Mean Square Error Beamforming in SatCom: Uplink-Downlink Duality with Per-Feed Constraints

Andreas Gründinger, Michael Joham, Andreas Barthelme, and Wolfgang Utschick
Associate Institute for Signal Processing, Technische Universität München, Germany
Email: {gruendinger,joham,a.barthelme,utschick}@tum.de

***Abstract*—Balancing the per-user average *mean square errors* (MSEs) of a *satellite communication* (SatCom) system under per-feed limitations—linear transmit power constraints—is the focus of this work. We propose an uplink-downlink duality for this optimization via Lagrangian multiplier theory. Strong duality is shown. Simple fixed point methods are used for the uplink power allocation and the worst-case noise covariance calculation. Simulation results are used to compare the performance for the SatCom channel with that of a standard Gaussian channel model.**

***Index Terms*—uplink-downlink MSE duality; imperfect CSI; MSE balancing; per-feed constraints; satellite communication**

I. Introduction

Multi-spotbeam *satellite communication* (SatCom) requires adaptive beamforming to manage the increasing intercell interference in the downlink [1], when narrowing down the cell width below the 3 dB coverage and increasing the frequency reuse in future systems. The goal is a reliable data service provision proportional to the receivers' demands.

The beamformer optimization must deal with the special channel model and needs to incorporate *per-feed power limitations* that differs from the standard sum power constraint in terrestrial wireless communications. Chrisopoulos et al. [2] model the per-feed limitations as general linear power constraints. Additional non-linear power constraints were included in [3] to represent saturation effects in the radio frequency amplifiers. Here, we restrict to linear per-feed constraints which is the basis for future work on non-linear power limitations.

Special forms of per-feed constraints, e.g., per-antenna constraints, can also be encountered in terrestrial systems when power sharing is impossible, e.g., due to a physical separation of the antennas. In [4], the ratios between achievable and target *signal-to-interference-and-noise-ratios* (SINRs) are maximized under such power restrictions. The solution is obtained via repeatedly solving related *second order cone* (SOC) programs using a convex optimization toolbox, e.g., CVX [5]. However, the number of spotbeams and antennas in SatCom, that can meet a few hundreds, leads to a large number of per-feed constraints. Therefore, this approach is not attractive and the use of uplink-downlink duality is favorable.

Uplink-downlink duality is an utmost useful tool for beamformer optimizations which are difficult to solve directly in the vector *broadcast channel* (BC). The problem is transformed to a dual *multiple access channel* (MAC) problem with the same achievable SINRs [6] or (average) MSEs [7]. The precoder design translates to an equalizer design and a power allocation. Fixed point methods for these tasks are efficient, reliable, and can be implemented without using optimization toolboxes [8].

The research for this work was supported by Deutsche Forschungsgemeinschaft (DFG) under fund Jo 724/1-2.

Due to delays, shadowing, and scattering effects in SatCom, the *channel state information* (CSI) at the transmitter (satellite) is imperfect. Therefore, we design the beamformers according to a min-max balancing of the average MSEs. Note that a lower bound for the rates is maximized via minimizing MSEs.

MSE dualities between the vector BC and the vector MAC were first revealed using SINR duality [9] for a sum power constraint. The SINR uplink-downlink duality was extended to linear power constraints via Lagrangian multiplier theory, by Yu and Lan [10] for the power minimization and in [11] for a max-min balancing formulation. However, similar *average* MSE dualities are missing to the best of our knowledge. Only for the average sum MSE minimization via alternating optimization, an uplink-downlink approach was presented by Bogale and Vandendorpe in [12].

We propose an uplink-downlink average MSE duality with linear per-feed power constraints via Lagrangian multiplier theory. In the dual uplink, the min-max average MSE problem has a weighted sum-power constraint and includes a search for the worst case noise covariance. Strong duality can be shown via a related power minimization problem formulation. Exploiting duality, we solve the min-max average MSE balancing problem in the uplink. Simulation results for a standard channel model and for a SatCom channel model are presented.

II. Downlink System Model

The downlink received signals, e.g., from a SatCom scenario, read as $y_k = \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{t}_k s_k + \boldsymbol{h}_k^{\mathrm{H}} \sum_{i \neq k}^{K} \boldsymbol{t}_i s_i + n_k$. The independent data signals $s_i \sim \mathcal{N}_{\mathbb{C}}(0,1)$ are linearly precoded with $\boldsymbol{t}_i$ at the transmitter and sent over the channels $\boldsymbol{h}_k \in \mathbb{C}^N$ to the K mobile receivers. The additive noise at mobile i is $n_i \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_i^2)$, $i = 1, \ldots, K$. The received signal is scaled with $f_k \in \mathbb{C}$, i.e., $\hat{s}_k = f_k y_k$, such that the MSE $\mathrm{E}\big[|s_k - \hat{s}_k|^2\big]$ for the k-th receiver's signal estimate reads as

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

A. Linear Per-Feed Power Constraints

The per-feed transmit power limitations are represented as

$$\sum_{i=1}^{K} \boldsymbol{t}_i^{\mathrm{H}} \boldsymbol{A}_{i,\ell} \boldsymbol{t}_i = \sum_{i=1}^{K} \|\boldsymbol{A}_{i,\ell}^{1/2} \boldsymbol{t}_i\|_2^2 \leq P_\ell, \quad \ell = 1, \ldots, L \quad (2)$$

where $\boldsymbol{A}_{i,\ell} = \boldsymbol{A}_{i,\ell}^{\mathrm{H}/2} \boldsymbol{A}_{i,\ell}^{1/2} \succeq \boldsymbol{0}$ with $\mathrm{rank}\big\{\sum_{\ell=1}^{L} \boldsymbol{A}_{i,\ell}\big\} = N$.

