A Complete Description of the QoS Feasibility Region in the Vector Broadcast Channel

Raphael Hunger, *Student Member, IEEE*, and Michael Joham, *Member, IEEE*

***Abstract*—We characterize the complete quality-of-service (QoS) feasibility region and present a simple feasibility test for given QoS requirements in the Gaussian vector broadcast channel. While most contributions in the literature recast the QoS constraints into requirements for the signal-to-interference-and-noise ratios (SINRs), we convert them into upper bounds for the minimum mean square errors (MMSEs) instead and test feasibility in the MMSE domain. Our main contribution is a complete description of the feasible MMSE region. Its closure is shown to be a polytope and we find the complete set of its bounding half-spaces noniteratively after a finite number of steps. The polytope can easily be converted into any other QoS domain like SINR or rate. However, the simple geometry of the MMSE domain is lost in other domains. Once the bounding half-spaces are determined, any target MMSE tuple can quickly be checked for feasibility by verifying its membership to the interior of the polytope. For nondegenerate channels, the only relevant bounding half-space is essentially the lower bound on the sum mean square error. No further computations are necessary contrary to existing feasibility checks which iteratively solve eigenproblems in an alternating optimization framework for every single QoS requirement to test. For two particular user/antenna configurations, we find a noniterative closed form solution for the optimum power allocation of the signal-to-interference ratio (SIR) balancing.**

***Index Terms*—Feasibility check, quality-of-service (QoS) region, vector broadcast channel.**

I. Introduction

POWER minimization under a set of quality-of-service (QoS) requirements is a frequently arising optimization problem, see, for example, [1]–[5]. Such QoS constraints are typically formulated as lower bounds on the signal-to-interference-and-noise ratios (SINRs), lower bounds on the data rate under Gaussian signaling, or as upper bounds on the minimum mean square error (MMSE). Likewise, any other metric which follows from a bijective mapping of the three aforementioned ones may be used. In the vector broadcast channel, where the base station is equipped with at least as many antennas as single antenna users are present in the system, any set of QoS requirements is feasible for full rank channels due to the existence of the zero-forcing filter. However, when more receivers shall be served than antennas are available at the transmitter, or when linearly dependent channels come into play, feasibility cannot be guaranteed for any set of QoS requirements. To ensure the convergence of any algorithm targeted at the power minimization, feasibility has to be detected first. Since the feasibility region was not known so far, one had to run an additional optimization to test the constraint set. In [2], infeasible QoS requirements are said to be reported by the second order cone program itself. A more clever test is to perform a balancing algorithm which fixes the ratios of the individual users' QoS metrics and then optimizes their common scalar. Feasibility holds if this scalar is larger than or equal to one. In contrast to the power minimization, the balancing algorithm is always feasible. Detecting feasibility by means of balancing with an additional constraint imposed on the maximum available transmit power was done in [3], [6], and [7]. The more general balancing without power limitation is treated in [8]–[13]. This signal-to-interference ratio (SIR) balancing delivers the separation plane between feasibility and infeasibility, when power is minimized for given SINR requirements.

Although the detection of feasibility by means of balancing works well, there are several disadvantages embedded in this methodology. First of all, the balancing algorithm has to be executed *for every single set* of SINR constraints. Assume the requirements have changed since the previous ones have turned out to be infeasible. Then, the complete balancing optimization must be solved again to find out whether the newly chosen constraints are valid or not. Second, distinguishing between feasibility and infeasibility becomes sensitive close to the separation plane. Remember that the decision criterion was that the common scalar of all SINR values, which represents the radius in polar coordinates, is either smaller than one indicating infeasibility or larger than or equal to one indicating feasibility. As the balancing algorithm is iterative, it is difficult to decide whether a common scalar close to but below one at any iteration step might grow beyond the critical value of one in further iterations or not. The same problem arises for the detection of feasibility by means of the second order cone program in [2]. A higher accuracy in knowing whether QoS requirements close to the feasibility boundary are feasible or not is directly connected to a larger complexity. Finally, little knowledge about the underlying geometrical structure of the feasibility region is gained by means of the balancing algorithm. In [6] and [13], the feasible region is defined by the set of SI(N)R tuples for which the spectral radius of the scaled interference coupling matrix is smaller than or equal to one. Only a few properties of the underlying geometric structure of the feasibility region can be deduced from this, see, for example, [14].

Manuscript received July 30, 2009; accepted March 07, 2010. Date of publication March 29, 2010; date of current version June 16, 2010. This work was supported in part by Deutsche Forschungsgemeinschaft (DFG) under Fund Ut 36/9-1. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Subhrakanti Dey.

The authors are with the Associate Institute for Signal Processing, Technische Universität München, Theresienstr. 90, D-80290 München, Germany (e-mail: hunger@tum.de; joham@tum.de).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2046895

Our contributions in this paper are as follows:

1) We show that the domain of choice for the description of the QoS feasibility region is the MMSE domain, where a simple description is available.
2) The closure of the feasible MMSE region is a polytope and any point belonging to the interior of this polytope is achievable with finite transmit power. Any point on the boundary of the polytope which is not the all-ones MMSE vector can only be achieved asymptotically with an infinite amount of power. This polytope in the MMSE domain can easily be translated to any other domain like SINR or rate by means of a bijective mapping.
3) The separating half-spaces of the polytope are shown to be simple inequalities for the sum of individual users' MMSEs. For nondegenerate, regular channels, only the sum MMSE of all users determines feasibility. Therefore, checking feasibility means to convert the QoS requirements by the bijective mapping into MMSE requirements and afterwards test their sum MMSE.
4) For the special case when at most as many users are in the system as the base station has antennas, we derive an upper bound on the sum power that is required to satisfy given QoS constraints.
5) In case of singular, linearly dependent channels, additional constraints on the sum MMSE of subsets of users also determine the polytope. We derive the complete geometric structure for this case as well.
6) For the two special cases when there is either one user more than antennas at the base, or when the base station is equipped with only a single antenna, we find a closed form expression for the asymptotically optimum power allocation which achieves the QoS requirements.
7) The presented MSE-based framework for single-antenna terminals also serves as the key technique when extending the feasible rate *region* to multi-antenna terminals, where it turns out that only single stream transmission per user is relevant for the feasibility check, see [15]. In contrast to the MAC *admission problem* considered in [16], an infinite amount of power is assumed for the feasibility test again.

Notation: Throughout this paper, matrices and vectors are denoted by upper and lower case bold letters, respectively. The operators $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, $\mathrm{rank}(\cdot)$, $\det(\cdot)$, and $\mathrm{E}(\cdot)$ stand for transposition, Hermitian transposition, rank, determinant of a matrix, and statistical expectation, respectively. The operator $\mathrm{int}(\cdot)$ returns the interior of a set. Finally, $\mathbf{I}_N$ and $\mathbf{1}$ denote the $N \times N$ identity matrix and the all-ones vector, respectively.

