

Maximizing Sum Rate and Minimizing MSE on Multiuser Downlink: Optimality, Fast Algorithms and Equivalence via Max-min SIR

Chee Wei Tan^{1,2}, Mung Chiang² and R. Srikant³

¹ California Institute of Technology, CA 91125 USA

² Princeton University, NJ 08544 USA

³ University of Illinois at Urbana-Champaign, IL 61801 USA

Abstract—Maximizing the minimum weighted SIR, minimizing the weighted sum MSE and maximizing the weighted sum rate in a multiuser downlink system are three important performance objectives in joint transceiver and power optimization, where all the users have a total power constraint. We show that, through connections with the nonlinear Perron-Frobenius theory, jointly optimizing power and beamformers in the max-min weighted SIR problem can be solved optimally in a distributed fashion. Then, connecting these three performance objectives through the arithmetic-geometric mean inequality and nonnegative matrix theory, we solve the weighted sum MSE minimization and weighted sum rate maximization in the low to moderate interference regimes using fast algorithms.

I. INTRODUCTION

In this paper, we focus on the downlink transmission, where the transmitter (at the base station) is equipped with an antenna array and each user has a single receive antenna. Full channel information is available at the transmitter to adapt the beamformers to minimize interference. All the users share the same bandwidth and meet a total power constraint.

We consider a joint optimization of power and transmit beamformer for the min-max (weighted) mean squared error (MSE) problem or, equivalently, the max-min (weighted) signal-to-interference ratio (SIR) problem. This problem is challenging to solve, because the transmit beamformers are coupled across users, making them hard to optimize in a distributed fashion. While previous algorithms in the literature require centralized computation of the eigenvalue and eigenvector of an extended coupling matrix, we propose a fast distributed algorithm that computes the optimal power and transmit beamformer in the max-min weighted SIR problem with geometric convergence rate. This is achieved by applying the nonlinear Perron-Frobenius theory in [1], [2], [3] and the uplink-downlink duality in [4], [5], [6], [7], [8], [9], wherein the uplink acts as an intermediate mechanism to optimize transmit beamformers in the downlink.

We next keep the beamformers fixed and study the nonconvex problems of, 1) minimizing the weighted sum MSE between the transmitted and estimated symbols, and, 2) maximizing the weighted sum rate. The max-min SIR problem is shown to be a special case of these two problems. Previous work in the literature, see e.g., [10], only solve these two

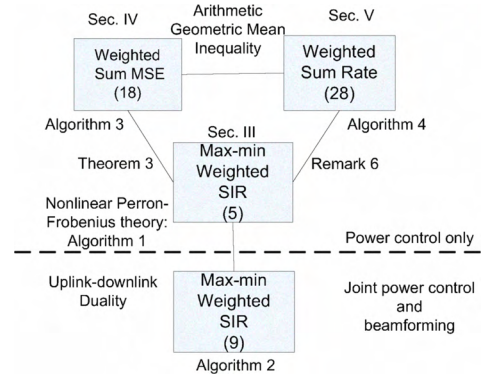


Fig. 1. Overview of the connection (solid lines) between the three optimization problems in the paper: i) Weighted sum MSE minimization in (18), ii) weighted sum rate maximization in (28), and iii) max-min weighted SIR in (5). The upper half of the dotted line considers power control only, while the lower half considers both power control and beamforming.

nonconvex problems suboptimally. We develop fast algorithms (independent of stepsize) to solve these two nonconvex problems optimally under low to medium interference conditions. We leverage the standard interference function approach in [11] to show that our algorithms converge under synchronous and asynchronous updates. Proofs can be found in [12].

We refer the readers to Figure 1 for an overview of the connection between the three main optimization problems in the paper. The following notations are used. Boldface upper-case letters denote matrices, boldface lowercase letters denote column vectors, italics denote scalars, and $\mathbf{u} \geq \mathbf{v}$ ($\mathbf{B} \geq \mathbf{F}$) denotes componentwise inequality between vectors \mathbf{u} and \mathbf{v} (matrices \mathbf{B} and \mathbf{F}). We let $(\mathbf{B}\mathbf{y})_l$ denote the l th element of $\mathbf{B}\mathbf{y}$. The Perron-Frobenius eigenvalue of a nonnegative matrix \mathbf{F} is denoted as $\rho(\mathbf{F})$, and the Perron (right) and left eigenvectors of \mathbf{F} associated with $\rho(\mathbf{F})$ are denoted by $\mathbf{x}(\mathbf{F})$ and $\mathbf{y}(\mathbf{F})$, respectively. The super-scripts $(\cdot)^\top$ and $(\cdot)^\dagger$ denote transpose and complex conjugate transpose respectively. We let \mathbf{e}_l denote the l th unit coordinate vector and \mathbf{I} denote the identity matrix. Let $\mathbf{x} \circ \mathbf{y}$ denote $\mathbf{x} \circ \mathbf{y} = [x_1 y_1, \dots, x_L y_L]^\top$.

