Weighted Sum Rate Maximization with Multiple Linear Conic Constraints

Hans H. Brunner†, Andreas Dotzler‡, Wolfgang Utschick‡, Josef A. Nossek†
†Institute for Circuit Theory and Signal Processing, ‡Associate Institute for Signal Processing
Technische Universität München, 80290 Munich, Germany
e-mail: {brunner|dotzler|utschick|josef.a.nossek}@tum.de

***Abstract*—In the *downlink* (DL) of a *multi-user multiple-input and multiple-output* (MU-MIMO) system, the maximization of the weighted sum rate with *dirty paper precoding* (DPC) is treated under multiple linear and linear conic constraints. By network duality, the problem is transformed to a minimax *uplink* (UL) problem. In the UL, the minimization of the utility with respect to the noise covariance and the maximization of the utility with respect to the transmit covariances is solved either jointly or alternately with the gradient-projection algorithm. The proposed algorithms do not only allow to find the maximum weighted sum rate with respect to conic constraints, they are also efficient implementations with respect to multiple linear constraints.**

I. Introduction

Maximizing the weighted sum rate in the DL of a MU-MIMO system with a single transmit power constraint typically consists of two steps: a transformation of the DL problem to the dual UL problem and solving the UL problem, which can be rewritten as a convex optimization problem. The dual transformation was introduced in [1] and [2] in parallel. Additionally to generic methods for convex optimizations, there are some methods specifically tailored to the UL problem. One possible approach to solve the problem is the steepest ascend algorithm. Hunger et al. showed that an orthogonal projection of the gradient onto the constraint set is required to find the steepest ascend direction [3]. With this projection, the deepest ascend algorithm has a very fast convergence behavior. The polite water-filling from Liu et al. is another possible approach for both, the dual transformation and solving the UL problem with very fast convergence behavior [4]. The references in [3] and [4] give a broad overview of the existing algorithms for solving the weighted sum rate maximization.

Multiple linear constraints for the weighted sum rate maximization in the UL were already addressed in [5], where Yu et al. investigated individual transmit power constraints for each user. However, multiple constraints in the DL could not be handled by the existing dual transformations. In [6], Yu introduced a minimax duality, where the UL problem is, on the one hand, a maximization of the weighted sum rate with respect to the transmit covariance matrices and, on the other hand, a minimization of the weighted sum rate with respect to the noise covariance matrix. With the minimax duality, constraints on the sum transmit covariance matrix in the DL are transformed to constraints on the noise covariance matrix in the UL. Yu et al. used this duality in [7] to tackle the weighted sum rate maximization with per antenna power constraints in the DL. To solve the resulting minimax problem, Yu et al. updated the transmit and noise covariance matrices simultaneously with an adopted Newton's method. Feasibility of the covariance matrices was assured with the interior-point method.

The weighted sum rate algorithm of Yu et al. was generalized by Huh et al. in [8] for general linear constraints. They alternately solved the maximization of the transmit and noise covariance matrices. The utility was minimized with respect to the noise covariance matrix by a subgradient method, while the maximization with respect to the transmit covariance matrices was done with the adopted Newton's method.

In [9], Zhang et al. relaxed the multiple linear constraints to a single weighted sum constraint. The resulting problem could than be transformed to an UL problem with the existing duality [1], [2]. Zhang et al. claimed that the optimum could be reached by alternately maximizing the UL problem with respect to the transmit covariance matrices and minimizing the DL problem with respect to the weights. The claim was proven by showing that both, the original and the changed problem, have the same Lagrangian function. The Lagragian multipliers of the original problem are the weights in the transformed problem.

Designed for more general interference networks, Liu et al. proposed two algorithms based on their polite water-filling to solve the weighted sum rate maximization with multiple linear constraints [10]. The cost function has to be maximized with respect to the transmit covariance matrices and minimized with respect to the Lagrangian multipliers of the constraints. The first algorithm alternates between polite water-filling for the transmit covariance matrices and a subgradient or an heuristic update algorithm for the Lagrangian multipliers. The polite water-filling itself is an alternating algorithm, which rotates between updating the transmit covariances in the UL and DL. The second algorithm includes the update of the Lagrangian multipliers into the alternating process of the polite water-filling.

Dotzler et al. introduced in [11] a minimax duality with linear conic constraints, where the UL noise covariance itself is the Lagrangian multiplier for the constraints on the DL covariance. In this paper, we extend the minimax duality for multiple linear conic constraints. We show that the UL problem can be solved with the rather simple but efficient gradient-projection algorithm. We give details to the required projections and

978-1-4673-6432-4/15/$31.00 ©2015 IEEE

propose to update the transmit and noise covariances either alternately or jointly. Additionally, we discuss the tangent cone projection of the gradients to find the steepest ascend or descend. The proposed algorithms are tailored for conic constraints, which cannot be handled by the existing methods. Nevertheless, the proposed algorithms are also very efficient for optimizations with multiple linear constraints.

Next to per antenna power constraints, multiple linear constraints can be used as interference temperatures, which limit the received interference at selected receivers to a given level [7], [12]. In a cellular network with local optimizations of the transmission strategy, a conic constraint, which shapes the sum transmit covariance of a transmitter to a scaled identity matrix, can be useful to remove uncertainty in the interference variance at disturbed users [13]. By combining the shaping constraint with a linear sum power constraint, a controlled uncertainty in the intercell interference variance at the users can be introduced [14]. An example for combining a conic and multiple linear constraints is the combination of the above mentioned shaping and interference temperature constraints.

II. System Model

In this paper, the *base station* (BS) is equipped with M antennas and serves K users with N receive antennas each. The matrix $\boldsymbol{H}_k \in \mathbb{C}^{N\times M}$ contains the channel coefficients between the antennas of the BS and user k. Perfect channel state information is assumed.