Important examples from terrestrial communications 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 $i=1,...,K$ 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 $L=N$.

Per-feed constraints are equal to per-antenna constraints if a single horn antenna sends the output of a high-power amplifier to the reflector at the satellite. However, when small phased arrays form the feeds, the matrices $\boldsymbol{A}_{i,\ell}$ have $\mathrm{rank}\{\boldsymbol{A}_{i,\ell}\} > 1$.

The per-feed constraints in (2) are a first step towards more advanced power constraints for the beamformer design of multi-beam SatCom systems. For example, non-linear power constraints could model the saturation effects in the radio frequency amplifiers in SatCom (cf. [3]).

B. Channel State Information Models

Imperfect CSI at the transmitter is the second issue in SatCom. We assume knowledge of the first and second order moment of the channels $\boldsymbol{h}_k$, i.e.,

$$\bar{\boldsymbol{h}}_k = \mathrm{E}[\boldsymbol{h}_k], \qquad \boldsymbol{R}_k = \mathrm{E}[\boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}}]. \tag{3}$$

The expressions for $\bar{\boldsymbol{h}}_k$ and $\boldsymbol{R}_k$ result from the employed fading model. For example, the Gaussian model $\boldsymbol{h}_k = \bar{\boldsymbol{h}}_k + \tilde{\boldsymbol{h}}_k$ with $\tilde{\boldsymbol{h}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_k)$ is common to represent fading effects and limited training and/or feedback capabilities in terrestrial systems. Therewith, the second moment is $\boldsymbol{R}_k = \bar{\boldsymbol{h}}_k \bar{\boldsymbol{h}}_k^{\mathrm{H}} + \boldsymbol{C}_k$.

Multi-spotbeam satellite mobile channels follow the fading model in [13] (see also [1]). The basis is Rician fading that models the line-of-sight characteristic of satellite channels, i.e.,

$$\boldsymbol{z}_k = \sqrt{\frac{\kappa}{\kappa+1}} \bar{\boldsymbol{z}}_k + \sqrt{\frac{1}{\kappa+1}} \tilde{\boldsymbol{z}}_k \tag{4}$$

with Rician factor κ, line-of-sight component $\bar{\boldsymbol{z}}$, and complex random $\tilde{\boldsymbol{z}}_k$. Since scatterers are mainly around the receivers and far from the transmitting satellite, we modeled $\tilde{\boldsymbol{z}}_k = w_k \bar{\boldsymbol{z}}_k$ and $w_k \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{w_k}^2)$ in previous work (e.g., [14]). Such a restriction is not imposed in this work, i.e., $\tilde{\boldsymbol{z}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{z}})$ may have a full rank covariance matrix, e.g., $\boldsymbol{C}_{\boldsymbol{z}} = \mathbf{I}_N$.

According to Loo's model [15], the line-of-sight component is subject to a multiplicative log-normal distributed factor ξ_k, i.e., $\bar{\boldsymbol{z}}_k = \xi_k \hat{\boldsymbol{z}}_k$ where $\ln(\xi_k) \sim \mathcal{N}(m_k, \sigma_{\xi_k}^2)$. This factor comprises rain fading, shadowing, and the receivers' mobility. The resulting fading vector in (4) is distorted by the beam gain characteristic $\boldsymbol{B}_k = \mathrm{diag}(b_{1,k}, \ldots, b_{N,k})$, that depends on the relative position of the receivers to the spotbeam centers. Here, $b_{j,k} = \sqrt{g_{j,k}}\, \mathrm{e}^{-\mathrm{j}\psi_{j,k}}$ contains the tapered-aperture antenna gain [16] from antenna j to user k with $g_{i,k} = \left(\frac{J_1(u_{i,k})}{2u_{i,k}} + 36\frac{J_3(u_{i,k})}{u_{i,k}^3}\right)^2$, where $J_1(\cdot)$ and $J_3(\cdot)$ are the first kind Bessel functions of order one and three, respectively, of $u_{i,k} = 2.07123\frac{\sin(\theta_{i,k})}{\sin(\theta_{3\mathrm{dB}})}$. The angle $\theta_{i,k}$ is between beamcenter i and user k as seen from the satellite and $\theta_{3\mathrm{dB}}$ is the one-sided half-power beamwidth. For an approximation of the small phase shifts $\psi_{j,k}$, we assumed that the antennas are oriented in a plane orthogonal to the central beam, that is directed to Munich.

