSE Relation

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

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

where r_k is the kth user-rate

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

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*

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

where

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

is the kth-user downlink 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 $r_k = -\ln\det\Big(\mathbf{E}_k\Big)$ and [16], [17], [18]

$$\boldsymbol{W}_k = \mathbf{E}_k^{-1} = \boldsymbol{I} + \boldsymbol{\mathcal{G}}_k^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}} (\sigma_n^2 \boldsymbol{I} + \sum_{j=1}^{K} \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_j \boldsymbol{\mathcal{G}}_j^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}})^{-1} \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_k, \qquad (6)$$

$$\boldsymbol{\mathcal{F}}_k = (\sigma_n^2 \boldsymbol{I} + \sum_{j=1}^{K} \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_j \boldsymbol{\mathcal{G}}_j^{\mathrm{H}} \boldsymbol{H}_k^{\mathrm{H}})^{-1} \boldsymbol{H}_k \boldsymbol{\mathcal{G}}_k. \qquad (7)$$

Now consider both (2) and (4), and let us introduce $\xi_k = \ln\det\big(\boldsymbol{W}_k\big) + d_k - r_k^{\triangle}$, the WMSE requirement, with target rate $r_k^{\triangle}$. Assume that we shall be able to concoct an optimization algorithm that ensures that at all times and for all users the matrix-weighted MSE (WMSE) satisfies $\epsilon_{w,k} = \mathrm{tr}\big(\boldsymbol{W}_k \mathbf{E}_k\big) \leq d_k$ and $\ln\det\big(\boldsymbol{W}_k\big) \geq r_k^{\triangle}$ or hence $\xi_k \geq d_k$. This leads $\forall k$ to

$$\begin{aligned} &\frac{\epsilon_{w,k}}{\xi_k} \leq 1 \\ \Longleftrightarrow\ & \ln\det\big(\boldsymbol{W}_k\big) + d_k - \mathrm{tr}\big(\boldsymbol{W}_k \mathbf{E}_k\big) \geq r_k^{\triangle} \\ \overset{(a)}{\Longrightarrow}\ & r_k / r_k^{\triangle} \geq 1 \end{aligned} \qquad (8)$$

where (a) follows from (4). To get to (8), what we can exploit in (2) is a scale factor t that can be chosen freely in the rate weights r_k° in (2). We shall take $t = \min_k r_k / r_k^{\circ}$, which allows to transform the rate weights r_k° into target rates $r_k^{\triangle} = t r_k^{\circ}$, and 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

$$\min_{\{\boldsymbol{G}, \boldsymbol{P}, \boldsymbol{F}, \boldsymbol{\beta}\}} \max_k \; \epsilon_{w,k} / \xi_k \quad \text{s.t. } \mathrm{tr}\big(\boldsymbol{P}\big) \leq P_{\max}, \qquad (9)$$

which needs to be complemented with an outer loop in which $\boldsymbol{W}_k = \mathbf{E}_k^{-1}$, $t = \min_k r_k / r_k^{\circ}$, $r_k^{\triangle} = t r_k^{\circ}$ and $\xi_k = d_k + r_k - r_k^{\triangle}$ get updated. The problem in (9) is still difficult to be handled directly.

IV. The Weighted User-MSE Optimization

In this section, the problem (9) with respect to the matrix weighted user-MSE is studied. Consider first the per stream MSE values $\boldsymbol{\varepsilon} = \text{diag}\big\{[\varepsilon_1 \ldots \varepsilon_{N_d}]\big\} = \boldsymbol{\varepsilon}$. The downlink power allocation achieving these MSEs is obtained by solving the MSE expressions (1) w.r.t. the powers,

