Close-to-Optimal Partial Decode-and-Forward Rate in the MIMO Relay Channel via Convex Programming

Thomas Wiegart, Christoph Hellings, and Wolfgang Utschick
Fachgebiet Methoden der Signalverarbeitung, Technische Universität München, 80290 Munich, Germany
Telephone: +49 89 289-28516, Fax: +49 89 289-28504, Email: {thomas.wiegart,hellings,utschick}@tum.de

Abstract—**We study the optimization of the data rate achievable with partial decode-and-forward in the Gaussian MIMO relay channel and propose a new algorithm to find close-to-optimal input covariance matrices. Although the optimization is a non-convex problem in its original formulation, we can show that a reformulated version with slightly modified constraint set is convex. In particular, we replace a positive-semidefiniteness constraint by a strict positive-definiteness constraint. This may introduce an inaccuracy, but we conjecture this inaccuracy to be small so that the obtained solutions are close to the globally optimal ones.**

I. Introduction

The concept of relay networks was introduced by [1] and is of great interest in recent research as the usage of a relay can increase the achievable data rate compared to direct transmission. However, neither the capacity of a relay channel nor the optimal strategy of the relay are known so far. In [2], several strategies such as decode-and-forward and compress-and-forward were proposed. Moreover, the authors of [2] derived the so-called cut-set bound (CSB), which is an upper bound to the capacity.

These general results were later transfered to the case of a Gaussian MIMO relay channel, where all terminals are equipped with multiple antennas, and where additive Gaussian noise is assumed at each receive antenna. In particular, it was shown in [3] that both the CSB and the decode-and-forward rate are maximized by Gaussian channel inputs, and it was shown in [4], [5] that the optimal input covariance matrices can be obtained by solving a convex program.

In this work, we consider the partial decode-and-foward (PDF) scheme (e.g., [6, Section 16.6], [7, Section 9.4.1], [8]–[10]). This extension of decode-and-forward belongs to a wider class of transmit strategies proposed in [2], where the relay decodes only a part of the source message and applies compress-and-forward to the other part. In the special case of PDF, which was first employed in [8], the second part is simply ignored by the relay (i.e., it is compressed to a constant value zero). Compared to classical decode-and-forward, PDF can achieve a higher data rate, especially if the source-relay link is the limiting factor [9], [7, Section 9.4.1]. However, PDF is more difficult to analyze and to optimize due to the second part of the source signal, which acts as an interfering signal at the relay.

The first application of PDF in a Gaussian MIMO relay channel can be found in [10] under the name transmit-side message splitting. Later, it was shown that circularly symmetric Gaussian signals are the optimal input distribution [11], [12] also for this transmit strategy. However, the problem of finding the optimal input covariance matrices is more involved than in the case of decode-and-forward since the mutual information expression becomes non-concave due to the interference at the relay.

The recent works [13], [14] establish upper bounds on the difference between the CSB and the rate achievable by PDF. In addition to such theoretical results, it would also be interesting to study this difference from a numerical perspective to see how small the gap can become in various scenarios if the optimal input covariances are used for PDF. However, there are only suboptimal approaches to maximize the achievable PDF rate (e.g., [15], [16]) and solutions for special cases (e.g., [17]), but there is still no way to find the optimal transmit covariance matrices for the PDF scheme in the general case. Thus, it is not clear whether the observed gap between existing solution methods and the CSB is due to the suboptimal choice of the covariance matrices or rather inherent to the PDF scheme. To answer this question, a globally optimal solution of the PDF rate maximization is needed.

In this work, we take an important step towards finding such a globally optimal solution. We decompose the non-convex PDF rate maximization into an outer optimization over a so-called *innovation covariance matrix* $\boldsymbol{C}$ (cf. [12]) and an inner problem to optimize the remaining parameters for a fixed innovation covariance matrix. By restricting the innovation covariance matrix to be strictly positive-definite (i.e., all eigenvalues have to be greater than or equal to a small positive constant ε instead of being greater than or equal to zero), we obtain an approximated problem, for which we can show that both the inner and outer problems can be solved in a globally optimal manner by means of convex programming. Even though the approximation introduces an inaccuracy in scenarios where the optimal distribution requires a rank-deficient innovation covariance matrix, we conjecture the error to be small when choosing a sufficiently small ε.

To obtain a convex reformulation of the rate optimization, we exploit a result from [12] which shows that an arising subproblem is mathematically equivalent to the maximization of the sum rate with dirty paper coding [18], [19] in a broadcast channel (BC) under a shaping constraint based on the innovation covariance matrix. This BC problem can be transformed into a convex minimax problem in a dual multiple access channel (MAC) [20], [21].

After introducing the system model and the problem formulation in Section II, we give the details of the proposed reformulation in Section III. Therein, we also study the properties of the objective function of the outer problem and show that we can obtain a convex program after a slight modification of the constraint set. In Section IV, we propose a transformation of the inner problem to a simplified equivalent problem, which can then be solved via the dual MAC. After introducing these reformulations and studying the properties of the outer and the inner problem, we turn our attention to the algorithmic aspects in Section V. For the outer problem of finding the optimal innovation covariance matrix, we derive a solution method based on the cutting plane algorithm [22, Section 6.3.3]. The inner problem can be tackled by an alternating gradient-projection method (cf., e.g., [23]). Finally, we discuss the obtained solutions in comparison to an existing suboptimal approach by means of a numerical simulation in Section VI.

Notation: We use $\mathbf{I}$ for the identity matrix of appropriate size. The order relations $\succeq$ and $\succ$ have to be understood in the sense of positive-semidefiniteness and positive-definiteness, respectively. We use the shorthand notation $(\bullet_k)_{\forall k}$ for $(\bullet_1, \dots, \bullet_K)$. The set $\mathbb{H}^N$ is the space of Hermitian matrices of size $N \times N$.

II. System Model

The Gaussian MIMO relay channel consists of a source S with N_S transmit antennas, a destination D with N_D receive antennas, and a relay R with N_R antennas. The source transmits data to the destination over a direct channel and with the help of the relay. The channel matrices of the links source-destination, source-relay, and relay-destination are given by $\boldsymbol{H}_\mathrm{SD} \in \mathbb{C}^{N_\mathrm{D} \times N_\mathrm{S}}$, $\boldsymbol{H}_\mathrm{SR} \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{S}}$, and $\boldsymbol{H}_\mathrm{RD} \in \mathbb{C}^{N_\mathrm{D} \times N_\mathrm{R}}$, respectively. We assume perfect channel knowledge as well as full-duplex transmission with perfect self-interference cancellation at the relay.

A. Partial Decode-and-Forward

Using the partial decode-and-forward strategy (e.g., [7, Section 9.4.1]), the transmit signal $\boldsymbol{x}_\mathrm{S}$ of the source is a superposition of two independent parts $\boldsymbol{u}$ and $\boldsymbol{v}$, where $\boldsymbol{u}$ denotes the part that is sent in cooperation with the relay, and $\boldsymbol{v}$ denotes the part that is directly transmitted without the help of the relay and causes interference at the relay. As proposed in [12], the cooperative part $\boldsymbol{u}$ can be further decomposed into a part $\boldsymbol{q}$ being independent of the relay transmit signal $\boldsymbol{x}_\mathrm{R}$ and a part $\boldsymbol{z}$ being linearly dependent on the relay transmit signal:

$$\boldsymbol{x}_\mathrm{S} = \boldsymbol{u} + \boldsymbol{v} = \boldsymbol{A}\boldsymbol{x}_\mathrm{R} + \boldsymbol{q} + \boldsymbol{v} = \boldsymbol{z} + \boldsymbol{q} + \boldsymbol{v}. \tag{1}$$

As $\boldsymbol{z}$ has linear dependence on the relay transmit signal, and the relay can, due to causality, only transmit data it has previously received,[1] $\boldsymbol{z}$ does not contain new information. The remaining parts $\boldsymbol{q}$ and $\boldsymbol{v}$, which contain new information, are then called *innovation*, and the covariance matrix $\boldsymbol{C} = \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}$ is called *innovation covariance matrix* [12], where $\boldsymbol{C}_{\boldsymbol{v}}$ and $\boldsymbol{C}_{\boldsymbol{q}}$ denote the covariance matrices of $\boldsymbol{v}$ and $\boldsymbol{q}$, respectively.

[1]Note that the rate expressions given in Section II-B are achievable using a block-Markov coding scheme [7, Section 9.4.1].

