Multi-Cell MIMO User Rate Balancing with Partial CSIT

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*—In this paper, we consider the problem of user rate balancing in the downlink of multi-cell multi-user (MU) Multiple-Input-Multiple-Output (MIMO) systems with partial Channel State Information at the Transmitter (CSIT). With MIMO leading to multiple streams per user, user rate balancing involves both aspects of balancing and sum rate optimization. We linearize the problem by introducing a rate minorizer and by formulating the balancing operation as constraints leading to a Lagrangian, allowing to transform rate balancing into weighted sum minimization with Perron Frobenius theory. We provide original analytical expressions for the Lagrange multipliers for the multiple power constraints which can also handle the case in which some power constraints are satisfied with inequality, as can arise in a multi-cell scenario. We introduce two partial CSIT formulations. One is based on the ergodic rate Mean Squared Error (EMSE) relation, the other involves an original rate minorizer in terms of the received interference plus noise covariance matrix, in the partial CSIT case applied to the Expected Signal and Interference Power (ESIP) rate. The simulation results exhibit the improved performance of the proposed techniques over naive partial CSIT beamforming based on perfect CSIT algorithms, and in particular illustrate the close to optimal performance of the ESIP approach.**

***Index Terms*—Inter-cell interference coordination (ICIC), Coordinated Beamforming (CoBF), Multi-User MIMO, Rate Balancing, Partial CSIT**

I. Introduction

Multi-user multiple-input multiple-output (MIMO) systems are considered as a promising technique for next generation cellular networks for their great potential to achieve high throughput [1]. In downlink communications, when a certain knowledge of the Channel State Information (CSI) at the transmitter is available, the system throughput can be maximized. In practical, obtaining CSI at the receiver is easily possible via training, whereas CSI at the transmitter (CSIT) acquires reciprocity or feedback from the receiver. Therefore, many works address the problem of optimizing the performance of MIMO systems with the presence of CSIT uncertainties, better known as partial CSIT. Among the different optimization criteria, we distinguish the transmit power minimization, and the max-min/min-max problems w.r.t. either signal-to-interference-plus-noise ratio (SINR) [2]–[6], Mean Square Error (MSE) [7]–[9] or user rate. The latter is the focus of this work. In particular, we study Multi-Cell MIMO User Rate Balancing with Partial CSIT.

In perfect CSIT case, [10] studies the balancing problem w.r.t. SINR for MISO system using uplink/downlink duality. In fact, most of max-min beamforming problems are transformed into the dual problem power minimization problem in the uplink. In [11], SINR balancing problem subject to multiple weighted-sum power constraints for MISO system is solved by exploiting Perron-Frobenius theory and uplink/downlink duality, and an iterative subgradient projection algorithm is used to satisfy the per-stream power constraints. Similarly, MSE duality, which states that the same MSE values are achievable in the downlink and the uplink with the same transmit power constraint, has been exploited to solve max-min beamforming problems w.r.t. MSE. In [9], three levels of MSE dualities are established between MIMO BC and MIMO MAC with the same transmit power constraint; these dualities are exploited to reduce the computational complexity of the sum-MSE and weighted sum-MSE minimization problems (with fixed weights) in a MIMO BC. On the other hand, we prove that 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 [12]–[14].

In contrast, due to the inevitability of channel estimation error, CSI can never be perfect. This motivates [15] to consider an MSE-based transceiver design problem where the channel knowledge is modeled in terms of channel mean and variance both at the transmitter and receivers. Then, an iterative algorithm is proposed to solve the expected MSE balancing problem by switching between the broadcast and the multiple access channels. Also, SINR balancing problem with imperfect CSIT is studied in [16] for multi-cell multi-user MISO system. Therein, the authors introduce an alternative biased SINR estimate to incorporate the knowledge of the channel estimation error, outperforming the unbiased maximum-likelihood estimate. In [17], CSI error matrix is represented as a bounded hyper-spherical region within some radius, leading to a robust max-min SINR problem for single-stream MIMO system. The latter is solved as semidefinite program problem, where robust transmit and receive beamformers are obtained using alternating optimization. Rate balancing problem is studied in [18], for broadcast MISO channel, where the case of erroneous CSI at the receiver is considered. The authors use duality w.r.t. SINR to solve the balancing problem: they transform the BC problem into dual MAC problem taking into consideration the erroneous receiver CSI. Actually, in the single stream per user case (e.g. in MISO systems), balancing w.r.t. SINR, MSE or user rate is equivalent (in the unweighted case). Another rate balancing work for MISO system is studied in [19], wherein the statistical properties of the channel are exploited and an algorithm for optimal downlink beamforming is derived using uplink/downlink duality.

In this work, we focus on ergodic user rate balancing, which corresponds to maximizing the minimum (weighted) per user expected rate in the network. We consider a multi-cell multi-user MIMO system with partial CSIT, which combines both channel estimates and channel (error) covariance information. In particular, we introduce a novel extension of [14] to partial CSIT, maximizing an expected rate lower bound in terms of expected MSE. Furthermore, we introduce a second algorithm by exploiting a better approximation of the expected rate as the Expected Signal and Interference Power (ESIP) rate. Whereas we have considered the ESIP approach in previous sum utility optimization work, the algorithm here is based on an original minorizer for every individual rate term, different from existing DC programming approaches in sum utility optimization. Both algorithms are based on a Lagrangian formulation introduced in [14] for perfect CSIT, in which utility balancing gets transformed into a weighted sum utility with known optimal beamformers.

II. System Model

We consider a MIMO system with C cells. Each cell c has one base station (BS) of M_c transmit antennas serving K_c users, with total number of users $\sum_c K_c = K$. We refer to the BS of user $k \in \{1, \ldots, K\}$ by b_k. Each user has N_k antennas. The channel between the kth user and the BS in cell c is denoted by $\boldsymbol{H}_{k,c} \in \mathbb{C}^{N_k \times M_c}$. We consider 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 $\mathbf{s}_c = [\mathbf{s}^{\mathrm{T}}_{K_{1:c-1}+1} \ldots \mathbf{s}^{\mathrm{T}}_{K_{1:c}}]^{\mathrm{T}}$, i.e., $\mathbb{E}\big[\mathbf{s}_c \mathbf{s}_c^{\mathrm{H}}\big] = \boldsymbol{I}$, 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 and $K_{1:c} = \sum_{i=1}^{c} K_i$. The latter is transmitted using the transmit filtering matrix $\boldsymbol{\mathcal{G}}_c = [\boldsymbol{\mathcal{G}}_{K_{1:c-1}+1} \ldots \boldsymbol{\mathcal{G}}_{K_{1:c}}] \in \mathbb{C}^{M_c \times N_c}$, with $\boldsymbol{\mathcal{G}}_k = p_k^{1/2} \boldsymbol{G}_k$, $\boldsymbol{G}_k$ being the (unit Frobenius norm) beamforming matrix, p_k is non-negative downlink power allocation of user k and $N_c = \sum_{k:b_k=c} d_k$ is the total number of streams in cell c. Each cell is constrained with $P_{\max,c}$, i.e., the total transmit power in c is limitted such that $\sum_{k:b_k=c} p_k \leq P_{\max,c}$. The received signal at user k in cell b_k is

$$\boldsymbol{y}_k = \underbrace{\boldsymbol{H}_{k,b_k} \boldsymbol{\mathcal{G}}_k \boldsymbol{s}_k}_{\text{signal}} + \underbrace{\sum_{\substack{i \neq k \\ b_i = b_k}} \boldsymbol{H}_{k,b_k} \boldsymbol{\mathcal{G}}_i \boldsymbol{s}_i}_{\text{intracel interf.}} + \underbrace{\sum_{j \neq b_k} \sum_{i: b_i = j} \boldsymbol{H}_{k,j} \boldsymbol{\mathcal{G}}_i \boldsymbol{s}_i}_{\text{intercell interf.}} + \boldsymbol{n}_k \tag{1}$$

Similarly, the receive filtering matrix for each user k is defined as $\boldsymbol{\mathcal{F}}_k^{\mathrm{H}} = p_k^{-1/2} \boldsymbol{F}_k^{\mathrm{H}} \in \mathbb{C}^{d_k \times N_k}$, composed of beamforming matrix $\boldsymbol{F}_k^{\mathrm{H}} \in \mathbb{C}^{d_k \times N_k}$. The received filter output is $\hat{\boldsymbol{s}}_k = \boldsymbol{\mathcal{F}}_k^{\mathrm{H}} \boldsymbol{y}_k$.

