Average Sum MSE Minimization in the Multi-User Downlink With Multiple Power Constraints

*Andreas Gründinger, *Michael Joham, †José Pablo González-Coma, †Luis Castedo, and *Wolfgang Utschick,
*Associate Institute for Signal Processing, Technische Universität München, Germany
†Department of Electronics and System, University of A Coruña, Spain
Email: *{gruendinger,joham,utschick}@tum.de and †{jose.gcoma,luis}@udc.es

***Abstract*—We consider a *sum mean square error* (SMSE) based transceiver design for the multi-user downlink with multiple linear transmit power constraints. Since the multi-antenna transmitter has only imperfect *channel state information* (CSI), the average SMSE is minimized via an *alternating optimization* (AO). For fixed equalizers, the average SMSE minimizing precoders are found via an uplink-downlink SMSE duality. The precoder update therewith transforms to an uplink max-min average SMSE problem, i.e., a *minimum mean square error* (MMSE) equalizer design and an outer worst-case noise search. This dual problem is optimally solved, e.g., by a gradient projection approach, and strong duality is shown.**

***Index Terms*—uplink-downlink sum MSE duality; imperfect CSI; general power constraints; per-antenna constraints**

I. Introduction

Uplink-downlink duality is an utmost useful tool to transform precoder optimizations in multi-antenna *broadcast channels* (BCs) to less complex filter designs and power allocations in the dual *multiple access channels* (MACs). We focus on an average SMSE transceiver design via AO in this context, where the downlink precoder update becomes a simple MMSE filter calculation in the dual uplink. This result was revealed in [1] using SINR uplink-downlink duality (e.g., see [2]). Simplified versions of this duality and AO approach were shown in [3] and depend on the required *mean square error* (MSE) term, i.e., SMSE, per-user MSE, or per-stream MSE.

Note that these references impose a sum power constraint and perfect *transmitter CSI* (TxCSI). In contrast, we explore an AO transceiver design for imperfect TxCSI and extend the work in [4] to multiple linear power constraints, including per-beam and per-antenna constraints as special cases. This formulation corresponds to the scenario, where the antennas cannot negotiate on their payload, e.g., either due to their distributed locations or due to the radio-frequency hardware.

A similar problem was considered by Bogale and Vandendorpe in [5], [6]. They introduced an uplink-downlink duality for the precoder design step within the AO based on an algorithmic argumentation and for special types of power constraints, e.g., per-antenna and per-precoder constraints. In contrast, we derive the uplink-downlink duality for the average SMSE precoder optimization problem concisely via Lagrangian duality theory and for generalized power constraints, which include above constraints as special cases. Due to strong duality, the computationally attractive uplink filter and power allocation solutions lead to an optimal precoder update for the locally optimal alternating convex search.

This work shall give the main idea of the SMSE precoder optimization via (Lagrangian) uplink-downlink duality. We show that the duality holds for single-antenna and multi-antenna receiver scenarios. Furthermore, we show that the average SMSE measure varies only a little for Gaussian channels if a per-antenna, per-array, or a sum power constraint are imposed in the optimization, even though the dynamic power range of each antenna is increasing in this order.

We remark that similar uplink-downlink duality based transceiver designs are applicable for other optimization problems as well, e.g., an average MSE balancing which we recently presented in [7]. Moreover, even though the shown AO considers only imperfect TxCSI, i.e., the receivers exploit perfect CSI to calculate their equalizers, the method is applicable when the receivers have also imperfect CSI (e.g., see [6]) and the transmitter knows the receivers' strategies.

II. System Model

First, we consider a K-user vector downlink with single-antenna receivers and an N-antenna transmitter. The k-th user's MSE in this system reads as

$$\mathrm{MSE}_k = 1 - 2\,\mathrm{Re}\{f_k^* \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{b}_k\} + \sum_{i=1}^{K} |f_k|^2 |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{b}_i|^2 + |f_k|^2 \sigma_k^2 \quad (1)$$

when mutually independent zero-mean unit-variance data signals are linearly precoded with $\boldsymbol{b}_i$, $i = 1, \ldots, K$, and sent over the channel $\boldsymbol{h}_k \in \mathbb{C}^N$. The receivers are subject to zero-mean additive noise with variance σ_k^2 and filter the obtained signal with $f_k = v\tilde{f}_k \in \mathbb{C}$.[1]

As mentioned above, we consider imperfect TxCSI. The transmitter shall only be aware of some channel statistics. For example, we may assume the error model

$$\boldsymbol{h}_k = \bar{\boldsymbol{h}}_k + \tilde{\boldsymbol{h}}_k \quad (2)$$

with the estimated channel mean $\bar{\boldsymbol{h}}_k$ and the Gaussian error $\tilde{\boldsymbol{h}}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{C}_k)$ with known covariance matrix $\boldsymbol{C}_k$.

***Acknowledgment*—This joint work has been funded by DFG under grant Jo 724/1-2 and by Xunta de Galicia, MINECO of Spain, and FEDER funds of the EU under grants 2012/287 and TEC2013-47141-C4-1-R.

[1] Here, v is a normalization variable for the precoder design in the AO.

A. Multi-Antenna Receivers

In the second scenario, the receivers have multiple antennas and support a multi-stream data transmission, such that the k-th user's MSE reads as

$$\begin{aligned}\mathrm{MSE}_k = M_k - 2\,\mathrm{Re}\{\mathrm{tr}(\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{B}_k)\} \\ + \sum_{i=1}^{K}\|\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{B}_i\|_{\mathrm{F}}^2 + \mathrm{tr}(\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{\Sigma}_k\boldsymbol{F}_k).\end{aligned} \tag{3}$$

To obtain (3), we assumed that M_k independent zero-mean unit variance data signals are linearly precoded by $\boldsymbol{B}_k$ and transmitted to the k-th receiver, which filters the signals with $\boldsymbol{F}_k = v\tilde{\boldsymbol{F}}_k$.[1] If we additionally substitute $\boldsymbol{b}_k = \mathrm{vec}(\boldsymbol{B}_k)$ and $\boldsymbol{f}_k = \mathrm{vec}(\boldsymbol{F}_k)$ in (3), we obtain the equivalent beamformer-like MSE representation [cf. (1)]

$$\begin{aligned}\mathrm{MSE}_k = M_k - 2\,\mathrm{Re}\{\boldsymbol{f}_k^{\mathrm{H}}(\mathbf{I}_{M_k}\otimes\boldsymbol{H}_k^{\mathrm{H}})\boldsymbol{b}_k\} \\ + \sum_{i=1}^{K}\|(\mathbf{I}_{M_i}\otimes\boldsymbol{F}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}})\boldsymbol{b}_i\|_2^2 + \boldsymbol{f}_k^{\mathrm{H}}(\mathbf{I}_{M_k}\otimes\boldsymbol{\Sigma}_k)\boldsymbol{f}_k.\end{aligned} \tag{4}$$

Similar to the vector channel case, the transmitter shall only be aware of the channel statistics for the multi-antenna receiver scenario. The matrix equivalent of the channel model in (2) is

$$\boldsymbol{H}_k = \bar{\boldsymbol{H}}_k + \tilde{\boldsymbol{H}}_k \tag{5}$$

with the channel mean $\bar{\boldsymbol{H}}_k$ and the Gaussian error $\tilde{\boldsymbol{h}}_k = \mathrm{vec}(\tilde{\boldsymbol{H}}_k) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_k)$ with covariance matrix $\boldsymbol{C}_k$.