The channel to mobile k reads finally as (cf. [3])

$$\boldsymbol{h}_k = \sqrt{g_{\mathrm{FSL},k}} \boldsymbol{B}_k \boldsymbol{z}_k \tag{5}$$

where $g_{\mathrm{FSL},k} = \left(\frac{\lambda}{4\pi}\right)^2 \frac{1}{d_k}$ models the *free space loss* (FSL) with wavelength λ and altitude d_k. The moments of (5) are

$$\begin{aligned} \bar{\boldsymbol{h}}_k &= \sqrt{\frac{g_{\mathrm{FSL},k}\kappa}{\kappa+1}}\, \mathrm{e}^{m_k + \sigma_{\xi_k}^2/2} \boldsymbol{B}_k \hat{\boldsymbol{z}}_k, \\ \boldsymbol{R}_k &= \mathrm{e}^{\sigma_{\xi_k}^2} \bar{\boldsymbol{h}}_k \bar{\boldsymbol{h}}_k^{\mathrm{H}} + \frac{g_{\mathrm{FSL},k}}{\kappa+1} \boldsymbol{B}_k \boldsymbol{C}_{\boldsymbol{z}_k} \boldsymbol{B}_k^{\mathrm{H}}. \end{aligned} \tag{6}$$

III. Average Mean Square Error Balancing

We design the beamformers to minimize the maximum average MSE. We remark that this is a conservative approach for average rate balancing. The rate $r_k = -\log_2(\mathrm{MMSE}_k)$ is a convex function of the *minimum MSE* (MMSE), i.e., the MSE with perfect CSI MMSE receive filters

$$f_{k,\mathrm{MMSE}} = \frac{\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{t}_k}{\sum_{i=1}^K |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{t}_i|^2 + \sigma_k^2}. \tag{7}$$

Therewith, Jensen's inequality provides the lower bound

$$\mathrm{E}[r_k] \geq -\log_2(\mathrm{E}[\mathrm{MMSE}_k]).$$

Balancing the average MSEs with the filters in (7), i.e.,

$$\mathrm{E}[\mathrm{MMSE}_k] = 1 - \mathrm{E}\left[|\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{t}_k|^2 \Big/ \Big(\sum_{i=1}^K |\boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{t}_i|^2 + \sigma_k^2\Big)\right], \tag{8}$$

remains difficult for imperfect transmitter CSI. The expectation in (8) is over a ratio of correlated random parameters. Even though closed form expressions may be found for this expectation, a direct minimization of $\max_k \mathrm{E}[\mathrm{MMSE}_k]$ w.r.t. the beamformers is still a non-convex optimization problem.

We resolve this issue with a beamformer optimization that considers the receivers to have the same imperfect channel knowledge as the transmitter. Note that any receive filter different from that in (7) results in an upper bound to the achievable instantaneous MMSE with perfect receiver CSI. Therefore, also an upper bound for the average MMSE in (8) is obtained. The joint minimization of these MSE upper bounds results in the maximization of lower bounds for the achievable rates. Whenever the average MSEs are balanced at a level $\hat{\varepsilon}$, the average rates are ensured to lie above $-\log_2(\hat{\varepsilon})$.

If the receivers have the same imperfect CSI as the transmitter, the min-max average MSE optimization with (linear) per-feed constraints reads as

$$\begin{aligned} &\min_{\boldsymbol{f},\boldsymbol{t}} \max_k \widehat{\mathrm{MSE}}_k \\ &\text{s. t.:} \sum_{i=1}^K \|\boldsymbol{A}_{i,\ell}^{1/2} \boldsymbol{t}_i\|_2^2 \leq P_\ell, \quad \forall\, \ell = 1, \ldots, L \end{aligned} \tag{9}$$

where $\boldsymbol{t} = [\boldsymbol{t}_1^{\mathrm{T}}, \ldots \boldsymbol{t}_K^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{f} = [f_1, \ldots, f_K]^{\mathrm{T}}$, and the average MSEs $\mathrm{E}[\mathrm{MSE}_k]$ are given by

$$\widehat{\mathrm{MSE}}_k = 1 - 2\,\mathrm{Re}\{\tilde{f}_k^* \bar{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{t}_k\} + \sum_{i=1}^K |\tilde{f}_k|^2 \boldsymbol{t}_i^{\mathrm{H}} \boldsymbol{R}_k \boldsymbol{t}_i + |\tilde{f}_k|^2 \sigma_k^2. \tag{10}$$

We provide a dual uplink formulation to solve problem (9), which is then a receive filter design and power allocation problem over a worst-case noise covariance matrix. This matrix may be found via a subgradient method, for example, and the power allocation is a simple fixed point algorithm.

IV. MSE Upper Bound Minimization

By inserting the MMSE filters for imperfect receiver CSI

$$\tilde{f}_{k,\text{MMSE}} = \frac{\bar{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{t}_k}{\sum_{i=1}^K \boldsymbol{t}_i^{\mathrm{H}} \boldsymbol{R}_k \boldsymbol{t}_i + \sigma_k^2}, \tag{11}$$

problem (9) becomes a quasiconvex program. The objective, i.e., $\max_k \widehat{\text{MMSE}}_k$ with

$$\widehat{\text{MMSE}}_k = 1 - \frac{|\bar{\boldsymbol{h}}_k^{\mathrm{H}} \boldsymbol{t}_k|^2}{\sum_{i=1}^K \boldsymbol{t}_i^{\mathrm{H}} \boldsymbol{R}_k \boldsymbol{t}_i + \sigma_k^2}, \tag{12}$$

is the pointwise maximum of quasiconvex functions, since the lower level set of the MMSE features a convex reformulation. Moreover, the per-feed power constraints in (9) are convex.