A. Downlink Rate

The achievable, normalized rate of user k, within the capacity region of the MIMO broadcast channel with DPC, can be expressed as

$$r_k^{\mathrm{DL}} = \ln \frac{\left|\mathbf{I} + \sum_{\hat{k}\geq k} \boldsymbol{H}_k \boldsymbol{Q}_{\hat{k}} \boldsymbol{H}_k^{\mathrm{H}}\right|}{\left|\mathbf{I} + \sum_{\hat{k}> k} \boldsymbol{H}_k \boldsymbol{Q}_{\hat{k}} \boldsymbol{H}_k^{\mathrm{H}}\right|}, \tag{1}$$

where $\boldsymbol{Q}_k \in \mathbb{C}^{M\times M}$ is the transmit covariance matrix for user k. $\sum_k \boldsymbol{Q}_k = \boldsymbol{Q} \in \mathbb{C}^{M\times M}$ will be used as the sum transmit covariance matrix. $\sum_{\hat{k}>k} \boldsymbol{H}_k \boldsymbol{Q}_{\hat{k}} \boldsymbol{H}_k^{\mathrm{H}}$ is the variance of the intracell interference with DPC. Without loss of generality it is assumed that the noise covariance matrix at each user is fixed to an identity matrix.

B. Dual Uplink Rate

In the dual multiple access channel, the rate of user k with successive interference cancellation can be found with flipped channels as

$$r_k^{\mathrm{UL}} = \ln \frac{\left|\boldsymbol{\Omega} + \sum_{\hat{k}\leq k} \boldsymbol{H}_{\hat{k}}^{\mathrm{H}} \boldsymbol{\Sigma}_{\hat{k}} \boldsymbol{H}_{\hat{k}}\right|}{\left|\boldsymbol{\Omega} + \sum_{\hat{k}< k} \boldsymbol{H}_{\hat{k}}^{\mathrm{H}} \boldsymbol{\Sigma}_{\hat{k}} \boldsymbol{H}_{\hat{k}}\right|}, \tag{2}$$

where $\boldsymbol{\Sigma}_k \in \mathbb{C}^{N\times N}$ is the transmit covariance matrix of user k. $\sum_{\hat{k}<k} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{\Sigma}_{\hat{k}} \boldsymbol{H}_k$ is the variance of the intracell interference with successive interference cancellation and $\boldsymbol{\Omega} \in \mathbb{C}^{M\times M}$ is the covariance matrix of the noise at the BS.

C. Linear and Linear Conic Constraints

Linear constraints are given as trace constraints on the positive semidefinite transmit covariance matrix:

$$\mathrm{tr}\left(\boldsymbol{Q}\boldsymbol{A}\right) \leq a. \tag{3}$$

This restricts the radiated sum power in the directions of $\boldsymbol{A}$ to a. The sum power constraint is $\boldsymbol{A} = \mathbf{I}$. Per antenna power constraints or interference temperatures can be implemented by selecting an appropriate $\boldsymbol{A}$ for each antenna or interference direction [7], [12].

Conic constraints consist of a shaping matrix $\boldsymbol{C}$ and a linear subspace $\mathcal{Z}$:

$$\exists \boldsymbol{Z} \in \mathcal{Z} : \boldsymbol{Q} \preceq \boldsymbol{C} + \boldsymbol{Z}. \tag{4}$$

The selection $\boldsymbol{C}^{\mathrm{eye}} = c\mathbf{I}$ and $\mathcal{Z}^{\mathrm{eye}} = \{\mathbf{0}\}$ shapes the transmit covariance to a scaled identity matrix with sum transmit power cM.

In the following, linear constraints are formulated as conic constraints. The transformation can be done with any positive definite matrix $\boldsymbol{C}^{\mathrm{lin}}$ that fulfills $a = \mathrm{tr}(\boldsymbol{C}^{\mathrm{lin}}\boldsymbol{A})$ and the subspace $\mathcal{Z}^{\mathrm{lin}} = \left\{\boldsymbol{Z}^{\mathrm{lin}} : \mathrm{tr}(\boldsymbol{A}\boldsymbol{Z}^{\mathrm{lin}}) = 0\right\}$ [11].

D. Problem Formulation

The weighted sum rate problem with L linear and/or linear conic constraints in the DL of an MU-MIMO system can be formulated as

$$\max_{\substack{\boldsymbol{Q}_k\succeq\mathbf{0},\,\forall k\\ \boldsymbol{Z}_l\in\mathcal{Z}_l,\,\forall l}} \sum_k r_k^{\mathrm{DL}} w_k, \quad \text{s.t. } \boldsymbol{Q} \preceq \boldsymbol{C}_l + \boldsymbol{Z}_l,\ \forall l, \tag{5}$$

where w_k is the weight for user k.

III. Uplink-Downlink Duality

The minimax duality with a conic constraint from [11] is extended for multiple conic constraints in this Section.[1]

Because of strong duality[2] and the variables being selected from closed sets, problem (5) can be replaced by its Lagrangian dual

$$\min_{\boldsymbol{\Omega}_l\succeq\mathbf{0},\,\forall l} \max_{\substack{\boldsymbol{Q}_k\succeq\mathbf{0},\,\forall k\\ \boldsymbol{Z}_l\in\mathcal{Z}_l,\,\forall l}} \sum_k r_k^{\mathrm{DL}} w_k - \sum_l \mathrm{tr}\left(\boldsymbol{\Omega}_l(\boldsymbol{Q} - \boldsymbol{C}_l - \boldsymbol{Z}_l)\right), \tag{6}$$

where $\boldsymbol{\Omega}_l$ is the Lagrangian multiplier of constraint l. The inner problem of (6) is unbounded, unless $\boldsymbol{\Omega}_l \in \mathcal{Z}_l^{\perp},\ \forall l$, where

$$\mathcal{Z}_l^{\perp} = \left\{\boldsymbol{\Omega}_l : \mathrm{tr}(\boldsymbol{\Omega}_l \boldsymbol{Z}_l) = 0, \forall \boldsymbol{Z}_l \in \mathcal{Z}_l\right\}, \tag{7}$$

[1]Note that the definition of $\mathcal{Z}_l$, $\mathcal{Z}_l^{\perp}$ is different from the one used in [11]