B. Achievable Data Rates

The achievable data rate with the partial decode-and-forward scheme and circularly symmetric Gaussian signals is given as the minimum of two mutual information expressions

$$R = \min\{R_\mathrm{A}, R_\mathrm{B}\} \tag{2}$$

[7, Section 9.4.1], where R_A and R_B can be expressed as [12]

$$R_\mathrm{A} = \log_2 \det(\mathbf{I}_{N_\mathrm{D}} + \boldsymbol{H}_\mathrm{SD}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{H}_\mathrm{SD}^\mathrm{H}) + \log_2 \frac{\det(\mathbf{I}_{N_\mathrm{R}} + \boldsymbol{H}_\mathrm{SR}(\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}})\boldsymbol{H}_\mathrm{SR}^\mathrm{H})}{\det(\mathbf{I}_{N_\mathrm{R}} + \boldsymbol{H}_\mathrm{SR}\boldsymbol{C}_{\boldsymbol{v}}\boldsymbol{H}_\mathrm{SR}^\mathrm{H})} \tag{3}$$

$$R_\mathrm{B} = \log_2 \det(\mathbf{I}_{N_\mathrm{D}} + \boldsymbol{H}_\mathrm{SD}(\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}})\boldsymbol{H}_\mathrm{SD}^\mathrm{H} + \boldsymbol{H}\boldsymbol{R}\boldsymbol{H}^\mathrm{H}). \tag{4}$$

We have assumed the noise covariances $\boldsymbol{C}_{\boldsymbol{\eta}_\mathrm{R}} = \mathbf{I}_{N_\mathrm{R}}$ and $\boldsymbol{C}_{\boldsymbol{\eta}_\mathrm{D}} = \mathbf{I}_{N_\mathrm{D}}$ w.l.o.g., and we use the joint channel matrix

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_\mathrm{SD} & \boldsymbol{H}_\mathrm{RD} \end{bmatrix}. \tag{5}$$

The joint covariance matrix of $\boldsymbol{z}$ and $\boldsymbol{x}_\mathrm{R}$ is denoted by

$$\boldsymbol{R} = \boldsymbol{C}_{\left[\begin{smallmatrix}\boldsymbol{z}\\ \boldsymbol{x}_\mathrm{R}\end{smallmatrix}\right]} = \mathrm{E}\left[\begin{bmatrix}\boldsymbol{z}\\ \boldsymbol{x}_\mathrm{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{z}\\ \boldsymbol{x}_\mathrm{R}\end{bmatrix}^\mathrm{H}\right]. \tag{6}$$

C. Problem Formulation

We aim at maximizing the achievable data rate under constraints on the transmit powers of the source and of the relay. The power constraints can be formulated as

$$\mathrm{E}\left[\|\boldsymbol{x}_\mathrm{S}\|_2^2\right] = \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}} + \boldsymbol{D}_\mathrm{S}\boldsymbol{R}\boldsymbol{D}_\mathrm{S}^\mathrm{H}) \leq P_\mathrm{S} \tag{7}$$

$$\mathrm{E}\left[\|\boldsymbol{x}_\mathrm{R}\|_2^2\right] = \mathrm{tr}(\boldsymbol{D}_\mathrm{R}\boldsymbol{R}\boldsymbol{D}_\mathrm{R}^\mathrm{H}) \leq P_\mathrm{R} \tag{8}$$

with the selection matrices

$$\boldsymbol{D}_\mathrm{S} = \begin{bmatrix}\mathbf{I}_{N_\mathrm{S}} & \mathbf{0}_{N_\mathrm{S}\times N_\mathrm{R}}\end{bmatrix} \text{ and } \boldsymbol{D}_\mathrm{R} = \begin{bmatrix}\mathbf{0}_{N_\mathrm{R}\times N_\mathrm{S}} & \mathbf{I}_{N_\mathrm{R}}\end{bmatrix}. \tag{9}$$

The optimization problem we consider is then given by

$$\begin{aligned} \max_{\substack{\boldsymbol{C}_{\boldsymbol{v}}\succeq\mathbf{0}, \boldsymbol{C}_{\boldsymbol{q}}\succeq\mathbf{0}\\ \boldsymbol{R}\succeq\mathbf{0}}} \; & \min\left\{R_\mathrm{A}(\boldsymbol{C}_{\boldsymbol{v}}, \boldsymbol{C}_{\boldsymbol{q}}), R_\mathrm{B}(\boldsymbol{C}_{\boldsymbol{v}}, \boldsymbol{C}_{\boldsymbol{q}}, \boldsymbol{R})\right\} \\ \text{s.t.} \; & \mathrm{tr}(\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}) + \mathrm{tr}(\boldsymbol{D}_\mathrm{S}\boldsymbol{R}\boldsymbol{D}_\mathrm{S}^\mathrm{H}) \leq P_\mathrm{S} \\ & \mathrm{tr}(\boldsymbol{D}_\mathrm{R}\boldsymbol{R}\boldsymbol{D}_\mathrm{R}^\mathrm{H}) \leq P_\mathrm{R} \end{aligned} \tag{10}$$

In (1), we have assumed that $\boldsymbol{z}$ and $\boldsymbol{x}_\mathrm{R}$ are fully correlated so that $\boldsymbol{z}$ can be expressed as $\boldsymbol{A}\boldsymbol{x}_\mathrm{R}$. This would impose an additional structural constraint on the joint covariance matrix $\boldsymbol{R}$ of $\boldsymbol{z}$ and $\boldsymbol{x}_\mathrm{R}$. However, this constraint can be neglected since it is automatically fulfilled in the optimum, which can be seen as follows. Assume that we have

$$\boldsymbol{R} = \underbrace{\mathrm{E}\left[\begin{bmatrix}\boldsymbol{A}\boldsymbol{x}_\mathrm{R}\\ \boldsymbol{x}_\mathrm{R}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}\boldsymbol{x}_\mathrm{R}\\ \boldsymbol{x}_\mathrm{R}\end{bmatrix}^\mathrm{H}\right]}_{\boldsymbol{R}'} + \begin{bmatrix}\boldsymbol{B} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} \tag{11}$$

with $\boldsymbol{B} \succeq \mathbf{0}$. Then, we could replace $\boldsymbol{R}$ by $\boldsymbol{R}'$ and $\boldsymbol{C}_{\boldsymbol{q}}$ by $\boldsymbol{C}'_{\boldsymbol{q}} = \boldsymbol{C}_{\boldsymbol{q}} + \boldsymbol{B}$. By doing so, the left side of the constraints as well as the rate R_B would be left unchanged (to see this, note the definition of $\boldsymbol{H}$ in (5)), and the rate R_{A} would be increased or left unchanged.

As already mentioned before, our approach is based on restricting the optimization to cases where the innovation covariance $\boldsymbol{C} = \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}$ is strictly positive-definite. Therefore, we will later consider a slightly modified problem with the additional constraint $\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}} \succeq \varepsilon \mathbf{I}$ for a given small value of ε.

III. Primal Decomposition

The optimization problem (10) is non-convex in its original form. However, we can reformulate the problem and show the reformulated version to be convex, as long as the innovation covariance matrix is strictly positive-definite. To do so, we apply the concept of primal decomposition [24] to (10) with the innovation covariance matrix $\boldsymbol{C}$ and the matrix $\boldsymbol{R}$ as coupling variables. A similar approach was pursued in [12] as part of a proof, but has not yet been considered for algorithm design. We obtain the optimization problem

$$\begin{aligned}\max_{\boldsymbol{C}\succeq\mathbf{0},\boldsymbol{R}\succeq\mathbf{0}} \quad & \min\left\{R_{\mathrm{A}}^{\star}(\boldsymbol{C}), R_{\mathrm{B}}(\boldsymbol{C},\boldsymbol{R})\right\} \\ \text{s.t.} \quad & \operatorname{tr}(\boldsymbol{C}) + \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{R}\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}} \\ & \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{R}\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}\end{aligned} \tag{12}$$

with

$$R_{\mathrm{A}}^{\star}(\boldsymbol{C}) = \max_{\boldsymbol{C}_{\boldsymbol{v}}\succeq\mathbf{0},\boldsymbol{C}_{\boldsymbol{q}}\succeq\mathbf{0}} \quad R_{\mathrm{A}} \quad \text{s.t.} \quad \boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}} \preceq \boldsymbol{C} \tag{13}$$

and

$$R_{\mathrm{B}}(\boldsymbol{C},\boldsymbol{R}) = \log_2\det(\mathbf{I}_{N_{\mathrm{D}}} + \boldsymbol{H}_{\mathrm{SD}}\boldsymbol{C}\boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H}\boldsymbol{R}\boldsymbol{H}^{\mathrm{H}}). \tag{14}$$

Due to the max-min-inequality, the reformulation (12) is an upper bound to the original problem (10), but the bound is tight [12] since R_{B} is no longer a function of $\boldsymbol{C}_{\boldsymbol{v}}$ and $\boldsymbol{C}_{\boldsymbol{q}}$ if $\boldsymbol{C}$ is given, i.e., the inner optimization over $\boldsymbol{C}_{\boldsymbol{v}}$ and $\boldsymbol{C}_{\boldsymbol{q}}$ is only needed for R_{A}.