III. Joint Mean and Covariance Gaussian CSIT

In this section we drop the user index k and BS index c for simplicity. Assume that the channel has a (prior) Gaussian distribution with zero mean and separable correlation model

$$\boldsymbol{H} = \boldsymbol{C}_r^{1/2} \boldsymbol{H}' \boldsymbol{C}_t^{1/2} \tag{2}$$

where $\boldsymbol{C}_r^{1/2}$, $\boldsymbol{C}_t^{1/2}$ are Hermitian square-roots of the Rx and Tx side covariance matrices

$$\begin{aligned} \mathrm{E}\, \boldsymbol{H}\, \boldsymbol{H}^H &= \mathrm{tr}\{\boldsymbol{C}_t\}\, \boldsymbol{C}_r \\ \mathrm{E}\, \boldsymbol{H}^H \boldsymbol{H} &= \mathrm{tr}\{\boldsymbol{C}_r\}\, \boldsymbol{C}_t \end{aligned} \tag{3}$$

Now, the Tx dispose of a (deterministic) channel estimate

$$\widehat{\boldsymbol{H}}_d = \boldsymbol{H} + \boldsymbol{C}_r^{1/2}\, \widetilde{\boldsymbol{H}}'_d\, \boldsymbol{C}_d^{1/2} \tag{4}$$

where again the elements of $\widetilde{\boldsymbol{H}}'_d$ are i.i.d. $\sim \mathcal{CN}(0,1)$, and typically $\boldsymbol{C}_d = \sigma^2_{\widetilde{\boldsymbol{H}}} I_M$. The combination of the estimate with the prior information leads to the (posterior) LMMSE estimate

$$\widehat{\boldsymbol{H}} = \mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}_d} \boldsymbol{H} = \widehat{\boldsymbol{H}}_d\, (\boldsymbol{C}_t + \boldsymbol{C}_d)^{-1} \boldsymbol{C}_t,\ \boldsymbol{C}_p = \boldsymbol{C}_d\, (\boldsymbol{C}_t + \boldsymbol{C}_d)^{-1} \boldsymbol{C}_t \tag{5}$$

where the estimation error on $\widehat{\boldsymbol{H}}$ can be modeled as $\widehat{\boldsymbol{H}} - \boldsymbol{H} = \boldsymbol{C}_r^{1/2}\, \widetilde{\boldsymbol{H}}'_p\, \boldsymbol{C}_p^{1/2}$ with $\widehat{\boldsymbol{H}}$ and $\widetilde{\boldsymbol{H}}'_p$ being independent (or decorrelated if not Gaussian). Note that we get for the MMSE estimate of a quadratic quantity of the form

$$\mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}_d} \boldsymbol{H}^H \boldsymbol{H} = \widehat{\boldsymbol{H}}^H \widehat{\boldsymbol{H}} + \mathrm{tr}\{\boldsymbol{C}_r\} \boldsymbol{C}_p = \boldsymbol{R}\,. \tag{6}$$

Let us emphasize that this MMSE estimate implies $\boldsymbol{R} = \arg\min_{\boldsymbol{T}}\ \mathrm{E}_{\boldsymbol{H}|\hat{\boldsymbol{H}}_d} ||\boldsymbol{H}^H \boldsymbol{H} - \boldsymbol{T}||^2$. It averages out to

$$\mathrm{E}_{\widehat{\boldsymbol{H}}_d} \boldsymbol{R} = \ \mathrm{E}_{\boldsymbol{H},\widehat{\boldsymbol{H}}_d} \boldsymbol{H}^H \boldsymbol{H} = \ \mathrm{E}_{\boldsymbol{H}} \boldsymbol{H}^H \boldsymbol{H} = \mathrm{tr}\{\boldsymbol{C}_r\} \boldsymbol{C}_t\,. \tag{7}$$

Hence, if we want the best estimate for $\boldsymbol{H}^H \boldsymbol{H}$ (which appears in the signal or interference powers), it is not sufficient to replace $\boldsymbol{H}$ by $\widehat{\boldsymbol{H}}$ but also the (estimation error) covariance information should be exploited. Other useful expressions are

$$\mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}_d} \boldsymbol{H}^H \boldsymbol{Q} \boldsymbol{H} = \widehat{\boldsymbol{H}}^H \boldsymbol{Q} \widehat{\boldsymbol{H}} + \mathrm{tr}\{\boldsymbol{C}_r \boldsymbol{Q}\} \boldsymbol{C}_p \tag{8}$$

and

$$\mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}_d} \boldsymbol{H} \boldsymbol{P} \boldsymbol{H}^H = \widehat{\boldsymbol{H}} \boldsymbol{P} \widehat{\boldsymbol{H}}^H + \mathrm{tr}\{\boldsymbol{C}_p \boldsymbol{P}\} \boldsymbol{C}_r\,. \tag{9}$$

Note that $\rho_P = \frac{\mathrm{tr}\{\widehat{\boldsymbol{H}}^H \widehat{\boldsymbol{H}}\}}{\mathrm{tr}\{\boldsymbol{C}_r\} \mathrm{tr}\{\boldsymbol{C}_p\}}$ is a form of Ricean factor that represents posterior channel estimation quality. It depends on the deterministic channel estimation quality $\rho_D = 1/\sigma^2_{\widetilde{\boldsymbol{H}}}$. Below we consider $\boldsymbol{C}_r = \boldsymbol{I}$, and the only covariance $\boldsymbol{C}$ we shall need is $\boldsymbol{C}_p$, hence we drop the subscript p. Perfect CSIT algorithms can be obtained by setting $\sigma^2_{\widetilde{\boldsymbol{H}}} = 0$, leading to $\widehat{\boldsymbol{H}} = \boldsymbol{H}$ and $\boldsymbol{C}_p = 0$.

IV. Expected Rate Balancing Problem

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

$$\begin{aligned} &\max_{\boldsymbol{G},p} \min_k \ r_k / r_k^\circ \\ &\text{s.t.} \sum_{k: b_k = c} p_k \leq P_{\max,c}, 1 \leq c \leq C \end{aligned} \tag{10}$$

where r_k is the kth user-rate

$$r_k = \ln\det\Big(\boldsymbol{I} + \boldsymbol{R}_{\overline{k}}^{-1} \boldsymbol{H}_{k,b_k} \boldsymbol{\mathcal{G}}_k \boldsymbol{\mathcal{G}}_k^{\mathrm{H}} \boldsymbol{H}_{k,b_k}^{\mathrm{H}}\Big) = \ln\det\Big(\boldsymbol{R}_{\overline{k}}^{-1} \boldsymbol{R}_k\Big), \tag{11}$$

$$\boldsymbol{R}_{\overline{k}} = \sigma_n^2 \boldsymbol{I} + \sum_{l \neq k} \boldsymbol{H}_{k,b_l} \boldsymbol{\mathcal{G}}_l \boldsymbol{\mathcal{G}}_l^{\mathrm{H}} \boldsymbol{H}_{k,b_l}^{\mathrm{H}}\,, \tag{12}$$

$$\boldsymbol{R}_k = \boldsymbol{R}_{\overline{k}} + \boldsymbol{H}_{k,b_k} \boldsymbol{\mathcal{G}}_k \boldsymbol{\mathcal{G}}_k^{\mathrm{H}} \boldsymbol{H}_{k,b_k}^{\mathrm{H}}\,, \tag{13}$$

$\boldsymbol{R}_{\overline{k}}$ and $\boldsymbol{R}_k$ are the interference plus noise and total received signal covariances, and r_k° is the rate priority (weight) for user k. Actually, in the presence of partial CSIT, we shall be interested in balancing the expected (or ergodic) rates

$$\max_{\boldsymbol{G},p} \min_{k} \ \overline{r}_k / r_k^{\circ}$$
$$\text{s.t.} \sum_{k:b_k=c} p_k \leq P_{\max,c},\ c=1,\ldots,C \tag{14}$$

where $\overline{r}_k = \ \mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\, r_k$. We shall need

$$\overline{\boldsymbol{S}}_{k,i} = \widehat{\boldsymbol{H}}_{k,b_i}\boldsymbol{G}_i\boldsymbol{G}_i^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_i}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{G}_i^{\mathrm{H}}\boldsymbol{C}_{k,b_i}\boldsymbol{G}_i\}\boldsymbol{I},\ \overline{\boldsymbol{S}}_k = \overline{\boldsymbol{S}}_{k,k} \tag{15}$$
$$\overline{\boldsymbol{R}}_{\overline{k}} = \mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\boldsymbol{R}_{\overline{k}} = \sigma_n^2\boldsymbol{I} + \sum_{i\neq k} p_i\overline{\boldsymbol{S}}_{k,i}\ ,\ \ \overline{\boldsymbol{R}}_k = \overline{\boldsymbol{R}}_{\overline{k}} + p_k\overline{\boldsymbol{S}}_k \tag{16}$$

However, the problem presented in (14) is complex and can not be solved directly.