[2]The strong duality for multiple constraints can be shown with the same methods presented in [11] for a single constraint

and can be rearranged as

$$\min_{\substack{\boldsymbol{\Omega}_l \succeq \mathbf{0},\, \forall l \\ \boldsymbol{\Omega}_l \in \mathcal{Z}_l^{\perp},\, \forall l}} \max_{\boldsymbol{Q}_k \succeq \mathbf{0},\, \forall k} \sum_k r_k^{\mathrm{DL}} w_k - \mathrm{tr}\left(\sum_l \boldsymbol{\Omega}_l \boldsymbol{Q}\right) + \sum_l \mathrm{tr}\left(\boldsymbol{\Omega}_l \boldsymbol{C}_l\right). \tag{8}$$

Due to the complementary slackness condition of all constraints,

$$\sum_l \mathrm{tr}\left(\boldsymbol{\Omega}_l \boldsymbol{Q}\right) = \sum_l \mathrm{tr}\left(\boldsymbol{\Omega}_l \boldsymbol{C}_l\right) \tag{9}$$

has to be fulfilled. A joint scaling of all dual variables does not change the solution. Without loss of generality, both sides of Equation (9) can be fixed to some value P:

$$\min_{\substack{\boldsymbol{\Omega}_l \succeq \mathbf{0},\, \boldsymbol{\Omega}_l \in \mathcal{Z}_l^{\perp},\, \forall l \\ \sum_l \mathrm{tr}(\boldsymbol{\Omega}_l \boldsymbol{C}_l) = P}} \max_{\substack{\boldsymbol{Q}_k \succeq \mathbf{0},\, \forall k \\ \mathrm{tr}(\boldsymbol{\Omega} \boldsymbol{Q}) = P}} \sum_k r_k^{\mathrm{DL}} w_k, \tag{10}$$

where $\boldsymbol{\Omega} = \sum_l \boldsymbol{\Omega}_l$ is the sum of all Lagrangian multipliers.

The inner maximization of problem (10) is a weighted sum rate optimization with a single linear constraint on the transmit covariance matrix $\boldsymbol{Q}$ in the DL. As shown in [9] with the duality from [6], the inner problem can be transformed to a weighted sum rate optimization in the UL, where $\boldsymbol{\Omega}$ becomes the noise covariance matrix (see Equation (2)) and P the sum transmit power:

$$\min_{\substack{\boldsymbol{\Omega}_l \succeq \mathbf{0},\, \boldsymbol{\Omega}_l \in \mathcal{Z}_l^{\perp},\, \forall l \\ \sum_l \mathrm{tr}(\boldsymbol{\Omega}_l \boldsymbol{C}_l) = P}} \max_{\substack{\boldsymbol{\Sigma}_k \succeq 0,\, \forall k \\ \sum_k \mathrm{tr}(\boldsymbol{\Sigma}_k) = P}} \sum_k r_k^{\mathrm{UL}} w_k. \tag{11}$$

In this paper, it is assumed that $\boldsymbol{\Omega}$ is invertible. Handling cases with singularities in the UL noise are discussed in [11]. $\boldsymbol{\Omega}$ might be rank deficient, if directions exist, which cannot be reached by any of the channel matrices.

IV. Joint and Alternating Gradient-Projection

It is assumed, without loss of generality, that the users are handled in the optimal decoding order, i.e., the weights w_k are in a non ascending order. By selecting $\alpha_k = w_k - w_{k+1}$, $\alpha_0 = -w_1$, and $\alpha_K = w_K$ problem (11) can be written as

$$\min_{\substack{\boldsymbol{\Omega}_l \succeq \mathbf{0},\, \boldsymbol{\Omega}_l \in \mathcal{Z}_l^{\perp},\, \forall l \\ \sum_l \mathrm{tr}(\boldsymbol{\Omega}_l \boldsymbol{C}_l)) = P}} \max_{\substack{\boldsymbol{\Sigma}_k \succeq 0,\, \forall k \\ \sum_k \mathrm{tr}(\boldsymbol{\Sigma}_k) = P}} \Psi, \tag{12}$$

$$\Psi = \sum_k \alpha_k \ln \left| \boldsymbol{\Omega} + \sum_{\hat{k} \le k} \boldsymbol{H}_{\hat{k}}^{\mathrm{H}} \boldsymbol{\Sigma}_{\hat{k}} \boldsymbol{H}_{\hat{k}} \right| + \alpha_0 \ln |\boldsymbol{\Omega}|. \tag{13}$$

As shown in [3], the inner maximization of problem (12) can be solved efficiently with the iterative scaled gradient algorithm with an orthogonal projection onto the constraint set. We propose to solve the complete minimax problem (12) with the joint gradient algorithm listed as Algorithm 1. In each iteration of the algorithm, the transmit and noise covariance matrices are updated with a steepest ascend/descend step. Lines 6 to 18 describe the update of the transmit covariance matrices according to Section IV-A. The description of the noise covariance matrix update in lines 19 to 32 can be found in Section IV-B.

We also suggest an alternating gradient-projection algorithm that puts the updates of the transmit and noise covariances in additional inner loops, respectively. The algorithm alternates between finding the optimal transmit covariances for a fixed noise covariance and vice versa. It has the same alternating structure as the algorithms for multiple linear constraints from the literature, but the inner loops run methods known to be efficient.