B. Transmit Power Limitations

In either of the two scenarios, the transmitter shall be subject to general power limitations of the form

$$\sum_{i=1}^{K}\boldsymbol{b}_i^{\mathrm{H}}\boldsymbol{A}_{i,\ell}\boldsymbol{b}_i = \sum_{i=1}^{K}\|\boldsymbol{A}_{i,\ell}^{1/2}\boldsymbol{b}_i\|_2^2 \le P_\ell, \quad \ell = 1,\ldots,L \tag{6}$$

where $\boldsymbol{A}_{i,\ell} \in \mathbb{C}^{N\times N}$, $\boldsymbol{A}_{i,\ell} \succeq \mathbf{0}$ has the square root representation $\boldsymbol{A}_{i,\ell} = \boldsymbol{A}_{i,\ell}^{\mathrm{H}/2}\boldsymbol{A}_{i,\ell}^{1/2}$ and $\mathrm{rank}\{\sum_{\ell=1}^{L}\boldsymbol{A}_{i,\ell}\} = N$. Important examples from wireless communications in vector channels are a sum power constraint, per-beam constraints, and per-antenna constraints. Depending on the imposed transmit power constraint(s), the matrices $\boldsymbol{A}_{i,\ell}$ have different forms:

- *sum power:* $\boldsymbol{A}_{i,\ell} = \mathbf{I}_N$ for all $i = 1,\ldots,K$ and $L = 1$;
- *per-beam:* $\boldsymbol{A}_{i,i} = \mathbf{I}_N$ and $\boldsymbol{A}_{i,\ell} = \mathbf{0}_{N\times N}$, $\ell \neq i$ with $L = K$;
- *per-antenna:* $\boldsymbol{A}_{i,\ell} = \boldsymbol{e}_\ell\boldsymbol{e}_\ell^{\mathrm{T}}$ with $L = N$.

Besides these special cases, the formulation in (6) also includes power constraints per antenna array, e.g.,

- *per-array:* $\boldsymbol{A}_{i,\ell} = \mathrm{blockdiag}\{\mathbf{0}, \boldsymbol{A}'_\ell, \mathbf{0}\}$ and $L < N$,

where the matrices $\boldsymbol{A}'_\ell \succeq \mathbf{0}$ may have full rank, that is, $1 \le \mathrm{rank}\{\boldsymbol{A}_{i,\ell}\} \le N$. These per-array constraints are important for systems where several multi-antenna transmitters, e.g., WiFi stations, cooperatively supply several mobile devices.

We remark that the power constraints in (6) are also suitable for the scenario with multiple streams per user. Then, the matrices $\boldsymbol{A}_{i,\ell}$ for the sum, per-beam, per-antenna, and per-array constraints are blockdiagonal with replications of above matrices on the diagonal, i.e., $\boldsymbol{A}_{i,\ell} = \mathbf{I}_{M_k}\otimes\boldsymbol{A}'_{i,\ell}$.

III. Average Sum MSE Minimization

The goal is the minimization of the average downlink SMSE $\overline{\mathrm{SMSE}}_{\mathrm{DL}} = \sum_{k=1}^{K}\mathrm{E}[\mathrm{MSE}_k]$ for perfect RxCSI, imperfect TxCSI, and subject to above transmit power constraints, i.e.,

$$\min_{\boldsymbol{b},\boldsymbol{f}}\overline{\mathrm{SMSE}}_{\mathrm{DL}} \ \text{s.t.:} \sum_{i=1}^{K}\boldsymbol{b}_i^{\mathrm{H}}\boldsymbol{A}_{i,\ell}\boldsymbol{b}_i \le P_\ell, \quad \ell = 1,\ldots,L \tag{7}$$

where we substitute $\boldsymbol{b} = [\boldsymbol{b}_1^{\mathrm{T}},\ldots,\boldsymbol{b}_K^{\mathrm{T}}]^{\mathrm{T}}$ and write $\boldsymbol{f} = [f_1,\ldots,f_k]^{\mathrm{T}}$ or $\boldsymbol{f} = [\boldsymbol{f}_1^{\mathrm{T}},\ldots,\boldsymbol{f}_k^{\mathrm{T}}]^{\mathrm{T}}$ for the single- or multi-antenna receiver scenario, respectively, to simplify expositions.

If the receivers have instantaneous CSI, they can employ MMSE filters. These filters are

$$f_{k,\mathrm{MMSE}} = v\tilde{f}_{k,\mathrm{MMSE}} = \frac{\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{b}_k}{\sum_{i=1}^{K}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{b}_i|^2 + \sigma_k^2} \tag{8}$$

for the case of vector channels resulting in the average SMSE

$$\overline{\mathrm{SMSE}}_{\mathrm{DL}} = K - \sum_{k=1}^{K}\mathrm{E}\left[\frac{|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{b}_k|^2}{\sum_{i=1}^{K}|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{b}_i|^2 + \sigma_k^2}\right]. \tag{9}$$

Unfortunately, this objective is non-convex in the precoders and contains an expectation whose closed-form expression is only known for very simple channel models, e.g., zero-mean Gaussian channels [8], and whose evaluation requires either numerical integration or Monte-Carlo methods for other standard models. The same properties hold for the matrix channels, where the MMSE equalizers read as

$$\boldsymbol{F}_{k,\mathrm{MMSE}} = v\tilde{\boldsymbol{F}}_{k,\mathrm{MMSE}} = \boldsymbol{Y}_k^{-1}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{B}_k \tag{10}$$

with $\boldsymbol{Y}_k = \sum_{i=1}^{K}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{B}_i\boldsymbol{B}_i^{\mathrm{H}}\boldsymbol{H}_k + \boldsymbol{\Sigma}_k$, and result in the average SMSE expression

$$\overline{\mathrm{SMSE}}_{\mathrm{DL}} = \sum_{k=1}^{K}M_k - \mathrm{E}\left[\mathrm{tr}\left(\boldsymbol{B}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{Y}_k^{-1}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{B}_k\right)\right]. \tag{11}$$

Since $\overline{\mathrm{SMSE}}_{\mathrm{DL}}$ in (11) is non-convex in $\boldsymbol{B}_k$, minimizing the average SMSE directly w.r.t. the precoders is difficult if the closed form MMSE equalizers are incorporated into (7).