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

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

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

The link between stream-wise and user-wise MSEs is simplified if $\boldsymbol{W}$ is diagonal. This can be obtained without loss in optimality as a unitary transformation will leave the spatially white transmit signal vectors white [19]. So, introduce the eigendecomposition $\mathbf{E}_k = \boldsymbol{V}_k\boldsymbol{\Sigma}_k\boldsymbol{V}_k^{\mathrm{H}}$. Considering the optimal $\mathbf{E}_k$ in Lemma 1, we can obtain diagonal $\mathbf{E}$ with a transformed transmit filter as

$$\breve{\mathbf{E}}_k = \boldsymbol{\Sigma}_k, \breve{\boldsymbol{\mathcal{G}}}_k = \boldsymbol{\mathcal{G}}_k\boldsymbol{V}_k = \boldsymbol{G}_k\boldsymbol{P}_k^{1/2}\boldsymbol{V}_k = \breve{\boldsymbol{G}}_k\breve{\boldsymbol{P}}_k^{1/2}. \tag{11}$$

Now, as explained in [19], it is not actually required to carry out explictly this diagonalization. Indeed if in (4) the matrices $\boldsymbol{W}_k$ are constrained to be diagonal, then this leads a priori to a minorization, but actually this minorization is tight, since it will lead to BF filters that will adjust themselves to lead to diagonal $\boldsymbol{E}_k$.

With diagonal weight matrices, we can furthermore consider the user weighted MSE problem directly in the DL and optimize the DL power allocation, unlike in [15] where we solve the equivalent UL problem and optimize the respective UL power allocation.

The matrix weighted per user MSE can be expressed as follows, with $l_k = \sum_{j=1}^{k-1} d_j + 1$ (d_j being the number of streams of user j)

$$\epsilon_{w,k} = \mathrm{tr}\big(\boldsymbol{W}_k\boldsymbol{E}_k\big) = \sum_{i=l_k}^{l_k+d_k-1} w_i\varepsilon_i, \forall k. \tag{12}$$

Collecting all layer MSEs in a vector, we get with $\boldsymbol{P} = \text{diag}\{\boldsymbol{p}\}$

$$\boldsymbol{\varepsilon}\mathbf{1}_{N_d} = \boldsymbol{P}^{-1}\Big[(\boldsymbol{D} + \boldsymbol{\beta}^2\boldsymbol{\Psi})\boldsymbol{p} + \sigma_n^2\boldsymbol{\beta}^2\mathbf{1}_{N_d}\Big]. \tag{13}$$

Note that $\boldsymbol{P}^{-1}\boldsymbol{D}\boldsymbol{p} = \boldsymbol{D}\mathbf{1}_{N_d}$. Now introduce the user powers $\tilde{\boldsymbol{p}}$, which relate to the stream powers as $\boldsymbol{P}_k = \bar{\boldsymbol{P}}_k\tilde{\boldsymbol{p}}_k$ with $\mathrm{tr}\{\bar{\boldsymbol{P}}_k\} = 1$ and $\bar{\boldsymbol{P}} = \text{blkdiag}(\bar{\boldsymbol{P}}_1, \ldots, \bar{\boldsymbol{P}}_K)$. Also consider the case of diagonal weighting matrices $\boldsymbol{W}_k$ and the overall diagonal $\boldsymbol{W} = \text{blkdiag}(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K)$. Consider now the weighted user MSE (WMSE) $\epsilon_{w,k}^{=}\mathrm{tr}\{\boldsymbol{W}_k\boldsymbol{E}_k\}$ where the diagonal of (the not necessarily diagonal) $\boldsymbol{E} = \text{blkdiag}(\boldsymbol{E}_1, \ldots, \boldsymbol{E}_K)$ is the set of stream MSEs $\boldsymbol{\varepsilon}$, and let $\boldsymbol{\epsilon}_w = \text{diag}(\epsilon_{w,1}, \ldots, \epsilon_{w,K})$ be the set of user WMSEs. We shall also need the per user stream distribution matrix $\mathbb{1} = \text{blkdiag}(\mathbf{1}_{d_1}, \ldots, \mathbf{1}_{d_K})$. Then we get from (13)

$$\boldsymbol{\epsilon}_w\,\mathbf{1}_K = \mathbb{1}^H\boldsymbol{W}\boldsymbol{\varepsilon} = \mathbb{1}^H\boldsymbol{W}\boldsymbol{P}^{-1}\Big[(\boldsymbol{D} + \boldsymbol{\beta}^2\boldsymbol{\Psi})\boldsymbol{p} + \sigma_n^2\boldsymbol{\beta}^2\mathbf{1}_{N_d}\Big]. \tag{14}$$