Note that the problem (13) for given $\boldsymbol{C}$ is mathematically equivalent to maximizing the sum rate in a two-user MIMO broadcast channel with dirty paper coding [18], [19] under a shaping constraint. In particular, the right side of the shaping constraint is determined by the innovation covariance matrix (cf. [12]), and we have identity matrices as noise covariance matrices at the receivers, i.e.,

$$\max_{\substack{(\boldsymbol{Q}_k\succeq\mathbf{0})\forall k \\ \sum_k \boldsymbol{Q}_k \preceq \boldsymbol{C}}} \quad \sum_{k=1}^{K} \log_2 \frac{\det\left(\mathbf{I} + \sum_{i=k}^{K}\boldsymbol{H}_k\boldsymbol{Q}_i\boldsymbol{H}_k^{\mathrm{H}}\right)}{\det\left(\mathbf{I} + \sum_{i=k+1}^{K}\boldsymbol{H}_k\boldsymbol{Q}_i\boldsymbol{H}_k^{\mathrm{H}}\right)} \tag{15}$$

with $K=2$. This mathematical equivalence to an optimization in a MIMO broadcast channel cannot only be exploited to evaluate $R_{\mathrm{A}}^{\star}$ for given $\boldsymbol{C}$ (see next section), but also to prove the following theorem.

A. Properties of $R_{\mathrm{A}}^{\star}$

Theorem 1. *For a strictly positive-definite innovation covariance matrix $\boldsymbol{C}$, the expression $R_{\mathrm{A}}^{\star}(\boldsymbol{C})$ from* (13) *is concave in $\boldsymbol{C}$, and a (concave) subgradient is given by the optimal Lagrangian multiplier $\boldsymbol{\Omega}$ for the shaping constraint $\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}} \preceq \boldsymbol{C}$.*

Proof: In the following, we perform a so-called sensitivity analysis [25, Section 5.6], which was already done in [26] for a sum rate maximization under a power constraint. We extend this approach to the sum rate maximization with a shaping constraint.

As a first step, we express $R_{\mathrm{A}}^{\star}(\boldsymbol{C})$ from (13) by its Lagrangian dual function:

$$R_{\mathrm{A}}^{\star}(\boldsymbol{C}) = \min_{\boldsymbol{\Omega}\succeq\mathbf{0}} \max_{\boldsymbol{C}_{\boldsymbol{v}}\succeq\mathbf{0},\boldsymbol{C}_{\boldsymbol{q}}\succeq\mathbf{0}} R_{\mathrm{A}} + \operatorname{tr}\left(\boldsymbol{\Omega}(\boldsymbol{C} - (\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}}))\right) \tag{16}$$

Equality in (16) holds for a full-rank innovation covariance $\boldsymbol{C}$ as $R_{\mathrm{A}}^{\star}$ is mathematically equivalent to the BC sum rate maximization with a shaping constraint (15), which has a zero duality gap provided that the shaping matrix at the right side of the shaping constraint has full rank [20].

The optimal value of the minimization in (16) can be bounded from above by

$$\begin{aligned} R_{\mathrm{A}}^{\star}(\boldsymbol{C}) &\leq \max_{\boldsymbol{C}_{\boldsymbol{v}}\succeq\mathbf{0},\boldsymbol{C}_{\boldsymbol{q}}\succeq\mathbf{0}} R_{\mathrm{A}} + \operatorname{tr}\left(\widetilde{\boldsymbol{\Omega}}\left(\boldsymbol{C} - (\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}})\right)\right) && (17) \\ &= \max_{\boldsymbol{C}_{\boldsymbol{v}}\succeq\mathbf{0},\boldsymbol{C}_{\boldsymbol{q}}\succeq\mathbf{0}} R_{\mathrm{A}} + \operatorname{tr}\left(\widetilde{\boldsymbol{\Omega}}\left(\widetilde{\boldsymbol{C}} - (\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}})\right)\right) \\ &\qquad + \operatorname{tr}\left(\widetilde{\boldsymbol{\Omega}}\left(\boldsymbol{C} - \widetilde{\boldsymbol{C}}\right)\right) && (18) \end{aligned}$$

where (17) is valid for all $\widetilde{\boldsymbol{\Omega}} \succeq \mathbf{0}$, and (18) arises from (17) by adding and substracting the term $\operatorname{tr}(\widetilde{\boldsymbol{\Omega}}\widetilde{\boldsymbol{C}})$ for a $\widetilde{\boldsymbol{C}} \succeq \mathbf{0}$.

Since (17) holds for all $\widetilde{\boldsymbol{\Omega}} \succeq \mathbf{0}$, it is also valid if the matrix $\widetilde{\boldsymbol{\Omega}}$ is chosen as the optimal Lagrangian multiplier corresponding to the shaping constraint $\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}} \preceq \widetilde{\boldsymbol{C}}$. Thus, (18) can be expressed as

$$R_{\mathrm{A}}^{\star}(\boldsymbol{C}) \leq \underbrace{R_{\mathrm{A}}^{\star}(\widetilde{\boldsymbol{C}}) + \left\langle \widetilde{\boldsymbol{\Omega}}, \boldsymbol{C} - \widetilde{\boldsymbol{C}} \right\rangle}_{\hat{R}_{\mathrm{A}}^{\star}(\boldsymbol{C};\widetilde{\boldsymbol{C}})} \tag{19}$$

and we can identify $\widetilde{\boldsymbol{\Omega}}$ to be a (concave) subgradient of $R_{\mathrm{A}}^{\star}(\widetilde{\boldsymbol{C}})$ for a strictly positive-definite $\widetilde{\boldsymbol{C}}$. Additionally, since $R_{\mathrm{A}}^{\star}(\boldsymbol{C})$ can be bounded from above by the linear approximation $\hat{R}_{\mathrm{A}}^{\star}(\boldsymbol{C};\widetilde{\boldsymbol{C}})$ around any $\widetilde{\boldsymbol{C}} \succ \mathbf{0}$, the expression $R_{\mathrm{A}}^{\star}(\boldsymbol{C})$ is concave in $\boldsymbol{C}$ for strictly positive-definite $\boldsymbol{C}$. ■

Remark: If $\boldsymbol{C}$ is not strictly positive-definite, but has also eigenvalues that are zero, the proof from [20] for the zero duality gap in (16) does not hold. Thus, we cannot extend the proof presented above to a rank-deficient positive-semidefinite $\boldsymbol{C}$.

B. Properties of R_{B}

A similar result can be obtained for $R_{\mathrm{B}}(\boldsymbol{C},\boldsymbol{R})$. As $\log\det(\boldsymbol{X})$ is concave and differentiable in $\boldsymbol{X}$ [25, Section 3.1.5], the rate expression $R_{\mathrm{B}}(\boldsymbol{C},\boldsymbol{R})$ is jointly concave in

$\boldsymbol{C}$ and $\boldsymbol{R}$, and the gradient matrices $\frac{\partial R_{\mathrm{B}}}{\partial \boldsymbol{C}^{\mathrm{T}}}$ and $\frac{\partial R_{\mathrm{B}}}{\partial \boldsymbol{R}^{\mathrm{T}}}$ form a (concave) subgradient. Thus, a linear approximation of R_{B} around $(\widetilde{\boldsymbol{C}}, \widetilde{\boldsymbol{R}})$ is given by

$$R_{\mathrm{B}}(\boldsymbol{C}, \boldsymbol{R}) \leq \hat{R}_{\mathrm{B}}(\boldsymbol{C}, \boldsymbol{R}; \widetilde{\boldsymbol{C}}, \widetilde{\boldsymbol{R}}) \tag{20}$$

$$= R_{\mathrm{B}}(\widetilde{\boldsymbol{C}}, \widetilde{\boldsymbol{R}}) + \left\langle \frac{\partial R_{\mathrm{B}}}{\partial \boldsymbol{C}^{\mathrm{T}}}\Big|_{\widetilde{\boldsymbol{C}}}, \boldsymbol{C} - \widetilde{\boldsymbol{C}} \right\rangle + \left\langle \frac{\partial R_{\mathrm{B}}}{\partial \boldsymbol{R}^{\mathrm{T}}}\Big|_{\widetilde{\boldsymbol{R}}}, \boldsymbol{R} - \widetilde{\boldsymbol{R}} \right\rangle \tag{21}$$

where the gradient matrices $\frac{\partial R_{\mathrm{B}}}{\partial \boldsymbol{C}^{\mathrm{T}}}$ and $\frac{\partial R_{\mathrm{B}}}{\partial \boldsymbol{R}^{\mathrm{T}}}$ are given by

$$\frac{\partial R_{\mathrm{B}}}{\partial \boldsymbol{C}^{\mathrm{T}}} = \frac{1}{\ln 2} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} \boldsymbol{D}^{-1} \boldsymbol{H}_{\mathrm{SD}} \tag{22}$$

$$\frac{\partial R_{\mathrm{B}}}{\partial \boldsymbol{R}^{\mathrm{T}}} = \frac{1}{\ln 2} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{D}^{-1} \boldsymbol{H} \tag{23}$$

with

$$\boldsymbol{D} = \mathbf{I}_{N_{\mathrm{D}}} + \boldsymbol{H}_{\mathrm{SD}} \boldsymbol{C} \boldsymbol{H}_{\mathrm{SD}}^{\mathrm{H}} + \boldsymbol{H} \boldsymbol{R} \boldsymbol{H}^{\mathrm{H}}. \tag{24}$$