However, as the original MSE expressions in (1) and (3) are convex in the precoders for fixed equalizers and vice versa, i.e., the MSEs and also the SMSEs are biconvex functions [9] in $\boldsymbol{b}$ and $\boldsymbol{f}$, an alternating convex search of the precoders and filters converges to locally optimal transceivers in the multi-user downlink (cf. [10]). The following two steps would be performed in each iteration of the alternating convex search for $\boldsymbol{f}$ and $\boldsymbol{b}$ if perfect CSI were available (cf. [1]):

1) The equalizers in $\tilde{\boldsymbol{f}}$ (from $\boldsymbol{f} = v\tilde{\boldsymbol{f}}$) are first found in the downlink for fixed precoders in $\boldsymbol{b}$ as in (8) or (10).
2) Second, the downlink precoders in $\boldsymbol{b}$ are optimized as equalizers in the dual uplink system for given $\tilde{\boldsymbol{f}}$.

Note that the employed uplink-downlink duality in Step 2 of the AO can reduce the complexity for implementation and computations. Otherwise, the precoder update requires a convex solver for including the generalized constraints.

Also average SMSE minimizations can be based on uplink-downlink duality [4]. Then, the resulting precoders in $\boldsymbol{b}$ minimize an upper bound for the minimum achievable average

SMSE for fixed $\tilde{\boldsymbol{f}}$, which is reduced in each AO step and hence converges (cf. [9, Theorem 4.5]). For a further extension to the generalized power constraints in (6), we derive the uplink formulations of the precoder design optimization in Step 2 of the AO concisely via Lagrangian duality in what follows.

IV. Precoder Design via Uplink-Downlink Duality

If the downlink equalizers $\boldsymbol{f} = v\tilde{\boldsymbol{f}}$ are fixed up to v, the downlink precoder optimization step within the AO can be written for both antenna scenarios in the generalized form

$$\min_{\boldsymbol{b},v} \overline{\mathrm{SMSE}}_{\mathrm{DL}} \ \mathrm{s.t.}\!: \boldsymbol{b}^{\mathrm{H}}\boldsymbol{A}_\ell\boldsymbol{b} \leq P_\ell,\ \ell = 1,\dots,L \tag{12}$$

where the generalized average SMSE reads as

$$\overline{\mathrm{SMSE}}_{\mathrm{DL}} = M - 2\,\mathrm{Re}\{v^*\bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{b}\} + |v|^2\boldsymbol{b}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{b} + |v|^2\sigma^2 \tag{13}$$

and the generalized power constraints are equivalent to those in (7) if $\boldsymbol{A}_\ell = \mathrm{blockdiag}\{\boldsymbol{A}_{1,\ell},\dots,\boldsymbol{A}_{K,\ell}\}$. To obtain (13), we introduced the effective channel mean

$$\bar{\boldsymbol{h}} = \mathrm{E}[\hat{\boldsymbol{H}}\tilde{\boldsymbol{f}}], \tag{14}$$

the effective noise variance

$$\sigma^2 = \mathrm{E}[\tilde{\boldsymbol{f}}^{\mathrm{H}}\boldsymbol{\Sigma}\tilde{\boldsymbol{f}}], \tag{15}$$

and the second-order matrix

$$\boldsymbol{R} = \mathbf{I}_M \otimes \mathrm{E}[\boldsymbol{H}\tilde{\boldsymbol{F}}\tilde{\boldsymbol{F}}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}], \tag{16}$$

which shall have full-rank.[2] The expressions for the substitutes M, $\hat{\boldsymbol{H}}$, $\boldsymbol{H}$, $\tilde{\boldsymbol{F}}$, and $\boldsymbol{\Sigma}$ depend on the considered scenario:

- for single-antenna receivers, we have $M = K$,

$$\begin{aligned}\hat{\boldsymbol{H}} &= \mathrm{blockdiag}\{\boldsymbol{h}_1,\dots,\boldsymbol{h}_K\},\\ \boldsymbol{H} &= [\boldsymbol{h}_1,\dots,\boldsymbol{h}_K],\\ \tilde{\boldsymbol{F}} &= \mathrm{diag}\{\tilde{f}_1,\dots,\tilde{f}_K\},\\ \boldsymbol{\Sigma} &= \mathrm{diag}\{\sigma_1^2,\dots,\sigma_K^2\}\end{aligned} \tag{17}$$

- and for multi-antenna receivers, we have $M = \sum_{k=1}^K M_i$,

$$\begin{aligned}\hat{\boldsymbol{H}} &= \mathrm{blockdiag}\{\mathbf{I}_{M_1}\otimes\boldsymbol{H}_1,\dots,\mathbf{I}_{M_K}\otimes\boldsymbol{H}_K\},\\ \boldsymbol{H} &= [\boldsymbol{H}_1,\dots,\boldsymbol{H}_K],\\ \tilde{\boldsymbol{F}} &= \mathrm{blockdiag}\{\tilde{\boldsymbol{F}}_1,\dots,\tilde{\boldsymbol{F}}_K\},\\ \boldsymbol{\Sigma} &= \mathrm{blockdiag}\{\mathbf{I}_{M_1}\otimes\boldsymbol{\Sigma}_1,\dots,\mathbf{I}_{M_K}\otimes\boldsymbol{\Sigma}_K\}.\end{aligned} \tag{18}$$

Note that the expectations to compute $\bar{\boldsymbol{h}}$, $\boldsymbol{R}$, and σ^2 require numerical evaluations for the considered imperfect TxCSI, i.e., with known channel mean and error statistics. We computed the expected values in (14)-(16) via the corresponding sample means for the simulations.

Before deriving the dual uplink problem of (12), we first present a primal solution based on consecutive power minimizations. The power minimization problem is required for proving strong (uplink-downlink) duality in Section IV-B.

[2]If $\boldsymbol{R}$ is degenerate, we can find an equivalent representation of (13) and the power constraints in (12) via replacing $\boldsymbol{b}$ by $\tilde{\boldsymbol{b}} = \boldsymbol{V}\boldsymbol{b}$, where $\boldsymbol{V}$ is an orthonormal basis for the span of $\boldsymbol{R}$. The SMSE minimization (13) needs then to be performed w.r.t. $\tilde{\boldsymbol{b}}$ instead of $\boldsymbol{b}$.