A. Transmit Covariance Gradient Update

Following [3], the gradient with respect to the individual transmit covariance matrices is

$$\frac{\partial \Psi}{\partial \boldsymbol{\Sigma}_k^{\mathrm{T}}} = \sum_{\hat{k} \ge k} \alpha_{\hat{k}} \boldsymbol{H}_k \left(\boldsymbol{\Omega} + \sum_{\check{k} \le \hat{k}} \boldsymbol{H}_{\check{k}}^{\mathrm{H}} \boldsymbol{\Sigma}_{\check{k}} \boldsymbol{H}_{\check{k}} \right)^{-1} \boldsymbol{H}_k^{\mathrm{H}} \tag{14}$$

and the unconstrained updates of the covariance matrices are calculated as

$$\hat{\boldsymbol{\Sigma}}_k^{(i)} = \boldsymbol{\Sigma}_k^{(i)} + p_{\mathrm{t}}^{(i)} s_{\mathrm{t}}^{(i)} \left. \frac{\partial \Psi}{\partial \boldsymbol{\Sigma}_k} \right|_i. \tag{15}$$

The preconditioning scalar

$$p_{\mathrm{t}}^{(i)} = \frac{P}{\sum_k \mathrm{tr}\left(\left. \frac{\partial \Psi}{\partial \boldsymbol{\Sigma}_k^{\mathrm{T}}} \right|_i \right)} \tag{16}$$

normalizes the sum of the gradient traces to the chosen UL transmit power P. This makes the gradient almost independent of the selection of P. $s_{\mathrm{t}}^{(i)}$ is the joint stepsize for all transmit covariance matrix updates in iteration (i). We use stepsize control with diminishing stepsize to ensure convergence.

The steepest ascend update, which gives the transmit covariance matrices in the next iteration $(i+1)$, is found with the joint projection of all unconstrained updates onto the constraints set.

$$\left\{ \boldsymbol{\Sigma}_1^{(i+1)}, \ldots, \boldsymbol{\Sigma}_K^{(i+1)} \right\} = \left(\left\{ \hat{\boldsymbol{\Sigma}}_1^{(i)}, \ldots, \hat{\boldsymbol{\Sigma}}_K^{(i)} \right\} \right)_{\perp} \tag{17}$$

The orthogonal projection $(\bullet)_{\perp}$ onto the constraint set is done with the water-spilling algorithm from [3]. The required water-spilling is exactly the same as the water-spilling for the orthogonal projection of the noise covariance matrix onto the constraint set with a single scaled identity shaping constraint as explained in Section V-A.

B. Noise Covariance Gradient Update

The gradient update of the noise covariance matrix follows the same steps as the update of the transmit covariance matrices except that it is a minimization instead of a maximization. The gradient of Ψ with respect to the noise covariance matrix, or each of its summands, is

$$\frac{\partial \Psi}{\partial \boldsymbol{\Omega}^{\mathrm{T}}} = \sum_k \alpha_k \left(\boldsymbol{\Omega} + \sum_{\hat{k} \le k} \boldsymbol{H}_{\hat{k}}^{\mathrm{H}} \boldsymbol{\Sigma}_{\hat{k}} \boldsymbol{H}_{\hat{k}} \right)^{-1} + \alpha_0 \boldsymbol{\Omega}^{-1}. \tag{18}$$

Algorithm 1 Joint Gradient-Projection Algorithm

Require: Accuracy ε, constraints $(C_l, \mathcal{Z}_l)$, $\forall l$

$\boldsymbol{\Sigma}_k \leftarrow \frac{1}{MK}\mathbf{I}$, $\forall k$ ▷ initialize transmit covariances
$\boldsymbol{\Omega}_l \leftarrow \frac{1}{\mathrm{tr}(\boldsymbol{C}_l)L}\left(\mathbf{I}\right)_{\perp}$, $\forall l$ ▷ initialize noise covariances
$d_\mathrm{t} \leftarrow 1$, $d_\mathrm{n} \leftarrow 1$, $i \leftarrow 1$ ▷ initialize step size
cost_n $\leftarrow \Psi(\boldsymbol{\Sigma}_{1\ldots K}, \sum_l \boldsymbol{\Omega}_l)$ ▷ evaluate objective (13)
repeat
  $\boldsymbol{G}_{\mathrm{t},k} \leftarrow \frac{\partial \Psi}{\partial \boldsymbol{\Sigma}_k^\mathrm{T}}$, $\forall k$ ▷ gradient computation (14)
  $p_\mathrm{t} \leftarrow \frac{1}{\sum_k \mathrm{tr}(\boldsymbol{G}_{\mathrm{t},k})}$ ▷ preconditioning (16)
  repeat
    $s_\mathrm{t} \leftarrow \frac{1}{d_\mathrm{t}\sqrt{i}}$ ▷ set step-size
    $\hat{\boldsymbol{\Sigma}}_k \leftarrow \boldsymbol{\Sigma}_k - p_\mathrm{t} s_\mathrm{t} \boldsymbol{G}_{\mathrm{t},k}$, $\forall k$ ▷ unconstr. update (15)
    $\tilde{\boldsymbol{\Sigma}}_{1\ldots K} \leftarrow \left(\hat{\boldsymbol{\Sigma}}_{1\ldots K}\right)_{\perp}$ ▷ joint projection (17)
    cost_t $\leftarrow \Psi(\tilde{\boldsymbol{\Sigma}}_{1\ldots K}, \sum_l \boldsymbol{\Omega}_l)$ ▷ evaluate objective
    cost_increase $\leftarrow$ cost_t $-$ cost_n
    if cost_increase ≤ 0 **then**
      $d_\mathrm{t} \leftarrow d_\mathrm{t} + 1$ ▷ decrease step-size
    end if
  until cost_increase > 0
  $\boldsymbol{\Sigma}_k \leftarrow \tilde{\boldsymbol{\Sigma}}_k$, $\forall k$ ▷ new covariances
  $\boldsymbol{G}_\mathrm{n} \leftarrow \frac{\partial \Psi}{\partial \boldsymbol{\Omega}^\mathrm{T}}$ ▷ gradient computation (18)
  $p_\mathrm{n} \leftarrow \frac{1}{\mathrm{tr}(\boldsymbol{G}_\mathrm{n} \sum_l \boldsymbol{C}_l)}$ ▷ preconditioning (21)
  $\boldsymbol{G}_{n,1\ldots L} \leftarrow (\boldsymbol{G}_n)_{\perp}$ ▷ tangent cone projection (19)
  repeat
    $s_\mathrm{n} \leftarrow \frac{1}{d_\mathrm{n}\sqrt{i}}$ ▷ set step-size
    $\hat{\boldsymbol{\Omega}}_l \leftarrow \boldsymbol{\Omega}_l - p_\mathrm{n} s_\mathrm{n} \boldsymbol{G}_{\mathrm{n},l}$, $\forall l$ ▷ unconstr. update (20)
    $\tilde{\boldsymbol{\Omega}}_{1\ldots L} \leftarrow \left(\hat{\boldsymbol{\Omega}}_{1\ldots L}\right)_{\perp}$ ▷ joint projection (22)
    cost_n $\leftarrow \Psi(\tilde{\boldsymbol{\Sigma}}_{1\ldots K}, \sum_l \boldsymbol{\Omega}_l)$ ▷ evaluate objective
    cost_decrease $\leftarrow$ cost_n $-$ cost_t
    if cost_decrease ≥ 0 **then**
      $d_\mathrm{n} \leftarrow d_\mathrm{n} + 1$ ▷ decrease step-size
    end if
  until cost_decrease < 0
  $\boldsymbol{\Omega}_l \leftarrow \tilde{\boldsymbol{\Omega}}_l$, $\forall l$ ▷ new covariances
  $i \leftarrow i + 1$ ▷ iteration counter
until cost_increase $\leq \varepsilon$ **and** cost_decrease $\geq \varepsilon$
$\boldsymbol{Q}_{1\ldots K} \leftarrow$ uplink2downlink$(\boldsymbol{\Sigma}_{1\ldots K}, \boldsymbol{\Omega})$ ▷ [1], [2], [15]