C. Modified Optimization Problem

Due to the restriction to a strictly positive-definite innovation covariance matrix $\boldsymbol{C}$, we slightly modify the decomposed optimization problem (12) and introduce the additional constraint $\boldsymbol{C} \succeq \varepsilon \mathbf{I}$. This leads to the following problem.

$$\begin{aligned} \max_{\boldsymbol{C} \succeq \varepsilon \mathbf{I}, \boldsymbol{R} \succeq \mathbf{0}} \quad & \min \left\{ R_{\mathrm{A}}^{\star}(\boldsymbol{C}), R_{\mathrm{B}}(\boldsymbol{C}, \boldsymbol{R}) \right\} \\ \text{s.t.} \quad & \operatorname{tr}(\boldsymbol{C}) + \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}} \boldsymbol{R} \boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}} \\ & \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}} \boldsymbol{R} \boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}. \end{aligned} \tag{25}$$

Although this modification may lead to a suboptimal solution in the case where the optimal solution of the original problem has a rank-deficient $\boldsymbol{C}$, we conjecture the error we make by introducing the additional constraint $\boldsymbol{C} \succeq \varepsilon \mathbf{I}$ to be small for sufficiently small ε. However, the following important corollary holds for (25).

Corollary 1. *The modified optimization problem* (25) *is a convex program.*

Proof: The pointwise minimum of concave functions is concave [25, Section 3.2.3], R_{B} is jointly concave in $\boldsymbol{C}$ and $\boldsymbol{R}$, and $R_{\mathrm{A}}^{\star}$ is concave for $\boldsymbol{C} \succeq \varepsilon \mathbf{I}$, cf. Theorem 1 ∎

IV. Inner Problem: Broadcast Sum Rate Maximization with Shaping Constraint

What remains to be done is solving the inner problem of evaluating $R_{\mathrm{A}}^{\star}$ for fixed $\boldsymbol{C}$, which is mathematically equivalent to a sum rate maximization in the broadcast channel. Although the broadcast sum rate maximization (13) is a non-convex problem, it can be solved by transforming it into a convex minimax problem in the dual multiple access channel.

In the first part of this section, we apply the duality framework presented in [20], [21] to obtain a convex problem in the multiple access channel for a BC sum rate maximization problem under a shaping constraint with $\boldsymbol{C} = \mathbf{I}$ and regular channels. In the second part, we extend this to a general full-rank matrix $\boldsymbol{C}$ and to rank-deficient channel matrices by introducing an appropriate transformation. Finally, we discuss why the results cannot be easily extended to the case of a rank-deficient $\boldsymbol{C}$.

A. MAC-BC-Duality

The classical uplink downlink duality from [18] only holds under a sum power constraint. Therefore, we make instead use of the duality framework from [20], [21], which can handle problems in the form

$$\min_{\substack{(\boldsymbol{C}_{\boldsymbol{\eta}_k} \succ \mathbf{0})_{\forall k} \\ (\boldsymbol{C}_{\boldsymbol{\eta}_k})_{\forall k} \in \mathcal{Y}^{\perp} \\ \sum_k \operatorname{tr}(\boldsymbol{B}_k \boldsymbol{C}_{\boldsymbol{\eta}_k}) = \sigma^2}} \max_{\substack{(\boldsymbol{Q}_k \succeq \mathbf{0})_{\forall k} \\ \boldsymbol{Z} \in \mathcal{Z} \\ \sum_k \boldsymbol{Q}_k \preceq \boldsymbol{C} + \boldsymbol{Z}}} R_{\mathrm{BC}}((\boldsymbol{Q}_k, \boldsymbol{C}_{\boldsymbol{\eta}_k})_{\forall k}; (\boldsymbol{H}_k)_{\forall k}) \tag{26}$$

with

$$R_{\mathrm{BC}} = \sum_{k=1}^{K} \log_2 \frac{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_k} + \sum_{i=k}^{K} \boldsymbol{H}_k \boldsymbol{Q}_i \boldsymbol{H}_k^{\mathrm{H}}\right)}{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}_k} + \sum_{i=k+1}^{K} \boldsymbol{H}_k \boldsymbol{Q}_i \boldsymbol{H}_k^{\mathrm{H}}\right)} \tag{27}$$

denoting the BC sum rate depending on the transmit covariance matrices $(\boldsymbol{Q}_k)_{\forall k}$, the noise covariances $(\boldsymbol{C}_{\boldsymbol{\eta}_k})_{\forall k}$ at the BC receivers, and the channel matrices $(\boldsymbol{H}_k)_{\forall k}$. The number of users is denoted by K and quantities with an index k refer to quantities of the k-th user. The linear subsets $\mathcal{Z} \subseteq \mathbb{H}^N$ and $\mathcal{Y} \subseteq \mathbb{H}^{M_1} \times \cdots \times \mathbb{H}^{M_K}$ can be used to model specific constraints. Here, N refers to the number of antennas at the base station and M_k to the number of antennas at the k-th receiver.

The corresponding uplink problem achieving the same sum rate is given by

$$\min_{\substack{\boldsymbol{C}_{\boldsymbol{\eta}} \succ \mathbf{0} \\ \boldsymbol{C}_{\boldsymbol{\eta}} \in \mathcal{Z}^{\perp} \\ \operatorname{tr}(\boldsymbol{C} \boldsymbol{C}_{\boldsymbol{\eta}}) = \sigma^2}} \max_{\substack{(\boldsymbol{\Sigma}_k \succeq \mathbf{0})_{\forall k} \\ (\boldsymbol{Y}_k)_{\forall k} \in \mathcal{Y} \\ \boldsymbol{\Sigma}_k \preceq \boldsymbol{B}_k + \boldsymbol{Y}_k \ \forall k}} R_{\mathrm{MAC}}((\boldsymbol{\Sigma}_k)_{\forall k}, \boldsymbol{C}_{\boldsymbol{\eta}}; (\boldsymbol{H}_k)_{\forall k}) \tag{28}$$

with

$$R_{\mathrm{MAC}} = \log_2 \frac{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}} + \sum_{k=1}^{K} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{\Sigma}_k \boldsymbol{H}_k\right)}{\det\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)} \tag{29}$$

denoting the sum rate of the dual multiple access channel depending on the transmit covariance matrices $(\boldsymbol{\Sigma}_k)_{\forall k}$ of the users, the noise covariance matrix $\boldsymbol{C}_{\boldsymbol{\eta}}$ at the base station, and the channel matrices $(\boldsymbol{H}_k)_{\forall k}$ of the original broadcast channel. Note that (29) is convex in $\boldsymbol{C}_{\boldsymbol{\eta}}$ and concave in $(\boldsymbol{\Sigma}_k)_{\forall k}$ [21].

For the purpose of evaluating $R_{\mathrm{A}}^{\star}(\boldsymbol{C})$ in (13), we are only interested in the case $\boldsymbol{C}_{\boldsymbol{\eta}_1} = \boldsymbol{C}_{\boldsymbol{\eta}_2} = \mathbf{I}$. Therefore, we omit the variables $\boldsymbol{C}_{\boldsymbol{\eta}_k}$ of the function $R_{\mathrm{BC}}((\boldsymbol{Q}_k, \boldsymbol{C}_{\boldsymbol{\eta}_k})_{\forall k}; (\boldsymbol{H}_k)_{\forall k})$ from now on. The broadcast sum rate maximization problem with a shaping constraint is then given by

$$\max_{(\boldsymbol{Q}_k \succeq \mathbf{0})_{\forall k}} R_{\mathrm{BC}}((\boldsymbol{Q}_k)_{\forall k}; (\boldsymbol{H}_k)_{\forall k}) \quad \text{s.t.} \quad \sum_{k=1}^{K} \boldsymbol{Q}_k \preceq \boldsymbol{C} \tag{30}$$

where $K = 2$. If we in addition assume $\boldsymbol{C} = \mathbf{I}$, the application of the duality framework leads to the dual uplink problem

$$\min_{\substack{\boldsymbol{C}_{\boldsymbol{\eta}} \succ \mathbf{0} \\ \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}}) = P}} \max_{\substack{(\boldsymbol{\Sigma}_k \succeq \mathbf{0})_{\forall k} \\ \sum_k \operatorname{tr}(\boldsymbol{\Sigma}_k) = P}} R_{\mathrm{MAC}}((\boldsymbol{\Sigma}_k)_{\forall k}, \boldsymbol{C}_{\boldsymbol{\eta}}; (\boldsymbol{H}_k)_{\forall k}) \tag{31}$$

where both the transmit covariance matrices and the noise covariance matrix are subject to a power constraint with $P = N_{\mathrm{D}} + N_{\mathrm{R}}$ [21].