II. SYSTEM MODEL

We consider a single cell multiuser system with N antennas at the base station and L decentralized users, each equipped with a single receive antenna, operating in a frequency-flat fading channel. The downlink channel can be modeled as a vector Gaussian broadcast channel:

$$\mathbf{y}_l = \mathbf{h}_l^\dagger \mathbf{x} + \mathbf{z}_l, \quad l = 1, \dots, L, \quad (1)$$

where $\mathbf{y}_l \in \mathbb{C}^{1 \times 1}$ is the received signal of the l th user, $\mathbf{h}_l \in \mathbb{C}^{N \times 1}$ is the channel matrix between the base station and the l th user, $\mathbf{x} \in \mathbb{C}^{N \times 1}$ is the transmitted signal vector, and \mathbf{z}_l 's are the i.i.d. additive complex Gaussian noise vectors with variance $n_l/2$ on each of its real and imaginary components. We assume that the multiuser system adopts a linear transmission and reception strategy. In transmit beamforming, the base station transmits a signal \mathbf{x} in the form of $\mathbf{x} = \sum_{l=1}^L d_l \hat{\mathbf{w}}_l$, where $\hat{\mathbf{w}}_l \in \mathbb{C}^{N \times 1}$ is the transmit beamformer that carries the information signal d_l of the l th user. We assume a total power constraint at the transmit antennas, i.e., $\mathbb{E}[\mathbf{x}^\dagger \mathbf{x}] = \bar{P}$. From (1), the received signal for the l th user can be expressed as $\mathbf{y}_l = (\mathbf{h}_l^\dagger \hat{\mathbf{w}}_l) d_l + \sum_{j \neq l} (\mathbf{h}_l^\dagger \hat{\mathbf{w}}_j) d_j + \mathbf{z}_l$. Next, we write $\hat{\mathbf{w}}_l = p_l \mathbf{u}_l$, where p_l is the downlink transmit power and \mathbf{u}_l is the normalized transmit beamformer, i.e., $\mathbf{u}_l^\dagger \mathbf{u}_l = 1$, of the l th user. Now, the received SIR of the l th user in the downlink transmission can be given in terms of \mathbf{p} and $\mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_L]$:

$$\text{SIR}_l(\mathbf{p}, \mathbf{U}) = \frac{p_l |\mathbf{h}_l^\dagger \mathbf{u}_l|^2}{\sum_{j \neq l} p_j |\mathbf{h}_l^\dagger \mathbf{u}_j|^2 + n_l}. \quad (2)$$

We define the matrix \mathbf{G} with entries $G_{lj} = |\mathbf{h}_l^\dagger \mathbf{u}_j|^2$ in the downlink transmission. In terms of the beamforming matrix \mathbf{U} , we also define the (cross channel interference) matrix $\mathbf{F}(\mathbf{U})$ with entries:

$$F_{lj}(\mathbf{U}) = \begin{cases} 0, & \text{if } l = j \\ \frac{G_{lj}(\mathbf{U})}{G_{ll}(\mathbf{U})}, & \text{if } l \neq j \end{cases} \quad (3)$$

and

$$\mathbf{v}(\mathbf{U}) = \left(\frac{n_1}{G_{11}(\mathbf{U})}, \frac{n_2}{G_{22}(\mathbf{U})}, \dots, \frac{n_L}{G_{LL}(\mathbf{U})} \right)^\top. \quad (4)$$

For brevity, we omit the dependency on \mathbf{U} when we fix the beamformers.

III. MAX-MIN WEIGHTED SIR MAXIMIZATION

In this section, we first consider optimizing only power before we consider a joint optimization between power and transmit beamformers. Let β be a positive vector, where the l th entry β_l is assigned by the network to the l th link (to reflect some long-term priority). We first consider the following max-min weighted SIR problem:

$$\begin{aligned} & \text{maximize} \quad \min_l \frac{\text{SIR}_l(\mathbf{p})}{\beta_l} \\ & \text{subject to} \quad \mathbf{1}^\top \mathbf{p} \leq \bar{P}, \quad \mathbf{p} \geq \mathbf{0}, \\ & \text{variables:} \quad \mathbf{p}. \end{aligned} \quad (5)$$

Note that (5) is equivalent to the min-max weighted MSE problem: $\min_{\mathbf{p}} \max_l \beta_l / (1 + \text{SIR}_l(\mathbf{p}))$ subject to $\mathbf{1}^\top \mathbf{p} \leq \bar{P}$.