Note that $\boldsymbol{P}\mathbf{1}_{N_d} = \boldsymbol{p} = \bar{\boldsymbol{P}}\mathbb{1}\tilde{\boldsymbol{p}}$ and $\mathbf{1}_{N_d} = \mathbb{1}\,\mathbf{1}_K$. By multiplying both sides of (14) with $\text{diag}\{\tilde{\boldsymbol{p}}\}$, we get

$$\begin{aligned}&\boldsymbol{\epsilon}_w\,\tilde{\boldsymbol{p}} = \boldsymbol{A}\,\tilde{\boldsymbol{p}} + \sigma_n^2\,\boldsymbol{C}\,\mathbf{1}_K \quad \text{with}\\ &\boldsymbol{A} = \mathbb{1}^H\boldsymbol{W}(\boldsymbol{D} + \boldsymbol{\beta}^2\bar{\boldsymbol{P}}^{-1}\boldsymbol{\Psi}^{\mathrm{T}}\bar{\boldsymbol{P}})\mathbb{1},\\ &\boldsymbol{C} = \mathbb{1}^H\boldsymbol{W}\boldsymbol{\beta}^2\bar{\boldsymbol{P}}^{-1}\mathbb{1}.\end{aligned} \tag{15}$$

Let $\boldsymbol{\xi} = \text{diag}(\xi_1, \ldots, \xi_K)$, then

$$\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w\,\tilde{\boldsymbol{p}} = \boldsymbol{\xi}^{-1}\boldsymbol{A}\,\tilde{\boldsymbol{p}} + \sigma_n^2\,\boldsymbol{\xi}^{-1}\boldsymbol{C}\,\mathbf{1}_K. \tag{16}$$

On the other hand we have the power constraint $\mathbf{1}_K^H\boldsymbol{p}^{\mathrm{opt}} = P_{\max}$. Reparameterize $\tilde{\boldsymbol{p}} = \frac{P_{\max}}{\mathbf{1}_K^H\tilde{\boldsymbol{p}}'}\tilde{\boldsymbol{p}}'$ where now $\tilde{\boldsymbol{p}}'$ is unconstrained. This allows us to write (16) as (rewriting $\tilde{\boldsymbol{p}}'$ as $\tilde{\boldsymbol{p}}$)

$$\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w\,\tilde{\boldsymbol{p}} = \boldsymbol{\Lambda}\,\tilde{\boldsymbol{p}} \text{ with } \boldsymbol{\Lambda} = \boldsymbol{\xi}^{-1}\boldsymbol{A} + \frac{\sigma_n^2}{P_{\max}}\boldsymbol{\xi}^{-1}\boldsymbol{C}\,\mathbf{1}_K\mathbf{1}_K^H. \tag{17}$$

Now with (17), the WMSE balancing problem of (9) becomes

$$\min_{\tilde{\boldsymbol{p}}}\max_k \frac{\epsilon_{w,k}}{\xi_k} = \min_{\tilde{\boldsymbol{p}}}\max_k \frac{[\boldsymbol{\Lambda}\,\tilde{\boldsymbol{p}}]_k}{\tilde{\boldsymbol{p}}_k} \tag{18}$$

According to the Collatz–Wielandt formula [20, Chapter 8], the above expression corresponds to the Perron-Frobenius (maximal) eigenvalue Δ of $\boldsymbol{\Lambda}$ and the optimal $\tilde{\boldsymbol{p}}$ is the corresponding Perron-Frobenius (right) eigenvector

$$\boldsymbol{\Lambda}\tilde{\boldsymbol{p}} = \Delta\,\tilde{\boldsymbol{p}}. \tag{19}$$

Note that this implies the equality $\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w = \Delta\,\boldsymbol{I}$.