B. General Shaping Constraint and Rank-Deficient Channels

In order to handle general shaping constraints with a full-rank matrix $\boldsymbol{C} \neq \mathbf{I}$, we perform a transformation of the broadcast problem in the following and include the matrix $\boldsymbol{C}$ into the channel matrices. Furthermore, we perform transformations to eliminate row rank deficiencies of the channels $(\boldsymbol{H}_k)_{\forall k}$ and to eliminate column rank deficiencies of the joint channel matrix $\left[\boldsymbol{H}_1^{\mathrm{T}}, \boldsymbol{H}_2^{\mathrm{T}}, \ldots, \boldsymbol{H}_K^{\mathrm{T}}\right]^{\mathrm{T}}$.

While eliminating row rank deficiencies just reduces the dimensionality of the expressions, the elimination of the column rank deficiencies is crucial for the algorithm that is applied to solve (31). The reason for this is that a rank-deficient joint channel matrix would lead to a rank-deficient noise covariance matrix in the dual MAC.

In the following, we use $\boldsymbol{\Lambda}_k$ to denote the Lagrangian multiplier for the constraint $\boldsymbol{Q}_k \succeq \mathbf{0}$ in (30), and $\boldsymbol{\Omega}$ is the Lagrangian multiplier for the shaping constraint $\sum_k \boldsymbol{Q}_k \preceq \boldsymbol{C}$.

Proposition 1. *Let $\boldsymbol{C}_{\boldsymbol{\eta}_k} = \mathbf{I}\ \ \forall k$, and $\boldsymbol{C} = \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{C}^{\frac{\mathrm{H}}{2}} \succ \mathbf{0}$ with square $\boldsymbol{C}^{\frac{1}{2}}$. Furthermore, let*

$$\begin{bmatrix}\boldsymbol{U}_k & \boldsymbol{U}_k^{\perp}\end{bmatrix} \begin{bmatrix}\boldsymbol{S}_k & \\ & \mathbf{0}\end{bmatrix} \boldsymbol{V}_k^{\mathrm{H}} = \boldsymbol{H}_k \boldsymbol{C}^{\frac{1}{2}} \quad \forall k \tag{32}$$

and

$$\boldsymbol{U} \begin{bmatrix}\boldsymbol{S} & \\ & \mathbf{0}\end{bmatrix} \begin{bmatrix}\boldsymbol{V} & \boldsymbol{V}^{\perp}\end{bmatrix}^{\mathrm{H}} = \begin{bmatrix}\boldsymbol{U}_1^{\mathrm{H}} \boldsymbol{H}_1 \boldsymbol{C}^{\frac{1}{2}} \\ \cdots \\ \boldsymbol{U}_K^{\mathrm{H}} \boldsymbol{H}_K \boldsymbol{C}^{\frac{1}{2}}\end{bmatrix} \tag{33}$$

be singular value decompositions where $(\boldsymbol{S}_k)_{\forall k}$ and $\boldsymbol{S}$ have full rank. Then, the optimum of (30) *can be found by solving a transformed problem*

$$\max_{(\boldsymbol{Q}_k' \succeq \mathbf{0})_{\forall k}} R_{\mathrm{BC}}((\boldsymbol{Q}_k')_{\forall k}); (\boldsymbol{H}_k')_{\forall k}) \quad \text{s.t.} \quad \sum_{k=1}^{K} \boldsymbol{Q}_k' \preceq \mathbf{I} \tag{34}$$

with channels

$$\boldsymbol{H}_k' = \boldsymbol{U}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V}. \tag{35}$$

In particular, the optimal primal and dual variables of (30) *are given by*

$$\boldsymbol{Q}_k = \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V} \boldsymbol{Q}_k' \boldsymbol{V}^{\mathrm{H}} \boldsymbol{C}^{\frac{\mathrm{H}}{2}} \quad \forall k \tag{36}$$

$$\boldsymbol{\Lambda}_k = \boldsymbol{C}^{-\frac{\mathrm{H}}{2}} \boldsymbol{V} \boldsymbol{\Lambda}_k' \boldsymbol{V}^{\mathrm{H}} \boldsymbol{C}^{-\frac{1}{2}} \quad \forall k \tag{37}$$

$$\boldsymbol{\Omega} = \boldsymbol{C}^{-\frac{\mathrm{H}}{2}} \boldsymbol{V} \boldsymbol{\Omega}' \boldsymbol{V}^{\mathrm{H}} \boldsymbol{C}^{-\frac{1}{2}} \tag{38}$$

where $(\boldsymbol{Q}_k')_{\forall k}$, $(\boldsymbol{\Lambda}_k')_{\forall k}$ and $\boldsymbol{\Omega}'$ are the optimal primal and dual variables of (34).

Proof: See Appendix A ■

C. Rank-Deficient Shaping Matrix

So far, all derivations have been based on the assumption that $\boldsymbol{C}$ has full rank. For instance, a matrix inversion $\boldsymbol{C}^{-\frac{1}{2}}$ is performed in the transformation rule (38), and it is also required at some points in the proof of Proposition 1 that $\boldsymbol{C}$ is invertible. For this reason, the case of a singular matrix $\boldsymbol{C}$ is explicitly excluded in Theorem 1. However, we now want to discuss whether the approach can be extended to a singular matrix $\boldsymbol{C}$.

If $\boldsymbol{C}$ has eigenvalues equal to zero, the corresponding eigenvectors can be interpreted as forbidden directions in which no signal power can be used. This leads to the following result.

Proposition 2. *Let*

$$\begin{bmatrix}\boldsymbol{W}_1 & \boldsymbol{W}_2\end{bmatrix} \begin{bmatrix}\boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\boldsymbol{W}_1^{\mathrm{H}} \\ \boldsymbol{W}_2^{\mathrm{H}}\end{bmatrix} = \boldsymbol{C} \tag{39}$$

be the eigenvalue decomposition of $\boldsymbol{C}$ with a full-rank diagonal matrix $\boldsymbol{\Psi}$. Then, (30) *has the same optimal value as a transformed problem*

$$\max_{(\bar{\boldsymbol{Q}}_k \succeq \mathbf{0})_{\forall k}} R_{BC}((\bar{\boldsymbol{Q}}_k)_{\forall k}); (\bar{\boldsymbol{H}}_k)_{\forall k}) \quad s.t. \quad \sum_{k=1}^{K} \bar{\boldsymbol{Q}}_k \preceq \boldsymbol{\Psi} \tag{40}$$

with channels $(\bar{\boldsymbol{H}}_k = \boldsymbol{H}_k \boldsymbol{W}_1)_{\forall k}$ and shaping matrix $\boldsymbol{\Psi}$. The covariance matrices

$$\boldsymbol{Q}_k = \boldsymbol{W}_1 \bar{\boldsymbol{Q}}_k \boldsymbol{W}_1^{\mathrm{H}} \quad \forall k \tag{41}$$

are optimizers of the original problem (30) *if and only if the matrices $(\bar{\boldsymbol{Q}}_k)_{\forall k}$ are optimizers of the transformed problem.*

Proof: The transformation cuts away the dimensions of $\boldsymbol{C}$ with eigenvalues equal to zero. In these forbidden dimensions, the transmit covariance matrices are forced to zero, as $\mathbf{0} \preceq \sum_k \boldsymbol{Q}_k \preceq \boldsymbol{C}$. Thus, we can perform the optimization only in the remaining dimensions, and we can then construct a solution of the original problem by filling up the forbidden dimensions with zero power. ■

This means that we can solve the optimization (40) by means of Proposition 1 and obtain a solution to (30) with rank-deficient $\boldsymbol{C}$ by using the transformation in (41). A similar result can be found in [20]. Therein, the authors do not further study the original problem with rank-deficient $\boldsymbol{C}$, but they only point out that the transformed problem (40), which has the same optimal value, can be solved instead.

However, this point of view does not help for our purpose since for the linear approximation (19), we need the zero-duality gap property and the numerical value of the optimal dual variable $\boldsymbol{\Omega}$. From just knowing the primal solution, neither of them can be obtained. At the first glance, we might think that the only missing element is a transformation that allows us to construct a dual solution of (30) from a dual solution of (40), but in fact, this is not true. It turns out that the problem is much more fundamental.