Next, let us define the matrix

$$\mathbf{B} = \mathbf{F} + (1/\bar{P})\mathbf{v}\mathbf{1}^\top. \quad (6)$$

By exploiting a connection between the nonlinear Perron-Frobenius theory in [1], [2] and the algebraic structure of (5), we can give a closed form solution to (5).¹

Lemma 1: The optimal objective and solution of (5) is given by $1/\rho(\text{diag}(\beta)\mathbf{B})$ and $(\bar{P}/\mathbf{1}^\top \mathbf{x}(\text{diag}(\beta)\mathbf{B}))\mathbf{x}(\text{diag}(\beta)\mathbf{B})$ respectively.

The following algorithm computes the optimal power of (5) given in Lemma 1. We let k index discrete time slots.

Algorithm 1 (Max-min weighted SIR):

1) Update power $\mathbf{p}(k+1)$:

$$p_l(k+1) = \left(\frac{\beta_l}{\text{SIR}_l(\mathbf{p}(k))} \right) p_l(k) \quad \forall l. \quad (7)$$

2) Normalize $\mathbf{p}(k+1)$:

$$\mathbf{p}(k+1) \leftarrow \mathbf{p}(k+1) \cdot \bar{P}/\mathbf{1}^\top \mathbf{p}(k+1). \quad (8)$$

Corollary 1: Starting from any initial point $\mathbf{p}(0)$, $\mathbf{p}(k)$ in Algorithm 1 converges geometrically fast to the optimal solution of (5), $(\bar{P}/\mathbf{1}^\top \mathbf{x}(\text{diag}(\beta)\mathbf{B}))\mathbf{x}(\text{diag}(\beta)\mathbf{B})$.

Remark 1: Interestingly, (7) in Algorithm 1 is simply the Distributed Power Control (DPC) algorithm in [13], where the l th user has a virtual SIR threshold of β_l in both the (virtual) uplink and downlink transmission. However, due to (8), the standard interference function approach in [11] cannot be used to prove the convergence of Algorithm 1.

Next, we consider the joint optimization of power and transmit beamformer in the following max-min weighted SIR problem:

$$\begin{aligned} & \text{maximize} \quad \min_l \frac{\text{SIR}_l(\mathbf{p}, \mathbf{U})}{\beta_l} \\ & \text{subject to} \quad \sum_{l=1}^L p_l \leq \bar{P}, \quad p_l \geq 0, \quad \mathbf{u}_l^\dagger \mathbf{u}_l = 1 \quad \forall l, \\ & \text{variables:} \quad \mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_L], \quad \mathbf{p}. \end{aligned} \quad (9)$$

We first review the notion of uplink-downlink duality. The duality theory states that, under a same total power constraint and additive white noise for all users, the achievable SIR region for a downlink transmission with joint transmit beamforming and power control optimization is equivalent to that of an uplink transmission with joint receive beamforming and power control optimization. Further, the optimal *receive* beamforming vectors in the uplink is also the optimal *transmit* beamforming vectors in the downlink. Since joint power control and beamforming optimization in the uplink does not have the beamformer coupling difficulty associated with the downlink (hence easier to solve), the (dual) uplink problem can be first used to obtain the optimal transmit beamformers in the downlink. The optimal downlink transmit power is then computed by keeping the transmit beamformers fixed. In the

¹A closed-form solution to (5) was first obtained in [8] using a nonnegative (increased dimension) matrix totally different from \mathbf{B} . As such, the algorithmic solution to (5) in [8] is different and is mainly centralized. On the other hand, our solution exploits the DPC algorithm in [13] and is distributed.

case where the noise is different for each user, a virtual uplink transmission (assuming that all users have the same noise, i.e., $n_l = 1$ for all l) is constructed as an intermediary step to compute the optimal transmit beamforming vector.

Let the virtual uplink power be given by \mathbf{q} . Now, suppose there exists positive values γ (optimal max-min weighted SIR) and q_l for all l such that the virtual uplink $\hat{\text{SIR}}_l$ satisfies

$$\hat{\text{SIR}}_l(\mathbf{p}, \mathbf{U}) = \frac{q_l |\mathbf{h}_l^\dagger \mathbf{u}_l|^2}{\sum_{j \neq l} q_j |\mathbf{h}_j^\dagger \mathbf{u}_l|^2 + 1} \geq \beta_l \gamma \quad (10)$$

for all l . Since $\hat{\text{SIR}}_l$ in (10) only depends on the beamforming vector \mathbf{u}_l , the receive beamforming optimization, with the power fixed at \mathbf{q} , is solved by

$$\mathbf{u}_l^* = \arg \min_{\mathbf{u}_l} \sum_{j \neq l} \frac{G_{jl}(\mathbf{U})}{G_{ll}(\mathbf{U})} q_j + \frac{1}{G_{ll}(\mathbf{U})}, \quad (11)$$