A. Algorithmic Solution of the Primal Problem

If the SMSE minimizing $v = \boldsymbol{b}^{\mathrm{H}}\bar{\boldsymbol{h}}/(\boldsymbol{b}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{b} + \sigma^2)$ is substituted in (13), the downlink objective, which reads as

$$\overline{\mathrm{SMSE}}_{\mathrm{DL}} = M - \frac{|\bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{b}|^2}{\boldsymbol{b}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{b} + \sigma^2}, \tag{19}$$

becomes quasiconvex in $\boldsymbol{b}$. The lower level set

$$\mathcal{S} = \{\boldsymbol{b} \in \mathbb{C}^{MN} | \varepsilon \geq \overline{\mathrm{SMSE}}_{\mathrm{DL}}\}$$

is convex, i.e., $\varepsilon \geq \overline{\mathrm{SMSE}}_{\mathrm{DL}}$ has the convex *second order cone* (SOC) representation $\sqrt{M-\varepsilon}\|[\boldsymbol{b}^{\mathrm{H}}\boldsymbol{R}^{\mathrm{H}/2},\sigma]^{\mathrm{H}}\|_2 \leq \bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{b}$, where w.l.o.g. we restrict $\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{b}_k$ to be real and positive. Since, moreover, the power constraints have the SOC form $\|\boldsymbol{A}_{i,\ell}^{1/2}\boldsymbol{b}\|_2 \leq \sqrt{P_\ell}$, (12) can equivalently be rewritten as

$$\begin{aligned}\min_{\boldsymbol{b},\varepsilon}\ \varepsilon \quad \mathrm{s.t.}\!: &\|\boldsymbol{A}_\ell^{1/2}\boldsymbol{b}\|_2 \leq \sqrt{P_\ell}, \quad \ell = 1,\dots,L,\\ &\|[\boldsymbol{b}^{\mathrm{H}}\boldsymbol{R}^{\mathrm{H}/2},\sigma]^{\mathrm{H}}\|_2 \leq \tfrac{\bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{b}}{\sqrt{M-\varepsilon}}.\end{aligned} \tag{20}$$

This problem may be solved in the downlink via a bisection over $\varepsilon \in [0, M]$, where it is tested in each iteration whether there exists a feasible $\boldsymbol{b}$ that achieves the target ε.

We test feasibility using the strictly monotonically decreasing function $\psi_{\mathrm{DL}}(\varepsilon)$, which we define as the power factor that minimizes the convex SOC power balancing problem

$$\begin{aligned}\min_{\boldsymbol{b},\alpha}\ \alpha^2 \quad \mathrm{s.t.}\!: &\|\boldsymbol{A}_\ell^{1/2}\boldsymbol{b}\|_2 \leq \alpha\sqrt{P_\ell}, \quad \ell = 1,\dots,L,\\ &\|[\boldsymbol{b}^{\mathrm{H}}\boldsymbol{R}^{\mathrm{H}/2},\sigma]^{\mathrm{H}}\|_2 \leq \tfrac{\bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{b}}{\sqrt{M-\varepsilon}}.\end{aligned} \tag{21}$$

Note that this function is the inverse function to $\varepsilon_{\mathrm{DL}}(\psi)$ that denotes the minimum of (20) if we substitute P_ℓ with ψP_ℓ, that is, $\varepsilon_{\mathrm{DL}}(\psi_{\mathrm{DL}}(\varepsilon)) = \varepsilon$. Therefore, if $\psi_{\mathrm{DL}}(\varepsilon) < 1$, the corresponding precoder $\boldsymbol{b}$ can attain the target ε without violating the power constraints in (20), and if we finally achieve $\psi_{\mathrm{DL}}(\varepsilon) = 1$, the corresponding ε and $\boldsymbol{b}$ are the solutions of the SMSE minimization in (20).

B. Dual Uplink SMSE Problem

Based on Lagrangian duality, we next show that the downlink precoder design problem in (12) can be transformed into a dual uplink max-min average SMSE optimization.

Theorem 1. *The strongly dual uplink SMSE optimization of the downlink SMSE minimization in* (12) *reads as*

$$\max_{\boldsymbol{\mu}\geq\mathbf{0}}\ \min_{\boldsymbol{u},\lambda\geq 0}\ \overline{\mathrm{SMSE}}_{\mathrm{UL}} \ \mathrm{s.t.}\!: \sigma^2\lambda \leq 1,\ \sum_{\ell=1}^L \mu_\ell P_\ell \leq 1 \tag{22}$$

where the uplink SMSE is given by

$$\overline{\mathrm{SMSE}}_{\mathrm{UL}} = M - 2\sqrt{\lambda}\,\mathrm{Re}\{\bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{u}\} + \boldsymbol{u}^{\mathrm{H}}\Big(\lambda\boldsymbol{R} + \sum_{\ell=1}^L \mu_\ell\boldsymbol{A}_\ell\Big)\boldsymbol{u},$$

i.e., an inner equalizer design and power allocation and an outer worst-case noise covariance search w.r.t. the dual variables $\boldsymbol{\mu} = [\mu_1,\dots,\mu_L]^{\mathrm{T}}$ *corresponding to the downlink power constraints in* (12).

The strong duality proof between the SMSE minimizations in (12) and (22) is based on the downlink power balancing problem in (21) and its dual uplink SMSE formulation.

Lemma 1. *The strongly dual uplink max-min problem to the downlink power balancing in* (21) *can be written as*

$$\max_{\boldsymbol{\mu}\geq\mathbf{0}} \min_{\boldsymbol{u},\lambda\geq 0} \lambda\sigma^2 \ \text{s.t.:} \sum_{\ell=1}^{L} \mu_\ell P_\ell \leq 1,\ \varepsilon \geq \overline{\text{SMSE}}_{\text{UL}}. \tag{23}$$

In the appendix, this dual uplink problem is derived via similar steps as in [11], which is based on SINR constraints.

Proof of Theorem 1. We show that the optimal value of the primal SMSE minimization in (12) is also the optimum of the dual problem (22). Let ε denote the primal optimal average SMSE value. Then, we know that the optimum value of (21) is one [cf. Subsection IV-A], which is also the optimum of (23) due to zero duality gap (cf. Lemma 1), i.e., $\psi_{\text{DL}}(\varepsilon) = \psi_{\text{UL}}(\varepsilon) = 1$. Finally, it remains to prove that $\psi_{\text{UL}}(\varepsilon) = 1$ is an identifier to declare ε as the solution for (22).

If $\psi_{\text{UL}}(\varepsilon) = 1$, the corresponding power allocation

$$\lambda^\star = \frac{1}{\sigma^2} \tag{24}$$

and the corresponding variables $\boldsymbol{\mu}^\star$ satisfy the SMSE constraint $\varepsilon \leq \overline{\text{SMSE}}_{\text{UL}}$ in (23) with equality [cf. Appendix A]. Next, we prove that these parameters are also optimal for (22) in order to conclude that the optimum of (22) is ε.

To show that (24) is optimal for (22), we substitute the SMSE minimizing uplink equalizer

$$\boldsymbol{u}_{\text{SMSE}} = \sqrt{\lambda}\Big(\lambda\boldsymbol{R} + \sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_\ell\Big)^{-1}\bar{\boldsymbol{h}}, \tag{25}$$

which is an optimizer for both problems (23) and (22), into the uplink SMSE. The resulting uplink SMSE reads as [cf. (38)]

$$\min_{\boldsymbol{u}} \overline{\text{SMSE}}_{\text{UL}} = M - \lambda\bar{\boldsymbol{h}}^{\text{H}}\Big(\lambda\boldsymbol{R} + \sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_\ell\Big)^{-1}\bar{\boldsymbol{h}}, \tag{26}$$

which is decreasing in $\lambda \geq 0$. Therefore, the SMSE is minimized by maximizing λ, which is upper bounded by $\sigma^2\lambda \leq 1$ in (22). This shows that (24) is also optimal for (22).