V. Algorithmic Solution via Lagrangian Duality

A. Algorithm

The max-min weighted user rate optimization problem (2) can be reformulated as

$$\begin{aligned}&\min_{t,\boldsymbol{G},\boldsymbol{P}} -t\\ &\text{s.t. } t\,r_k^{\circ} - r_k \leq 0,\ \mathrm{tr}(\boldsymbol{P}) - P_{\max} \leq 0\,.\end{aligned} \tag{20}$$

Introducing Lagrange multipliers to augment the cost function with the constraints leads to the Lagrangian

$$\begin{aligned}&\max_{\lambda',\mu}\min_{t,\boldsymbol{G},\boldsymbol{P}} \mathcal{L}\\ &\mathcal{L} = -t + \sum_k \lambda_k'(t\,r_k^{\circ} - r_k) + \mu(\mathrm{tr}(\boldsymbol{P}) - P_{\max})\end{aligned} \tag{21}$$

Integrating the result (4), we get a modified Lagrangian

$$\begin{aligned}&\max_{\lambda',\mu}\min_{t,\boldsymbol{G},\boldsymbol{P},\boldsymbol{\mathcal{F}},\boldsymbol{W}} \mathcal{L}\\ &\mathcal{L} = -t + \sum_k \lambda_k'(\mathrm{tr}(\boldsymbol{W}_k\mathbf{E}_k) - \xi_k) + \mu(\mathrm{tr}(\boldsymbol{P}) - P_{\max})\end{aligned} \tag{22}$$

Introducing $\lambda_k = \lambda_k'\xi_k$, we can rewrite as

$$\begin{aligned}&\max_{\lambda,\mu}\min_{t,\boldsymbol{G},\boldsymbol{P},\boldsymbol{\mathcal{F}},\boldsymbol{W}} \mathcal{L}\\ &\mathcal{L} = -t + \sum_k \lambda_k(\frac{\mathrm{tr}(\boldsymbol{W}_k\mathbf{E}_k)}{\xi_k} - 1) + \mu(\mathrm{tr}(\boldsymbol{P}) - P_{\max})\end{aligned} \tag{23}$$

We shall solve this saddlepoint condition for $\mathcal{L}$ by alternating optimization. As far as the dependence on $\lambda, \boldsymbol{G}, \boldsymbol{P}, \boldsymbol{\mathcal{F}}$ is concerned, we have (omitting the power constraint)

$$\max_{\lambda}\min_{\boldsymbol{G},\boldsymbol{P},\boldsymbol{\mathcal{F}}} \sum_k \lambda_k\,\frac{\mathrm{tr}(\boldsymbol{W}_k\mathbf{E}_k)}{\xi_k} \tag{24}$$

which is of the form Weighted Sum MSE (WSMSE). Optimizing w.r.t. Rxs $\boldsymbol{\mathcal{F}}$ leads to the MMSE solution mentioned

TABLE I

PSEUDO CODE OF THE PROPOSED ALGORITHM

1. initialize: $\boldsymbol{G}_k^{(0,0)} = (\boldsymbol{I}_{d_k} : \mathbf{0})^{\mathrm{T}}$, $\bar{\boldsymbol{P}}_k^{(0,0)} = \frac{1}{d_k}\boldsymbol{I}_{d_k}$, $\tilde{\boldsymbol{p}}_k^{(0,0)} = \tilde{\boldsymbol{q}}_k^{(0,0)} = \frac{P_{\max}}{K}$, $\boldsymbol{P}_k^{(0,0)} = \tilde{\boldsymbol{p}}_k^{(0,0)}\bar{\boldsymbol{P}}_k^{(0,0)}$, $m = n = 0$ and fix $n_{\max}, m_{\max}$ and $r_k^{\circ(0)}$, initialize $\boldsymbol{W}_k^{(0)} = \boldsymbol{I}_{d_k}$ and $\xi_k^{(0)} = d_k$
2. initialize $\boldsymbol{F}^{(0,0)}$, $\boldsymbol{\beta}^{(0,0)}$ in $\boldsymbol{\mathcal{F}}^{(0,0)} = \boldsymbol{F}\boldsymbol{\beta}^{(0,0)}\boldsymbol{P}^{(0,0)-1/2}$ from (7)
3. **repeat**
   - 3.1. $m \leftarrow m+1$
   - 3.2. **repeat**
     - $n \leftarrow n+1$
     - *i* update $\boldsymbol{G}_k$, $\bar{\boldsymbol{P}}_k$ in $\boldsymbol{G}_k\tilde{\boldsymbol{p}}_k^{1/2}\bar{\boldsymbol{P}}_k^{1/2} = \boldsymbol{\mathcal{G}}_k$ from (25)
     - *ii* update $\boldsymbol{F}$, $\boldsymbol{\beta}$ in $\boldsymbol{\mathcal{F}} = \boldsymbol{F}\boldsymbol{\beta}\boldsymbol{P}^{-1/2}$ from (7)
     - *iii* compute Λ, update $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{q}}$ as right and left Perron Frobenius eigen vectors of Λ, and update $\boldsymbol{P}_k = \tilde{p}_k\bar{\boldsymbol{P}}_k$
   - 3.3 **until** required accuracy is reached or $n \geq n_{\max}$
   - 3.4 compute $\boldsymbol{E}_k^{(m)}$ and update $\boldsymbol{W}_k^{(m)} = (\mathrm{diag}(\boldsymbol{E}_k^{(m)}))^{-1}$
   - 3.5 determine $t = \min_k \frac{r_k^{(m)}}{r_k^{\circ(m-1)}}$, $r_k^{\circ(m)} = t\, r_k^{\circ(m-1)}$, and $\xi_k^{(m)} = d_k + r_k^{(m)} - r_k^{\circ(m)}$
   - 3.6 set $n \leftarrow 0$ and set $(.)^{(n_{\max},m-1)} \rightarrow (.)^{(0,m)}$ in order to re-enter the inner loop
4. **until** required accuracy is reached or $m \geq m_{\max}$