Lemma 1. *The user k rate in (11) is lower bounded as [20]*

$$\overline{r}_k = \ E_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \max_{\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k} \left[\ln\det\big(\boldsymbol{W}_k\big) - \mathrm{tr}\big(\boldsymbol{W}_k\mathbf{E}_k\big) + d_k\right] \tag{17}$$
$$\geq \overline{r}_k^l = \max_{\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k} \underline{f}_k^l\,,\ \underline{f}_k^l = \ln\det\big(\boldsymbol{W}_k\big) - \mathrm{tr}\big(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k\big) + d_k \tag{18}$$

where $\underline{f}_k^l = \underline{f}_k^l(\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k)\,,\quad \overline{\boldsymbol{E}}_k = \mathbb{E}\Big[(\hat{\mathbf{s}}_k - \mathbf{s}_k)(\hat{\mathbf{s}}_k - \mathbf{s}_k)^{\mathrm{H}}\Big]$

$$= \boldsymbol{I} - \boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_k}\boldsymbol{\mathcal{G}}_k - \boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k + \sigma_n^2\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k$$
$$+ \sum_{l=1}^{K}\boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\big(\widehat{\boldsymbol{H}}_{k,b_l}\boldsymbol{\mathcal{G}}_l\boldsymbol{\mathcal{G}}_l^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_l}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{\mathcal{G}}_l^{\mathrm{H}}\boldsymbol{C}_{k,b_l}\boldsymbol{\mathcal{G}}_l\}\boldsymbol{I}\big)\boldsymbol{\mathcal{F}}_k \tag{19}$$

is the k^{th} user downlink Expected MSE (EMSE) matrix between the decision variable $\hat{\mathbf{s}}_k$ and the transmit signal $\mathbf{s}_k$, and $\{\boldsymbol{W}_k\}_{1\leq k\leq K}$ are auxiliary weight matrix variables with optimal solution $\boldsymbol{W}_k^{\mathrm{opt}} = \overline{\boldsymbol{E}}_k^{-1}$ and the optimal receivers are

$$\boldsymbol{\mathcal{F}}_k = \overline{\boldsymbol{R}}_k^{-1}\widehat{\boldsymbol{H}}_{k,b_k}\boldsymbol{\mathcal{G}}_k. \tag{20}$$

with rate lower bound

$$\overline{r}_k^l = -\ln\det(\boldsymbol{I} - \boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}}\overline{\boldsymbol{R}}_k^{-1}\widehat{\boldsymbol{H}}_{k,b_k}\boldsymbol{\mathcal{G}}_k)\ . \tag{21}$$

Note that $\underline{f}_k^l$ is a lower bound for any $\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k$ and so is $\overline{r}_k^l = \max_{\boldsymbol{W}_k,\boldsymbol{\mathcal{F}}_k}\underline{f}_k^l$. Now consider both (14) and (18), and let us introduce $\xi_k = \ln\det\big(\boldsymbol{W}_k\big) + d_k - r_k^{\triangle}$, the matrix-weighted EMSE (WEMSE) 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 WEMSE satisfies $\epsilon_{w,k} = \mathrm{tr}\big(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k\big) \leq d_k$ and $\ln\det\big(\boldsymbol{W}_k\big) \geq r_k^{\triangle}$ or hence $\xi_k \geq d_k$. This leads $\forall k$ to

$$\frac{\epsilon_{w,k}}{\xi_k} \leq 1 \iff \ln\det\big(\boldsymbol{W}_k\big) + d_k - \mathrm{tr}\big(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k\big) \geq r_k^{\triangle} \tag{22}$$
$$\overset{(a)}{\Longrightarrow} \overline{r}_k^l / r_k^{\triangle} \geq 1$$

where (a) follows from (18). To get to (22), what we can exploit in (14) is a scale factor t that can be chosen freely in the rate weights r_k° in (14). We shall take $t = \min_k \overline{r}_k^l/r_k^{\circ}$, which allows to transform the rate weights r_k° into target rates $r_k^{\triangle} = tr_k^{\circ}$, and at the same time allows to interpret the WEMSE weights ξ_k as target WEMSE values.

Doing so, the initial rate balancing optimization problem (14) can be transformed into a WEMSE balancing problem expressed as follows

$$\min_{\boldsymbol{G},p,\boldsymbol{\mathcal{F}}} \max_{k} \ \epsilon_{w,k}/\xi_k$$
$$\text{s.t.} \sum_{k:b_k=c} p_k \leq P_{\max,c}, 1\leq c\leq C \tag{23}$$

which needs to be complemented with an outer loop in which $\boldsymbol{W}_k = \overline{\boldsymbol{E}}_k^{-1}$, $t = \min_k \overline{r}_k^l/r_k^{\circ}$, $r_k^{\triangle} = tr_k^{\circ}$ and $\xi_k = d_k + \overline{r}_k^l - r_k^{\triangle}$ get updated. The problem in (23) is still difficult to be handled directly.

V. The Weighted User EMSE Optimization

In this section, the problem (23) with respect to the matrix weighted user EMSE is studied. The per user matrix WEMSE can be expressed as follows

$$\epsilon_{w,k} = \mathrm{tr}\big(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k\big) \tag{24}$$
$$= \mathrm{tr}\big(\boldsymbol{W}_k\big) - 2\,\mathrm{tr}\big(\boldsymbol{W}_k\boldsymbol{G}_k^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}}\boldsymbol{F}_k\big) + \sigma_n^2 p_k^{-1}\mathrm{tr}\big(\boldsymbol{W}_k\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{F}_k\big)$$
$$+ p_k^{-1}\sum_{l=1}^{K} p_l\mathrm{tr}\big(\boldsymbol{W}_k\boldsymbol{F}_k^{\mathrm{H}}\big(\widehat{\boldsymbol{H}}_{k,b_l}\boldsymbol{G}_l\boldsymbol{G}_l^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_l}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{G}_l^{\mathrm{H}}\boldsymbol{C}_{k,b_l}\boldsymbol{G}_l\}\boldsymbol{I}\big)\boldsymbol{F}_k\big)$$

Define the diagonal matrix $\boldsymbol{D}$ of signal WEMSE contributions

$$[\boldsymbol{D}]_{ii} = \mathrm{tr}\big(\boldsymbol{W}_i\big) - 2\,\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{G}_i^{\mathrm{H}}\widehat{\boldsymbol{H}}_{i,b_i}^{\mathrm{H}}\boldsymbol{F}_i\big) \tag{25}$$
$$+ \mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\big(\widehat{\boldsymbol{H}}_{i,b_i}\boldsymbol{G}_i\boldsymbol{G}_i^{\mathrm{H}}\widehat{\boldsymbol{H}}_{i,b_i}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{G}_i^{\mathrm{H}}\boldsymbol{C}_{i,b_i}\boldsymbol{G}_i\}\boldsymbol{I}\big)\boldsymbol{F}_i\big),$$

and the matrix of weighted interference powers

$$[\boldsymbol{\Psi}]_{ij} = \begin{cases}\mathrm{tr}\{\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\big(\widehat{\boldsymbol{H}}_{i,b_j}\boldsymbol{G}_j\boldsymbol{G}_j^{\mathrm{H}}\widehat{\boldsymbol{H}}_{i,b_j}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{G}_j^{\mathrm{H}}\boldsymbol{C}_{i,b_j}\boldsymbol{G}_j\}\boldsymbol{I}\big)\boldsymbol{F}_i\}, & i\neq j\\ 0, & i=j.\end{cases}$$

We can rewrite (24) as, with $\boldsymbol{p} = [p_1\cdots p_K]^{\mathrm{T}}$

$$\epsilon_{w,i} = [\boldsymbol{D}]_{ii} + p_i^{-1}[\boldsymbol{\Psi}\boldsymbol{p}]_i + \sigma_n^2 p_i^{-1}\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\boldsymbol{F}_i\big) \tag{26}$$

Collecting all user WEMSEs in a vector $\boldsymbol{\epsilon}_w = \mathrm{diag}(\epsilon_{w,1},\ldots,\epsilon_{w,K})$, we get

$$\boldsymbol{\epsilon}_w\boldsymbol{1}_K = \mathrm{diag}(\boldsymbol{p})^{-1}\left[(\boldsymbol{D}+\boldsymbol{\Psi})\mathrm{diag}(\boldsymbol{p})\boldsymbol{1}_K + \boldsymbol{\sigma}\right] \tag{27}$$

where the $K\times 1$ vector $\boldsymbol{\sigma}$ is defined as

$$\boldsymbol{\sigma}_i = \sigma_n^2\,\mathrm{tr}\big(\boldsymbol{W}_i\boldsymbol{F}_i^{\mathrm{H}}\boldsymbol{F}_i\big).$$