whose solution is the linear minimum mean squared error (LMMSE) receiver given by (optimal up to a scaling factor): $\mathbf{u}_l^* = \left(\sum_{j \neq l} q_j \mathbf{h}_j \mathbf{h}_j^\dagger + \mathbf{I} \right)^+ \mathbf{h}_l$, where $(\cdot)^+$ denotes pseudo-inversion. Using this LMMSE receiver, the SIR constraint in (10) is always met with equality, i.e., $\hat{\text{SIR}}_l(\mathbf{p}, \mathbf{U}) = \beta_l \gamma$. By the uplink-downlink duality, the LMMSE receiver is also the optimal transmit beamformer in the downlink max-min weighted SIR problem given by (9).

Now, we are ready to use Algorithm 1 to solve the joint power control and beamforming problem in (9) in a fast and distributed fashion. The following algorithm computes the optimal power and transmit beamformer in (9):

Algorithm 2 (Max-min SIR-Power Control & Beamforming):

1) Update (virtual) uplink power $\mathbf{q}(k+1)$:

$$q_l(k+1) = \left(\frac{\beta_l}{\hat{\text{SIR}}_l(\mathbf{q}(k), \mathbf{U}(k))} \right) q_l(k) \quad \forall l. \quad (12)$$

2) Normalize $\mathbf{q}(k+1)$:

$$\mathbf{q}(k+1) \leftarrow \mathbf{q}(k+1) \cdot \bar{P}/\mathbf{1}^\top \mathbf{q}(k+1). \quad (13)$$

3) Update transmit beamforming matrix $\mathbf{U}(k) = [\mathbf{u}_1(k) \dots \mathbf{u}_L(k)]$:

$$\mathbf{u}_l(k) = \left(\sum_{j \neq l} q_j(k) \mathbf{h}_j \mathbf{h}_j^\dagger + \mathbf{I} \right)^+ \mathbf{h}_l \quad \forall l. \quad (14)$$

4) Update downlink power $\mathbf{p}(k+1)$:

$$p_l(k+1) = \left(\frac{\beta_l}{\hat{\text{SIR}}_l(\mathbf{p}(k), \mathbf{U}(k))} \right) p_l(k) \quad \forall l. \quad (15)$$

5) Normalize $\mathbf{p}(k+1)$:

$$\mathbf{p}(k+1) \leftarrow \mathbf{p}(k+1) \cdot \bar{P}/\mathbf{1}^\top \mathbf{p}(k+1). \quad (16)$$

Theorem 1: Let the optimal power and beamforming matrix in (9) be \mathbf{p}^* and \mathbf{U}^* respectively. Then, starting from any initial point $\mathbf{q}(0)$ and $\mathbf{p}(0)$, $\mathbf{p}(k)$ in Algorithm 2 converges

geometrically fast to $\mathbf{p}^* = \mathbf{x}(\mathbf{B}(\mathbf{U}^*))$ (unique up to a scaling constant).

Remark 2: Note that (12) and (15) of Algorithm 2 use the DPC algorithm in [13], where the l th user has a virtual SIR threshold of β_l in both the (virtual) uplink and downlink transmission. In the case where n_l 's are equal for all l , \mathbf{q} in (12) is the exact uplink transmit power, and only computing $\mathbf{1}^\top \mathbf{q}(k+1)$ in (13) requires a global coordination at the base station. Compared to previous centralized solution in [4], [8], our solution has less complexity and provable geometric convergence rate.

IV. WEIGHTED SUM MSE MINIMIZATION

In this section, we study minimizing the weighted sum of the MSE's of individual data streams under a sum power constraint. We assume that all the receivers use the LMMSE filter for estimating the received symbols of all users. The weighted sum MSE at the output of the LMMSE receiver is given by [14]:

$$\sum_{l=1}^L w_l \frac{1}{1 + \text{SIR}_l(\mathbf{p})}, \quad (17)$$

where w_l is some positive weight assigned by the network to the l th link (to reflect some long-term priority). Without loss of generality, we assume that \mathbf{w} is a probability vector. The weighted sum MSE minimization problem is given by

$$\begin{aligned} & \text{minimize} && \sum_{l=1}^L w_l \frac{1}{1 + \text{SIR}_l(\mathbf{p})} \\ & \text{subject to} && \sum_{l=1}^L p_l \leq \bar{P}, \quad p_l \geq 0 \quad \forall l, \\ & \text{variables:} && p_l \quad \forall l. \end{aligned} \quad (18)$$

We denote the optimal power vector to (18) by \mathbf{p}^* . We can rewrite (18) as

$$\begin{aligned} & \text{minimize} && \sum_{l=1}^L w_l \frac{\sum_{k \neq l} G_{lk} p_k + n_l}{\sum_k G_{lk} p_k + n_l} \\ & \text{subject to} && \sum_{l=1}^L p_l \leq \bar{P}, \quad p_l \geq 0 \quad \forall l, \\ & \text{variables:} && p_l \quad \forall l. \end{aligned} \quad (19)$$

It can be shown that the total power constraint in (18) and (19) are tight at optimality, which we exploit to transform (19) in the variables \mathbf{p} into another optimization problem that can be used to solve (19) optimally. To proceed further, we need to introduce the notion of quasi-invertibility of a nonnegative matrix in [15], which will be useful in solving (19) optimally.