It can be seen, that the gradient is the same for all summands of the noise covariance matrix, which results in a slow convergence behavior for multiple constraints. In general, the steepest descend needs a projection of the gradient onto the tangent cone $(\bullet)_{\perp}$:

$$\left\{\boldsymbol{G}_{\mathrm{n},1}^{(i)}, \ldots, \boldsymbol{G}_{\mathrm{n},L}^{(i)}\right\} = \left(\left.\frac{\partial \Psi}{\partial \boldsymbol{\Omega}}\right|_i\right)_{\perp}. \tag{19}$$

For optimizing the transmit covariances, the effect of the tangent cone projection is negligible [16] and, therefore, omitted in equation (15). The tangent cone projection of the noise covariance gradient with a single shaping constraint can also be ignored. Due to page limitations the tangent cone projection is only discussed for multiple linear constraints in Appendix A.

The unconstrained updates of the noise covariance matrix summands are

$$\hat{\boldsymbol{\Omega}}_l^{(i)} = \boldsymbol{\Omega}_l^{(i)} - p_\mathrm{n}^{(i)} s_\mathrm{n}^{(i)} \boldsymbol{G}_{\mathrm{n},l}^{(i)}. \tag{20}$$

The preconditioning scalar $p_\mathrm{n}^{(i)}$ can be found as

$$p_\mathrm{n}^{(i)} = -\frac{P}{\mathrm{tr}\left(\left.\frac{\partial \Psi}{\partial \boldsymbol{\Omega}^\mathrm{T}}\right|_i \sum_l \boldsymbol{C}_l\right)} \tag{21}$$

and the joint stepsize for the noise covariance matrix updates $s_\mathrm{n}^{(i)}$ follows the same rules as the stepsize for the transmit covariance updates.

The steepest descend update of the noise covariance matrix summands

$$\left\{\boldsymbol{\Omega}_1^{(i+1)}, \ldots, \boldsymbol{\Omega}_K^{(i+1)}\right\} = \left(\left\{\hat{\boldsymbol{\Omega}}_1^{(i)}, \ldots, \hat{\boldsymbol{\Omega}}_K^{(i)}\right\}\right)_{\perp}. \tag{22}$$

is found with a joint orthogonal projection step, which is done with a generalized water-spilling as presented in the next Section.

V. Orthogonal Projection of the Uplink Noise

The orthogonal projection onto the constraint set has to minimizes the Euclidean distance between all the unconstrained update steps and the constraint set simultaneously. It can be found with the optimization

$$\left\{\boldsymbol{\Omega}_1^{(i+1)}, \ldots, \boldsymbol{\Omega}_K^{(i+1)}\right\} = \underset{\substack{\boldsymbol{\Omega}_l \succeq \mathbf{0},\, \boldsymbol{\Omega}_l \in \mathcal{Z}_l^{\perp},\, \forall l \\ \sum_l \mathrm{tr}(\boldsymbol{\Omega}_l \boldsymbol{C}_l) = P}}{\operatorname{argmin}} \sum_l \left\|\boldsymbol{\Omega}_l - \hat{\boldsymbol{\Omega}}_l^{(i)}\right\|_\mathrm{F}^2, \tag{23}$$

where $\|\bullet\|_\mathrm{F}$ is the Frobenius norm.

The Lagrangian function can be constructed with the dual variables $\boldsymbol{S}_l$, $\forall l$, for the positive semidefiniteness constraints, $\boldsymbol{T}_l$, $\forall l$, for the subspace constraints, and μ for the joint trace constraint:

$$L = \sum_l \mathrm{tr}\left(\left(\boldsymbol{\Omega}_l - \hat{\boldsymbol{\Omega}}_l^{(i)}\right)\left(\boldsymbol{\Omega}_l - \hat{\boldsymbol{\Omega}}_l^{(i)}\right)^\mathrm{H}\right) - \sum_l \mathrm{tr}\left(\boldsymbol{\Omega}_l \boldsymbol{S}_l\right) - \sum_l \mathrm{tr}\left(\boldsymbol{\Omega}_l \boldsymbol{T}_l\right) + \mu\left(\sum_l \mathrm{tr}\left(\boldsymbol{\Omega}_l \boldsymbol{C}_l\right) - P\right). \tag{24}$$

Setting the derivation of the Lagrangian function with respect to $\boldsymbol{\Omega}_l^\mathrm{T}$ to zero yields

$$\boldsymbol{\Omega}_l = \hat{\boldsymbol{\Omega}}_l^{(i)} + \boldsymbol{S}_l + \boldsymbol{T}_l - \mu \boldsymbol{C}_l \tag{25}$$

For the subspace constraints, the complementary slackness conditions $\mathrm{tr}(\boldsymbol{\Omega}_l \boldsymbol{T}_l) = 0$, $\forall l$, have to hold. As $\boldsymbol{\Omega}_l$ has to be element of the set $\mathcal{Z}_l^{\perp}$, $\boldsymbol{T}_l$ has to be from the orthogonal subspace $\mathcal{Z}_l$.