Now, we prove that the optimizer $\boldsymbol{\mu}^\star$ for (23), which exactly meets ε, is also the SMSE maximizer in (22). To prove achievability, we remark that $\boldsymbol{\mu}^\star$ is feasible for (22), i.e., it satisfies the constraint $\sum_{\ell=1}^{L}\mu_\ell^\star P_\ell \leq 1$. Therefore, the optimum ε_{UL} of (22) satisfies $\varepsilon_{\text{UL}} \geq \varepsilon$. For the converse, we assume that there is another SMSE maximizing $\boldsymbol{\mu}'$ for (22), which satisfies the constraint set and achieves an SMSE $\varepsilon_{\text{UL}} > \varepsilon$ for the power allocation in (24) and the equalizer in (25). However, since this $\boldsymbol{\mu}'$ would also be feasible for (23), it contradicts the optimality of $\lambda^\star$ for (23), which was assumed to achieve the SMSE target ε. Therefore, we have $\varepsilon_{\text{UL}} \leq \varepsilon$ for all feasible $\boldsymbol{\mu}$ in (22). Together with the achievability result, this converse proof indicates that $\boldsymbol{\mu}^\star$ is indeed an SMSE maximizer for (22) if and only if it is an optimizer of (23) and that the max-min SMSE value is ε for (22) if $\psi_{\text{UL}}(\varepsilon) = 1$. □

If we know the optimal $\boldsymbol{u}$ and λ, the downlink variables are obtained via a scaling of these variables, i.e., (cf. [3])

$$\boldsymbol{b} = \sqrt{\beta}\boldsymbol{u}, \quad v = \sqrt{\lambda}/\sqrt{\beta}. \tag{27}$$

The squared transformation factor reads as

$$\beta = \frac{\lambda\sigma^2}{\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{u}^{\text{H}}\boldsymbol{A}_\ell\boldsymbol{u}} \tag{28}$$

and follows from inserting (27) into (13) and equating the SMSE terms, i.e., $\overline{\text{SMSE}}_{\text{UL}} = \overline{\text{SMSE}}_{\text{DL}}$. How to obtain $\boldsymbol{u}$ and λ together with the dual variable $\boldsymbol{\mu}$ is detailed next.

C. Algorithmic Solution of the Dual Problem

In contrast to the complex $NM \times 1$ complex-valued beamformer search presented in Section IV-A, which is suited for the primal SMSE minimization problem (12), we next present a *semidefinite programming* (SDP) solution and a *gradient projection* (GP) approach for the dual problem (22) that search only for the L non-negative reals in $\boldsymbol{\mu}$.

The solutions for the power allocation and the SMSE minimizing equalizer for the inner minimization in (22) are given in (24) and (25), respectively. Inserting these expressions, the outer maximization in (22) transforms to the convex problem

$$\max_{\boldsymbol{\mu}\geq\mathbf{0}} M - \bar{\boldsymbol{h}}^{\text{H}}\Big(\boldsymbol{R} + \sigma^2\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_\ell\Big)^{-1}\bar{\boldsymbol{h}} \ \ \text{s.t.:} \sum_{\ell=1}^{L}\mu_\ell P_\ell \leq 1. \tag{29}$$

Note that the constraint in (29) is tight at the optimal point since the objective is elementwise increasing in $\boldsymbol{\mu}$.

We may solve the worst-case noise search in (29) by means of a standard interior-point solver (e.g., CVX [12] with [13]). Using Schur's complement, (29) can be rewritten into the SDP problem formulation

$$\begin{aligned} \max_{\varepsilon,\boldsymbol{\mu}\geq\mathbf{0}} \ \varepsilon \quad \text{s.t.:} & \begin{bmatrix} M-\varepsilon & \bar{\boldsymbol{h}}^{\text{H}} \\ \bar{\boldsymbol{h}} & \boldsymbol{R}+\sigma^2\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_\ell \end{bmatrix} \succeq \mathbf{0}, \\ & \sum_{\ell=1}^{L}\mu_\ell P_\ell \leq 1 \end{aligned} \tag{30}$$

with a semidefiniteness constraint that is affine in $\boldsymbol{\mu}$.

Alternatively, we find the optimizer of (29) iteratively, e.g., with a GP algorithm. The ℓ-th component of the gradient $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_L]^{\text{T}}$ is

$$\delta_\ell = \sigma^2\bar{\boldsymbol{h}}_k^{\text{H}}\boldsymbol{X}^{-1}\boldsymbol{A}_\ell\boldsymbol{X}^{-1}\bar{\boldsymbol{h}}_k \tag{31}$$

where $\boldsymbol{X} = \boldsymbol{R} + \sigma^2\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_\ell$, and a GP step reads as

$$\boldsymbol{\mu}^{(j+1)} \leftarrow \mathcal{P}_{\mathcal{C}}\big(\boldsymbol{\mu}^{(j)} + a_j\boldsymbol{\delta}\big)$$

where a_j denotes the step size in iteration j. The orthogonal projection onto $\mathcal{C} = \big\{\boldsymbol{\mu} \in \mathbb{R}_+^L \big| \sum_{\ell=1}^{L}\mu_\ell P_\ell \leq 1\big\}$ can be expressed as

$$\mathcal{P}_{\mathcal{C}}(\boldsymbol{x}) = [\boldsymbol{x} - \gamma\boldsymbol{p}]^+ \tag{32}$$

where $\boldsymbol{p} = [P_1, \ldots, P_L]^{\text{T}}$, $[\cdot]^+$ performs an elementwise projection onto the non-negative orthant, i.e., $[z]^+ = \max(0, z)$ for a scalar $z \in \mathbb{R}$, and γ is chosen to satisfy the constraint in (30) with equality (cf. [14]).

Bogale and Vandendorpe proposed a *fixed-point* (FP) update in [5] and [6] to find the optimizer $\boldsymbol{\mu}^\star$, which leads to the primal optimal beamformer update in the AO. In our notation, this update rule reads as

$$\mu_\ell^{(j+1)} \leftarrow \frac{\boldsymbol{p}^{\mathrm{T}}\boldsymbol{\mu}^{(j)}}{\boldsymbol{\delta}^{(j),\mathrm{T}}\boldsymbol{\mu}^{(j)}} \frac{\delta_\ell^{(j)}}{P_\ell} \mu_\ell^{(j)}. \tag{33}$$

The normalized power iteration may be motivated via the complementary slackness conditions for the power constraints in (12), i.e., $\mu_\ell P_\ell = \mu_\ell \boldsymbol{b}^{\mathrm{H}} \boldsymbol{A}_\ell \boldsymbol{b}$ for $\ell = 1, \ldots, L$, that need to be satisfied for any locally optimal KKT point of (12).