Definition 1 (Quasi-invertibility): A square nonnegative matrix \mathbf{B} is a quasi-inverse of a square nonnegative matrix $\tilde{\mathbf{B}}$ if $\mathbf{B} - \tilde{\mathbf{B}} = \mathbf{B}\tilde{\mathbf{B}} = \tilde{\mathbf{B}}\mathbf{B}$. Furthermore, $(\mathbf{I} - \tilde{\mathbf{B}})^{-1} = \mathbf{I} + \mathbf{B}$ [15].

Using \mathbf{B} in (6), we next study the existence of $\tilde{\mathbf{B}}$, which can interestingly be associated with the SNR regime. In the case where the total maximum power is very large, i.e., $\bar{P} \rightarrow \infty$ (high SNR regime) or when interference (off-diagonals of \mathbf{F}) is very large, it is deduced in the following that $\tilde{\mathbf{B}}$ does not exist.

Lemma 2: $\tilde{\mathbf{B}}$ does not exist when $\mathbf{B} = \mathbf{F}$, where $F_{lj} > 0$ for all l, j and $l \neq j$.

However, when $\mathbf{F} = \mathbf{0}$ (no interference) such that $\mathbf{B} = \mathbf{v}\mathbf{1}^\top/\bar{P}$ or when \bar{P} is sufficiently small (low SNR regime)

such that $\mathbf{B} \approx \mathbf{v}\mathbf{1}^\top / \bar{P}$, then $\tilde{\mathbf{B}}$ always exists, as shown by the following lemma.

Lemma 3: For any nonnegative vector \mathbf{v} , $\tilde{\mathbf{B}} = 1/(1 + \mathbf{1}^\top \mathbf{v})\mathbf{v}\mathbf{1}^\top$ when $\mathbf{B} = \mathbf{v}\mathbf{1}^\top$.

In a numerical example for a ten-user IEEE 802.11b network, we experiment with the maximum power constraint of 33mW and 1W (the largest possible value allowed in IEEE 802.11b). Averaging over 10,000 random channel coefficient instances, the percentage of instances where $\tilde{\mathbf{B}}$ exists is 99% and 65% corresponding to the maximum power constraint of 33mW and 1W, respectively.

For the rest of the paper, we focus on the case when $\tilde{\mathbf{B}}$ exists. We next solve (19) in the following. Let us define

$$\mathbf{z} = (\mathbf{I} + \mathbf{B})\mathbf{p}. \quad (20)$$

Note that $G_{ll}z_l$ is the total received (desired and interfering) signal power plus the additive white noise at the l th receiver.

Then, we can rewrite (19) in terms of \mathbf{z} as

$$\begin{aligned} & \text{minimize} && \sum_{l=1}^L w_l \frac{(\tilde{\mathbf{B}}\mathbf{z})_l}{z_l} \\ & \text{subject to} && z_l \geq (\tilde{\mathbf{B}}\mathbf{z})_l, \quad l = 1, \dots, L, \\ & \text{variables:} && z_l \quad \forall l, \end{aligned} \quad (21)$$

where the constraints in (21) are due to the nonnegativity of \mathbf{p} , since, using Definition 1, $\mathbf{p} = (\mathbf{I} + \mathbf{B})^{-1}\mathbf{z} = (\mathbf{I} - \tilde{\mathbf{B}})\mathbf{z} \geq \mathbf{0}$.

The following result provides a condition under which the optimal solution to (21), \mathbf{z}^* , can be transformed to yield the optimal solution to (19) or equivalently (18).

Theorem 2: The optimal solution to (18) is given by $\mathbf{p}^* = (\mathbf{I} + \mathbf{B})^{-1}\mathbf{z}^*$, where \mathbf{z}^* is the optimal solution to (21), if $\tilde{\mathbf{B}}$ is the quasi-inverse of a nonnegative matrix $\tilde{\mathbf{B}}$, where $\rho(\tilde{\mathbf{B}}) < 1$.

Lemma 4: If $\tilde{\mathbf{B}}$ exists, then $\tilde{\mathbf{B}}$ has the spectral radius

$$\rho(\tilde{\mathbf{B}}) = \frac{\rho(\mathbf{B})}{1 + \rho(\mathbf{B})}, \quad (22)$$

with the corresponding left and Perron eigenvectors of \mathbf{B} .