By multiplying both sides of (27) with $\mathrm{diag}(\boldsymbol{p})$, we get

$$\boldsymbol{\epsilon}_w\boldsymbol{p} = (\boldsymbol{D}+\boldsymbol{\Psi})\boldsymbol{p} + \boldsymbol{\sigma}\,. \tag{28}$$

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

$$\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w\,\boldsymbol{p} = \boldsymbol{\xi}^{-1}(\boldsymbol{D}+\boldsymbol{\Psi})\boldsymbol{p} + \boldsymbol{\xi}^{-1}\boldsymbol{\sigma}\,. \tag{29}$$

Actually, problem (23) always has a global minimizer $\boldsymbol{p}$ characterized by the equality $\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w(\boldsymbol{p}) = \Delta\boldsymbol{I}$, i.e.,

$$\Delta\,\boldsymbol{p} = \boldsymbol{\xi}^{-1}(\boldsymbol{D}+\boldsymbol{\Psi})\boldsymbol{p} + \boldsymbol{\xi}^{-1}\boldsymbol{\sigma}\,. \tag{30}$$

Now, consider the following problem

$$\max_{\boldsymbol{G},p,\boldsymbol{\mathcal{F}}} \min_{k} \ \overline{r}_k/r_k^{\circ}$$
$$\text{s.t.} \sum_{c=1}^{C}\theta_c\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} \leq \sum_{c=1}^{C}\theta_c P_{\max,c} \tag{31}$$

where $\boldsymbol{c}_c$ is a column vector with $\boldsymbol{c}_c(j) = 1$ for $K_{1:c-1}+1 \leq j \leq K_{1:c}$, and 0 elsewhere. This problem formulation is a relaxation of (14), and $\boldsymbol{\theta} = [\theta_1\cdots\theta_C]^{\mathrm{T}}$can be interpreted as the weights on the individual power constraints in the relaxed problem. The power constraint in (31) can be interpreted as a single weighted power constraint

$$(\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}})\,\boldsymbol{p} \leq \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max} \tag{32}$$

with $\boldsymbol{C}_C = [\boldsymbol{c}_1\cdots\boldsymbol{c}_C] \in \mathbb{R}_+^{K_{1:C}\times C}$ and $\boldsymbol{p}_{\max} = [P_{\max,1}\cdots P_{\max,C}]^{\mathrm{T}}$. Reparameterize $\boldsymbol{p} = \frac{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max}}{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}}\boldsymbol{p}'}\boldsymbol{p}'$ where now $\boldsymbol{p}^{'}$ is unconstrained, which allows us to write (30) as follows (rewriting $\boldsymbol{p}^{'}$ as $\boldsymbol{p}$)

$$\Delta\,\boldsymbol{p} = \boldsymbol{\Lambda}\,\boldsymbol{p} \text{ with } \boldsymbol{\Lambda} = \boldsymbol{\xi}^{-1}(\boldsymbol{D}+\boldsymbol{\Psi}) + \frac{1}{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max}}\boldsymbol{\xi}^{-1}\boldsymbol{\sigma}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}}. \quad (33)$$

Now with (33), the WEMSE balancing problem of (23) becomes

$$\min_{\boldsymbol{p}} \max_{k} \frac{\epsilon_{w,k}}{\xi_k} = \min_{\boldsymbol{p}} \max_{k} \frac{[\boldsymbol{\Lambda}\,\boldsymbol{p}]_k}{p_k} \quad (34)$$

According to the Collatz–Wielandt formula [21, Chapter 8], the above expression corresponds to the Perron-Frobenius (maximal) eigenvalue Δ of $\boldsymbol{\Lambda}$ and the optimal $\boldsymbol{p}$ is the corresponding Perron-Frobenius (right) eigenvector

$$\boldsymbol{\Lambda}\boldsymbol{p} = \Delta\,\boldsymbol{p}. \quad (35)$$

Note that this implies the equality $\boldsymbol{\xi}^{-1}\boldsymbol{\epsilon}_w = \Delta\,\boldsymbol{I}$ as announced in (30).

VI. Algorithmic Solution via Lagrangian Duality

The max-min weighted user rate optimization problem (14) can be reformulated as

$$\begin{aligned} &\min_{t,\boldsymbol{G},\boldsymbol{p}} -t \\ &\text{s.t. } t\,r_k^{\circ} - \overline{r}_k^{l} \leq 0,\ \boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c} \leq 0\,, \forall k, c. \end{aligned} \quad (36)$$

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

$$\begin{aligned} &\max_{\lambda^{'},\mu} \min_{t,\boldsymbol{G},\boldsymbol{p}} \mathcal{L} \\ &\mathcal{L} = -t + \sum_k \lambda_k^{'}(t\,r_k^{\circ} - \overline{r}_k^{l}) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \end{aligned} \quad (37)$$

Integrating the result (22), we get a modified Lagrangian

$$\begin{aligned} &\max_{\lambda^{'},\mu} \min_{t,\boldsymbol{G},\boldsymbol{p},\boldsymbol{\mathcal{F}},\boldsymbol{W}} \mathcal{L} \\ &\mathcal{L} = -t + \sum_k \lambda_k^{'}(\mathrm{tr}(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k) - \xi_k) + \sum_c \mu_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \end{aligned} \quad (38)$$

We get $\mu_c = \mu_o\theta_c$, $\sum_c \theta_c = 1$, where μ_o is the Lagrange multiplier associated with the constraint in (31). Introducing $\lambda_k = \lambda_k^{'}\xi_k$, we can rewrite (with some abuse of notation since actually $\min_{\boldsymbol{W}}$ continues to apply to $\mathrm{tr}(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k) - \xi_k(\boldsymbol{W}_k)$),

$$\begin{aligned} &\max_{\lambda,\mu} \min_{t,\boldsymbol{G},\boldsymbol{p},\boldsymbol{F},\boldsymbol{W}} \mathcal{L} \\ &\mathcal{L} = -t + \sum_k \lambda_k(\frac{\mathrm{tr}(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k)}{\xi_k} - 1) + \mu_o\sum_c \theta_c(\boldsymbol{c}_c^{\mathrm{T}}\boldsymbol{p} - P_{\max,c}) \end{aligned} \quad (39)$$

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

$$\begin{aligned} \max_{\lambda} \min_{\boldsymbol{G},\boldsymbol{p},\boldsymbol{\mathcal{F}}} &\sum_k \frac{\lambda_k}{\xi_k}\,\mathrm{tr}(\boldsymbol{W}_k\overline{\boldsymbol{E}}_k) \\ &+ \sum_c \mu_c\,\big(\sum_{i:b_i=c} \mathrm{tr}\{\boldsymbol{\mathcal{G}}_i^{\mathrm{H}}\boldsymbol{\mathcal{G}}_i\} - P_{max,c}\big) \end{aligned} \quad (40)$$

which is of the form Weighted Sum EMSE (WSEMSE). Optimizing w.r.t. Txs $\boldsymbol{\mathcal{G}}_k$:

$$\begin{aligned} &\tfrac{\partial\mathcal{L}}{\partial\boldsymbol{\mathcal{G}}_k^*} = 0 = -\tfrac{\lambda_k}{\xi_k}\widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k\boldsymbol{W}_k + \mu_{b_k}\boldsymbol{\mathcal{G}}_k + \\ &\big(\textstyle\sum_i \tfrac{\lambda_i}{\xi_i}\big(\widehat{\boldsymbol{H}}_{i,b_k}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_i\boldsymbol{W}_i\boldsymbol{\mathcal{F}}_i^{\mathrm{H}}\widehat{\boldsymbol{H}}_{i,b_k} + \mathrm{tr}\{\boldsymbol{\mathcal{F}}_i\boldsymbol{W}_i\boldsymbol{\mathcal{F}}_i^{\mathrm{H}}\}\boldsymbol{C}_{i,b_k}\big)\big)\,\boldsymbol{\mathcal{G}}_k \end{aligned} \quad (41)$$