II. System Model

Given perfect channel state information at the transmitter, the vector broadcast channel (BC) and the dual vector multiple-access channel (MAC) share the same MSE region under a sum power constraint, e.g., [17] and [18]. Thus, we may describe the maximum feasible MMSE region of the BC in its dual MAC and feasibility can be detected in the dual MAC as well. The main advantage of the dual MAC is its simple description by the powers of the transmitting users, since the optimum receive beamformers can be computed independently and are known to minimize the MSE and maximize the receive SINR. Hence, the K user system is described by only K nonnegative real-valued scalars in the dual MAC, whereas K complex-valued N-dimensional vectors are needed in the BC for the transmit beamformers. Here, N denotes the number of antennas at the base station, and we define the set of user indices $\mathcal{K} := \{1, \ldots, K\}$.

The transmit powers of the K users in the dual MAC are denoted by $p_1, \ldots, p_K$ and the respective K data symbols $s_1, \ldots, s_K$ are zero mean with unit variance. A frequency flat channel is assumed for the transmission from the users to the base station and the channel matrix $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K] \in \mathbb{C}^{N \times K}$ contains the channel vectors of all K users. At the base station (receiver in the MAC), zero-mean Gaussian noise $\boldsymbol{\eta} \in \mathbb{C}^N$ with covariance matrix $\mathrm{E}[\boldsymbol{\eta}\boldsymbol{\eta}^{\mathrm{H}}] = \sigma^2 \mathbf{I}_N$ is added and the receive filter for user k is denoted by $\boldsymbol{g}_k^{\mathrm{T}} \in \mathbb{C}^{1 \times N}$. With above definition, the mean square error $\breve{\varepsilon}_k$ of user k is defined as the expected squared magnitude of the difference between the symbol s_k and its estimate, i.e.,

$$\breve{\varepsilon}_k := \mathrm{E}\left[\left| s_k - \boldsymbol{g}_k^{\mathrm{T}} \left(\sum_{\ell \in \mathcal{K}} \boldsymbol{h}_\ell \sqrt{p_\ell} s_\ell + \boldsymbol{\eta} \right) \right|^2\right]. \tag{1}$$

Minimizing above expression with respect to $\boldsymbol{g}_k$ leads to the *minimum* mean square error ε_k, for which the well-known relation to the *maximum* SINR γ_k holds via [19]

$$\varepsilon_k = \frac{1}{1 + \gamma_k}. \tag{2}$$

Note that the receive filter minimizing (1) also leads to the maximum SINR. Due to this bijective mapping between the minimum MSE and the maximum SINR, describing a system in terms of the MMSE instead of the maximum SINR is fully equivalent and does not entail any loss of information or interpretation.

III. Geometry of the Feasible MSE Region

Consider a two user scenario where the two channel vectors $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are colinear (which happens for example if the base station has only $N = 1$ antenna). It can easily be shown that the upper boundary of the maximum feasible SINR region is given by [11]

$$\gamma_1 \gamma_2 = 1 \tag{3}$$

where γ_k denotes the SINR of user k. This boundary is only asymptotically achieved when the ratios of both powers of the two users and the noise variance goes to infinity, i.e., when $p_1/\sigma^2 \to \infty$ and $p_2/\sigma^2 \to \infty$. A generalization to more than two users is not straightforward in the SINR domain and does not feature such a simple relation as in (3). For example, when a third user is added to the single-antenna system, the maximum feasible SINR region is governed by

$$2\gamma_1\gamma_2\gamma_3 + \gamma_1\gamma_2 + \gamma_1\gamma_3 + \gamma_2\gamma_3 < 1 \tag{4}$$

which will be derived in the next section, see (16).

In the following, we will derive the exact description of the feasible region in the MSE domain for an arbitrary number of users. In Section III-A, we assume the regular case in which the channel vectors of any user subset $\mathcal{I} \subseteq \mathcal{K}$ with cardinality $1 \leq |\mathcal{I}| \leq K$ satisfy the rank relation

$$\mathrm{rank}(\boldsymbol{H}_{\mathcal{I}}) = \min\{|\mathcal{I}|, N\}, \quad \forall \mathcal{I} \subseteq \mathcal{K} \tag{5}$$

whereas the general (singular) case with linearly dependent channel vectors allowed is discussed afterwards in Section III-B. In (5), the matrix $\boldsymbol{H}_{\mathcal{I}}$ consists of the channel vectors of all users belonging to $\mathcal{I}$.

A. Regular Channels Scenario

Under the reasonable assumption of the channel properties defined in (5), we will first state and afterwards prove the following theorem:

Theorem III.1: The closure of the feasible MMSE region in the vector broadcast channel with regular full rank channels satisfying (5) is a polytope $\mathcal{P} = \{\boldsymbol{\varepsilon} | 0 \leq \varepsilon_k \leq 1 \ \forall k \in \mathcal{K} \wedge \sum_{k \in \mathcal{K}} \varepsilon_k \geq K - N\}$ whose bounding half-spaces are the individual box constraints $0 \leq \varepsilon_k \leq 1 \ \forall k \in \mathcal{K}$ and the sum MMSE constraint $\sum_{k \in \mathcal{K}} \varepsilon_k \geq K - N$. By means of a positive power allocation with finite sum power, any point belonging to the interior of the polytope can be achieved. For MMSEs equal to one no power is allocated to the respective user.

From the above theorem, we directly conclude the next corollary, which results from the existence of the zero-forcing filter in the considered case.

Corollary III.2: In the K-user vector broadcast channel obeying (5) with an $N \geq K$ antenna base station, arbitrary QoS requirements satisfying $0 < \varepsilon_k \leq 1 \ \forall k$ are feasible with finite sum power.

For the proof of Theorem III.1, we define $\boldsymbol{P} = \mathrm{diag}\{p_k\}_{k=1}^{K}$ as the diagonal matrix containing the nonnegative powers of all users in the set $\mathcal{K}$. Then, the MMSE receive filter for user k that minimizes (1) reads as

$$\boldsymbol{g}_k^{\mathrm{T}} = \sqrt{p_k} \boldsymbol{h}_k^{\mathrm{H}} (\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}} + \sigma^2 \mathbf{I}_N)^{-1} \tag{6}$$

and achieves the MMSE

$$\begin{aligned} \varepsilon_k &= 1 - p_k \boldsymbol{h}_k^{\mathrm{H}} (\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}} + \sigma^2 \mathbf{I}_N)^{-1} \boldsymbol{h}_k \\ &= \left[\left(\mathbf{I}_K + \sigma^{-2} \boldsymbol{P}^{\frac{1}{2}} \boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \boldsymbol{P}^{\frac{1}{2}} \right)^{-1} \right]_{k,k}. \end{aligned} \tag{7}$$