To see this, we first reformulate (9) with (12) as

$$\begin{aligned}\min_{\hat{\varepsilon},\boldsymbol{t}}\ \hat{\varepsilon} \quad \text{s.t.:}\ & \sum_{i=1}^K \|\boldsymbol{A}_{i,\ell}^{1/2}\boldsymbol{t}_i\|_2^2 \leq P_\ell,\ \forall \ell = 1,\dots,L, \\ \frac{|\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{t}_k|^2}{(1-\hat{\varepsilon})} \geq & \sum_{i=1}^K \|\boldsymbol{R}_k^{1/2}\boldsymbol{t}_i\|_2^2 + \sigma_k^2,\ \forall k = 1,\dots,K\end{aligned} \tag{13}$$

where we introduced the balancing level $\hat{\varepsilon} \in [0,1]$ as a slack variable for the maximum of the K average MMSEs and $\boldsymbol{R}_k^{1/2}$ is the square root matrix of $\boldsymbol{R}_k = \boldsymbol{R}_k^{\mathrm{H}/2}\boldsymbol{R}_k^{1/2}$. Since the average MMSEs in (12) and the power limitations are independent w.r.t. a phase shift of the beamformers, we restrict $\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{t}_k$ to be real and positive in (13).[1] The convex lower level set representations for the MMSEs are then obtained via a square root operation on both sides of the inequality as is seen in (14).

A. Bisection Over Power Minimizations

We may find the optimizers of (13) via solving a series of convex problems. This is similar to SINR balancing with per-antenna constraints [4], where the balanced SINRs and the corresponding beamformers are found with a bisection. In each bisection step, a power minimization in SOC form is solved.

We rewrite the power minimization corresponding to (13) as

$$\begin{aligned}&\min_{\alpha,\boldsymbol{t}}\ \alpha^2 \\ &\text{s.t.:}\ \frac{\mathrm{Re}\{\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{t}_k\}}{\sqrt{1-\hat{\varepsilon}}} \geq \left\|[\boldsymbol{t}^{\mathrm{H}}(\mathbf{I}_K \otimes \boldsymbol{R}_k^{\mathrm{H}/2}), \sigma_k]\right\|_2, \\ &\qquad \mathrm{Im}\{\bar{\boldsymbol{h}}_k^{\mathrm{H}}\boldsymbol{t}_k\} = 0, \quad \forall\, k = 1,\dots,K \\ &\qquad \left\|\boldsymbol{A}_\ell^{1/2}\boldsymbol{t}\right\|_2 \leq \alpha\sqrt{P_\ell}, \quad \forall\, \ell = 1,\dots,L\end{aligned} \tag{14}$$

where $\boldsymbol{A}_\ell^{1/2}$ are blockdiagonal matrices with elements $\boldsymbol{A}_{i,\ell}^{1/2}$, $i = 1,\dots,K$, and the joint MSE level $\hat{\varepsilon}$ is fixed. Obviously, the minimum $\alpha_{\text{min}}^2(\hat{\varepsilon})$ of (14) is strictly monotonically decreasing in $\hat{\varepsilon}$. Similarly, the minimum $\hat{\varepsilon}_{\text{min}}(P_1,\dots,P_L)$ of (13) is strictly monotonically decreasing in α if $P_\ell = \alpha^2 P'_\ell$ and $P'_\ell > 0$ is fixed. Therefore, a simple line search via (14), e.g., a bisection, is able to find the optimizers of (13) if it meets $\alpha_{\text{min}}(\hat{\varepsilon}) = 1$ with a predefined accuracy. We used the disciplined convex programming toolbox CVX [5] to find $\alpha_{\text{min}}(\hat{\varepsilon})$ and check our numerical simulations.[2]

B. Uplink-Downlink MSE Duality

Alternatively, the average MSE balancing problem in (9) can be solved in the dual uplink. As can be inferred from the proof of Proposition 1, the dual uplink average MSE balancing optimization can be written as

$$\begin{aligned}&\max_{\boldsymbol{\mu}\geq\mathbf{0}}\ \min_{\boldsymbol{\lambda}\geq\mathbf{0},\boldsymbol{u}}\ \max_i\ \widehat{\text{MSE}}_{i,\text{UL}} \\ &\text{s.t.:}\ \sum_{i=1}^K \lambda_i\sigma_i^2 \leq \sum_{\ell=1}^L \mu_\ell P_\ell.\end{aligned} \tag{15}$$

The average uplink MSE that corresponds to user i reads as

$$\begin{aligned}\widehat{\text{MSE}}_{i,\text{UL}} = 1 - \sqrt{\lambda_i}2\,\mathrm{Re}\{\bar{\boldsymbol{h}}_i^{\mathrm{H}}\boldsymbol{u}_i\} \\ + \boldsymbol{u}_i^{\mathrm{H}}\Big(\sum_{k=1}^K \lambda_k\boldsymbol{R}_k + \sum_{\ell=1}^L \mu_\ell\boldsymbol{A}_{i,\ell}\Big)\boldsymbol{u}_i.\end{aligned} \tag{16}$$

The uplink power allocation vector $\boldsymbol{\lambda} = [\lambda_1,\dots,\lambda_K]^{\mathrm{T}} \geq \mathbf{0}$ comprises the dual variables associated with the MMSE constraints in (13). The vector $\boldsymbol{\mu} = [\mu_1,\dots,\mu_L]^{\mathrm{T}} \geq \mathbf{0}$, that defines the worst-case noise covariance matrix $\sum_{\ell=1}^L \mu_\ell\boldsymbol{A}_{i,\ell}$ in (16), contains the dual variables for the per-feed constraints.