This leads to

$$\begin{aligned} \boldsymbol{\mathcal{G}}_k' = &\big(\sum_{l=1}^{K}\big(\widehat{\boldsymbol{H}}_{l,b_k}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_l\boldsymbol{W}_l'\boldsymbol{\mathcal{F}}_l^{\mathrm{H}}\widehat{\boldsymbol{H}}_{l,b_k} + \mathrm{tr}\{\boldsymbol{\mathcal{F}}_l\boldsymbol{W}_l'\boldsymbol{\mathcal{F}}_l^{\mathrm{H}}\}\boldsymbol{C}_{l,b_k}\big) + \mu_{b_k}\boldsymbol{I}\big)^{-1} \\ &\times \widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k\boldsymbol{W}_k'\,,\ \boldsymbol{\mathcal{G}}_k = \sqrt{p_k}\,\boldsymbol{G}_k\,,\ \boldsymbol{G}_k = \frac{1}{\sqrt{\mathrm{tr}\{\boldsymbol{\mathcal{G}}_k'^{\mathrm{H}}\boldsymbol{\mathcal{G}}_k'\}}}\boldsymbol{\mathcal{G}}_k' \end{aligned} \quad (42)$$

where $\boldsymbol{W}_k' = \lambda_k/\xi_k\boldsymbol{W}_k$, and accounting for the fact that the user powers are actually optimized by the Perron-Frobenius theory. Note that we can solve for μ_c by multiplying (41) from the left by $\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}$ and summing over the users in cell c:

$$\begin{aligned} \mu_c = &\frac{1}{P_{max,c}}\lfloor\mathrm{tr}\{\sum_{k:b_k=c}\big[\frac{\lambda_k}{\xi_k}\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_k\boldsymbol{W}_k - \\ &\boldsymbol{\mathcal{G}}_k^{\mathrm{H}}\big(\sum_i\big(\widehat{\boldsymbol{H}}_{i,b_k}^{\mathrm{H}}\boldsymbol{\mathcal{F}}_i\boldsymbol{W}_i'\boldsymbol{\mathcal{F}}_i^{\mathrm{H}}\widehat{\boldsymbol{H}}_{i,b_k} + \mathrm{tr}\{\boldsymbol{\mathcal{F}}_i\boldsymbol{W}_i'\boldsymbol{\mathcal{F}}_i^{\mathrm{H}}\}\boldsymbol{C}_{i,b_k}\big)\big)\boldsymbol{\mathcal{G}}_k\big]\}\rfloor_+ \end{aligned} \quad (43)$$

where we noted that $\boldsymbol{\mathcal{F}}_k = \boldsymbol{\mathcal{F}}_k\boldsymbol{W}_k\overline{\boldsymbol{E}}_k = \boldsymbol{\mathcal{F}}_k\boldsymbol{W}_k(\boldsymbol{I} - \boldsymbol{\mathcal{F}}_k^{\mathrm{H}}\widehat{\boldsymbol{H}}_{k,b_k}\boldsymbol{\mathcal{G}}_k)$ and $\lfloor x\rfloor_+ = x$ if $x \geq 0$ and is zero otherwise. This nonnegativity constraint on μ_c stems from the fact that $\mu_c = -\frac{\partial\mathcal{L}}{\partial P_{max,c}} \geq 0$ since indeed the WSMSE can only get smaller if we allow a larger power budget. We then get $\theta_c = \mu_c/\sum_{c'}\mu_{c'}$.

The Perron-Frobenius theory also allows for the optimization of the Lagrange multipliers λ_k. With (34), we can reformulate (40) as

$$\Delta = \max_{\lambda:\sum_k\lambda_k=1}\min_{\boldsymbol{p}}\sum_k \lambda_k\,\frac{[\boldsymbol{\Lambda}\,\boldsymbol{p}]_k}{p_k} \quad (44)$$

which is the Donsker–Varadhan–Friedland formula [21, Chapter 8] for the Perron Frobenius eigenvalue of Λ. A related formula is the Rayleigh quotient

$$\Delta = \max_{\boldsymbol{q}}\min_{\boldsymbol{p}}\frac{\boldsymbol{q}^T\boldsymbol{\Lambda}\,\boldsymbol{p}}{\boldsymbol{q}^T\boldsymbol{p}} \quad (45)$$

where $\boldsymbol{p}, \boldsymbol{q}$ are the right and left Perron Frobenius eigenvectors. Comparing (45) to (44), then apart from normalization factors, we get $\lambda_k/p_k = q_k$ or hence $\lambda_k = p_k q_k$.

The proposed optimization framework is summarized in Table I. Superscripts refer to iteration numbers. The algorithm in Table I is based on a double loop. The inner loop solves the WEMSE balancing problem in (23) whereas the outer loop iteratively transforms the WEMSE balancing problem into the original rate balancing problem in (14). The proof of convergence of this transformation is similar to the one in [12].

VII. ESIP Rate Balancing

Now we follow another approximation of the expected rate expression The following approach will use a rate minorizer for every r_k, similar but not identical to what is used as in the DC programming approach which for the optimization of $\boldsymbol{G}_k$ keeps r_k and linearizes the $r_{\overline{k}}$. The approach does not require the introduction of Rxs. We consider again the (ergodic) rate balancing problem (14) where $\overline{r}_k = \mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\,r_k$

TABLE I: WEMSE based User Rate Balancing

1. initialize: $\boldsymbol{G}_k^{(0,0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})^{\mathrm{T}}$, $\boldsymbol{p}_k^{(0,0)} = \boldsymbol{q}_k^{(0,0)} = \frac{P_{\max,c}}{K_c}$, $m = n = 0$ and fix $n_{\max}, m_{\max}$ and $r_k^{\circ(0)} = r_k^{\circ}$, initialize $\boldsymbol{W}_k^{(0)} = \boldsymbol{I}_{d_k}$ and $\xi_k^{(0)} = d_k$
2. initialize $\boldsymbol{F}_k^{(0,0)}$ in $\boldsymbol{\mathcal{F}}_k^{(0,0)} = p_k^{(0,0)-1/2}\boldsymbol{F}_k$ from (20)
3. **repeat**
   - 3.1. $m \leftarrow m+1$
   - 3.2. **repeat**
     - $n \leftarrow n+1$
     - *i* update $\boldsymbol{G}_k$, $\boldsymbol{\mathcal{G}}_k$, μ_c from (42),(43)
     - *ii* update $\boldsymbol{F}_k = p_k^{1/2}\boldsymbol{\mathcal{F}}_k$ from (20)
     - *iii* update $\boldsymbol{p}$ and $\boldsymbol{q}$ using (45)
   - 3.3 **until** required accuracy is reached or $n \geq n_{\max}$
   - 3.4 compute $\overline{\boldsymbol{E}}_k^{(m)}$ and update $\boldsymbol{W}_k^{(m)} = (\overline{\boldsymbol{E}}_k^{(m)})^{-1}$
   - 3.5 determine $t = \min_k \frac{\overline{r}_k^{l\,(m)}}{r_k^{\circ(m-1)}}$, $r_k^{\circ(m)} = t\, r_k^{\circ(m-1)}$, and $\xi_k^{(m)} = d_k + \overline{r}_k^{l\,(m)} - r_k^{\circ(m)}$
   - 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 $m \geq m_{\max}$

is now approximated by the Expected Signal and Interference Power (ESIP) rate

$$
\begin{aligned}
\overline{r}_k = &\ \mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \,\mathrm{lndet}\Big(\boldsymbol{I}+p_k\, \boldsymbol{G}_k^{\mathrm{H}}\, \boldsymbol{H}_{k,b_k}^{\mathrm{H}} \boldsymbol{R}_{\overline{k}}^{-1} \boldsymbol{H}_{k,b_k} \boldsymbol{G}_k\Big) \\
\approx &\ \mathrm{lndet}\Big(\boldsymbol{I}+p_k\, \boldsymbol{G}_k^{\mathrm{H}}\ \mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\{\boldsymbol{H}_{k,b_k}^{\mathrm{H}} (\,\mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}} \boldsymbol{R}_{\overline{k}})^{-1} \boldsymbol{H}_{k,b_k}\}\boldsymbol{G}_k\Big) \\
= &\ \overline{r}_k^s = f_k^s(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}) = \ \mathrm{lndet}\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\, \overline{\boldsymbol{B}}_k(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})\, \boldsymbol{G}_k\Big)\,, \qquad (46)
\end{aligned}
$$

$$
\overline{\boldsymbol{B}}_k(\overline{\boldsymbol{T}}_k) = \widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}}\, \overline{\boldsymbol{T}}_k^{-1}\, \widehat{\boldsymbol{H}}_{k,b_k} + \mathrm{tr}\{\overline{\boldsymbol{T}}_k^{-1}\}\, \boldsymbol{C}_{k,b_k} \qquad (47)
$$

where the $\overline{r}_k$ approximation $\overline{r}_k^s$ in (46) in general is neither an upper nor lower bound but in the Massive MIMO limit becomes a tight upper bound.