The second line can be obtained by applying the matrix inversion lemma. From the first line of (7), we can observe that applying MMSE receive filters according to (6) automatically leads to the box constraints

$$0 \leq \varepsilon_k \leq 1 \quad \forall k \in \mathcal{K} \tag{8}$$

which contribute to the polytope. A necessary condition for a particular user i to achieve the lower bound $\varepsilon_i = 0$ asymptotically with equivalence is that his power goes to infinity, i.e., $p_i \to \infty$. Given an antenna configuration with $N < K$, not all MMSEs can be chosen arbitrarily small *simultaneously*. With the individual MMSEs from (7), we can express the sum MMSE for arbitrary N via (see, for example, [17] and [18])

$$\sum_{k \in \mathcal{K}} \varepsilon_k = K - N + \mathrm{tr}\left[(\mathbf{I}_N + \sigma^{-2} \boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^{\mathrm{H}})^{-1} \right] \tag{9}$$

resulting from the first line of (7). As the inverse matrix in (9) is positive definite, its trace is lower bounded by $N - \mathrm{rank}(\boldsymbol{H})$. Thus, any power allocation with *finite* sum power satisfies

$$\sum_{k \in \mathcal{K}} \varepsilon_k > K - \mathrm{rank}(\boldsymbol{H}) \quad \text{for} \quad \|\boldsymbol{p}\|_1 < \infty \tag{10}$$

for nonzero variance $\sigma^2 > 0$ *with strict inequality*. Letting *all*[1] powers $p_1, \ldots, p_K$ raise to infinity, i.e., $p_k \to \infty \ \forall k \in \mathcal{K}$, we asymptotically achieve the lower bound in (10) *with equality*

$$\sum_{k \in \mathcal{K}} \varepsilon_k = K - \mathrm{rank}(\boldsymbol{H}) \quad \text{for} \quad p_k \to \infty, \quad \forall k \in \mathcal{K}. \tag{11}$$

Under the regular channels assumption (5), the rank of the channel matrix is $\mathrm{rank}(\boldsymbol{H}) = \min\{N, K\}$, and the antenna configuration $N \geq K$ does not entail any limitations on the sum MSE since the resulting lower bound $\sum_{k \in \mathcal{K}} \varepsilon_k \geq 0$ from (11) is already included in the box constraints (8). The feasible MMSE region is then completely described by (8). Therefore, the case $N \geq K$ with regular channels allows for arbitrary QoS requirements, which nevertheless have to meet (8). In particular, any nonnegative SINR tuple is feasible since the maximum SINR γ_k and the minimum MSE ε_k are related via [19]

$$\gamma_k = \frac{1}{\varepsilon_k} - 1 \tag{12}$$

and ε_k can be made arbitrarily close to zero. However, when $K > N$, the polytope defined by (8) is shortened by the lower bound on the sum MMSE in (11), which in conjunction with $\mathrm{rank}(\boldsymbol{H}) = \min\{N, K\} = N$ simplifies to

$$\sum_{k \in \mathcal{K}} \varepsilon_k > K - N \quad \text{for} \quad \|\boldsymbol{p}\|_1 < \infty. \tag{13}$$

This obviously limits the set of feasible QoS requirements. Note that the lower bound $K - N$ in (13) can be achieved with equality when $p_k \to \infty \ \forall k \in \mathcal{K}$.

So far, we have proven that any positive power allocation with finite sum power achieves an MMSE tuple inside the polytope, see (13) and (8) with strict inequality for finite sum power. To complete the proof for Theorem III.1, we also have to show the converse, namely that there exists a power allocation $\boldsymbol{p}$ for any desired MMSE tuple belonging to $\mathrm{int}(\mathcal{P})$. Because only then, the mapping from the powers to the MMSEs in (7) is surjective in the interior of the polytope $\mathcal{P}$ defined in Theorem III.1. The proof goes as follows: First, we rewrite the mapping from the MMSE tuple $\varepsilon_1, \ldots, \varepsilon_K$ to the powers $\boldsymbol{p}$ as a fixed point equation at the optimum power allocation. Since the arising function in the fixed point equation is increasing and concave, there is at most one fixed point according to [20], [21]. If the fixed point does exist, the mapping from powers to MMSEs is then not only surjective, but also injective and thus bijective. For any target MMSE tuple for which a fixed point exists, there is a power allocation $\boldsymbol{p}$ obtaining it. Fortunately, Kennan also defines some sufficient conditions for the *existence* of a fixed point in [20]. We show that these conditions are met if and only if the

[1]For $K > N$, it suffices that at least N powers go to infinity.

target MMSE tuple is taken from the interior of the polytope $\mathcal{P}$, which will then complete the proof of Theorem III.1. The rather technical derivation showing the existence of a power allocation vector $\boldsymbol{p}$ resulting in the desired MMSE tuple can be found in Appendix A.

With the result from Theorem III.1, we can infer that the smallest *common* MMSE $\bar{\varepsilon} := \varepsilon_1 = \ldots = \varepsilon_K$ reads as

$$\bar{\varepsilon} = \begin{cases} \frac{K-N}{K} = 1 - \frac{N}{K}, & \text{for } K \geq N \\ 0, & \text{for } K \leq N \end{cases}$$

and leads via (12) to the largest common SIR

$$\bar{\gamma} = \frac{1}{\bar{\varepsilon}} - 1 = \begin{cases} \frac{N}{K-N}, & \text{for } K \geq N \\ \infty, & \text{for } K \leq N \end{cases}$$

which has already been observed in [19] in the CDMA context and was later shown in [4].

For the case $K \leq N$, a simple upper bound on the required sum power to achieve the target MMSEs $\boldsymbol{\varepsilon}^{\text{target}}$ can be stated:

Theorem III.3: The sum power $\|\boldsymbol{p}\|_1$ needed to satisfy positive target MMSEs $\boldsymbol{\varepsilon}^{\text{target}} > \mathbf{0}$ in the vector broadcast channel with $K \leq N$ users and regular channels defined in (5) is upper bounded by

$$\|\boldsymbol{p}\|_1 \leq \sigma^2 \sum_{k \in \mathcal{K}} \left(\frac{1}{\varepsilon_k^{\text{target}}} - 1 \right) \left[\left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \right]_{k,k} \tag{14}$$

or, equivalently, by

$$\|\boldsymbol{p}\|_1 \leq \sigma^2 \sum_{k \in \mathcal{K}} \gamma_k^{\text{target}} \left[\left(\boldsymbol{H}^{\mathrm{H}} \boldsymbol{H} \right)^{-1} \right]_{k,k}. \tag{15}$$

The proof immediately follows from (41) and (34). Note that the upper bound in (15) corresponds to the sum power that is needed to achieve the given SINR requirements when zero-forcing receive filters are assumed in the dual MAC.

Examples: In the single antenna case mentioned in (3), the SINRs γ_1 and γ_2 of the two users have to satisfy $\gamma_1\gamma_2 < 1$ in case of finite sum power, as shown in [11]. Using Theorem III.1, we are now able to extend the feasible SINR region to the case of $K = 3$ users and $N = 1$ antenna at the base station. The bijective mapping (12) converts the condition $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 > K - N = 2$ into

$$2\gamma_1\gamma_2\gamma_3 + \gamma_1\gamma_2 + \gamma_1\gamma_3 + \gamma_2\gamma_3 < 1. \tag{16}$$

As shown in Fig. 1, the closure of the feasible MMSE region simplifies to a pyramid for the given values $K = 3$ and $N = 1$, and the feasible SINR tuples are visualized in Fig. 2. Adding a second antenna to the base station, the MMSEs have to satisfy $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 > K - N = 1$, see Fig. 3. In the SINR domain

$$\gamma_1\gamma_2\gamma_3 - (\gamma_1 + \gamma_2 + \gamma_3) < 2$$

has to hold for finite sum power, see Fig. 4. For $N = 2$, any positive SINR pair is feasible when the third SINR goes to zero.