V. Numerical Results

The main simulation setup consist of an N antenna transmitter and $K = 2$ users, each equipped with either a single antenna or $M_1 = M_2 = 2$ antennas for the multi-user vector and MIMO channel examples, respectively. The noise variances are fixed to $\sigma_k^2 = 1$ and $\boldsymbol{\Sigma}_k = \mathbf{I}_M$ and the power is varied between $-10\,\mathrm{dB}$ and $20\,\mathrm{dB}$.

We created the channels' means for a Gaussian model and for an exponential channel model.[3] For the previous model, we have drawn channel mean realizations $\bar{\boldsymbol{h}}_k$ and $\bar{\boldsymbol{h}}_k = \mathrm{vec}(\bar{\boldsymbol{H}}_k)$ from a standard Gaussian distribution and the channels covariance matrices are $\boldsymbol{C}_k = \frac{\kappa}{N}\mathbf{I}_N$ and $\boldsymbol{C}_k = \frac{\kappa}{M_k N}\mathbf{I}_{M_k N}$ for vector channels and matrix channels, respectively, with $\kappa = -10\,\mathrm{dB}$. For each of the channels' means, we have drawn another 1.000 channel realizations to compute the average SMSE for imperfect TxCSI and perfect RxCSI results with the alternating optimization for an accuracy of $\epsilon \leq 10^{-4}$.

For the latter exponential channel model, we used an exponential power profile, i.i.d. uniformly-distributed phase coefficients, and i.i.d. slow log-normal shadow-fading, i.e., $\bar{\boldsymbol{h}}_k = \sqrt{\tau_k}\boldsymbol{t}_k$ where $[\boldsymbol{t}_k]_i = \mathrm{e}^{\mathrm{j}\phi_{i,k}}\sqrt{\rho^{|i-k|}}$, $\phi_{i,k} \in [0, 2\pi)$ and $\ln(\tau_k^{\mathrm{dB}}) \sim \mathcal{N}(0,1)$ and $\boldsymbol{C}_k = \frac{\kappa\tau_k}{N}\,\mathrm{diag}(\rho^{|1-k|}, \ldots, \rho^{|N-k|})$ for the vector channels. Due to the exponential model, the transmitter prefers to serve the k-th user with the k-th antenna element, while the shadow fading factors τ_k are a measure for the users' link quality. This model is used to analyze multi-spotbeam SatCom setups, where the antenna directivity plays an important role (e.g., see [15]).

An extreme case for the exponential model is $\rho = 0$. Each transmit antenna exactly serves one user if $K = N$ and the channels are orthogonal. Then, per-antenna constraints and a sum power constraint result in the same SMSEs when $\tau_k = \tau$ for all $k = 1, \ldots, K$. In contrast, per-antenna constraints result in larger SMSEs than a sum-power constraint, e.g., when $\tau_k \gg \tau_j$ for at least one receiver pair k and j. For the numerical simulations we used $\rho = 0.1$.

In Fig. 1, we plotted the average achievable minimum SMSE for Gaussian channel mean realizations in the vector and MIMO case and for the exponential vector channel model with $N = 4$. Four lines can be seen for each of these models that correspond to a sum power limitation P (solid line with square marks), per-array power constraints for antennas $1,2$ and $3,4$ with $P_\ell = P/2$ (dashed line), and per-antenna constraints with $P_\ell = P/N$ (dash-dotted line). For the lines for the per-array and per-antenna constraint, we used the AO with

[3]An exponential power profile is amongst others used for analyzing intercell effects in multi-spotbeam SatCom communications (e.g., see [15]).

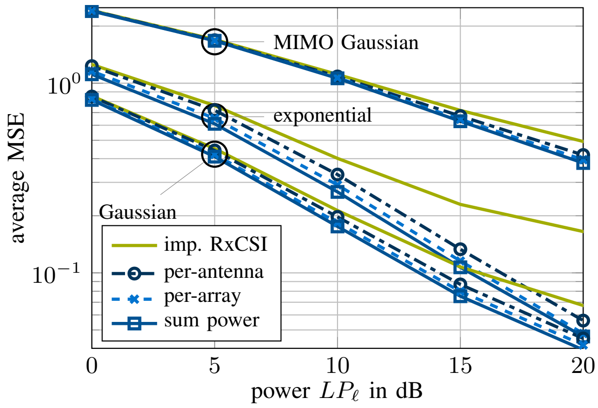

Figure 1: average SMSE vs. SNR for per-antenna, per-array, and a sum power constraint in a vector and MIMO BC with $N = 4$ transmit antennas and $K = 2$ users (with $M_k = 2$ receive antennas per user) for $\kappa = -10\,\mathrm{dB}$

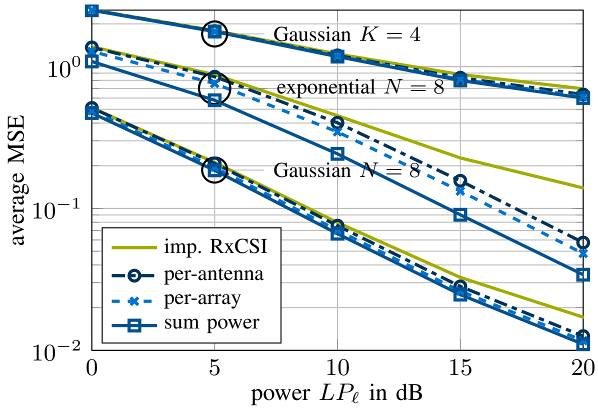

Figure 2: average SMSE vs. SNR for per-antenna, per-array, and a sum power constraint in a vector BC with $N = 4$ and $K = 4$ users and for $N = 8$ and $K = 2$ users for $\kappa = -10\,\mathrm{dB}$

the GP based dual beamformer update. The other beamformer update methods lead to the same SMSE performance curves.

With the choices for P_ℓ, the per-antenna constraints are stricter than the per-array constraints and the sum power constraint, resulting in an increased average SMSE. This is best visible for the exponential vector channel model, but results in only slight SMSE differences for the Gaussian vector model. In the Gaussian MIMO channel model, the multi-stream equalizers additionally compensate for the multiple transmit power restrictions with per-array and per-antenna constraints. In other words, the sum power limitation gives a tight lower bound for the achievable average SMSE with per-array or per-antenna constraints, which require more complex beamformer designs. An upper bound is obtained by assuming that the RxCSI equals the imperfect TxCSI (green solid line).