Note that the optimal filters in $\boldsymbol{u} = [\boldsymbol{u}_1^{\mathrm{T}},\dots,\boldsymbol{u}_K^{\mathrm{T}}]^{\mathrm{T}}$ of (15) are

$$\boldsymbol{u}_i = \Big(\sum_{k=1}^K \lambda_k\boldsymbol{R}_k + \sum_{\ell=1}^L \mu_\ell\boldsymbol{A}_{i,\ell}\Big)^{\dagger}\bar{\boldsymbol{h}}_i\sqrt{\lambda_i}. \tag{17}$$

Inserting (17) into (16), (15) may be written as a max-min MMSE balancing problem with the MMSEs

$$\widehat{\text{MMSE}}_{i,\text{UL}} = 1 - \lambda_i\bar{\boldsymbol{h}}_i^{\mathrm{H}}\Big(\sum_{k=1}^K \lambda_k\boldsymbol{R}_k + \sum_{\ell=1}^L \mu_\ell\boldsymbol{A}_{i,\ell}\Big)^{\dagger}\bar{\boldsymbol{h}}_i, \tag{18}$$

which are independent w.r.t. a common scaling of $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$.

Proposition 1. *The duality gap between* (9) *and* (15) *is zero.*

Proof. To prove the strong duality result, we create an inverse power minimization to the uplink max-min MSE balancing problem in (15). This power minimization problem is moreover strongly dual to the convex power minimization problem in (14). Therefore, the same transmit power is required to achieve the same MSE for all users in the uplink and the downlink. Since the power minimization in (14) is again inverse to the downlink MSE balancing problem in (9), the balanced MSE in the uplink and the downlink is the same.

To find the uplink power minimization, we remark that (15) is independent w.r.t. a common scaling of $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ and the power constraint will be satisfied with equality in the optimum. Therefore, we may replace the sum power constraint in (15) by the two constraints $\sum_{\ell=1}^L \mu_\ell P_\ell \leq 1$ and $\sum_{i=1}^K \lambda_i\sigma_i^2 \leq 1$

[1] Under this restriction, all possible solutions of (9) result from the possible solutions of (13) via $\boldsymbol{t}'_i = \mathrm{e}^{\mathrm{j}\phi_i}\,\boldsymbol{t}_i$, $\phi_i \in [0, 2\pi)$.

[2] Note that that the constraints in (14) may not be attainable, e.g., when $\hat{\varepsilon} < \frac{K-N}{K}$ even if all $\boldsymbol{R}_k$'s are rank-one [17]. For matrices $\boldsymbol{R}_k$ with a rank larger than one, the attainable $\hat{\varepsilon}$ can lie far below this bound. We set $\alpha_{\text{min}}(\hat{\varepsilon})$ to infinity in this cases.

without changing the solution. Keeping the former of the two constraints and changing the latter one to $\sum_{i=1}^{K}\lambda_i\sigma_i^2 \leq P$, the minimum balanced uplink MSE $\hat{\varepsilon}_{\mathrm{UL}}(P)$ becomes a strictly monotonically decreasing function in $P\geq 0$. The corresponding inverse function reads as

$$P_{\min}(\hat{\varepsilon}) = \max_{\boldsymbol{\mu}\geq\mathbf{0}} \min_{\boldsymbol{\lambda}\geq\mathbf{0},\boldsymbol{u}} \sum_{i=1}^{K}\lambda_i\sigma_i^2 \tag{19}$$
$$\text{s.t.:} \sum_{\ell=1}^{L}\mu_\ell P_\ell \leq 1,\ \hat{\varepsilon}\geq \widehat{\mathrm{MSE}}_{i,\mathrm{UL}},\ \forall i=1,\dots,K.$$

It remains to show that (19) is dual to the convex optimization in (14). This proof directly follows the steps from Yu and Lan in [10]. Note that duality can be based on the quadratic constraints from (13) instead of those in (14) since the resulting KKT conditions are equivalent [18, Appendix A]. Hence, we can write the Lagrangian function of (14) as

$$L(\alpha,\boldsymbol{t},\boldsymbol{\lambda},\boldsymbol{\mu}) = \sum_{i=1}^{K}\lambda_i\sigma_i^2 + \alpha^2\Big(1-\sum_{\ell=1}^{L}\mu_\ell P_\ell\Big) \tag{20}$$
$$+\sum_{i=1}^{K}\boldsymbol{t}_i\Big(\boldsymbol{Y}_i - \frac{\lambda_i}{1-\hat{\varepsilon}}\bar{\boldsymbol{h}}_i\bar{\boldsymbol{h}}_i^{\mathrm{H}}\Big)\boldsymbol{t}_i$$

where $\boldsymbol{Y}_i=\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_{i,\ell}+\sum_{k=1}^{K}\lambda_k\boldsymbol{R}_k$.