A. Single Scaled Identity Shaping Constraint

For a single shaping constraint, where $\mathcal{Z}^\mathrm{shape}$ contains only the all zero matrix. The set $\mathcal{Z}_l^{\perp,\mathrm{shape}}$ contains any hermitian matrix of appropriate size. Therefore, $\boldsymbol{T}^\mathrm{shape}$ is zero and Equation (25) becomes

$$\boldsymbol{\Omega}^\mathrm{shape} = \left(\hat{\boldsymbol{\Omega}}^{(i)} - \mu \boldsymbol{C}^\mathrm{shape}\right)_{+}, \tag{26}$$

where the influence of $\boldsymbol{S}^{\text{shape}}$ is replaced by the operation $(\bullet)_+$, which sets all negative eigenvalues to zero. In Equation (26), the water level μ and the eigenvectors of $\boldsymbol{\Omega}^{\text{shape}}$ depend on each other and can only be found iteratively for arbitrary $\boldsymbol{C}^{\text{shape}}$.

If $\boldsymbol{C}^{\text{shape}} = \boldsymbol{C}^{\text{eye}} = c\mathbf{I}$ is a scaled identity matrix, $\boldsymbol{\Omega}^{\text{eye}}$ and $\hat{\boldsymbol{\Omega}}^{(i)}$ have to have the same eigenbasis $\boldsymbol{U}^{\text{eye}}$. Therefore, Equation (26) can be diagonalized:

$$\boldsymbol{\Lambda}^{\text{eye}} = \left(\hat{\boldsymbol{\Lambda}}^{(i)} - \mu c\mathbf{I}\right)_+, \tag{27}$$

where $\boldsymbol{\Lambda}^{\text{eye}}$ and $\hat{\boldsymbol{\Lambda}}^{(i)}$ are diagonal matrices with the eigenvalues $(\lambda_1^{\text{eye}}, \dots, \lambda_M^{\text{eye}})$ and $(\hat{\lambda}_1^{(i)}, \dots, \hat{\lambda}_M^{(i)})$ of $\boldsymbol{\Omega}^{\text{eye}}$ and $\hat{\boldsymbol{\Omega}}^{(i)}$ as their diagonal elements, respectively. The water level μ^{eye} is found by plugging Equation (27) into the trace constraint

$$\operatorname{tr}\left(\boldsymbol{\Omega}^{\text{eye}}\boldsymbol{C}^{\text{eye}}\right) = c\sum_m \left(\hat{\lambda}_m^{(i)} - \mu^{\text{eye}} c\right)_+ = P. \tag{28}$$

Without loss of generality, it can be assumed that the eigenvalues of $\hat{\boldsymbol{\Omega}}^{(i)}$ are sorted in non-increasing order $\hat{\lambda}_1^{(i)} \geq \dots \geq \hat{\lambda}_M^{(i)}$. The water level is

$$\mu^{\text{eye}} = \frac{\sum_{m=1}^{\hat{M}} \hat{\lambda}_m^{(i)} c - P}{\sum_{m=1}^{\hat{M}} c^2}, \tag{29}$$

where the number of non negative eigenvalues $\hat{M}$ of $\boldsymbol{\Omega}^{\text{eye}}$ has to be found. $\hat{M}$ is initialized with M and reduced by one until the termination criterion $\hat{\lambda}_{\hat{M}}^{(i)} - \mu^{\text{eye}} c > 0$ holds. Finally, the orthogonal projected noise covariance matrix reads as

$$\boldsymbol{\Omega}^{(i+1),\text{eye}} = \boldsymbol{U}^{\text{eye}}\left(\hat{\boldsymbol{\Lambda}}^{(i)} - \mu^{\text{eye}} c\mathbf{I}\right)_+ \boldsymbol{U}^{\text{eye,H}}. \tag{30}$$

These steps are exactly the water-spilling algorithm from Hunger et al. presented in [3] for the transmit covariance matrices.

B. Multiple Linear Constraints

For linear constraints, $\mathcal{Z}_l^{\perp}$ is

$$\mathcal{Z}_l^{\perp,\text{lin}} = \left\{\boldsymbol{\Omega}_l^{\text{lin}} : \boldsymbol{\Omega}_l^{\text{lin}} = \omega_l \boldsymbol{A}_l,\ \forall \omega_l \in \mathbb{R}_0^+\right\}. \tag{31}$$

$\boldsymbol{\Omega}_l^{\text{lin}} = \omega_l \boldsymbol{A}_l$ has to be a scaled version of the constraint matrix $\boldsymbol{A}_l$ with the scaling variable ω_l. The same is true for $\hat{\boldsymbol{\Omega}}_l^{(i)} = \hat{\omega}_l^{(i)} \boldsymbol{A}_l$ as can be seen from Appendix A. $\boldsymbol{T}_l^{\text{lin}}$ has to fulfill $\operatorname{tr}(\boldsymbol{A}_l \boldsymbol{T}_l^{\text{lin}}) = 0$. Equation (25) can be multiplied with $\boldsymbol{A}_l$ on both sides from the left. Taking the trace, solving the result for ω_l, and multiplying both sides with $\boldsymbol{A}_l$ again gives

$$\boldsymbol{\Omega}_l^{(i+1),\text{lin}} = \omega_l \boldsymbol{A}_l = \left(\hat{\omega}_l^{(i)} - \mu^{\text{lin}} a_l\right)_+ \boldsymbol{A}_l \tag{32}$$

where $a_l = \operatorname{tr}(\boldsymbol{A}_l \boldsymbol{C}_l^{\text{lin}})$. Without loss of generality, it can be assumed that $\operatorname{tr}(\boldsymbol{A}_l^2) = 1$, $\forall l$. Both sides of Equation (3) can be scaled arbitrarily without changing the constraint.