in Lemma 1. Optimizing w.r.t. Txs $\boldsymbol{\mathcal{G}}$ leads to dual MMSE solutions as in [19] with weights $\boldsymbol{W}_k$ replaced by $\frac{\lambda_k}{\xi_k}\boldsymbol{W}_k$, namely

$$\bar{\boldsymbol{\mathcal{G}}} = (\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\mathcal{F}}\boldsymbol{W'}\boldsymbol{\mathcal{F}}^{\mathrm{H}}\boldsymbol{H} + \sigma_n^2\frac{\mathrm{tr}(\boldsymbol{W'}\boldsymbol{\mathcal{F}}^{\mathrm{H}}\mathbf{F})}{P_{\max}}\boldsymbol{I})^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{\mathcal{F}}\boldsymbol{W'}$$

$$\boldsymbol{\mathcal{G}}_k = \sqrt{\frac{\tilde{p}_k}{\mathrm{tr}(\bar{\boldsymbol{\mathcal{G}}}_k^{\mathrm{H}}\bar{\boldsymbol{\mathcal{G}}}_k)}}\bar{\boldsymbol{\mathcal{G}}}_k \tag{25}$$

where $\boldsymbol{W'} = \mathrm{blkdiag}\{\boldsymbol{W'_1}, \ldots, \boldsymbol{W'_K}\}$ and $\boldsymbol{W'_k} = \lambda_k/\xi_k\boldsymbol{W}_k$. The optimization of the Lagrange multipliers λ follows from Perron Frobenius theory. With (18), we can reformulate (24) as

$$\Delta = \max_{\lambda:\sum\lambda_k=1}\min_{\tilde{\boldsymbol{p}}}\sum_k \lambda_k\frac{[\boldsymbol{\Lambda}\,\tilde{\boldsymbol{p}}]_k}{\tilde{\boldsymbol{p}}_k} \tag{26}$$

which is the Donsker–Varadhan–Friedland formula [20, Chapter 8] for the Perron Frobenius eigenvalue of Λ. A related formula is the Rayleigh quotient

$$\Delta = \max_{\tilde{\boldsymbol{q}}}\min_{\tilde{\boldsymbol{p}}}\frac{\tilde{\boldsymbol{q}}^T\boldsymbol{\Lambda}\,\tilde{\boldsymbol{p}}}{\tilde{\boldsymbol{q}}^T\tilde{\boldsymbol{p}}} \tag{27}$$

where $\tilde{\boldsymbol{p}}$, $\tilde{\boldsymbol{q}}$ are the right and left Perron Frobenius eigenvectors. Comparing (27) to (26), then apart from normalization factors, we get $\lambda_k/\tilde{\boldsymbol{p}}_k = \tilde{\boldsymbol{q}}_k$ or hence $\lambda_k = \tilde{\boldsymbol{p}}_k\tilde{\boldsymbol{q}}_k$.