In the following, we derive useful lower bounds to (18), investigate special cases, and finally characterize the exact solution to (18).

A. A Lower Bound to Weighted Sum MSE Minimization

By exploiting the eigenspace of $\tilde{\mathbf{B}}$ and the Friedland-Karlin inequality in nonnegative matrix theory [16] (see [12] for its extension), the following result gives a lower bound on the weighted sum MSE problem in (18).

Theorem 3: If $\tilde{\mathbf{B}}$ exists,

$$\sum_{l=1}^L w_l \frac{1}{1 + \text{SIR}_l(\mathbf{p})} \geq \left(\frac{1}{1 + 1/\rho(\mathbf{B})} \right)^{\|\mathbf{w}\|_\infty^{\mathbf{x}(\mathbf{B}) \circ \mathbf{y}(\mathbf{B})}} \quad (23)$$

for all feasible \mathbf{p} in (18).

Equality is achieved if and only if $\mathbf{w} = \mathbf{x}(\mathbf{B}) \circ \mathbf{y}(\mathbf{B})$. Thus, $\text{SIR}_l(\mathbf{p}^*) = (1/\rho(\mathbf{B}))1$. In this case, $\mathbf{p}^* = \mathbf{x}(\mathbf{B})$ solves (18).

Interestingly, Theorem 3 shows that solving (5) with $\beta = 1$ can be seen as an approximation method to solving (18) sub-optimally, but with an approximation guarantee. In particular, by taking the logarithm of the objective function of (18), $\|\mathbf{w}\|_\infty^{\mathbf{x}(\mathbf{B}) \circ \mathbf{y}(\mathbf{B})}$ can be interpreted as the approximation ratio of Algorithm 1 with $\beta = 1$ in solving (18).

Remark 3: Based on Theorem 3, we obtain a connection between the min-max MSE problem and the weighted sum MSE optimization. Suppose we consider min-max MSE: $\min_{\mathbf{p}} \max_l 1/(1 + \text{SIR}_l(\mathbf{p}))$. Then, by the max-min characterization of $\rho(\tilde{\mathbf{B}})$:

$$\max_{\mathbf{z} > \mathbf{0}} \min_l \frac{(\tilde{\mathbf{B}}\mathbf{z})_l}{z_l} = \min_{\mathbf{z} > \mathbf{0}} \max_l \frac{(\tilde{\mathbf{B}}\mathbf{z})_l}{z_l} = \rho(\tilde{\mathbf{B}}), \quad (24)$$

the optimal objective of min-max MSE is simply $\rho(\tilde{\mathbf{B}}) = 1/(1 + 1/\rho(\mathbf{B}))$. It immediately follows from Theorem 3 that (5) with $\beta = 1$ yields the equivalent power allocation as min-max MSE and the optimal sum MSE with $\mathbf{w} = \mathbf{x}(\mathbf{B}) \circ \mathbf{y}(\mathbf{B})$.

B. Exact Solution to Weighted Sum MSE Minimization

The existence of $\tilde{\mathbf{B}}$ allows us to delineate cases of (18) that can be solved optimally from the general problem. The following result gives the exact closed-form solution to (18), which motivates a fast algorithm (Algorithm 3 below) to compute the optimal solution.

Theorem 4: If $\tilde{\mathbf{B}}$ exists, then the optimal solution to (18) is given by $\mathbf{p}^* = (\mathbf{I} - \tilde{\mathbf{B}})\mathbf{z}^* \geq \mathbf{0}$, where \mathbf{z}^* is given by

$$z_l^* = \sqrt{\frac{w_l \sum_{j \neq l} \tilde{B}_{lj} z_j^*}{\sum_{j \neq l} w_j \tilde{B}_{jl} / z_j^*}} \quad (25)$$

for all l and satisfies $\mathbf{1}^\top \mathbf{z}^* - \mathbf{1}^\top \tilde{\mathbf{B}}\mathbf{z}^* = \bar{P}$.

Now, (25) in Theorem 4 can be written in the form of $\mathbf{z} = I(\mathbf{z})$, where I is a homogeneous function. We will leverage the standard interference function results in [11] to propose the following (step size free) algorithm that computes \mathbf{z}^* in Theorem 4, and implicitly, the optimal transmit power of (18).