In the forbidden directions, the combination of inequalities $\boldsymbol{Q}_k \succeq \mathbf{0}\ \forall k$ and $\sum_k \boldsymbol{Q}_k \preceq \boldsymbol{C}$ corresponds to an equality constraint that forces the corresponding eigenvalues of the matrices $\boldsymbol{Q}_k$ to zero. This means that the constraint set does not have a non-empty interior. As a result, Slater's constraint qualifications are not fulfilled (see, e.g., [27, Ch. 5], or [28, Ch. 1] for the corresponding concept in semidefinite programming), and the KKT conditions might not even be necessary conditions for an optimum.

This conclusion can be supported by numerical evidence as follows. If the primal solution is known, it is usually possible to find the corresponding dual variables by a convex feasibility problem where the constraints are the KKT conditions with the primal solution plugged in. To avoid numerical problems, this feasibility test can instead be formulated as a distance minimization whose optimal value is zero apart from numerical inaccuracies. However, in our experiments, we came across many cases in which it was not possible to find a solution with a distance close to zero. This indicates that the optimum of the original problem can indeed be attained at a point which does not fulfill the KKT conditions.

As a result, there does not seem to be an obvious way to extend Theorem 1 to points where $\boldsymbol{C}$ is singular, and we leave this question open for future research. If such an extension could be found, the overall method proposed in this paper would become a globally optimal solution to the PDF rate maximization problem.

V. Proposed Algorithm

In this section, we give the details of the proposed algorithmic solutions for both the outer problem (25) and the inner problem (13).

A. Outer Problem

For finding the optimal $\boldsymbol{C}$ and $\boldsymbol{R}$ in (25), we use the cutting plane algorithm [22, Section 6.3.3]. The algorithm successively refines linear approximations of a concave function. To apply the algorithm, we reformulate (25) as follows:

$$\max_{(\boldsymbol{C},\boldsymbol{R})\in\mathcal{P}} z \quad \text{s.t.} \quad z \leq R_{\mathrm{A}}^{\star}(\boldsymbol{C}),\; z \leq R_{\mathrm{B}}(\boldsymbol{C},\boldsymbol{R}) \tag{42}$$

with

$$\mathcal{P} = \big\{(\boldsymbol{C} \succeq \varepsilon\mathbf{I}, \boldsymbol{R} \succeq \mathbf{0}) : \operatorname{tr}(\boldsymbol{C}) + \operatorname{tr}(\boldsymbol{D}_{\mathrm{S}}\boldsymbol{R}\boldsymbol{D}_{\mathrm{S}}^{\mathrm{H}}) \leq P_{\mathrm{S}} \\ \wedge \operatorname{tr}(\boldsymbol{D}_{\mathrm{R}}\boldsymbol{R}\boldsymbol{D}_{\mathrm{R}}^{\mathrm{H}}) \leq P_{\mathrm{R}}\big\}. \tag{43}$$

The optimal value of this problem can be bounded from above by replacing $R_{\mathrm{A}}^{\star}(\boldsymbol{C})$ and $R_{\mathrm{B}}(\boldsymbol{C},\boldsymbol{R})$ in the constraints of (42) by their linear approximations (19) and (21) around a finite number of points, which leads to

$$\max_{(\boldsymbol{C},\boldsymbol{R})\in\mathcal{P}} z \quad \text{s.t.} \quad \left.\begin{aligned} z &\leq \hat{R}_{\mathrm{A}}^{\star}(\boldsymbol{C};\widetilde{\boldsymbol{C}}) \\ z &\leq \hat{R}_{\mathrm{B}}(\boldsymbol{C},\boldsymbol{R};\widetilde{\boldsymbol{C}},\widetilde{\boldsymbol{R}}) \end{aligned}\right\} \forall(\widetilde{\boldsymbol{C}},\widetilde{\boldsymbol{R}}) \in \bar{\mathcal{P}} \tag{44}$$

where $\bar{\mathcal{P}} \subseteq \mathcal{P}$ is the set of points around which the linearizations are performed. The cutting plane algorithm is initialized with an initial set $\bar{\mathcal{P}}$ with a small number of elements, and (44) is solved. The optimal solution $(\boldsymbol{C}^{\star},\boldsymbol{R}^{\star})$ is then added to $\bar{\mathcal{P}}$ and the optimization problem (44) is solved again. This procedure is repeated until the desired accuracy is reached.

B. Inner Problem

To solve the subproblem of evaluating $R_{\mathrm{A}}^{\star}(\boldsymbol{C})$ from (13), we first perform the transformation from Proposition 1 and then solve the dual MAC problem (31) of the transformed problem. For this, we use an alternating gradient projection algorithm as proposed by [23]. This means, we perform gradient steps and projections onto the constraint set for the inner rate maximization and for the worst-case noise optimization in an alternating manner until convergence.

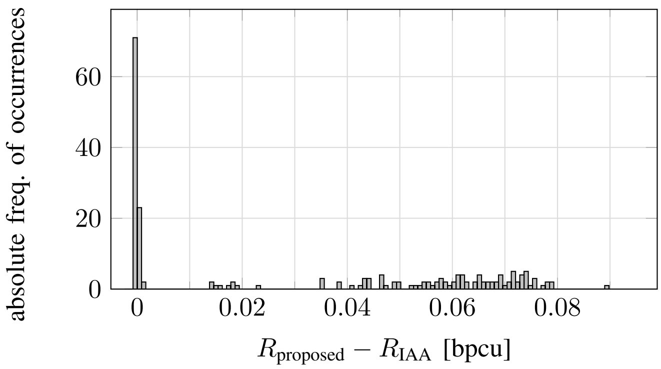

Fig. 1. Histogram of rate gain over IAA [15] for $N_{\mathrm{S}} = N_{\mathrm{R}} = N_{\mathrm{D}} = 2$, $P_{\mathrm{S}} = 100$, $P_{\mathrm{R}} = 10$, $d = 0.8$ and $\varepsilon = 10^{-5}P_{\mathrm{S}}$.

Furthermore, solving the dual MAC problem does not only lead to the optimal transmit covariance matrices, but also directly gives the Lagrangian multiplier for the shaping constraint $\boldsymbol{C}_{\boldsymbol{v}} + \boldsymbol{C}_{\boldsymbol{q}} \preceq \boldsymbol{C}$:

$$\boldsymbol{\Omega} = \mu\boldsymbol{C}_{\boldsymbol{\eta}}^{\star} \tag{45}$$

where $\boldsymbol{C}_{\boldsymbol{\eta}}^{\star}$ is the optimal worst-case noise from (31) and μ is the Lagrangian multiplier corresponding to the power constraint $\operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}}) \leq P$ [20]. The Lagrangian multiplier matrix $\boldsymbol{\Omega}$ is needed to compute the linear approximation of $R_{\mathrm{A}}^{\star}(\boldsymbol{C})$ as given in (19).

VI. Results and Conclusion

To evaluate the performance of the proposed algorithm, we compare the results to the inner approximation approach (IAA) from [15] and to the cut-set bound. As in [15], we assume a line network, where the relay lies on a line between source and destination. The distances source-relay, relay-destination, and source-destination are given by $d_{\mathrm{SR}} = d, d \in (0,1)$, $d_{\mathrm{RD}} = 1 - d$, and $d_{\mathrm{SD}} = 1$, respectively. The channel matrices are given by $\boldsymbol{H}_{\mathrm{AB}} = d_{\mathrm{AB}}^{-\gamma/2}\tilde{\boldsymbol{H}}_{\mathrm{AB}}$ with $\gamma = 4$ and $\mathrm{A},\mathrm{B} \in \{\mathrm{S},\mathrm{R},\mathrm{D}\}$. The individual elements of each $\tilde{\boldsymbol{H}}_{\mathrm{AB}}$ are independent and circularly symmetric complex Gaussian distributed with zero mean and unit variance.