We further remark that the larger K (for constant N) the more likely the Gaussian channel characteristics does not prefer certain (groups of) antennas and the tighter the SMSE with a sum power constraint approximates the SMSE with the stricter per-array or per-antenna constraints (see Fig. 2). However, if we increase N (for constant K), the difference between sum-power-constrained minimum SMSE and the per-

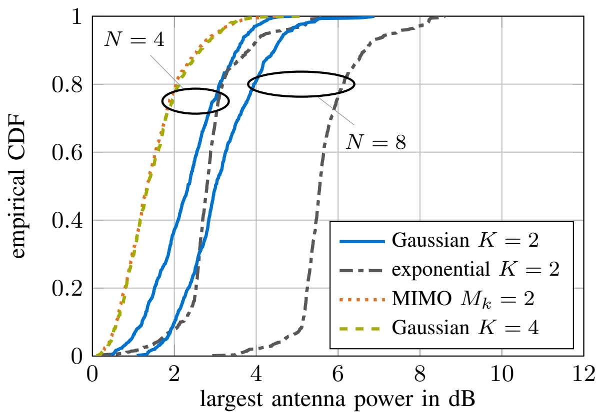

Figure 3: empirical CDF of largest per-antenna power in dB (normalized to P_ℓ) of the SMSE minimizing results with a sum power constraint and $\kappa = -10\,\mathrm{dB}$

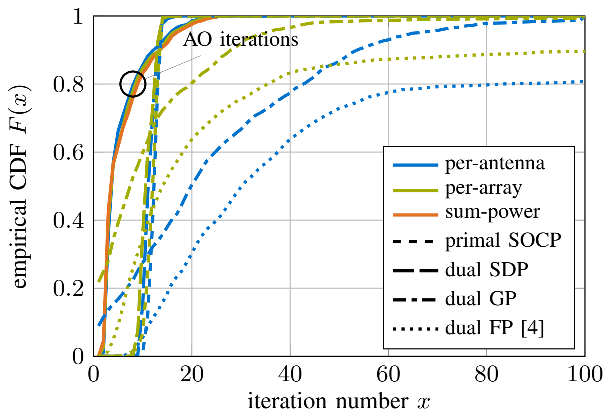

Figure 4: empirical CDF of required number of outer and inner iterations of the average SMSE minimization for the Gaussian channel model with $N = 4$, $K = 2$, and $\kappa = -10\,\mathrm{dB}$

antenna and per-array constrained minimum SMSEs increases for the exponential channel model but remains the same or even decreases for the Gaussian channel model.

Even though above SMSE results imply only a minor loss when we use more realistic per-antenna or per-array constraints compared to the less complex sum power constraints, we dramatically decrease the dynamic range of the per-antenna or per-array gains. In Fig. 3, we depict the empirical CDF of the largest per-antenna power for the given vector channel models when minimizing the average SMSE with a sum power constraint. The antenna power in dB is normalized such that an associated per-antenna power limitation $P_\ell = P/N$ represents $0\,\mathrm{dB}$.

For the vector BC with $N = 4$ and $K = 2$ users, a double of the per-antenna bound P_ℓ (i.e., $3\,\mathrm{dB}$) is surpassed in more than 20% of the channel realizations for the exponential and the Gaussian model. This 20% bound decreases to about $2\,\mathrm{dB}$ for an increasing number of receive antennas (or users) and fixed $N = 4$. This observation is in accordance with the conclusion that the SMSE with a sum-power constraint well approximates the achievable SMSE performance with per-antenna constraints for large K (see also Fig. 2).

In contrast, the normalized 20% largest per-antenna powers of the sum-power constrained optimization increase when increasing the number of transmit antennas to $N = 8$. While the increase is only to about $4\,\mathrm{dB}$ for the Gaussian model, $6\,\mathrm{dB}$ are surpassed with the exponential model. Because it prefers to serve the two users with only two of the eight antennas, the sum-power constraint SMSE optimization focuses most of the power to these two antennas. Since, moreover, the channels are subject to shadow fading, their gains are likely to differ, which results in an unbalanced power distribution for the users' main serving antennas. This unbalanced power is finally the reason for the increased SMSE gap between the sum power and the per-antenna constrained optimization results in Fig. 2.

To obtain an impression about the average SMSE precoder optimization via AO and the complexity of the dual beamformer updates, we plotted the empirical CDF for the required number of iterations until convergence of the outer AO and the inner (dual) beamformer design for the average SMSE minimization in the Gaussian channel setup with $N = 4$, $K = 2$, and $\kappa = -10\,\mathrm{dB}$ in Fig. 4. For the chosen setup, the number of AO iterations is almost independent with respect to the power constraints, i.e., four per-antenna constraints, two per-array constraint, or one sum power constraint, but the complexity of the inner beamformer updates increases with the number of power constraints L.

Both interior-point based solution approaches (dashed lines), i.e., the primal beamformer update with a bisection over SOCPs and the dual worst-case noise SDP, require almost the same number of interior-point iterations for per-antenna and per-array constraints (see Fig. 4). The SOCP based interior-point problem has a slightly lower worst-case complexity than the dual SDP method, but a sequence of these SOCPs needs to be solved in contrast to one SDP. Therefore, we experienced the dual SDP based beamformer update to be much faster than the SOCP based beamformer update. Note that the worst-case complexity of the conic programs is increasing in the number of power constraints L [16]. The number of SOCs and non-negative variables is $L + 1$ in (21) and in (30), respectively.

For the GP approach with an Armijo step size rule and the *fixed-point* (FP) update from [6], the number of iterations and, therewith, the computational complexity of the worst-case noise search increases from the per-array to the per-antenna constraint scenario (see Fig. 4). For example, while the GP converges in less than 20 iterations for 80% of the performed simulations with the two per-array constraints, more than 40 iterations are required for the per-antenna constraints. This effect is even more dramatic for the FP update rule, whose convergence speed is much slower than that of the GP with the Armijo step size rule. Especially, when some power constraints are inactive at the optimal point, the FP may require far more than a hundred iterations if convergence can be detected at all. While more than a hundred iterations are required in 10% of the simulations for per-array constraints, this number increases to about 19% for the per-antenna constraints. For this reason, we experienced that the GP based beamformer update is faster on average than the FP method [6], even though each FP iteration is less complex than a GP iteration with an Armijo step size rule.

VI. Conclusion

This paper provided an uplink-downlink duality for the (average) SMSE based precoder design with generalized power constraints by means of an alternating optimization. Similar to the standard SMSE minimizing AO approach, the precoder update in each AO step is transformed to a dual power allocation and filter design, which, however, also includes additionally a worst-case noise covariance search. Therewith, the complex search for an NM dimensional precoder is reduced to a worst case noise search over L non-negative reals that are the dual variables to the imposed power constraints. The solution is found via an SDP and a GP approach. The computational complexity of these approaches is increasing with the number of power constraints.