The dual objective results from the unconstrained minimization of (20) w.r.t. α and $\boldsymbol{t}$, i.e., $g(\boldsymbol{\lambda},\boldsymbol{\mu})=\min_{\alpha,\boldsymbol{t}} L(\alpha,\boldsymbol{t},\boldsymbol{\lambda},\boldsymbol{\mu})$. Since α and the $\boldsymbol{t}_i$'s are unconstrained, $g(\boldsymbol{\lambda},\boldsymbol{\mu})\to-\infty$ unless

$$\sum_{\ell=1}^{L}\mu_\ell P_\ell \leq 1 \tag{21}$$

and $\boldsymbol{Y}_i-\frac{\lambda_i}{1-\hat{\varepsilon}}\bar{\boldsymbol{h}}_i\bar{\boldsymbol{h}}_i^{\mathrm{H}}\succeq\mathbf{0}_{N\times N}$. With Schur's complement [19, A.5.5], we can recast the latter condition as (cf. [10])

$$(1-\hat{\varepsilon})-\lambda_i\bar{\boldsymbol{h}}_i^{\mathrm{H}}\boldsymbol{Y}_i^{\dagger}\bar{\boldsymbol{h}}_i\geq 0. \tag{22}$$

Equivalence follows since $\boldsymbol{Y}_i\succeq\mathbf{0}_{N\times N}$, $1-\hat{\varepsilon}>0$, and $(\mathbf{I}_N-\boldsymbol{Y}_i\boldsymbol{Y}_i^{\dagger})\bar{\boldsymbol{h}}_i=\mathbf{0}_N$ as $\boldsymbol{Y}_i\succeq\lambda_i\boldsymbol{R}_i=\lambda_i\,\mathrm{E}[\boldsymbol{h}_i\boldsymbol{h}_i^{\mathrm{H}}]\succ\lambda_i\bar{\boldsymbol{h}}_i\bar{\boldsymbol{h}}_i^{\mathrm{H}}$.

With (22) and (21), the dual problem of (14) reads as

$$\max_{\boldsymbol{\mu},\boldsymbol{\lambda}\geq\mathbf{0}} \sum_{i=1}^{K}\lambda_i\sigma_i^2 \quad \text{s.t.:} \sum_{\ell=1}^{L}\mu_\ell P_\ell\leq 1, \tag{23}$$
$$\lambda_i\leq\frac{1-\hat{\varepsilon}}{\bar{\boldsymbol{h}}_i^{\mathrm{H}}\big(\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_{i,\ell}+\sum_{k=1}^{K}\lambda_k\boldsymbol{R}_k\big)^{\dagger}\bar{\boldsymbol{h}}_i},\ \forall i=1,\dots,K.$$

The right hand side of the MMSE constraint in (23) is positive and sublinearly monotonically increasing in $\boldsymbol{\lambda}\geq\mathbf{0}$ when $\boldsymbol{\mu}$ is fixed. In other words, these right hand sides define a standard interference function [20] that is parametrized in $\boldsymbol{\mu}$ and there is a unique $\boldsymbol{\lambda}^{\star}\geq\mathbf{0}$ that satisfies all MMSE constraints with equality and minimizes the objective if $\hat{\varepsilon}$ is attainable. Hence, reversing the maximization over $\boldsymbol{\lambda}$ into a minimization and the direction of the inequality in the MMSE constraints does not affect the solution. Moreover, since

$$\hat{\varepsilon}\geq 1-\lambda_i\bar{\boldsymbol{h}}_i^{\mathrm{H}}\boldsymbol{Y}_i^{\dagger}\bar{\boldsymbol{h}}_i=\widehat{\mathrm{MMSE}}_{i,\mathrm{UL}} \tag{24}$$

as can be seen in (18), we indeed obtain the power minimization formulation in (19). This proves that a solution of (15) results in a solution for (9). □

The downlink beamformers and receive filters follow from the uplink filters in (17) and powers in $\boldsymbol{\lambda}$ by (cf. [7])

$$\boldsymbol{t}_i=\sqrt{\beta_i}\boldsymbol{u}_i,\quad \tilde{f}_i=\sqrt{\lambda_i}/\sqrt{\beta_i},\quad \forall i=1,\dots,K \tag{25}$$

when the balanced MSEs of (9) and (15) are equal, i.e.,

$$\widehat{\mathrm{MSE}}_i=\widehat{\mathrm{MSE}}_{i,\mathrm{UL}},\quad i=1,\dots,K. \tag{26}$$

Inserting (25) into the downlink MSEs from (10), the equation system in (26) can be rewritten as

$$\boldsymbol{\Psi}\boldsymbol{\beta}=\boldsymbol{\Sigma}\boldsymbol{\lambda} \tag{27}$$

where $\boldsymbol{\Sigma}=\mathrm{diag}(\sigma_1^2,\dots,\sigma_K^2)$, $\boldsymbol{\beta}=[\beta_1,\dots,\beta_K]^{\mathrm{T}}$, and

$$[\boldsymbol{\Psi}]_{i,j}=\begin{cases}\sum_{k\neq i}\lambda_k\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{R}_k\boldsymbol{u}_i,+\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{u}_i^{\mathrm{H}}\boldsymbol{A}_{i,\ell}\boldsymbol{u}_i & i=j,\\ -\lambda_i\boldsymbol{u}_j^{\mathrm{H}}\boldsymbol{R}_i\boldsymbol{u}_j, & i\neq j.\end{cases}$$

As $\boldsymbol{\Psi}$ is column-wise diagonally dominant with positive diagonal elements and non-positive off-diagonal elements, its inverse exists and has non-negative elements. That means, we can solve (27) for positive $\boldsymbol{\beta}$, and therewith calculate the downlink beamformers and filters in (25).