Algorithm 3 (Sum MSE Minimization):

- 1) Initialize an arbitrarily small $\epsilon > 0$.
- 2) Update auxiliary variable $\mathbf{z}(k+1)$:

$$z_l(k+1) = \sqrt{\frac{w_l \sum_{j \neq l} \tilde{B}_{lj} z_j(k)}{\sum_{j \neq l} w_j \tilde{B}_{jl} / z_j(k)}} + \epsilon \quad \forall l. \quad (26)$$

- 3) Update $\mathbf{p}(k+1)$:

$$p_l(k+1) = \frac{\text{SIR}_l(\mathbf{p}(k))}{1 + \text{SIR}_l(\mathbf{p}(k))} z_l(k+1) \quad \forall l. \quad (27)$$

- 4) Normalize $\mathbf{p}(k+1)$: $\mathbf{p}(k+1) \leftarrow \mathbf{p}(k+1) \cdot \bar{P} / (\mathbf{1}^\top \mathbf{p}(k+1))$.
-

The following theorem shows that Algorithm 3 converges to the optimal solution \mathbf{z}^* and \mathbf{p}^* in Theorem 4.

Theorem 5: If $\tilde{\mathbf{B}}$ exists, for arbitrarily small $\epsilon > 0$, Algorithm 3 converges to the unique fixed point \mathbf{z}^* and \mathbf{p}^* in Theorem 4 from any initial point $\mathbf{z}(0)$ under synchronous and asynchronous updates.

Remark 4: Algorithm 3 requires a complexity of $O(L^3)$ to compute $\tilde{\mathbf{B}}$. Step 2 of Algorithm 3 can be implemented by distributed message passing. Transforming from $\mathbf{z}(k+1)$ to $\mathbf{p}(k+1)$ in (27) is performed locally by each user, and the normalization at Step 4 is performed at the base station.

Remark 5: The convergence of Algorithm 3 leverages the standard interference function approach in [11] by introducing ϵ in (26) of Algorithm 3.

V. WEIGHTED SUM RATE MAXIMIZATION

In this section, we consider the weighted sum rate of all users as a performance metric to be optimized. We then quantify the connection of the weighted sum rate maximization and the weighted sum MSE minimization.

A. Exact Solution to Weighted Sum Rate Maximization

The weighted sum rate maximization problem is given by

$$\begin{aligned} & \text{maximize} && \sum_{l=1}^L w_l \log(1 + \text{SIR}_l(\mathbf{p})) \\ & \text{subject to} && \sum_{l=1}^L p_l \leq \bar{P}, \quad p_l \geq 0 \quad \forall l, \\ & \text{variables:} && p_l \quad \forall l. \end{aligned} \quad (28)$$

We can rewrite (28) to be equivalent to

$$\begin{aligned} & \text{minimize} && \prod_{l=1}^L \left(\frac{\sum_{k \neq l} G_{lk} p_k + 1}{\sum_k G_{lk} p_k + 1} \right)^{w_l} \\ & \text{subject to} && \sum_{l=1}^L p_l \leq \bar{P}, \quad p_l \geq 0 \quad \forall l, \\ & \text{variables:} && p_l \quad \forall l. \end{aligned} \quad (29)$$

Similar to Section IV, if $\tilde{\mathbf{B}}$ is the quasi-inverse of $\tilde{\mathbf{B}}$, we can rewrite (29) as

$$\begin{aligned} & \text{minimize} && \prod_{l=1}^L \left(\frac{(\tilde{\mathbf{B}}\mathbf{z})_l}{z_l} \right)^{w_l} \\ & \text{subject to} && z_l \geq (\tilde{\mathbf{B}}\mathbf{z})_l, \quad l = 1, \dots, L, \\ & \text{variables:} && z_l \quad \forall l, \end{aligned} \quad (30)$$

where \mathbf{z} is given by (20).

Similar to Theorem 2, the following Theorem 6 gives the condition under which (28) is solved optimally.

Theorem 6: If $\tilde{\mathbf{B}}$ exists, the optimal solution to (28) is given by $\mathbf{p}^* = (\mathbf{I} + \tilde{\mathbf{B}})^{-1} \mathbf{z}^*$, where \mathbf{z}^* is the optimal solution of (30).

Next, we connect the three optimization problems given in (28), (18) and (5) with $\beta = 1$.

Remark 6: Applying the arithmetic-geometric mean inequality to connect (21) and (30), we deduce that the weighted sum rate maximization has the same optimal power as the weighted sum MSE minimization when $\mathbf{w} = \mathbf{x}(\mathbf{B}) \circ \mathbf{y}(\mathbf{B})$.

Theorem 6 is used to give the following solution of (28).

Theorem 7: If $\tilde{\mathbf{B}}$ exists, then the optimal solution to (28) is given by $\mathbf{p}^* = (\mathbf{I} - \tilde{\mathbf{B}})\mathbf{z}^* \geq \mathbf{0}$, where \mathbf{z}^* is given by

$$z_l^* = \frac{w_l}{\sum_j w_j \tilde{B}_{jl} / (\tilde{\mathbf{B}}\mathbf{z}^*)_j} \quad (31)$$

for all l and satisfies $\mathbf{1}^\top \mathbf{z}^* - \mathbf{1}^\top \tilde{\mathbf{B}}\mathbf{z}^* = \bar{P}$.