If all constraints are linear constraints, the water level μ is found by plugging $\boldsymbol{\Omega}_l^{\text{lin}}$ into the sum trace constraint.

$$\mu^{\text{lin}} = \frac{\sum_{l=1}^{\hat{L}} \hat{\omega}_l^{(i)} a_l - P}{\sum_{l=1}^{\hat{L}} a_l^2}, \tag{33}$$

where all $\frac{\hat{\omega}_l^{(i)}}{a_l}$ are sorted in non-increasing order. The number of active constraints $\hat{L}$ is initialized with L and decreased by one until $\hat{\omega}_l^{(i)} - \mu^{\text{lin}} a_l > 0$ holds.

C. Scaled Identity and Multiple Linear Constraints

The scaled identity constraint and multiple linear constraints combined need to fulfill

$$\begin{aligned} P &= \operatorname{tr}\left(\boldsymbol{\Omega}^{\text{eye}}\boldsymbol{C}^{\text{eye}}\right) + \sum_l \operatorname{tr}\left(\boldsymbol{\Omega}_l^{\text{lin}}\boldsymbol{C}_l^{\text{lin}}\right) \\ &= \sum_m \left(\hat{\lambda}_m^{(i)} c - \mu^{\text{comb}} c^2\right)_+ + \sum_l \left(\hat{\omega}_l^{(i)} a_l - \mu^{\text{comb}} a_l^2\right)_+. \end{aligned} \tag{34}$$

The Equations for $\boldsymbol{\Omega}_l^{(i+1),\text{comb}}$ can directly be taken from (30) and (32) for the shaping and linear constraints, respectively. The combined water level is

$$\mu^{\text{comb}} = \frac{\sum_{\theta=1}^{\hat{\Theta}} \hat{\xi}_\theta^{(i)} \psi_\theta - P}{\sum_{\theta=1}^{\hat{\Theta}} \psi_\theta^2}. \tag{35}$$

where $\hat{\xi}_\theta^{(i)}$ is from the joint set of all eigenvalues $\hat{\lambda}_m^{(i)}$ and scaling variables $\hat{\omega}_l^{(i)}$. ψ_θ is from the joint set of all corresponding factors c and a_l. All $\frac{\hat{\xi}_\theta^{(i)}}{\psi_\theta}$ are sorted in non-increasing order. $\hat{\Theta}$ is initialized with $M+L$ and decreased by one until $\hat{\xi}_\theta^{(i)} - \mu^{\text{comb}} \psi_\theta > 0$ holds.

VI. Simulations

Like the proposed alternating algorithm, most existing algorithms for the weighted sum rate maximization with multiple linear constraints are alternating optimizations. In their context, an iteration step is a complete run of the outer loop of our alternating optimization. Comparing the performance of all the different algorithms would require a detailed analysis of their complexity. It has been proven in [3], that the gradient-projection method is very efficient for finding the optimal transmit covariances with fixed noise. We expect it to be also very efficient for finding the optimal noise covariance with fixed transmit covariances. Therefore, we assume that the performance of the existing algorithms cannot exceed the performance of our alternating optimization.

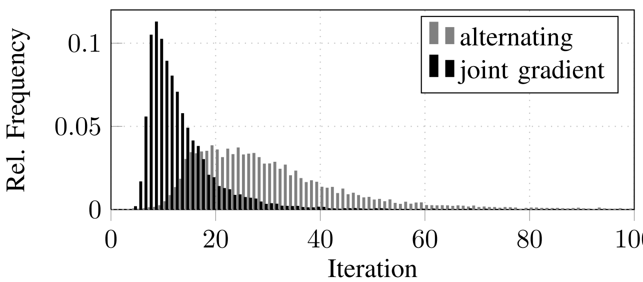

Figure 1. Relative frequency of the number of iterations required to find the maximum weighted sum rate with two linear constraints (sum power and one forbidden direction). The algorithms are stopped, if a sum-rate larger than $(1-\varepsilon)$ times the maximum sum-rate was achieved and each constraint was hurt by less than ε $(\varepsilon = 10^{-3})$.

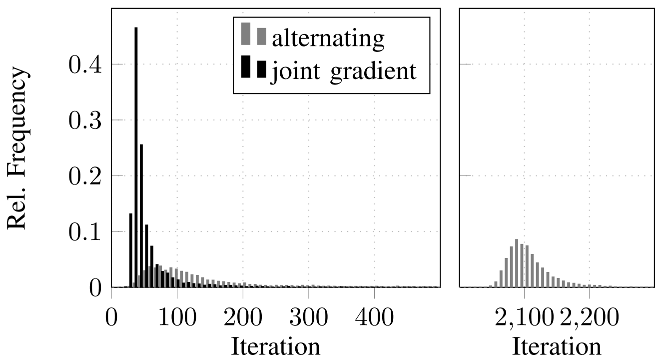

Figure 2. Relative frequency of the number of iterations required to find the maximum weighted sum rate with the scaled identity constraint. The algorithms are stopped, if a sum-rate larger than $(1-\varepsilon)$ times the maximum sum-rate was achieved and the squared Frobenius norm of the constraint error matrix was less than ε ($\varepsilon = 10^{-3}$).

As our joint gradient-projection algorithm does not have inner loops, we count the required iterations of each optimization as the number of times a transmit covariance update step is done. Histograms of the iterations required with the alternating and joint gradient-projection method are compared. Figures 1 and 2 depict two linear constraints and scaled identity shaping, respectively.

In the setup, the BS has $M = 4$ antennas and serves $K = 4$ single antennas users ($N = 1$) with sum power $P = 10$ and arbitrary weights $\boldsymbol{w} = [1, 2, 3, 4]^{\mathrm{T}}$. The simulations are averaged over 10000 i.i.d. channel realizations where each entry of the channel vector $\boldsymbol{H}_k$ has a zero-mean complex Gaussian distribution with variance one. The first linear constraint is the sum power constraint. The second linear constraints forbids to transmit anything in the direction of an additionally generated channel. The maximum sum-rate was found by running the joint algorithm with 1000 iterations, respectively.