C. Iterative Uplink MSE Balancing

An iterative solution for the uplink MSE balancing problem consists of two nested loops. The inner loop solves the power allocation (and equalizer optimization) in the MAC, while the outer loop controls the uplink noise covariance (cf. [11]).

To balance the MSEs and satisfy the sum-power constraint in (15) with equality, we use the globally convergent update

$$\lambda_i^{(n+1)}\leftarrow\frac{1-\hat{\varepsilon}^{(n+1)}}{\bar{\boldsymbol{h}}_i^{\mathrm{H}}\big(\sum_{k=1}^{K}\lambda_k^{(n)}\boldsymbol{R}_k+\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_{i,\ell}\big)^{\dagger}\bar{\boldsymbol{h}}_i}, \tag{28}$$

which follows from the constraint formulation in (23). The normalization with $1-\hat{\varepsilon}^{(n+1)}$, where [cf. constraint in (15)]

$$\hat{\varepsilon}^{(n+1)}\leftarrow 1-\frac{\sum_{\ell=1}^{L}\mu_\ell P_\ell}{\sum_{i=1}^{K}\sigma_i^2/\bar{\boldsymbol{h}}_i^{\mathrm{H}}(\sum_{k=1}^{K}\lambda_k^{(n)}\boldsymbol{R}_k+\sum_{\ell=1}^{L}\mu_\ell\boldsymbol{A}_{i,\ell})^{\dagger}\bar{\boldsymbol{h}}_i},$$

is to exploit full transmit power. This fixed-point update is less complex than the eigenvector calculation from [8], but more iterations are required until convergence (cf. [11]).

To find $\boldsymbol{\mu}$, a subgradient projection method similar to [21] can be employed in the outer loop. The ℓ-th component of the subgradient $\boldsymbol{\delta}=[\delta_1,\dots,\delta_L]^{\mathrm{T}}$ is $\delta_\ell=-P_\ell+\sum_{i=1}^{K}\boldsymbol{t}_i^{\mathrm{H}}\boldsymbol{A}_{i,\ell}\boldsymbol{t}_i$, where the $\boldsymbol{t}_i$'s are the beamformers from the previous iteration. Therewith, a subgradient projection step reads as

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

The projection shall w.l.o.g. be onto the simplex $\mathcal{C}=\big\{\boldsymbol{\mu}\in\mathbb{R}_+^L|\sum_{\ell=1}^{L}\mu_\ell=1\big\}$, since a scaling of $\boldsymbol{\mu}$ does not change the optimum of (15) and a_j denotes the step size in iteration j.

Another update rule, that is used in the literature, is (cf. [11])

$$\tilde{\mu}_\ell\leftarrow\frac{\mu_\ell^{(m)}}{P_\ell}\sum_{i=1}^{K}\boldsymbol{t}_i^{\mathrm{H}}\boldsymbol{A}_{i,\ell}\boldsymbol{t}_i,\quad \mu_\ell^{(m+1)}\leftarrow\frac{\tilde{\mu}_\ell}{\sum_{\ell=1}^{L}\tilde{\mu}_\ell}, \tag{29}$$

which ensures strong duality in the convergence point, i.e, $\mu_\ell\big(P_\ell-\sum_{i=1}^{K}\boldsymbol{t}_i^{\mathrm{H}}\boldsymbol{A}_{i,\ell}\boldsymbol{t}_i\big)=0$. We remark that (29) increases/decreases those μ_ℓ that correspond to violated/satisfied

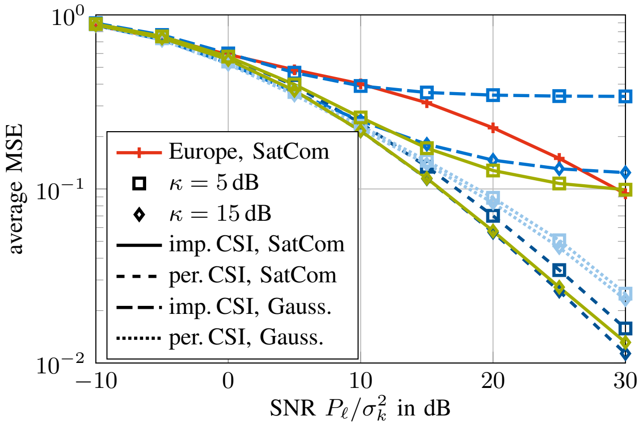

Figure 1: average MSE vs. SNR

Parameter	**Value**
satellite configuration	GEO; Ka-band; reuse 1
beamwidth $\theta_{3\text{dB}}$ (in degree)	0.2
number of beams (cluster 7/Europe)	7/128
log-normal fading $m_k/\sigma^2_{\xi_k}$ [22]	-3.06 dB/1.51 dB
max satellite antenna gain	52 dBi
max user antenna gain	40 dBi
base receive noise power; approx. FSL	-118 dBW; 210 dB
SNR P_ℓ/σ_k^2	-10,…,30 dB

Table I: Link budget Parameters in SatCom

power constraints, respectively, which is a necessary requirement for convergence to a local maximizer.