Lemma 2. *The approximate $\overline{r}_k$, $\overline{r}_k^s$, can be obtained as $f_k^s(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}) = \min_{\overline{\boldsymbol{T}}_k} \underline{f}_k^s(\overline{\boldsymbol{T}}_k, \frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})$, with $\underline{f}_k^s(\overline{\boldsymbol{T}}_k, \frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})$:*

$$
\underline{f}_k^s = \mathrm{lndet}\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k(\overline{\boldsymbol{T}}_k)\, \boldsymbol{G}_k\Big)+\mathrm{tr}\{\breve{\boldsymbol{W}}_k(\overline{\boldsymbol{T}}_k-\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})\} \quad (48)
$$

where

$$
\breve{\boldsymbol{W}}_k = \overline{\boldsymbol{T}}_k^{-1}\big(\widehat{\boldsymbol{H}}_{k,b_k} \boldsymbol{X}_k\, \widehat{\boldsymbol{H}}_{k,b_k}^{\mathrm{H}} + \mathrm{tr}\{\boldsymbol{X}_k\, \boldsymbol{C}_{k,b_k}\}\boldsymbol{I}\big)\overline{\boldsymbol{T}}_k^{-1} \quad (49)
$$

$$
\textit{with } \boldsymbol{X}_k = \boldsymbol{G}_k\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k(\overline{\boldsymbol{T}}_k)\, \boldsymbol{G}_k\Big)^{-1} \boldsymbol{G}_k^{\mathrm{H}} \quad (50)
$$

The optimizer is $\overline{\boldsymbol{T}}_k = \frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}$. Also, $\underline{f}_k^s$ is a minorizer for $f_k^s(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}})$ as a function of $\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}$.

Indeed, since $f_k^s(.)$ is a convex function, it gets minorized by its tangent at any point:

$$
f_k^s(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}) \geq \underline{f}_k^s = f_k^s(\overline{\boldsymbol{T}}_k)+\mathrm{tr}\{\frac{\partial f_k^s(\overline{\boldsymbol{T}}_k)}{\partial \overline{\boldsymbol{T}}_k}\,(\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}-\overline{\boldsymbol{T}}_k)\} \quad (51)
$$

and $\breve{\boldsymbol{W}}_k = -\frac{\partial f_k^s(\overline{\boldsymbol{T}}_k)}{\partial \overline{\boldsymbol{T}}_k}$. Note that for the Perron-Frobenius theory, we need a function that is linear in $\frac{\boldsymbol{p}_{\overline{k}}}{p_k}$, hence we need to work with $\frac{1}{p_k}\overline{\boldsymbol{R}}_{\overline{k}}$ instead of $\overline{\boldsymbol{R}}_{\overline{k}}$.

The modifications in the Lagrangian formulation in Section VI are now

$$
\begin{aligned}
&\sum_k \breve{\lambda}_k^{'}(t\, r_k^o - \underline{f}_k^s) \\
&= -\sum_k \breve{\lambda}_k^{'}\big(\,\mathrm{lndet}\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k\, \boldsymbol{G}_k\Big) - \frac{1}{p_k}\,\mathrm{tr}\{\breve{\boldsymbol{W}}_k\overline{\boldsymbol{R}}_{\overline{k}}\} \qquad (52) \\
&+\mathrm{tr}\{\breve{\boldsymbol{W}}_k\overline{\boldsymbol{T}}_k\}-t\, r_k^o\big) = \sum_k \breve{\lambda}_k(\frac{1}{p_k\, \breve{\xi}_k}\,\mathrm{tr}\{\breve{\boldsymbol{W}}_k\overline{\boldsymbol{R}}_{\overline{k}}\}-1) \qquad (53)
\end{aligned}
$$

$$
\text{where } \breve{\xi}_k = \mathrm{tr}\{\breve{\boldsymbol{W}}_k\overline{\boldsymbol{T}}_k\} + \mathrm{lndet}\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k\, \boldsymbol{G}_k\Big) - t\, r_k^o \qquad (54)
$$

and $\breve{\lambda}_k^{'} = \breve{\lambda}_k/\breve{\xi}_k$, $\overline{\boldsymbol{B}}_k = \overline{\boldsymbol{B}}_k(\overline{\boldsymbol{T}}_k)$. The balancing of the rates in (14) or equivalently the weighted interference plus noise powers in (52) now leads to the same problem formulation as in (34) with this time

$$
\breve{\boldsymbol{\Lambda}} = \breve{\boldsymbol{\xi}}^{-1}\breve{\boldsymbol{\Psi}} + \frac{1}{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{p}_{\max}}\breve{\boldsymbol{\xi}}^{-1}\breve{\boldsymbol{\sigma}}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{C}_C^{\mathrm{T}} \text{ with} \qquad (55)
$$

$$
[\breve{\boldsymbol{\Psi}}]_{ij} = \begin{cases} \mathrm{tr}\{\breve{\boldsymbol{W}}_i(\widehat{\boldsymbol{H}}_{i,b_j}\boldsymbol{G}_j\boldsymbol{G}_j^{\mathrm{H}}\widehat{\boldsymbol{H}}_{i,b_j}^{\mathrm{H}}+\mathrm{tr}\{\boldsymbol{G}_j^{\mathrm{H}}\boldsymbol{C}_{i,b_j}\boldsymbol{G}_j\}\boldsymbol{I})\}, & i \neq j \\ 0, & i = j \end{cases} \qquad (56)
$$

$$
\boldsymbol{\sigma}_i = \sigma_n^2\, \mathrm{tr}\{\breve{\boldsymbol{W}}_i\}, \breve{\boldsymbol{\xi}} = \mathrm{diag}(\breve{\xi}_1, \ldots, \breve{\xi}_K)\,. \qquad (57)
$$

The Tx BF and stream power optimization will be based on $\sum_i \frac{\breve{\lambda}_i}{\breve{\xi}_i}\underline{f}_i^s$, which from (52) becomes (apart from noise terms)

$$
\sum_k \frac{\breve{\lambda}_k}{\breve{\xi}_k}\underline{f}_k^s = \sum_k \frac{\breve{\lambda}_k}{\breve{\xi}_k}\,\mathrm{lndet}\Big(\boldsymbol{I}+\boldsymbol{G}_k^{\mathrm{H}}\overline{\boldsymbol{B}}_k\, \boldsymbol{G}_k\Big) - \sum_k \mathrm{tr}\{p_k\boldsymbol{G}_k^{H}\overline{\boldsymbol{A}}_k\boldsymbol{G}_k\} \qquad (58)
$$

$$
\text{with } \overline{\boldsymbol{A}}_k = \sum_{i\neq k} \frac{\breve{\lambda}_i}{p_i\, \breve{\xi}_i}\big(\widehat{\boldsymbol{H}}_{i,b_k}^{\mathrm{H}}\breve{\boldsymbol{W}}_i\widehat{\boldsymbol{H}}_{i,b_k}+\mathrm{tr}\{\breve{\boldsymbol{W}}_i\}\boldsymbol{C}_{i,b_k}\big)\,. \qquad (59)
$$

For the optimal Tx BF $\boldsymbol{G}_k$, the gradient of $\sum_i \frac{\breve{\lambda}_i}{\breve{\xi}_i}\underline{f}_i^s - \mu_{b_k}\sum_{i:b_i=b_k} p_i\mathrm{tr}\{\boldsymbol{G}_i^{\mathrm{H}}\boldsymbol{G}_i\}$ with (58) (or (46)) yields

$$
\frac{\breve{\lambda}_k}{p_k\, \breve{\xi}_k}\overline{\boldsymbol{B}}_k\boldsymbol{G}_k\,(\boldsymbol{I}+\boldsymbol{G}_k^{H}\overline{\boldsymbol{B}}_k\boldsymbol{G}_k)^{-1} - (\overline{\boldsymbol{A}}_k + \mu_{b_k}\boldsymbol{I})\,\boldsymbol{G}_k = 0\,. \qquad (60)
$$