As in the previous, (31) in Theorem 7 can be expressed as $\mathbf{z} = I(\mathbf{z})$, where I is a homogeneous function. Using the standard interference function approach in [11], the following algorithm computes the optimal solution of (28).

Algorithm 4 (Sum Rate Maximization):

- 1) Initialize an arbitrarily small $\epsilon > 0$.
- 2) Update auxiliary variable $\mathbf{z}(k+1)$:

$$z_l(k+1) = \frac{w_l}{\sum_j w_j \tilde{B}_{jl} / (\tilde{\mathbf{B}}\mathbf{z}(k))_j} + \epsilon \quad \forall l. \quad (32)$$

- 3) Update $\mathbf{p}(k+1)$:

$$p_l(k+1) = \frac{\text{SIR}_l(\mathbf{p}(k))}{1 + \text{SIR}_l(\mathbf{p}(k))} z_l(k+1) \quad \forall l. \quad (33)$$

- 4) Normalize $\mathbf{p}(k+1)$: $\mathbf{p}(k+1) \leftarrow \mathbf{p}(k+1) \cdot \bar{P} / (\mathbf{1}^\top \mathbf{p}(k+1))$.

The following result shows that Algorithm 4 converges to the optimal solution \mathbf{z}^* and \mathbf{p}^* in Theorem 7.

Theorem 8: If $\tilde{\mathbf{B}}$ exists, for arbitrarily small $\epsilon > 0$, Algorithm 4 converges to the unique fixed point \mathbf{z}^* and \mathbf{p}^* in Theorem 7 from any initial point $\mathbf{z}(0)$ under synchronous and asynchronous updates.

ACKNOWLEDGEMENT

We acknowledge helpful discussions with Steven Low at Caltech and Kevin Tang at Cornell. This research has been supported in part by ONR grant N00014-07-1-0864 and NSF CNS 0720570.

REFERENCES

- [1] V. D. Blondel, L. Ninove, and P. Van Dooren. An affine eigenvalue problem on the nonnegative orthant. *Linear Algebra and its Applications*, 404:69–84, 2005.
- [2] U. Krause. Concave Perron-Frobenius theory and applications. *Nonlinear analysis*, 47(2001):1457–1466, 2001.
- [3] C. W. Tan, M. Chiang, and R. Srikant. Fast algorithms and performance bounds for sum rate maximization in wireless networks. *Proc. of IEEE Infocom*, 2009.
- [4] M. Schubert and H. Boche. Solution of the multiuser downlink beamforming problem with individual SINR constraints. *IEEE Trans. on Vehicular Technology*, 53(1):18–28, 2004.
- [5] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas. Transmit beamforming and power control for cellular wireless systems. *IEEE Journal on Selected Areas in Communications*, 16(8):1437–1450, 1998.
- [6] P. Viswanath and D. N. C. Tse. Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. *IEEE Trans. on Information Theory*, 49(8):1912–1921, 2003.
- [7] W. Yu. Uplink-downlink duality via minimax duality. *IEEE Trans. on Information Theory*, 52(2):361–374, 2006.
- [8] W. Yang and G. Xu. Optimal downlink power assignment for smart antenna systems. *Proc. of IEEE ICASSP*, 1998.
- [9] M. Chiang, P. Hande, T. Lan, and C. W. Tan. Power control in wireless cellular networks. *Foundations and Trends in Networking*, 2(4):381–533, 2008.
- [10] M. Codreanu, A. Tolli, M. Juntti, and M. Latva-aho. Joint design of Tx-Rx beamformers in MIMO downlink channel. *IEEE Trans. on Signal Processing*, 55(9):4639–4655, 2007.
- [11] R. D. Yates. A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1348, 1995.
- [12] C. W. Tan. Nonconvex power control in multiuser communication systems. *Ph.D. Thesis, Princeton University*, November 2008.
- [13] G. J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans. on Vehicular Technology*, 42(4):641–646, 1993.
- [14] P. Viswanath, V. Anantharam, and D. N. C. Tse. Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers. *IEEE Trans. on Information Theory*, 45(6):1968–1983, 1999.
- [15] Y. K. Wong. Some mathematical concepts for linear economic models. *Chapter in Economic Activity Analysis*, Edited by O. Morgenstern, John Wiley & Sons, Inc.:283–339, 1954.
- [16] S. Friedland and S. Karlin. Some inequalities for the spectral radius of non-negative matrices and applications. *Duke Mathematical Journal*, 42(3):459–490, 1975.