B. Singular Channels Scenario

We have seen in Theorem III.1 that in the case of regular channels satisfying (5), the closure of the feasible MMSE region is a polytope. Besides the obvious MMSE box constraints

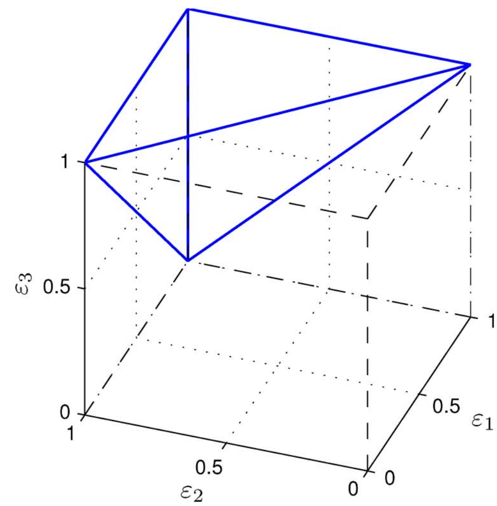

Fig. 1. Polytope structure of the closure of the feasible MMSE region for $K = 3$ users and an $N = 1$ antenna base station.

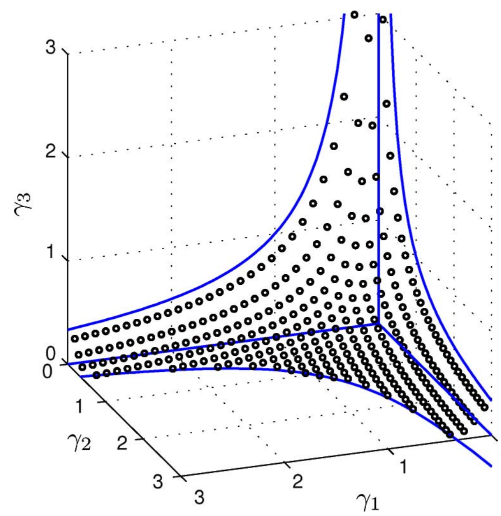

Fig. 2. Separating manifold in the SINR domain for $K = 3$ users and an $N = 1$ antenna base station.

$0 \leq \varepsilon_k \leq 1 \; \forall k$, the relevant boundary of the polytope follows from the half-space constraint $\sum_{k \in \mathcal{K}} \varepsilon_k \geq K - N$ when $K > N$. While probably all realistic channels fulfill (5), we drop this linear independence constraint in this section and allow for channels with arbitrary linear dependencies.

Fortunately, the basic structure of the feasible MSE region does not change much as the polytope type structure remains. In particular, the box constraints in (8) and the sum MMSE half-space constraints (10) and (11) remain valid. However, additional half-space constraints for the sum MMSE of subsets $\mathcal{I} \subset \mathcal{K}$ furthermore diminish the set of feasible MMSE tuples. For every subset $\mathcal{I} \subset \mathcal{K}$ for which

$$\operatorname{rank}(\boldsymbol{H}_{\mathcal{I}}) < \min\{N, |\mathcal{I}|\} \tag{17}$$

holds, i.e., for which the rank of the respective reduced channel matrix is smaller than the minimum of the antennas N and the cardinality $|\mathcal{I}|$ of the subset, an additional constraint on the sum MMSE of all users in $\mathcal{I}$ is imposed. This leads us to the following theorem.

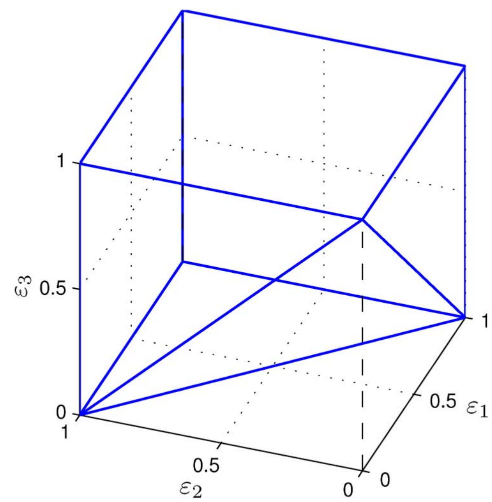

Fig. 3. Polytope structure of the closure of the feasible MMSE region for $K = 3$ users and an $N = 2$ antenna base station.

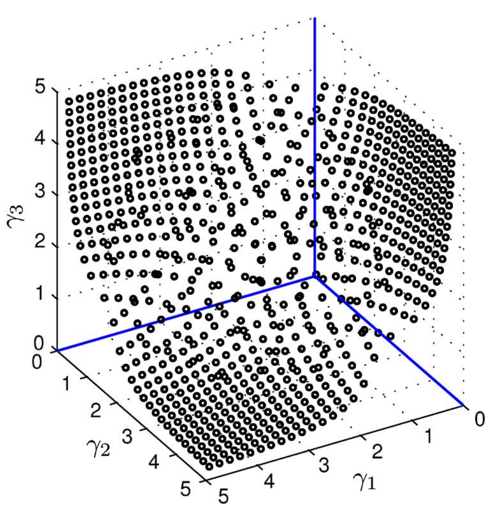

Fig. 4. Separating manifold in the SINR domain for $K = 3$ users and an $N = 2$ antenna base station.

Theorem III.4: The closure of the feasible MMSE region in the vector broadcast channel is a polytope $\mathcal{P}$ whose bounding half-spaces are the individual box constraints $0 \leq \varepsilon_k \leq 1\ \forall k \in \mathcal{K}$, the sum MMSE constraint $\sum_{k\in\mathcal{K}} \varepsilon_k \geq K - \mathrm{rank}(\boldsymbol{H})$, and additional constraints for the sum MMSE of user subsets whose channels are not full rank. Strictly speaking, a half-space constraint $\sum_{k\in\mathcal{I}} \varepsilon_k \geq |\mathcal{I}| - \mathrm{rank}(\boldsymbol{H}_{\mathcal{I}})$ is imposed for every subset $\mathcal{I} \subset \mathcal{K}$ for which $\mathrm{rank}(\boldsymbol{H}_{\mathcal{I}}) < \min\{N, |\mathcal{I}|\}$.

Note that some of these constraints might be redundant since they are already included in others, or they may even render the sum MMSE constraint over all users superfluous because they are stricter than it.