The simulations showed that the achievable SMSE with an imposed sum power constraint can give a good lower bound approximation of the performance with per-array or per-antenna power constraints for Gaussian channels. This good match of the SMSE is surprising since the dynamic range of the required per-antenna powers for the sum power constrained optimization is much larger compared to the restrictions with per-array or per-antenna constraints. A remarkable difference in the SMSEs is visible for the used exponential channel model, when the number of antennas is much larger than the number of users in the system and the channels are subject to different multiplicative fading, i.e. they have different norms.

Appendix

A. Proof of Lemma 1

Note that $\xi' = \xi(\|\boldsymbol{x}\|_2 + z)$ with $\xi \geq 0$ can be used as a Lagrangian multiplier for the convex SOC constraint $\|\boldsymbol{x}\|_2 \leq z$ without loss of optimality and strong duality [17, Appendix A]. Hence, is the Lagrangian function of the convex power balancing problem in (21) can be written as

$$L(\alpha, \boldsymbol{b}, \lambda, \boldsymbol{\mu}) = \lambda\sigma^2 + \alpha^2\Big(1 - \sum_{\ell=1}^{L}\mu_\ell P_\ell\Big) + \boldsymbol{b}^{\mathrm{H}}\Big(\boldsymbol{Y} - \tfrac{\lambda}{M-\varepsilon}\bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^{\mathrm{H}}\Big)\boldsymbol{b} \tag{34}$$

where $\boldsymbol{Y} = \lambda\boldsymbol{R} + \sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_\ell$ and $\lambda \geq 0$ and $\mu_\ell \geq 0$, $\ell = 1, \ldots, L$, are the multipliers to the SMSE and power constraints in (21), respectively.

The dual objective results from the unconstrained minimization of (34) w.r.t. α and $\boldsymbol{b}$, i.e., $g(\lambda, \boldsymbol{\mu}) = \min_{\alpha,\boldsymbol{b}} L(\alpha, \boldsymbol{b}, \lambda, \boldsymbol{\mu})$. Since α and $\boldsymbol{b}$ are unconstrained, $g(\lambda, \boldsymbol{\mu}) \to -\infty$ unless

$$\sum_{\ell=1}^{L}\mu_\ell P_\ell \leq 1 \tag{35}$$

and $\boldsymbol{Y} - \frac{\lambda}{M-\varepsilon}\bar{\boldsymbol{h}}\bar{\boldsymbol{h}}^{\mathrm{H}} \succeq \mathbf{0}_{N\times N}$. With Schur's complement [18, Appendix A.5.5], we can recast the latter condition as (cf. [11])

$$(M - \varepsilon) - \lambda\bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{Y}^{-1}\bar{\boldsymbol{h}} \geq 0. \tag{36}$$

Therewith, the dual problem of (21) reads as

$$\max_{\boldsymbol{\mu}\geq\mathbf{0},\lambda\geq 0} \lambda\sigma^2 \;\; \text{s.t.:} \sum_{\ell=1}^{L}\mu_\ell P_\ell \leq 1,\; \varepsilon \leq M - \lambda\bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{Y}^{-1}\bar{\boldsymbol{h}}. \tag{37}$$

The SMSE constraint upper bounds the objective in (37) since its right hand side is decreasing in λ when $\boldsymbol{\mu}$ is fixed. In other words, $\lambda^\star$ satisfies the SMSE constraint with equality if ε is attainable. Hence, jointly reversing the maximization over λ into a minimization and the direction of the inequality in the SMSE constraint does not affect the solution. Since, moreover,

$$M - \lambda\bar{\boldsymbol{h}}^{\mathrm{H}}\boldsymbol{Y}^{-1}\bar{\boldsymbol{h}} = \min_{\boldsymbol{u}} \overline{\mathrm{SMSE}}_{\mathrm{UL}} \tag{38}$$

with the SMSE minimizing equalizer $\boldsymbol{u}$ given in (25), (37) is equivalent to the uplink power minimization in (23).

We remark that we started from a convex formulation of the primal power balancing problem and, therefore, preserved strong duality within the derivation of the dual problem [18].

References

[1] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: Duality and Sum-MSE Minimization," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5436–5446, Nov. 2007.

[2] M. Schubert and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem with Individual SINR Constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[3] R. Hunger, M. Joham, and W. Utschick, "On the MSE-Duality of the Broadcast Channel and the Multiple Access Channel," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 698–713, Feb. 2009.

[4] M. Joham, M. Vonbun, and W. Utschick, "MIMO BC/MAC MSE Duality with Imperfect Transmitter and Perfect Receiver CSI," in *Proc. SPAWC 2010*, May 2010, pp. 1–5.

[5] T.E. Bogale and L. Vandendorpe, "Sum MSE Optimization for Downlink Multiuser MIMO Systems with Per Antenna Power Constraint: Downlink-Uplink Duality Approach," in *Proc. PIMRC, 2011*, Sept. 2011, pp. 2035–2039.

[6] T.E. Bogale and L. Vandendorpe, "Robust Sum MSE Optimization for Downlink Multiuser MIMO Systems With Arbitrary Power Constraint: Generalized Duality Approach," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1862–1875, Apr. 2012.

[7] A. Gründinger, A. Barthelme, M. Joham, and W. Utschick, "Mean Square Error Beamforming in SatCom: Uplink-Downlink Duality with Per-Feed Constraints," in *Proc. ISWCS 2014*, Aug. 2014, pp. 600–605.

[8] A. Gründinger, M. Joham, and W. Utschick, "Stochastic Transceiver Design in Multi-Antenna Channels with Statistical Channel State Information," in *Proc. ICASSP 2011*, Prague, Czech Republic, May 2011.

[9] J. Gorski, F. Pfeuffer, and K. Klamroth, "Biconvex Sets and Optimization with Biconvex Functions: A Survey and Extensions," *Math. Meth. of OR*, vol. 66, no. 3, pp. 373–407, May 2007.

[10] R. Hunger, W. Utschick, D. A. Schmidt, and M. Joham, "Alternating Optimization for MMSE Broadcast Precoding," in *Proc. ICASSP 2006*, Toulouse, France, Oct. 2003, vol. 4, pp. 757–760.

[11] 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, May 2007.

[12] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming," Feb. 2009.

[13] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB Toolbox for Optimization over Symmetric Cones," *Optimization Methods and Software*, vol. 11, no. 1, pp. 625–653, Jan. 1999.

[14] R. Hunger, D. A. Schmidt, M. Joham, and W. Utschick, "A General Covariance-Based Optimization Framework Using Orthogonal Projections," in *Proc. SPAWC 2008*, July 2008, pp. 76–80.

[15] N. Letzepis and A.J. Grant, "Capacity of the Multiple Spot Beam Satellite Channel With Rician Fading," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5210–5222, Nov. 2008.

[16] Y. Ye, "Interior-Point Algorithm: Theory and Application," Tech. Rep., Department of Management Sciences, University of Iowa City, 1995.

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

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, NY, USA, 1st edition, Mar. 2004.