Both updates for $\boldsymbol{\mu}$ converged to the same MSEs in our simulations. However, the fixed-point search in (29) is less complex than the subgradient method (cf. [11]).

V. Numerical Results

We computed results for a standard Gaussian fading model and a SatCom model. For the standard fading model, the channel means $\bar{\boldsymbol{h}}_k$ are drawn from a standard Gaussian distribution and scaled to have the same norm as the satellite channel means, and the covariances are $\boldsymbol{C}_k = \frac{1}{N}\mathbf{I}_N$. The main parameters for the considered SatCom scenarios are shown in Table I. Per-antenna constraints are imposed, i.e., one antenna per feed. The users are randomly placed within the 3 dB area of the beams, i.e., $N = K$ and one user per spotbeam. The balanced average MSEs are calculated for a 7 cell system with 100 different user placements and one user realization for an 128 cell system that represents the coverage of Europe.

In Fig. 1, the (average) balanced MSEs are depicted vs. P_ℓ/σ_k^2. For perfect CSI, the MSEs decrease unbounded while the imperfect CSI curves (MSE bound) saturate. The higher the Rician factor κ, the lower the saturation level. For the SatCom channels with $\kappa = 15$ dB, the multi-path scattering may be neglected. No saturation is visible in the given SNR regime. The exemplary 128 cell curve (Europe) decreases slower than the 7 cell curves (MSE bound) for $\kappa = 15$ dB. While the antenna characteristics of the SatCom channel sufficiently separates the users in a 7 cell system, the 128 cell system apparently suffers from the increased interference.

Note that the curves for the satellite channel differ from those of the Gaussian channel model. For perfect CSI, the Gaussian channel model results only in a slightly worse average performance than for the SatCom model. However, the *MSE bound* curves of the Gaussian channel model saturate earlier than those for the SatCom model. This is a consequence of the SatCom beam gain characteristic, which deforms the channel mean and error covariance alike [see (5) and (6)].

References

[1] P.-D. Arapoglou, K. Liolis, M. Bertinelli, A. Panagopoulos, P. Cottis, and R. De Gaudenzi, "MIMO Over Satellite: A Review," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 1, pp. 27–51, May 2011.

[2] D. Christopoulos, S. Chatzinotas, G. Zheng, J. Grotz, and B. Ottersten, "Linear and Nonlinear Techniques for Multibeam Joint Processing in Satellite Communications," *EURASIP J. Wirel. Commun. and Netw.*, vol. 1, no. 162, May 2012.

[3] G. Zheng, S. Chatzinotas, and B. Ottersten, "Generic Optimization of Linear Precoding in Multibeam Satellite Systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2308–2320, June 2012.

[4] A. Tölli, M. Codreanu, and M. Juntti, "Linear Multiuser MIMO Transceiver Design With Quality of Service and Per-Antenna Power Constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3049–3055, July 2008.

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

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

[7] 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.

[8] M. Schubert and H. Boche, "A Generic Approach to QoS-Based Transceiver Optimization," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1557–1566, Aug. 2007.

[9] 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.

[10] 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.

[11] G. Dartmann, Xitao Gong, W. Afzal, and G. Ascheid, "On the Duality of the Max-Min Beamforming Problem With Per-Antenna and Per-Antenna-Array Power Constraints," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 606–619, Feb. 2013.

[12] 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.

[13] D. Christopoulos, S. Chatzinotas, M. Matthaiou, and B. Ottersten, "Capacity Analysis of Multibeam Joint Decoding Over Composite Satellite Channels," in *Proc. Asilomar 2011*, Nov. 2011, pp. 1795–1799.

[14] Andreas Gründinger, Michael Joham, and Wolfgang Utschick, "Bounds on Optimal Power Minimization and Rate Balancing in the Satellite Downlink," in *Proc. ICC 2012*, June 2012, pp. 3600–3605.

[15] C. Loo, "A Statistical Model for a Land Mobile Satellite Link," *IEEE Trans. Veh. Technol.*, vol. 34, no. 3, pp. 122–127, Aug 1985.

[16] M.A. Diaz, N. Courville, C. Mosquera, Gianluigi Liva, and G.E. Corazza, "Non-Linear Interference Mitigation for Broadband Multimedia Satellite Systems," in *Proc. IWSSC 2007*, Sept. 2007, pp. 61–65.

[17] R. Hunger and M. Joham, "A Complete Description of the QoS Feasibility Region in the Vector Broadcast Channel," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 3870–3878, Dec. 2010.

[18] 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.

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

[20] R. D. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sept. 1995.

[21] D.W.H. Cai, T.Q.S. Quek, Chee-Wei Tan, and S.H. Low, "Max-Min Weighted SINR in Coordinated Multicell MIMO Downlink," in *Proc. WiOpt 2011*, May 2011, pp. 286–293.

[22] S. Enserink and M.P. Fitz, "Constrained Capacities of DVB-S2 Constellations in Log-Normal Channels at Ka Band," in *Proc. SPACOMM, 2010*, June 2010, pp. 93–99.