The proof for Theorem III.4 goes as follows: For any set $\mathcal{I}$ that fulfills (17), we assume that all users not belonging to $\mathcal{I}$ do not transmit anything at all, i.e., $p_k = 0\ \forall k \in \mathcal{K} \setminus \mathcal{I}$, whereas all users in $\mathcal{I}$ are active transmitters. This setup can then be interpreted as a $|\mathcal{I}|$-user scenario which of course also has to satisfy (10) and (11) with $\mathcal{K}$ and $\boldsymbol{H}$ replaced by $\mathcal{I}$ and $\boldsymbol{H}_{\mathcal{I}}$, respectively

$$\begin{aligned} &\sum_{k\in\mathcal{I}} \varepsilon_k > |\mathcal{I}| - \mathrm{rank}(\boldsymbol{H}_{\mathcal{I}}) \quad \text{for} \quad \|\boldsymbol{p}\|_1 < \infty \\ &\sum_{k\in\mathcal{I}} \varepsilon_k = |\mathcal{I}| - \mathrm{rank}(\boldsymbol{H}_{\mathcal{I}}) \quad \text{for} \quad p_k \to \infty \quad \forall k \in \mathcal{I}. \end{aligned} \tag{18}$$

Obviously, only those sets $\mathcal{I}$ have to be taken into account for which $\mathrm{rank}(\boldsymbol{H}_{\mathcal{I}}) < \min\{|\mathcal{I}|, N\}$ holds, i.e., for which (17) is valid. If $\mathrm{rank}(\boldsymbol{H}_{\mathcal{I}}) = \min\{|\mathcal{I}|, N\} = |\mathcal{I}|$, (18) would imply that $\sum_{k\in\mathcal{I}} \varepsilon_k \geq 0$, which is already included in the box constraints (8) and hence superfluous. The second case $\mathrm{rank}(\boldsymbol{H}_{\mathcal{I}}) = \min\{|\mathcal{I}|, N\} = N$ does not lead to an additional constraint on the sum MMSE of all users in $\mathcal{I}$ because the sum MMSE constraint $\sum_{k\in\mathcal{K}} \varepsilon_k \geq K - \mathrm{rank}(\boldsymbol{H})$ of *all* users is stricter and therefore includes it. Hence, let us assume that the inequality in (17) is strict. Letting the previously passive users in $\mathcal{K} \setminus \mathcal{I}$ actively transmit with constant positive power does not have any impact on the validity of (18) since their interference can be regarded as colored noise. However, colored noise does not change the asymptotic behavior in the second line of (18) since $p_k \to \infty\ \forall k \in \mathcal{I}$ is implied there. This completes the proof of Theorem III.4.

Examples: Assume a three user scenario and a base station with $N = 2$ antennas. The two channel vectors $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ shall be linearly dependent, and $\mathrm{rank}(\boldsymbol{H}) = 2$ is assumed implying that $\boldsymbol{h}_3$ is linearly independent of $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$. Besides the box constraints (8), the half-space defined by the lower bound on the sum MMSE reads as $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \geq K - \mathrm{rank}(\boldsymbol{H}) = 1$. In addition, the subset $\mathcal{I} = \{1, 2\}$ satisfies (17). So another half-space constraint contributes to the feasible MMSE region polytope, namely, $\varepsilon_1 + \varepsilon_2 \geq |\mathcal{I}| - \mathrm{rank}(\boldsymbol{H}_{\mathcal{I}}) = 1$. Note that this constraint is stricter than the sum MMSE constraint of all users, since $\varepsilon_1 + \varepsilon_2 \geq 1$ clearly implies $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \geq 1$ for $\varepsilon_3 \geq 0$, which is required by (8). Summing up, the closure of the feasible MMSE region for this particular channel is given by (8) and $\varepsilon_1 + \varepsilon_2 \geq 1$.

In a second example, $K = 5$ users are served by an $N = 3$ antenna base station. The three channel vectors $\boldsymbol{h}_1$, $\boldsymbol{h}_2$, and $\boldsymbol{h}_3$ shall lie in a two-dimensional plane, however, every channel vector pair out of those three vectors shall be linearly independent. In addition, no further linear dependencies shall exist. Then, $\mathrm{rank}(\boldsymbol{H}) = 3$ implies the sum MMSE constraint $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4 + \varepsilon_5 \geq 2$. The only user set satisfying (17) is $\mathcal{I} = \{1, 2, 3\}$, and this leads to the additional condition $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \geq 1$.

IV. Optimum Power Allocation

In this section, we consider the two cases when there is exactly one more user than antennas at the base station, i.e., $K = N + 1$, and when the base station has only a single antenna and more than one user shall be served, i.e., $N = 1$ and $K > 1$. For those two cases, we derive the optimum power allocation for the SIR balancing, With these power ratios, one could initialize an iterative SINR balancing algorithm which includes a sum power constraint and nonzero noise variance. In principle, one could modify the fixed point algorithm in [21] by setting $\sigma^2 = 0$ in

(34) and normalizing all updated powers after the fixed point iteration such that their sum equals some iteration-index independent chosen constant. This way, the (modified) fixed point algorithm is not only applicable for SINR balancing, but also for SIR balancing, if all target SIRs are chosen such that their MMSEs sum up to $K-N$. Of course, the right hand side in (34) does not satisfy the scalability-axiom for $\sigma^2=0$ (it is rather scale invariant) such that the framework derived in [21] cannot be used to prove its convergence. However, rescaling the powers after the fixed point update led to quick convergence of the balancing powers in all our simulations. A more sophisticated approach is presented in [6], [9], and [13], which features superlinear convergence. However, all hitherto existing algorithms are iterative, either they solve the fixed point equation by means of an afterwards rescaled Picard iteration as in [21], or they alternate between the power allocation for fixed receivers and the receiver optimization for fixed powers until convergence [6], [9], [13]. Thereby, the power allocation step which is run *once per iteration* requires to find the principal eigenvector of a matrix belonging to the largest, but unknown eigenvalue. We show that the optimum power allocation has a very simple closed form solution for the two given cases when we take the optimum receivers for a given power allocation directly into account. An iterative algorithm is not necessary, and no principal eigenvectors or matrix inverses have to be computed. Note that we assume regular channels according to (5).

The SIR balancing naturally evolves from the SINR balancing, when we either let the noise variance σ^2 go down to zero or allocate an infinite amount of power. For $\sigma^2=0$, we assume a power allocation $\boldsymbol{p}=\boldsymbol{p}_0$, whereas for $\sigma^2>0$, we set $\boldsymbol{p}=\alpha\boldsymbol{p}_0$ with $\alpha\to\infty$ and $\boldsymbol{p}_0$ taken from the interior of the unit simplex $\mathcal{S}$ in (43). In either case, the *asymptotically* achieved mean square error from (7) can be expressed as the quotient of two homogeneous polynomials when we apply the *Cramer* rule to the second line in (7) and drop all monomials having less than N power variables to obtain the asymptotic result. Thus, we first express the MMSE ε_k of user k as the kth entry of $\boldsymbol{x}_k$, i.e., as the quotient of two determinants according to the *Cramer* rule, where $\boldsymbol{x}_k$ satisfies [see (7)]

$$\boldsymbol{e}_k=\left[\mathbf{I}_K+\sigma^{-2}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{P}^{\frac{1}{2}}\right]\boldsymbol{x}_k \tag{19}$$

and $\boldsymbol{e}_k$ is the kth canonical unit vector. Afterwards, we drop the monomials containing less than N power variables to obtain the asymptotic behavior induced by either $\sigma^2\to 0$ or $p_i\to\infty\ \forall i\in\mathcal{K}$. This yields

$$\varepsilon_k=\frac{\sum_{1\le i_1<\ldots<i_N\le K, i_\bullet\neq k} c_{i_1,\ldots,i_N}p_{0,i_1}\cdot\cdots\cdot p_{0,i_N}}{\sum_{1\le i_1<\ldots<i_N\le K} c_{i_1,\ldots,i_N}p_{0,i_1}\cdot\cdots\cdot p_{0,i_N}}. \tag{20}$$

Note that the numerator of ε_k consists of $\binom{K-1}{N}$ monomials of degree N, whereas the denominator has $\binom{K}{N}$ monomials of degree N each. Moreover, $c_{i_1,\ldots,i_N}$ is the determinant of $\boldsymbol{H}_{\mathcal{I}}^{\mathrm{H}}\boldsymbol{H}_{\mathcal{I}}$ with the set $\mathcal{I}=\{i_1,\ldots,i_N\}$. For example, if $K=3$ and $N=2$, the MMSE of the first user reads as

$$\varepsilon_1=\frac{c_{2,3}p_{0,2}p_{0,3}}{c_{1,2}p_{0,1}p_{0,2}+c_{1,3}p_{0,1}p_{0,3}+c_{2,3}p_{0,2}p_{0,3}}.$$

While (20) is valid for any user/antenna configuration with $K>N$, it can be simplified for the two special cases we focus on, namely, $N=1$ and $K=N+1$.