The solution is the d_k maximal generalized eigen vectors

$$
\boldsymbol{G}_k^{'} = V_{1:d_k}(\overline{\boldsymbol{B}}_k, \overline{\boldsymbol{A}}_k + \mu_{b_k}\boldsymbol{I}), \boldsymbol{G}_k = \boldsymbol{G}_k^{'}\overline{\boldsymbol{P}}_k^{1/2}, \boldsymbol{\mathcal{G}}_k = \boldsymbol{G}_k\sqrt{p_k} \qquad (61)
$$

where the $\overline{\boldsymbol{P}}_k = \mathrm{diag}(p_{k,1}, \ldots, p_{k,d_k})$, $\mathrm{tr}\{\overline{\boldsymbol{P}}_k\} = 1$, are the relative stream powers. Indeed, (60) represents the definition of generalized eigen vectors. Consider

$$
\Sigma_k^{(1)} = \boldsymbol{G}_k^{'\mathrm{H}}\overline{\boldsymbol{B}}_k\boldsymbol{G}_k^{'},\ \Sigma_k^{(2)} = \boldsymbol{G}_k^{'\mathrm{H}}\overline{\boldsymbol{A}}_k\boldsymbol{G}_k^{'} \qquad (62)
$$

then the generalized eigen vectors $\boldsymbol{G}_k^{'}$ of $\overline{\boldsymbol{B}}_k, \overline{\boldsymbol{A}}_k + \mu_{b_k}\boldsymbol{I}$ lead to diagonal matrices $\Sigma_k^{(1)}$, $\Sigma_k^{(2)} + \mu_{b_k}\boldsymbol{G}_k^{'\mathrm{H}}\boldsymbol{G}_k^{'}$. Note that the normalized $\boldsymbol{G}_k^{'}$ are not orthogonal. Then (60) represents the generalized eigen vector condition with associated generalized eigen values in the diagonal matrix $\frac{p_k\, \breve{\xi}_k}{\breve{\lambda}_k}(\boldsymbol{I} + \Sigma_k^{(1)}\overline{\boldsymbol{P}}_k)$. Also, plugging in generalized eigen vectors into (58) reveals that one should choose the eigen vectors associated to d_k maximal eigen values to maximize (58). Now, premultiplying both sides of (60) by $p_k\boldsymbol{G}_k^{\mathrm{H}}$, summing over all users $k : b_k = c$, taking trace and identifying the last term with $\sum_{k:b_k=c} p_k\mathrm{tr}\{\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{G}_k\} = P_{max,c}$ allows to solve for

TABLE II: ESIPrate based User Rate Balancing

1. initialize: $\boldsymbol{G}_k^{(0,0)} = (\boldsymbol{I}_{d_k} : \boldsymbol{0})^{\mathrm{T}}$, $\boldsymbol{p}_k^{(0,0)} = \boldsymbol{q}_k^{(0,0)} = \frac{P_{\max,c}}{K}$, $m = n = 0$ and fix $n_{\max}, m_{\max}$, r_k°, and $\breve{\boldsymbol{W}}_k^{(0)}$ from (49)
2. compute $\overline{r}_k^{s\,(0)} = \mathrm{lndet}\Big(\boldsymbol{I} + \boldsymbol{G}_k^{\mathrm{H}} \overline{\boldsymbol{B}}_k (\frac{1}{p_k} \overline{\boldsymbol{R}}_{\overline{k}})\ \boldsymbol{G}_k\Big)$, determine $t = \min_k \frac{\overline{r}_k^{s\,(0)}}{r_k^{\circ}}$, $r_k^{\circ(0)} = t\, r_k^{\circ}$, and $\breve{\xi}_k^{(0)}$ from (54)
3. **repeat**
   - 3.1. $m \leftarrow m+1$
   - 3.2. **repeat**
     - $n \leftarrow n+1$
     - *i* update $\overline{\boldsymbol{A}}_k$ from (59)
     - *ii* update μ_c and $\boldsymbol{G}_k^{'}$ from (61),(63)
     - *iii* update $\overline{\boldsymbol{P}}_k$ from (65)
     - *iv* update $\boldsymbol{p}$ and $\boldsymbol{q}$ as maximal eigen vectors of $\breve{\boldsymbol{\Lambda}}$ in (55)
   - 3.3 **until** required accuracy is reached or $n \geq n_{\max}$
   - 3.4 compute $\overline{\boldsymbol{B}}_k(\overline{\boldsymbol{T}}_k)$ and update $\breve{\boldsymbol{W}}_k$ from (49)
   - 3.5 compute $\overline{r}_k^{s\,(m)} = \mathrm{lndet}\Big(\boldsymbol{I} + \boldsymbol{G}_k^{\mathrm{H}} \overline{\boldsymbol{B}}_k (\frac{1}{p_k} \overline{\boldsymbol{R}}_{\overline{k}})\ \boldsymbol{G}_k\Big)$ and determine $t = \min_k \frac{\overline{r}_k^{s\,(m)}}{r_k^{\circ(m-1)}}$, $r_k^{\circ(m)} = t\, r_k^{\circ(m-1)}$, and update $\breve{\xi}_k$ from (54)
   - 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 $m \geq m_{\max}$

$$\mu_c = \frac{1}{P_{max,c}} \left\lfloor \sum_{k:b_k=c} \mathrm{tr}\{ \frac{\breve{\lambda}_k}{\breve{\xi}_k} \Sigma_k^{(1)} \overline{\boldsymbol{P}}_k (\boldsymbol{I} + \Sigma_k^{(1)} \overline{\boldsymbol{P}}_k)^{-1} - p_k \Sigma_k^{(2)} \overline{\boldsymbol{P}}_k \} \right\rfloor_+ . \tag{63}$$

The $\overline{\boldsymbol{P}}_k$ are themselves found from an interference leakage aware water filling (ILAWF) operation. Substituting $\boldsymbol{G}_k^{'}$ into term k of (58), dividing by p_k, and accounting for the constraint $\mathrm{tr}\{\overline{\boldsymbol{P}}_k\} = 1$ by Lagrange multiplier ν_k, we get the Lagrangian

$$\begin{aligned} &\frac{\breve{\lambda}_k}{p_k\, \breve{\xi}_k} \ln\det\big(\boldsymbol{I} + \Sigma_k^{(1)} \overline{\boldsymbol{P}}_k\big) - \mathrm{tr}\{(\Sigma_k^{(2)} + \nu_k \boldsymbol{I}) \overline{\boldsymbol{P}}_k\} = \\ &\frac{\breve{\lambda}_k}{p_k\, \breve{\xi}_k} \ln\det\big(\boldsymbol{I} + \Sigma_k^{(1)} \overline{\boldsymbol{P}}_k\big) - \mathrm{tr}\{(\mathrm{diag}(\Sigma_k^{(2)}) + \nu_k \boldsymbol{I}) \overline{\boldsymbol{P}}_k\}. \end{aligned} \tag{64}$$

Maximizing w.r.t. $\overline{\boldsymbol{P}}_k$ leads to the ILAWF

$$\overline{\boldsymbol{P}}_k = \left\lfloor \frac{\breve{\lambda}_k}{p_k\, \breve{\xi}_k} (\mathrm{diag}(\Sigma_k^{(2)}) + \nu_k \boldsymbol{I})^{-1} - \Sigma_k^{-(1)} \right\rfloor_+ \tag{65}$$

where the Lagrange multiplier ν_k is adjusted (e.g. by bisection) to satisfy $\mathrm{tr}\{\overline{\boldsymbol{P}}_k\} = 1$. Elements in $\overline{\boldsymbol{P}}_k$ corresponding to zeros in $\Sigma_k^{(1)}$ should also be zero. This completes the ESIP rate balancing algorithm derivation (Table II).

VIII. SIMULATION RESULTS

In this section, we numerically evaluate the performance of the proposed algorithms. For the multipath channel model,

$$\boldsymbol{C}_t = \sum_{n=1}^{N_p} \frac{\alpha_i}{\boldsymbol{v}_i^H \boldsymbol{v}_i} \boldsymbol{v}_i \boldsymbol{v}_i^H \tag{66}$$

with $\mathrm{tr}\{\boldsymbol{C}_t\} = \sum_{n=1}^{N_p} \alpha_i = M_c$, $\alpha_i = c^{i-1}\alpha_1$ (we use $c = 0.3$) and the $\boldsymbol{v}_i$ are i.i.d. vectors of M_c i.i.d. elements $\mathcal{CN}(0,1)$. We take $N_p = M_c/K$. For all simulations, we take $n_{\max} = 20$, though typically 2-3 inner loop iterations suffice. The algorithm converges after 4-5 (or 13-15) (outer) iterations of m at SNR = 15dB (or 40dB). For all (partial CSIT) algorithms, we evaluate the actual expected rate $\overline{r}_k = \mathrm{E}_{\boldsymbol{H}|\widehat{\boldsymbol{H}}}\, r_k$ by Monte Carlo averaging over 500 channel realizations. The partial CSIT algorithms evaluated are the proposed WEMSE and ESIPrate, and also Naive Partial CSIT which corresponds to perfect CSIT by assuming the channel estimates to be the true channels. Perfect CSIT algorithms are obtained from WEMSE or ESIPrate by setting $\rho_D = \infty$. We also evaluate LB WEMSE, which considers $\ln(\det(\boldsymbol{W}_k \overline{\boldsymbol{E}}_k))$, and UB ESIPrate, which corresponds to (46) (Lower/Upper Bounds).

Figure 1 represents the average attained rate using the proposed algorithms for different configurations of the system (single and multi-cell). We can see that ESIPrate outperforms WEMSE and suffers little loss compared to perfect CSIT, and that the UB ESIPrate provides a tight upper bound. Note also that for fixed ρ_D as considered here, Naive saturates at high SNR, whereas WEMSE appears to suffer Degree-of-Freedom (DoF) (slope) loss.