The histogram in Figure 1 shows the difference $R_{\text{proposed}} - R_{\text{IAA}}$ for 200 i.i.d. channel realizations with two antennas at each terminal and distance parameter $d = 0.8$. It can be seen that the IAA and the proposed algorithm converge to the same value in many cases. However, there are also cases in which the proposed algorithm achieves a higher rate, meaning that the solution found by the IAA method is not the global optimum in these cases. Figure 2 shows the results for the same scenario with various values of d. By using the proposed method as a benchmark, we can conclude that the IAA has a close-to-optimal performance on average, which had not been clear in the first place since the IAA is only a local method. On the other hand, we can observe that the gap to the cut-set bound

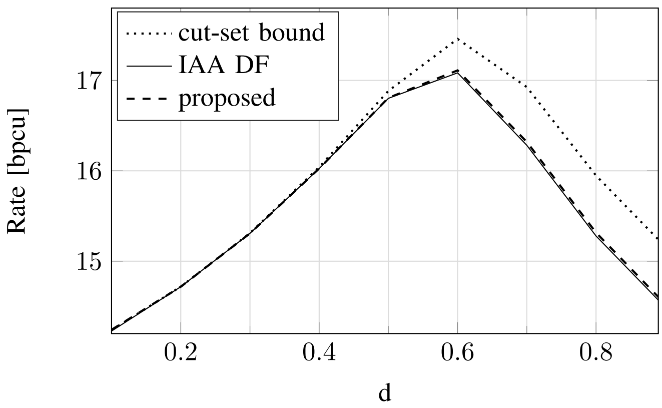

Fig. 2. Average rate compared to IAA [15] and cut-set bound for $N_\text{S} = N_\text{R} = N_\text{D} = 2$, $P_\text{S} = 100$, $P_\text{R} = 10$, and $\varepsilon = 10^{-5} P_\text{S}$.

(as seen in Figure 2 for $d \geq 0.5$) cannot be closed with the proposed algorithm.

Unlike existing suboptimal approaches to solve the PDF rate maximization problem, the proposed algorithm is not a local approach. Instead, we solve a slightly modified problem in a globally optimal manner. Thus, the proposed algorithm finds the globally optimal solution in cases where the optimal innovation covariance matrix has full rank, and we conjecture that it finds a close-to-optimal solution in the other cases if a sufficiently small ε is chosen. If this conjecture holds, the results reveal that the gap to the cut-set bound in Figure 2 is not due to the potentially suboptimal choice of the covariance matrices in existing algorithms, but inherent to the PDF scheme or inherent to the fact that the CSB might not be a tight bound to the capacity of the relay channel in general.

To settle this conjecture, it should be studied in future research whether Theorem 1 can be extended to the rank-deficient case in order to derive a method to find the globally optimal PDF rate without the approximation used in this paper.

Appendix A

Proof of Proposition 1: Since $\boldsymbol{C}$ and $\begin{bmatrix}\boldsymbol{V} & \boldsymbol{V}^\perp\end{bmatrix}$ are invertible by assumption, any feasible $\boldsymbol{Q}_k$ can be parametrized by

$$\boldsymbol{Q}_k = \boldsymbol{C}^{\frac{1}{2}} \begin{bmatrix}\boldsymbol{V} & \boldsymbol{V}^\perp\end{bmatrix} \begin{bmatrix}\boldsymbol{Q}'_k & \boldsymbol{X}_k \\ \boldsymbol{Y}_k & \boldsymbol{Z}_k\end{bmatrix} \begin{bmatrix}\boldsymbol{V} & \boldsymbol{V}^\perp\end{bmatrix}^\text{H} \boldsymbol{C}^{\frac{\text{H}}{2}}. \quad (46)$$

Note that $\boldsymbol{H}_k \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V}^\perp = \boldsymbol{0}\ \forall k$ by construction, and that only products of the form $\boldsymbol{H}_k \boldsymbol{Q}_i \boldsymbol{H}_k^\text{H}$ are relevant for the objective function R_BC given in (27). Therefore, we can restrict ourselves to the case where $\boldsymbol{X}_k = \boldsymbol{0}$, $\boldsymbol{Y}_k = \boldsymbol{0}$, and $\boldsymbol{Z}_k = \boldsymbol{0}$. This means that we can use the parameterization $\boldsymbol{Q}_k = \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V} \boldsymbol{Q}'_k \boldsymbol{V}^\text{H} \boldsymbol{C}^{\frac{\text{H}}{2}}$ without loss of generality.

Moreover, we have $\boldsymbol{U}_k \boldsymbol{U}_k^\text{H} \boldsymbol{H}_k = \boldsymbol{H}_k\ \forall k$ by construction. Using this equality in combination with $\det(\mathbf{I} + \boldsymbol{AB}) = \det(\mathbf{I} + \boldsymbol{BA})$ [29], we obtain

$$R_\text{BC} = \sum_{k=1}^{K} \log_2 \frac{\det\left(\mathbf{I} + \sum_{i=k}^{K} \boldsymbol{U}_k^\text{H} \boldsymbol{H}_k \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V} \boldsymbol{Q}'_i \boldsymbol{V}^\text{H} \boldsymbol{C}^{\frac{\text{H}}{2}} \boldsymbol{H}_k^\text{H} \boldsymbol{U}_k\right)}{\det\left(\mathbf{I} + \sum_{i=k+1}^{K} \boldsymbol{U}_k^\text{H} \boldsymbol{H}_k \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V} \boldsymbol{Q}'_i \boldsymbol{V}^\text{H} \boldsymbol{C}^{\frac{\text{H}}{2}} \boldsymbol{H}_k^\text{H} \boldsymbol{U}_k\right)}. \quad (47)$$

We can identify this as the sum rate $R_\text{BC}((\boldsymbol{Q}'_k)_{\forall k}; (\boldsymbol{H}'_k)_{\forall k})$ in a broadcast channel with channel matrices $\boldsymbol{H}'_k = \boldsymbol{U}_k^\text{H} \boldsymbol{H}_k \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V}$ and transmit covariance matrices $\boldsymbol{Q}'_k$.

In the following, we use $\boldsymbol{A} \succeq \boldsymbol{B} \Rightarrow \boldsymbol{SAS}^\text{H} \succeq \boldsymbol{SBS}^\text{H}$ for Hermitian $\boldsymbol{A}$ and $\boldsymbol{B}$ [29, Section 7.7] to reformulate the constraints of (30). Obviously, we have $\boldsymbol{Q}_k \succeq \boldsymbol{0}$ if $\boldsymbol{Q}'_k \succeq \boldsymbol{0}$. On the other hand, due to $\boldsymbol{V}^\text{H}\boldsymbol{V} = \mathbf{I}$, we can choose $\boldsymbol{S} = \boldsymbol{V}^\text{H} \boldsymbol{C}^{-\frac{1}{2}}$ to show that $\boldsymbol{Q}'_k = \boldsymbol{S} \boldsymbol{Q}_k \boldsymbol{S}^\text{H} \succeq \boldsymbol{0}$ if $\boldsymbol{Q}_k \succeq \boldsymbol{0}$.

If $\sum_k \boldsymbol{Q}'_k \preceq \mathbf{I}$, we have $\sum_k \boldsymbol{V} \boldsymbol{Q}'_k \boldsymbol{V}^\text{H} \preceq \boldsymbol{V}\boldsymbol{V}^\text{H} \preceq \mathbf{I}$ since each eigenvalue of a projector $\boldsymbol{V}\boldsymbol{V}^\text{H}$ is either 1 or 0 [29]. Thus, $\sum_k \boldsymbol{Q}_k = \sum_k \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V} \boldsymbol{Q}'_k \boldsymbol{V}^\text{H} \boldsymbol{C}^{\frac{\text{H}}{2}} \preceq \boldsymbol{C}$ if $\sum_k \boldsymbol{Q}'_k \preceq \mathbf{I}$. On the other hand, we can again use $\boldsymbol{S} = \boldsymbol{V}^\text{H} \boldsymbol{C}^{-\frac{1}{2}}$ to show that $\sum_k \boldsymbol{Q}'_k = \sum_k \boldsymbol{S} \boldsymbol{Q}_k \boldsymbol{S}^\text{H} \preceq \boldsymbol{SCS}^\text{H} = \mathbf{I}$ if $\sum_k \boldsymbol{Q}_k \preceq \boldsymbol{C}$.

We have shown that we can find an optimum of (30) by instead solving (34). It remains to be shown that $((\boldsymbol{Q}_k)_{\forall k}, (\boldsymbol{\Lambda}_k)_{\forall k}, \boldsymbol{\Omega})$ is a KKT point of (30) if $((\boldsymbol{Q}'_k)_{\forall k}, (\boldsymbol{\Lambda}'_k)_{\forall k}, \boldsymbol{\Omega}')$ is a KKT point of (34).[2]

Similar as above, it is easy to see that $\boldsymbol{\Lambda}'_k \succeq \boldsymbol{0} \Rightarrow \boldsymbol{\Lambda}_k \succeq \boldsymbol{0}$ and $\boldsymbol{\Omega}' \succeq \boldsymbol{0} \Rightarrow \boldsymbol{\Omega} \succeq \boldsymbol{0}$. Due to $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$, the complementary slackness condition $\operatorname{tr}(\boldsymbol{\Lambda}_k \boldsymbol{Q}_k) = 0\ \forall k$ is fulfilled if $\operatorname{tr}(\boldsymbol{\Lambda}'_k \boldsymbol{Q}'_k) = 0\ \forall k$ is fulfilled. Similarly, we find that $\operatorname{tr}\left(\boldsymbol{\Omega}' \left(\mathbf{I} - \sum_k \boldsymbol{Q}'_k\right)\right) = 0$ implies $\operatorname{tr}\left(\boldsymbol{\Omega} \left(\boldsymbol{C} - \sum_k \boldsymbol{Q}_k\right)\right) = 0$.