A. Single Antenna Base Station

For $N=1$, the MSE in (20) reads as

$$\varepsilon_k=\frac{\sum_{\ell\neq k}c_\ell p_{0,\ell}}{\sum_{\ell\in\mathcal{K}}c_\ell p_{0,\ell}}=1-\frac{c_k p_{0,k}}{\sum_{\ell\in\mathcal{K}}c_\ell p_{0,\ell}} \tag{21}$$

which leads to the following linear system of equations:

$$c_k p_{0,k}=\frac{1-\varepsilon_k}{\varepsilon_k}\sum_{\ell\neq k}c_\ell p_{0,\ell}\quad\forall k\in\mathcal{K}. \tag{22}$$

Defining the vector $\boldsymbol{c}:=[c_1,\ldots,c_K]^{\mathrm{T}}\in\mathbb{R}_+^K$, and the diagonal matrices $\boldsymbol{C}:=\operatorname{diag}\{c_k\}_{k=1}^K$ and $\boldsymbol{E}:=\operatorname{diag}\{\varepsilon_k\}_{k=1}^K$, (22) can be expressed as

$$(\boldsymbol{E}^{-1}-\mathbf{I})\boldsymbol{C}^{-1}(\mathbf{1}\boldsymbol{c}^{\mathrm{T}}-\boldsymbol{C})\boldsymbol{p}_0=\boldsymbol{p}_0. \tag{23}$$

Note that the matrix $(\boldsymbol{E}^{-1}-\mathbf{I})\boldsymbol{C}^{-1}(\mathbf{1}\boldsymbol{c}^{\mathrm{T}}-\boldsymbol{C})$ has only nonnegative entries, if $0<\varepsilon_k<1\ \forall k$. In addition, it is irreducible for $c_k>0\ \forall k$, which holds according to our assumption in (5). Thus, the Perron–Frobenius theory can be used to determine the unique positive vector $\boldsymbol{p}_0$. In the following, we show that the matrix in (23) has the spectral radius one, if and only if $\sum_{k\in\mathcal{K}}\varepsilon_k=K-N=K-1$. According to the Perron–Frobenius theory (e.g., [22]), the spectral radius is the only positive eigenvalue and there exists a strictly positive vector $\boldsymbol{p}_0$ satisfying (23).

To see that the matrix $\boldsymbol{Z}:=(\boldsymbol{E}^{-1}-\mathbf{I})\boldsymbol{C}^{-1}(\mathbf{1}\boldsymbol{c}^{\mathrm{T}}-\boldsymbol{C})$ has an eigenvalue which is one, we show that

$$\det[\boldsymbol{Z}-\mathbf{I}]=0 \tag{24}$$

when $\sum_{k\in\mathcal{K}}\varepsilon_k=K-1$. Using $\det(\mathbf{I}+\boldsymbol{A}\boldsymbol{B})=\det(\mathbf{I}+\boldsymbol{B}\boldsymbol{A})$ for any pair of rectangular matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of matching dimensions, we find

$$\begin{aligned}\det[\boldsymbol{Z}-\mathbf{I}]&=\det\left[(\boldsymbol{E}^{-1}-\mathbf{I})\boldsymbol{C}^{-1}\mathbf{1}\boldsymbol{c}^{\mathrm{T}}-\boldsymbol{E}^{-1}\right]\\&=\det\left[(\boldsymbol{E}^{-1}-\mathbf{I})\boldsymbol{C}^{-1}\mathbf{1}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{E}-\mathbf{I}\right]\det[\boldsymbol{E}^{-1}]\\&=\frac{(-1)^K}{\prod_{k\in\mathcal{K}}\varepsilon_k}\det\left[\mathbf{I}-(\boldsymbol{E}^{-1}-\mathbf{I})\boldsymbol{C}^{-1}\mathbf{1}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{E}\right]\\&=\frac{(-1)^K}{\prod_{k\in\mathcal{K}}\varepsilon_k}\left[1-\boldsymbol{c}^{\mathrm{T}}(\mathbf{I}-\boldsymbol{E})\boldsymbol{C}^{-1}\mathbf{1}\right]\\&=\frac{(-1)^K}{\prod_{k\in\mathcal{K}}\varepsilon_k}\left[1-\mathbf{1}^{\mathrm{T}}(\mathbf{I}-\boldsymbol{E})\mathbf{1}\right]\\&=\frac{(-1)^K}{\prod_{k\in\mathcal{K}}\varepsilon_k}\left[\sum_{k\in\mathcal{K}}\varepsilon_k-(K-1)\right]\end{aligned} \tag{25}$$

which vanishes if $\sum_{k\in\mathcal{K}}\varepsilon_k=K-1$. Therefore, the relative power allocation $\boldsymbol{p}_0$ is the principal eigenvector belonging to the eigenvalue 1 of the matrix $\boldsymbol{Z}$. Note that the geometric multiplicity of the eigenvalue 1 is equal to one since the arithmetic multiplicity is also one according to the Perron–Frobenius theory, see [22]. Since $\boldsymbol{Z}-\mathbf{I}=(\boldsymbol{E}^{-1}-\mathbf{I})\boldsymbol{C}^{-1}\mathbf{1}\boldsymbol{c}^{\mathrm{T}}-\boldsymbol{E}^{-1}$ is a rank-one perturbation of the diagonal matrix $-\boldsymbol{E}^{-1}$, the principal eigenvector $\boldsymbol{p}_0$ has a closed form expression

$$(\boldsymbol{Z}-\mathbf{I})\boldsymbol{p}_0=\mathbf{0}\Leftrightarrow \boldsymbol{p}_0=g(\mathbf{I}-\boldsymbol{E})\boldsymbol{C}^{-1}\mathbf{1}. \tag{26}$$

The positive real-valued scaling factor $g\in\mathbf{R}_+$ can in general be chosen arbitrarily. For $g=1/\mathbf{1}^{\mathrm{T}}(\mathbf{I}-\boldsymbol{E})\boldsymbol{C}^{-1}\mathbf{1}$, we get $\boldsymbol{p}_0\in\operatorname{int}(\mathcal{S})$. Rewriting (26) element-wise yields

$$p_{0,k}=g\frac{1-\varepsilon_k}{c_k}\quad \forall k\in\mathcal{K}, \tag{27}$$

which of course only holds for $\sum_{k\in\mathcal{K}}\varepsilon_k=K-1$, cf. (25).