The proposed optimization framework is summarized in Table I. Superscripts refer to iteration numbers. This algorithm is based on a double loop. The inner loop solves the WMSE balancing problem in (9) 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 $t = \min_k \frac{r_k^{(m)}}{r_k^{\circ(m-1)}} \geq 1$.

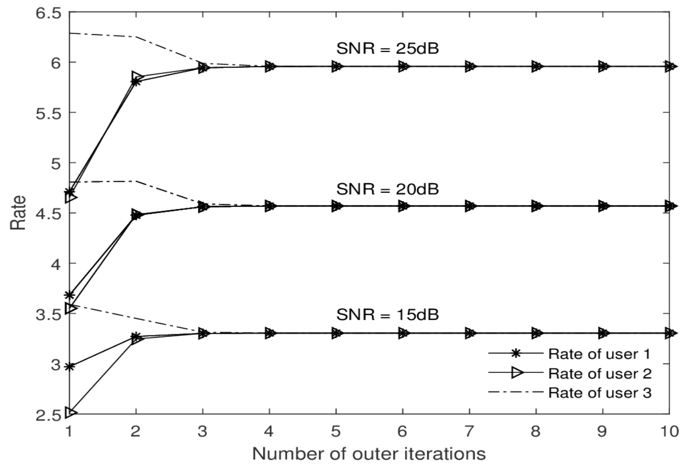

Fig. 2. User rate VS number of outer iterations for one loop ($n_{\max} = 1$) and two loops ($n_{\max} = 4$): $K = 3$, SNR= 15 dB, $M = 6, N_k = d_k = 2$.

By contradiction, if this was not the case, it can be shown to lead to $\frac{\mathrm{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)}\right)}{\xi_k^{(m-1)}} > 1,\ \forall k$ and hence $\Delta^{(m)} > 1$. But we have

$$\Delta^{(m)} = \frac{\mathrm{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)}\right)}{\xi_k^{(m-1)}},\ \forall k, = \max_k\frac{\mathrm{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)}\right)}{\xi_k^{(m-1)}}$$
$$\overset{(a)}{<} \max_k\frac{\mathrm{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m-1)}\right)}{\xi_k^{(m-1)}} = \max_k\frac{d_k}{\xi_k^{(m-1)}} \overset{(b)}{<} 1\,. \tag{28}$$

Let $\mathbf{E} = \{\mathbf{E}_k, k = 1, ..., K\}$ and $f^{(m)}(\mathbf{E}) = \max_k\frac{\mathrm{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k\right)}{\xi_k^{(m-1)}}$. 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 $\frac{\mathrm{tr}\left(\boldsymbol{W}_k^{(m-1)}\mathbf{E}_k^{(m)}\right)}{\xi_k^{(m-1)}} = \Delta^{(m)}, \forall k$. 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.

VI. 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

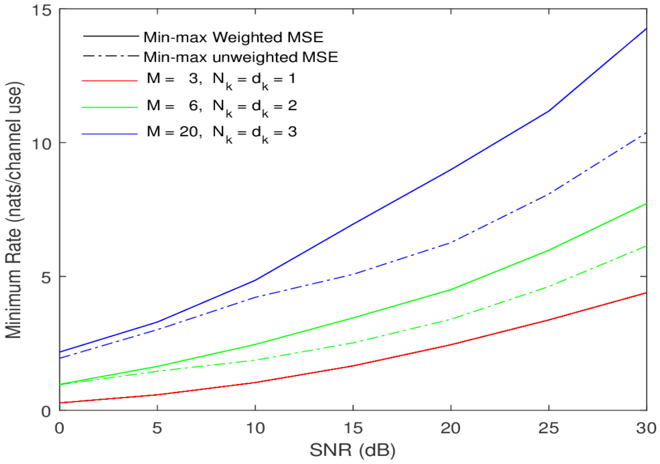

Fig. 3. Minimum rate in the system VS SNR: $K = 3$.