In Figure 2, we consider varying levels of channel estimation error $\sigma^2_{\widetilde{\boldsymbol{H}}}$. It is clear that when $\rho_D = 1/\sigma^2_{\widetilde{\boldsymbol{H}}}$ is proportional to SNR, all algorithms (only) suffer from varying SNR offset, but ESIPrate still outperforms WEMSE for which the signal link channel error covariance is accounted for in the interference power instead of in the signal power.

Figure 3 illustrates the convergence of the user rates w.r.t. the number of iterations for one channel realization, with SNR = 20dB and $\rho_D = 10$. We observe that the rates obtained using (46) and $\ln(\det(\boldsymbol{W}_k \overline{\boldsymbol{E}}_k))$ are balanced. Both approaches converge after about 5 iterations of the outer loop, while the inner loop converges after 2-3 iterations. Of course, due to the CSIT imperfections, the actual rates exhibit some randomness.

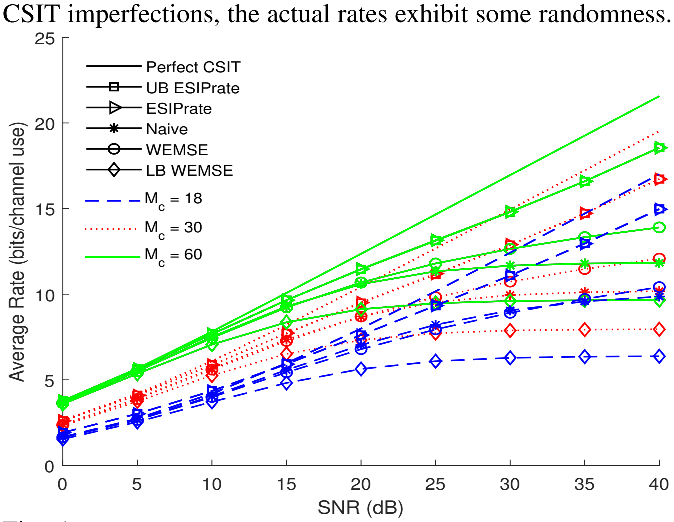

Fig. 1: Average Rate with Partial CSIT vs. SNR, $C = 2, K = 3, \rho_D = 10, N_r = d_k = 2$.

IX. CONCLUSIONS

In this work, we addressed the multiple streams per MIMO user case for which we considered user Erate (Expected rate) balancing, not stream Erate balancing, in a multicell downlink channel. Actually, we optimized the Erate distribution

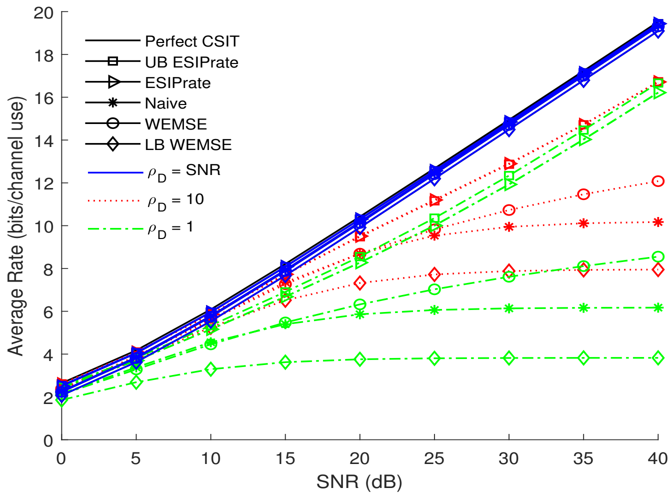

Fig. 2: Average Rate with Partial CSIT vs. SNR for different instances of ρ_D, $C = 2, K = 3, M_c = 30, N_k = d_k = 2$.

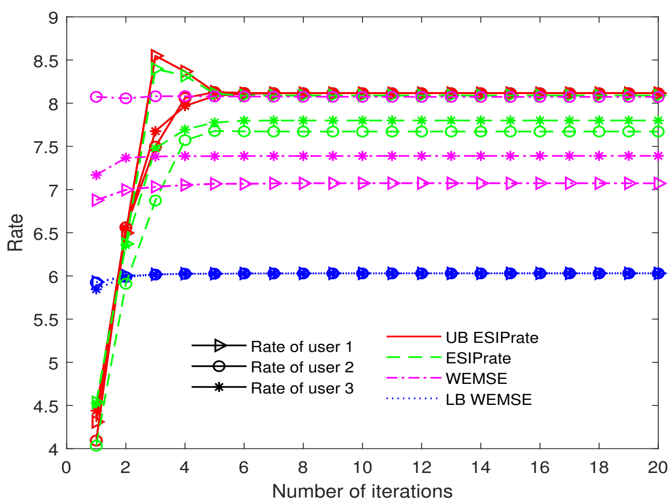

Fig. 3: User Rates with Partial CSIT vs. number of outer iterations, $C = 1, K = 3, M_c = 12, N_k = d_k = 2, \rho_D = 10$.

over the streams of a user, within the user Erate balancing under per cell power constraints. We transformed the max-min Erate optimization problem into a min-max weighted EMSE optimization problem which itself was shown to be related to a weighted sum EMSE minimization via Lagrangian duality, involving linearizing the EMSE balancing problem by transforming to EMSE constraints. The associated Lagrange multipliers and user powers get found as left and right eigen vectors of a weighted interference matrix in the Perron-Frobenius theory. We also considered the ESIP Erate approximation, for which we introduced an original minorizer, judiciously chosen to be amenable to the Perron-Frobenius theory. We furthermore introduced original explicit power constraint Lagrange multiplier solutions, which can handle the case in which some cell power constraints are met with inequality, as can happen in a multi-cell scenario. The simulation results exhibit the different SNR behavior of the Erate lower bound vs. actual Erate, showed that the upper bound is a quite tight approximation, and that the ESIP partial CSIT approach with LMMSE channel estimation leads to very limited performance loss compared to perfect CSIT. In the multi-cell case, the proposed algorithms can handle scenarios in which the CSIT quality could be very different between intracell and intercell links.

Acknowledgements

EURECOM's research is partially supported by its industrial members: ORANGE, BMW, Symantec, SAP, Monaco Telecom, iABG, and by the project SPOTLIGHT (EU ITN).

References

[1] A. Goldsmith, *Wireless Communications*. Combridge University Press, 2005.
[2] D. Cai, T. Quek, C. W. Tan, and S. H. Low, "Max-Min SINR Coordinated Multipoint Downlink Transmission—Duality and Algorithms," in *IEEE Transactions on Signal Processing*, vol. 60, no. 10, Oct 2012.
[3] M. Schubert and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem with Individual SINR Constraints," *IEEE Trans. Vehic. Tech.*, Jun 2004.
[4] W. Yu and T. Lan, "Transmitter Optimization for the Multi-antenna Downlink with Per-antenna Power Constraints," in *IEEE Trans. Signal Processing*, June 2007.
[5] 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.
[6] 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.
[7] 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.
[8] ——, "Capacity Balancing for Multiuser MIMO Systems," in *Proc. IEEE ICASSP*, Apr 2007.
[9] 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.
[10] 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.
[11] D. W. H. Cai, T. Q. S. Quek, C. W. Tan, and S. H. Low, "Max-Min SINR Coordinated Multipoint Downlink Transmission—Duality and Algorithms," in *IEEE Trans. Signal Processing*, Oct. 2012.
[12] 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.
[13] ——, "Rate balancing for multiuser multicell downlink MIMO Systems," in *27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain*, 2019.
[14] ——, "MIMO user rate balancing in multicell networks with per cell power constraints," in *IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium*, 2020.
[15] T. E. Bogale, B. K. Chalise, and L. Vandendorpe, "Robust Transceiver Optimization for Downlink Multiuser MIMO Systems," *IEEE Trans. Signal Processing*, Jan. 2011.
[16] S. Malla and G. Abreu, "Transmission strategies under imperfect instantaneous csit," *IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA*, 2015.
[17] C. Li, C. He, L. Jiang, and F. Liu, "Robust beamforming design for max–min sinr in mimo interference channels," *IEEE Communications Letters*, April 2016.
[18] M. Joham, A. Gründinger, J. R. F. A. Pastore, and W. Utschick, "Rate balancing in the vector bc with erroneous csi at the receivers," *47th Annual Conference on Information Sciences and Systems (CISS), Baltimore, MD*, 2013.
[19] D. Neumann, A. Gründinger, M. Joham, and W. Utschick, "Rate-balancing in massive mimo using statistical precoding," *IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm*, 2015.
[20] 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.
[21] C. Meyer, *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.