B. One User More Than Antennas at the Base Station

When $K=N+1$, the numerator in (20) consists of a single monomial only, whereas the denominator has K monomials. Although all arising monomials have degree N, we can convert the quotient of the two homogeneous polynomials of degree N each into a rational fraction, where both numerator and denominator are linear in the *inverse* powers. This is done via dividing numerator and denominator of (20) by $\prod_{k\in\mathcal{K}}p_k$. Thus, we can solve the $K=N+1$ case again by means of an eigenvalue problem in the inverse powers and obtain

$$\varepsilon_k=\frac{d_k t_{0,k}}{\sum_{\ell\in\mathcal{K}}d_\ell t_{0,\ell}} \tag{28}$$

where $t_{0,k}:=1/p_{0,k}\ \forall k$ and $d_k:=c_{1,\dots,k-1,k+1,\dots,K}$ is the determinant $\det(\boldsymbol{H}_{\mathcal{I}_k}^{\mathrm{H}}\boldsymbol{H}_{\mathcal{I}_k})$ with $\mathcal{I}_k=\mathcal{K}\setminus\{k\}$, i.e., the determinant of the channel Gram where the kth column has been removed. Similar to (22), this implies the linear system of equations

$$d_k t_{0,k}=\frac{\varepsilon_k}{1-\varepsilon_k}\sum_{\ell\neq k}d_\ell t_{0,\ell}. \tag{29}$$

Equivalently, we define $\boldsymbol{d}:=[d_1,\dots,d_K]^{\mathrm{T}}\in\mathbb{R}_+^K$ and the diagonal matrix $\boldsymbol{D}:=\operatorname{diag}\{d_k\}_{k=1}^K$ and reformulate (29) for all users as

$$\boldsymbol{E}(\mathbf{I}-\boldsymbol{E})^{-1}\boldsymbol{D}^{-1}(\mathbf{1}\boldsymbol{d}^{\mathrm{T}}-\boldsymbol{D})\boldsymbol{t}_0=\boldsymbol{t}_0, \tag{30}$$

where $\boldsymbol{t}_0=[t_{0,1},\dots,t_{0,K}]^{\mathrm{T}}$. With the same reasoning as in Section IV-A, the matrix $\boldsymbol{Y}:=\boldsymbol{E}(\mathbf{I}-\boldsymbol{E})^{-1}\boldsymbol{D}^{-1}(\mathbf{1}\boldsymbol{d}^{\mathrm{T}}-\boldsymbol{D})$ has spectral radius one, if and only if $\sum_{k\in\mathcal{K}}\varepsilon_k=N-K=1$

$$\begin{aligned}\det[\boldsymbol{Y}-\mathbf{I}]&=\det\left[\boldsymbol{E}(\mathbf{I}-\boldsymbol{E})^{-1}(\mathbf{1}\boldsymbol{d}^{\mathrm{T}}-\boldsymbol{D})\boldsymbol{D}^{-1}-\mathbf{I}\right]\\&=\frac{1}{\prod_{k\in\mathcal{K}}(1-\varepsilon_k)}\det\left[\boldsymbol{E}(\mathbf{1}\mathbf{1}^{\mathrm{T}}-\mathbf{I})-\mathbf{I}+\boldsymbol{E}\right]\\&=\frac{1}{\prod_{k\in\mathcal{K}}(1-\varepsilon_k)}\det[\boldsymbol{E}\mathbf{1}\mathbf{1}^{\mathrm{T}}-\mathbf{I}]\\&=\frac{(-1)^K}{\prod_{k\in\mathcal{K}}(1-\varepsilon_k)}[1-\mathbf{1}^{\mathrm{T}}\boldsymbol{E}\mathbf{1}]\\&=\frac{(-1)^K}{\prod_{k\in\mathcal{K}}(1-\varepsilon_k)}\left[1-\sum_{k\in\mathcal{K}}\varepsilon_k\right].\end{aligned} \tag{31}$$

The optimum relative inverse power allocation $\boldsymbol{t}_0$ therefore corresponds to the principal eigenvector of the matrix $\boldsymbol{Y}$ belonging to the eigenvalue 1. Again, $\boldsymbol{t}_0$ is positive according to the Perron–Frobenius theory and has a closed form solution due to the particular structure of $\boldsymbol{Y}$:

$$(\boldsymbol{Y}-\mathbf{I})\boldsymbol{t}_0=\mathbf{0}\Leftrightarrow \boldsymbol{t}_0=g'\boldsymbol{E}\boldsymbol{D}^{-1}\mathbf{1}. \tag{32}$$

The scaling $g'\in\mathbb{R}_+$ may be chosen arbitrarily and the particular choice $g'=1/\mathbf{1}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{D}^{-1}\mathbf{1}$ leads to $\boldsymbol{t}_0\in\operatorname{int}(\mathcal{S})$. The element-wise solutions for $\boldsymbol{t}_0$ and the relative power allocation $\boldsymbol{p}_0$ read as

$$t_{0,k}=g'\frac{\varepsilon_k}{d_k}\quad\text{and}\quad p_{0,k}=g'^{-1}\frac{d_k}{\varepsilon_k}\quad\forall k\in\mathcal{K} \tag{33}$$

and require $\sum_{k\in\mathcal{K}}\varepsilon_k=1$, cf. (31).

V. Conclusion

We derived a complete description of the feasible QoS region in the vector broadcast channel. Irrespective of the number of users in the system, the closure of the feasible MMSE region was shown to be a polytope. For regular, full rank channels, the plane that separates feasibility from infeasibility turned out to be the half-space defined by the lower bound on the sum MMSE of all users. Additional half-space conditions which furthermore reduce the polytope of feasible MMSE tuples arise for singular, rank-deficient channels. Nonetheless, even for singular channels, the polytope structure remained. As a by-product, our contribution allowed us to confirm the result that any QoS constraints can be satisfied when the base station has at least as many antennas as users are present in the system and the channel matrix is regular, which already follows from the existence of the zero-forcing filter. Besides feasibility, we considered the optimum power allocation for the signal-to-interference balancing problem for two special cases. First, when the base station has only a single antenna, and second, when there is one more user in the system than antennas at the base station. In both scenarios, we derived a closed-form solution for the power allocation, which hitherto had to be computed by means of iterative algorithms. Finally, we presented a simple upper bound on the sum power which is needed to obtain given QoS requirements when at most as many users are served as antennas are deployed at the base station.