Since $R_\text{BC}((\boldsymbol{Q}'_k)_{\forall k}; (\boldsymbol{H}'_k)_{\forall k}) = R_\text{BC}((\boldsymbol{Q}'_k)_{\forall k}; (\boldsymbol{U}_k \boldsymbol{H}'_k)_{\forall k})$, we can base the following considerations on $\tilde{\boldsymbol{H}} = \boldsymbol{U}_k \boldsymbol{H}'_k = \boldsymbol{H}_k \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V}$ instead of on $\boldsymbol{H}'_k$ for the sake of simplicity. The derivative of the objective function is given by

$$\frac{\partial R_\text{BC}((\boldsymbol{Q}'_k)_{\forall k}; (\tilde{\boldsymbol{H}}_k)_{\forall k})}{\partial \boldsymbol{Q}_k'^\text{T}} = \frac{1}{\ln 2} \Bigg[\Big(\sum_{i=1}^{k} \tilde{\boldsymbol{H}}_i^\text{H} \big(\mathbf{I} + \tilde{\boldsymbol{H}}_i \sum_{j=i}^{K} \boldsymbol{Q}'_j \tilde{\boldsymbol{H}}_i^\text{H}\big)^{-1} \tilde{\boldsymbol{H}}_i \Big) - \Big(\sum_{i=1}^{k-1} \tilde{\boldsymbol{H}}_i^\text{H} \big(\mathbf{I} + \tilde{\boldsymbol{H}}_i \sum_{j=i+1}^{K} \boldsymbol{Q}'_j \tilde{\boldsymbol{H}}_i^\text{H}\big)^{-1} \tilde{\boldsymbol{H}}_i \Big) \Bigg]. \quad (48)$$

Since $\boldsymbol{H}_k \boldsymbol{Q}_i \boldsymbol{H}_k^\text{H} = \tilde{\boldsymbol{H}}_k \boldsymbol{Q}'_i \tilde{\boldsymbol{H}}_k^\text{H}$ for any k and i, we have that

$$\frac{\partial R_\text{BC}((\boldsymbol{Q}'_k)_{\forall k}; (\tilde{\boldsymbol{H}}_k)_{\forall k})}{\partial \boldsymbol{Q}_k'^\text{T}} = \boldsymbol{V}^\text{H} \boldsymbol{C}^{\frac{\text{H}}{2}} \frac{\partial R_\text{BC}((\boldsymbol{Q}_k)_{\forall k}; (\boldsymbol{H}_k)_{\forall k})}{\partial \boldsymbol{Q}_k^\text{T}} \boldsymbol{C}^{\frac{1}{2}} \boldsymbol{V}. \quad (49)$$

[2] Note that the problem has a non-empty interior, so that Slater's constraint qualifications are fulfilled (see, e.g., [27, Ch. 5], or [28, Ch. 1] for the corresponding concept in semidefinite programming). Thus, the KKT conditions are necessary for optimality.

By exploiting that $\boldsymbol{H}_k\boldsymbol{C}^{\frac{1}{2}}\boldsymbol{V}^{\perp} = \boldsymbol{0}\ \forall k$, i.e., $\boldsymbol{H}_k\boldsymbol{C}^{\frac{1}{2}}\boldsymbol{V}\boldsymbol{V}^{\mathrm{H}} = \boldsymbol{H}_k\boldsymbol{C}^{\frac{1}{2}}$, we get

$$\boldsymbol{C}^{-\frac{\mathrm{H}}{2}}\boldsymbol{V}\frac{\partial R_{\mathrm{BC}}((\boldsymbol{Q}'_k)_{\forall k}; (\tilde{\boldsymbol{H}}_k)_{\forall k})}{\partial \boldsymbol{Q}'^{\mathrm{T}}_k}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{C}^{-\frac{1}{2}} = \frac{\partial R_{\mathrm{BC}}((\boldsymbol{Q}_k)_{\forall k}; (\boldsymbol{H}_k)_{\forall k})}{\partial \boldsymbol{Q}^{\mathrm{T}}_k}. \quad (50)$$

Therefore, the stationarity condition

$$\frac{\partial R_{\mathrm{BC}}((\boldsymbol{Q}_k)_{\forall k}; (\boldsymbol{H}_k)_{\forall k})}{\partial \boldsymbol{Q}^{\mathrm{T}}_k} + \boldsymbol{\Lambda}_k - \boldsymbol{\Omega} = \boldsymbol{0}\ \forall k \quad (51)$$

of (30) is fulfilled by $\boldsymbol{\Lambda}_k$ and $\boldsymbol{\Omega}$ from (37) and (38) if the corresponding condition of the transformed problem (34) is fulfilled. ■

References

[1] E. C. van der Meulen, "Three-terminal communication channels," *Advances Appl. Probability*, vol. 3, no. 1, pp. 120–154, 1971.

[2] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[3] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[4] C. T. K. Ng and G. J. Foschini, "Transmit signal and bandwidth optimization in multiple-antenna relay channels," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 2987–2992, Jul. 2011.

[5] L. Gerdes and W. Utschick, "Optimized capacity bounds for the MIMO relay channel," in *Proc. Int. Conf. Acoust., Speech, and Signal Process. (ICASSP) 2011*, May 2011, pp. 3336–3339.

[6] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2012.

[7] G. Kramer, "Topics in multi-user information theory," *Foundations and Trends in Commun. and Inf. Theory*, vol. 4, no. 4–5, pp. 265–444, 2007.

[8] A. El Gamal and M. Aref, "The capacity of the semideterministic relay channel," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, pp. 536–536, May 1982.

[9] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[10] C. K. Lo, S. Vishwanath, and R. W. Heath, "Rate bounds for MIMO relay channels," *J. Commun. Netw.*, vol. 10, no. 2, pp. 194–203, Jun. 2008.

[11] L. Gerdes, C. Hellings, L. Weiland, and W. Utschick, "On the maximum achievable partial decode-and-forward rate for the Gaussian MIMO relay channel," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6751–6758, Dec. 2015.

[12] C. Hellings, L. Gerdes, L. Weiland, and W. Utschick, "On optimal Gaussian signaling in MIMO relay channels with partial decode-and-forward," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3153–3164, Jun. 2014.

[13] X. Jin and Y.-H. Kim, "Approximate capacity of the MIMO relay channel," in *2014 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2014, pp. 2102–2106.

[14] ——, "The approximate capacity of the MIMO relay channel," submitted to the *IEEE Trans. Inf. Theory*. Preprint available at http://arxiv.org/abs/1509.01931.

[15] L. Weiland, L. Gerdes, and W. Utschick, "Partial decode-and-forward rates for the Gaussian MIMO relay channel: Inner approximation of non-convex rate constraints," in *Proc. IEEE 14th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 540–544.

[16] L. Gerdes, L. Weiland, and W. Utschick, "A zero-forcing partial decode-and-forward scheme for the Gaussian MIMO relay channel," in *Proceedings of 2013 IEEE International Conference on Communications (ICC)*, Jun. 2013, pp. 3349–3354.

[17] ——, "Optimal partial decode-and-forward rates for the Gaussian MIMO relay channel using the GSVD," in *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2014, pp. 259–263.

[18] 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, Oct. 2003.

[19] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[20] A. Dotzler, M. Riemensberger, and W. Utschick, "Minimax duality for MIMO interference networks," 2014, to be published. Preprint available at http://www.msv.ei.tum.de/ando/MiniMaxDuality.pdf.

[21] A. Dotzler, M. Riemensberger, W. Utschick, and G. Dietl, "Interference robustness for cellular MIMO networks," in *Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2012, pp. 229–233.

[22] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Scientific, 2004.

[23] H. H. Brunner, A. Dotzler, W. Utschick, and J. A. Nossek, "Weighted sum rate maximization with multiple linear conic constraints," in *Proc. 2015 IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 6247–6252.

[24] S. Boyd, L. Xiao, A. Mutapcic, and J. Mattingley, *Notes on Decomposition Methods*. Stanford University, 2008, Lecture Notes.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*, 7th ed. Cambridge University Press, 2009.

[26] C. Hellings and W. Utschick, "Energy efficiency optimization in MIMO broadcast channels with fairness constraints," in *Proc. IEEE 14th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 599–603.

[27] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. John Wiley & Sons, 2006.

[28] A. Ben-Tal and A. Nemirovski, "Lectures on modern convex optimization," Israel Institute of Technology and Georgia Institute of Technology, Lecture Notes, Fall Semester 2013.

[29] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge University Press, 2012.