VII. Conclusion

It could be shown that the weighted sum rate maximization with multiple linear and/or linear conic constraints can be solved efficiently with a joint gradient-projection algorithm. The parallel update of the transmit and noise covariance matrices has a superior convergence behavior compared to an alternating optimization, which optimizes the transmit and noise covariance matrices in turns.

Appendix

A. Tangent Cone Projection with Multiple Linear Constraints

The steps in [16] are followed to find the tangent cone projection. For linear constraints, the summands of the noise covariance matrix are of the form $\boldsymbol{\Omega}_l^{(i),\mathrm{lin}} = \omega_l^{(i)} \boldsymbol{A}_l$. To lie on the tangent cone, the update $\boldsymbol{\Omega}_l^{(i),\mathrm{lin}} + \epsilon \boldsymbol{G}_l^{(i),\mathrm{lin}}$ with the direction $\boldsymbol{G}_l^{(i),\mathrm{lin}}$ has to fulfill all constraints on $\boldsymbol{\Omega}_l^{\mathrm{lin}}$ for an arbitrary ϵ. Therefore, the update direction has to be a scaled versions of the constraint matrix $\boldsymbol{A}_l$: $\boldsymbol{G}_l^{(i),\mathrm{lin}} = g_l^{(i)} \boldsymbol{A}_l$. All updates have to fulfill jointly the sum trace constraint

$$\sum_l \mathrm{tr}\left(\boldsymbol{C}_l^{\mathrm{lin}}(\boldsymbol{\Omega}_l^{(i),\mathrm{lin}} + \epsilon \boldsymbol{G}_l^{(i),\mathrm{lin}})\right) = P, \qquad \sum_l g_l^{(i)} a_l = 0. \tag{36}$$

The projection of the gradient onto the tangent cone is done by minimizing the distance between the gradient for all summands $\boldsymbol{G}^{(i),\mathrm{lin}}$ and the update for each summand $g_l^{(i)} \boldsymbol{A}_l$:

$$\left\{g_1^{(i)}, \ldots, g_L^{(i)}\right\} = \underset{\sum_l g_l a_l = 0}{\mathrm{argmin}} \sum_l \left\| g_l \boldsymbol{A}_l - \boldsymbol{G}^{(i),\mathrm{lin}} \right\|_{\mathrm{F}}^2. \tag{37}$$

Solving (37) yields

$$g_l = \mathrm{tr}\left(\boldsymbol{A}_l \boldsymbol{G}^{(i),\mathrm{lin}}\right) - \nu a_l, \forall l, \tag{38}$$

$$\nu = \left(\sum_l a_l \,\mathrm{tr}\left(\boldsymbol{A}_l \boldsymbol{G}^{(i),\mathrm{lin}}\right)\right)\left(\sum_l a_l^2\right)^{-1}, \tag{39}$$

where $\mathrm{tr}(\boldsymbol{A}_l^2) = 1$, $\forall l$.

References

[1] S. Vishwanath, N. Jindal, and A. Goldsmith, “Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels,” *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, 2003.

[2] P. Viswanath and D.N.C. Tse, “Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink-Downlink Duality,” *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Oct. 2003.

[3] R. Hunger, D.A. Schmidt, M. Joham, and W. Utschick, “A General Covariance-Based Optimization Framework Using Orthogonal Projections,” in *IEEE SPAWC*, July 2008.

[4] A. Liu, Y. Liu, H. Xiang, and W. Luo, “Iterative Polite Water-Filling for Weighted Sum-Rate Maximization in iTree Networks,” in *IEEE GLOBECOM*, Dec 2010.

[5] W. Yu, W. Rhee, S. Boyd, and J.M. Cioffi, “Iterative Water-Filling for Gaussian Vector Multiple-Access Channels,” *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.

[6] W. Yu, “Uplink-Downlink Duality via Minimax Duality,” *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 361–374, Feb. 2006.

[7] W. Yu and T. Lan, “Transmitter Optimization for the Multi-Antenna Downlink With Per-Antenna Power Constraints,” *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[8] H. Huh, H. Papadopoulos, and G. Caire, “MIMO Broadcast Channel Optimization under General Linear Constraints,” in *IEEE ISIT*, July 2009.

[9] L. Zhang, R. Zhang, Y. Liang, Y. Xin, and H.V. Poor, “On Gaussian MIMO BC-MAC Duality With Multiple Transmit Covariance Constraints,” *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2064–2078, Apr. 2012.

[10] A. Liu, Y. Liu, H. Xiang, and W. Luo, “Polite Water-Filling for Weighted Sum-Rate Maximization in MIMO B-MAC Networks Under Multiple Linear Constraints,” *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 834–847, Feb. 2012.

[11] A. Dotzler, M. Riemensberger, and W. Utschick, “Minimax Duality for MIMO Interference Networks,” *arXiv*, 2014.

[12] H.H. Brunner and J.A. Nossek, “Interference Temperature Analysis in a Large Scale Cellular System,” in *ITG WSA*, Mar. 2013.

[13] A. Dotzler, M. Riemensberger, W. Utschick, and G. Dietl, “Interference Robustness for Cellular MIMO Networks,” in *IEEE SPAWC*, June 2012.

[14] H.H. Brunner, A. Dotzler, W. Utschick, and J.A. Nossek, “Intercell Interference Robustness Tradeoff with Loosened Covariance Shaping,” in *ITG WSA*, Mar. 2014.

[15] A. Liu, Y. Liu, H. Xiang, and W. Luo, “Duality, Polite Water-filling, and Optimization for MIMO B-MAC Interference Networks and iTree Networks,” *http://arxiv.org/abs/1004.2484*, Apr. 2010.

[16] R. Hunger, *Analysis and Transceiver Design for the MIMO Broadcast Channel*, Foundations in Signal Processing, Communications and Networking. Springer, 2013.