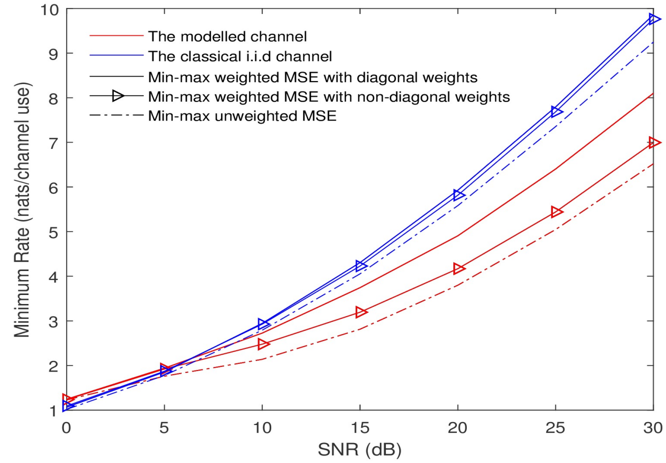

Fig. 4. Minimum rate in the system with respect to SNR for i.i.d. Gaussian channel VS modelled channel: $K = 3, M = 6, N_k = d_k = 2$.

have i.i.d. elements distributed as $\mathcal{CN}(0, 1)$; $\boldsymbol{\mathcal{U}}_k = \mu \boldsymbol{U}_k$, with the normalization parameter $\mu = \left(\text{trace}\left(\boldsymbol{U}_k\right)\right)^{-1/2}$ and $\boldsymbol{U}_k = \text{diag}\left\{1, \alpha, \alpha^2, \ldots, \alpha^{N_k - 1}\right\}$ ($\alpha \in \mathbb{R}$ being a scalar parameter). This model allows to control the rank profile of the MIMO channels. For all simulations, we fix $\alpha = 0.3$ and take 1000 channel realisations and $n_{\max} = 20$. The algorithm converges after 3-4 outer iterations with 4 to 10 inner iterations, see Figure 2.

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 [21] with respect to 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.

We observe, in Figure 4, the same behavior with the classical i.i.d. Gaussian channel, but with a smaller gap. Also, we can see that the balanced rate obtained using diagonal $\{\boldsymbol{W}_k\}$ and with DL power optimization outperforms the balanced rate derived with non-diagonal weight matrices and UL power optimization [15].

In Figure 5, 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.

Figure 6 considers a case where the number of transmit antennas does not support the total number of streams; we can see that the unsupported streams are off.

VII. Conclusions

In this work, we addressed the case of multiple streams per user (MIMO links) for which we considered user rate balancing, not stream rate balancing. Actually, we optimized

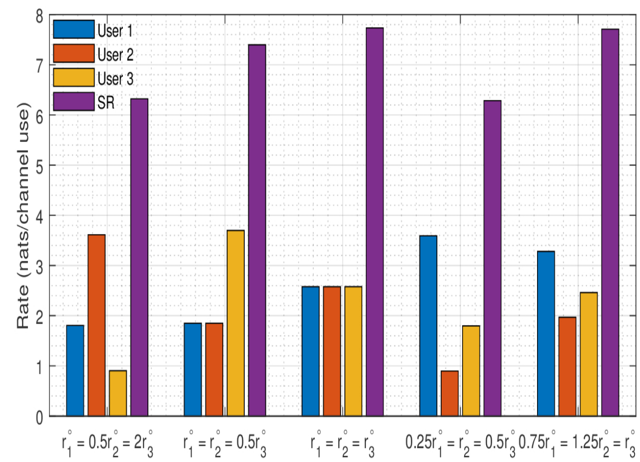

Fig. 5. Rate distribution among users: $K = 3$, SNR= 10 dB, $M = 6, N_k = d_k = 2$.

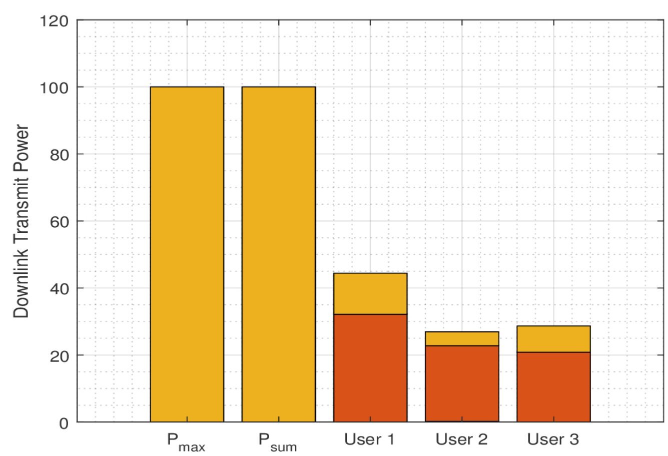

Fig. 6. Downlink transmit power distribution among users: $K = 3$, SNR= 20 dB, $M = 6, N_k = d_k = 3$.

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 which itself was shown to be related to a weighted sum MSE minimization via Langrangian duality. We also compared our diagonal DL matrix-weighted MSE approach with *i)* the min-max unweighted MSE optimization, and *ii)* non-diagonal UL matrix-weighted MSE. Simulation results showed that our solution outperforms the latters.

ACKNOWLEDGEMENTS

This work has been performed in the framework of the Horizon 2020 project ONE5G (ICT-760809) receiving funds from the European Union. The authors would like to acknowledge the contributions of their colleagues in the project, although the views expressed in this contribution are those of the authors and do not necessarily represent the project. EURECOM's research is partially supported by its industrial members: ORANGE, BMW, Symantec, SAP, Monaco Telecom, iABG, and by the projects DUPLEX (French ANR) and MASS-START (French FUI).

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 Transactions on Mobile Computing*, Jan. 2018.

[2] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear Precoding via Conic Optimization for Fixed MIMO Receivers," in *IEEE Transactions on Signal Processing*, vol. 54, no. 1, Jan 2006, pp. 161–176.

[3] M. Bengtsson and B. Ottersten, "Optimal and suboptimal Transmit Beamforming," *Handbook of 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," in *IEEE Transactions on Signal Processing*, vol. 59, no. 8, Aug 2011, p. 3850 –3862.

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

[6] G. Montalbano and D. Slock, "Matched Filter Bound Optimization For Multiuser Downlink Transmit Beamforming," in *Universal Personal Communications, 1998. ICUPC '98. IEEE 1998 International Conference*, vol. 1, Oct 1998, p. 677 –681.

[7] F. Negro, M. Cardone, I. Ghauri, and D. Slock, "SINR Balancing and Beamforming for the MISO Interference Channel," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep 2011.

[8] W. Yu and T. Lan, "Transmitter Optimization for the Multi-antenna Downlink with Per-antenna Power Constraints," in *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*, 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," in *IEEE Trans. Signal Process*, vol. 55, no. 11, Nov 2007.

[12] ——, "Capacity Balancing for Multiuser MIMO Systems," in *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," in *IEEE Trans. Signal Processing*, vol. 57, no. 2, Feb 2009, p. 698–713.

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

[15] I. Ghamnia, D. Slock, and Y. Yuan-Wu, "Rate Balancing for Multiuser MIMO Systems," in *IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Cannes, France, Jul 2019.

[16] 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," in *IEEE Trans. Signal Processing*, Sept. 2011.

[17] F. Negro, I. Ghauri, and D. T. Slock, "Sum Rate Maximization in the Noisy MIMO Interfering Broadcast Channel with Partial CSIT via the Expected Weighted MSE," in *Proc. Int'l Symp. on Wireless Communication Systems (ISWCS)*, Paris, France, 2012.

[18] F. Negro, I. Ghauri, and D. T. M. Slock, "Deterministic Annealing Design and Analysis of the Noisy MIMO Interference Channel," in *Proc. IEEE Inf. Theo. and Applic. Workshop (ITA)*, San Diego, CA, USA, 2011.

[19] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted Sum-Rate Maximization using Weighted MMSE for MIMO-BC Beamforming Design," in *Weighted Sum-Rate Maximization using Weighted MMSE for MIMO-BC Beamforming Design*, Dec 2008.

[20] C. Meyer, *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.

[21] 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.