Appendix A
Proof for the Existence of a Power Allocation Obtaining Certain MMSE Targets

Given MMSE targets $\varepsilon_k^{\mathrm{target}}$ with $k\in\mathcal{K}$, we assume that the power allocation $\check{\boldsymbol{p}}$ achieves these targets. Rewriting the first line of (7) with the help of the matrix inversion lemma leads to the well-known result

$$\varepsilon_k=\frac{1}{1+p_k\boldsymbol{h}_k^{\mathrm{H}}\left(\sigma^2\mathbf{I}+\sum_{\ell\neq k}\boldsymbol{h}_\ell\boldsymbol{h}_\ell^{\mathrm{H}}p_\ell\right)^{-1}\boldsymbol{h}_k}.$$

Equating $\varepsilon_k=\varepsilon_k^{\mathrm{target}}$ at the hypothetical fixed point $\check{\boldsymbol{p}}$, we obtain

$$\check{p}_k=f_k(\check{\boldsymbol{p}};\boldsymbol{\varepsilon}^{\mathrm{target}}):=\frac{\frac{1}{\varepsilon_k^{\mathrm{target}}}-1}{\boldsymbol{h}_k^{\mathrm{H}}\left(\sigma^2\mathbf{I}+\sum_{\ell\neq k}\boldsymbol{h}_\ell\boldsymbol{h}_\ell^{\mathrm{H}}\check{p}_\ell\right)^{-1}\boldsymbol{h}_k}, \tag{34}$$

where $f_k(\boldsymbol{p};\boldsymbol{\varepsilon}^{\text{target}})$ is well known from SINR balancing and satisfies the interference function properties defined in [21]. In particular, $f_k(\boldsymbol{p};\cdot)$ is positive, (quasi-)increasing, and the function $f_k(\boldsymbol{p};\cdot)-p_k$ is strictly radially quasiconcave for all k (see [20] for the definition). Hence, there is at most one fixed point according to [20, Corollary 1]. For the existence of a unique fixed point, Theorem 3 in [20] requires

$$\boldsymbol{f}(\boldsymbol{0};\boldsymbol{\varepsilon}^{\text{target}}) \geq \boldsymbol{0}, \tag{35}$$

$$\exists \boldsymbol{a} > \boldsymbol{0} \quad \text{with} \quad \boldsymbol{f}(\boldsymbol{a};\boldsymbol{\varepsilon}^{\text{target}}) > \boldsymbol{a}, \tag{36}$$

$$\exists \boldsymbol{b} > \boldsymbol{a} \quad \text{with} \quad \boldsymbol{f}(\boldsymbol{b};\boldsymbol{\varepsilon}^{\text{target}}) < \boldsymbol{b}. \tag{37}$$

The first requirement (35) immediately follows from (34):

$$f_k(\boldsymbol{0};\boldsymbol{\varepsilon}^{\text{target}}) = \sigma^2 \frac{\frac{1}{\varepsilon_k^{\text{target}}}-1}{\|\boldsymbol{h}_k\|_2^2} \geq 0 \quad \text{for} \quad 0 \leq \varepsilon_k^{\text{target}} \leq 1.$$

If there is a user whose target MMSE is equal to one, simply no power has to be allocated to that user and he can be discarded for the power computation of the remaining users. Hence, we assume $\varepsilon_k^{\text{target}} < 1\, \forall k \in \mathcal{K}$ in the following. For the second requirement (36), we choose $\boldsymbol{a} = a\mathbf{1}$ as the scaled all-ones vector. Using (34), we find

$$f_k(\boldsymbol{p};\boldsymbol{\varepsilon}^{\text{target}}) \geq \sigma^2 \frac{\frac{1}{\varepsilon_k^{\text{target}}}-1}{\|\boldsymbol{h}_k\|_2^2} \quad \text{for} \quad \boldsymbol{p} \geq \boldsymbol{0}$$

from which we find the upper bound $\bar{a}$ for a.

$$\bar{a} = \sigma^2 \min_{k\in\mathcal{K}} \frac{\frac{1}{\varepsilon_k^{\text{target}}}-1}{\|\boldsymbol{h}_k\|_2^2} \tag{38}$$

Note that $\bar{a} > 0$ for $\varepsilon_k^{\text{target}} < 1$. Thus, choosing $a < \bar{a}$ and setting $\boldsymbol{a} = a\mathbf{1}$ satisfies $\boldsymbol{f}(\boldsymbol{a};\boldsymbol{\varepsilon}^{\text{target}}) > \boldsymbol{a}$.

For the third requirement (37), we need to find a power vector $\boldsymbol{b}$ for which $\boldsymbol{f}(\boldsymbol{b};\boldsymbol{\varepsilon}^{\text{target}}) < \boldsymbol{b}$. The cases $K \leq N$ and $K > N$ will be treated separately. When $K \leq N$, the function $f_k(\boldsymbol{b};\boldsymbol{\varepsilon}^{\text{target}})$ can be upper bounded by lower bounding its denominator. Defining the set $\mathcal{I}_k := \mathcal{K} \setminus \{k\}$, we introduce

$$\mathbf{\Pi}_k = \mathbf{I}_N - \boldsymbol{H}_{\mathcal{I}_k}\left(\boldsymbol{H}_{\mathcal{I}_k}^{\text{H}}\boldsymbol{H}_{\mathcal{I}_k}\right)^{-1}\boldsymbol{H}_{\mathcal{I}_k}^{\text{H}}, \tag{39}$$

as the projector into the null-space of all channel vectors except the kth one that features $\operatorname{rank}(\mathbf{\Pi}_k) = N-K+1 \geq 1$ due to $K \leq N$. Herewith, we lower bound the denominator of $f_k(\boldsymbol{b};\boldsymbol{\varepsilon}^{\text{target}})$ via

$$\boldsymbol{h}_k^{\text{H}}\left(\sigma^2\mathbf{I} + \sum_{\ell\neq k}\boldsymbol{h}_\ell\boldsymbol{h}_\ell^{\text{H}}b_\ell\right)^{-1}\boldsymbol{h}_k \geq \sigma^{-2}\boldsymbol{h}_k^{\text{H}}\mathbf{\Pi}_k\boldsymbol{h}_k \tag{40}$$

which is valid for all $\boldsymbol{b} \geq \boldsymbol{0}$ and equality only holds for $b_k \to \infty\ \forall k \in \mathcal{I}_k$. Note that $\boldsymbol{h}_k^{\text{H}}\mathbf{\Pi}_k\boldsymbol{h}_k = 1/[(\boldsymbol{H}^{\text{H}}\boldsymbol{H})^{-1}]_{k,k}$ is larger than zero due to (5). Hence, an upper bound for $f_k(\boldsymbol{b};\boldsymbol{\varepsilon}^{\text{target}})$ is given by

$$f_k(\boldsymbol{b};\boldsymbol{\varepsilon}^{\text{target}}) \leq \sigma^2\left(\frac{1}{\varepsilon_k^{\text{target}}}-1\right)\left[\left(\boldsymbol{H}^{\text{H}}\boldsymbol{H}\right)^{-1}\right]_{k,k} \tag{41}$$

and when $K \leq N$ and if $\boldsymbol{b}$ is chosen such that

$$b_k > \sigma^2\left(\frac{1}{\varepsilon_k^{\text{target}}}-1\right)\left[\left(\boldsymbol{H}^{\text{H}}\boldsymbol{H}\right)^{-1}\right]_{k,k} \quad \forall k \in \mathcal{K} \tag{42}$$

the third requirement (37) is satisfied. Since $\boldsymbol{a} = a\mathbf{1}$ with $0 < a < \bar{a}$ and $\bar{a}$ defined in (38), choosing $\boldsymbol{b}$ according to (42) satisfies $\boldsymbol{b} > \boldsymbol{a}$, which is also required in (37). Since the three conditions (35)–(37) are fulfilled, Theorem III.1 is proven for $N \geq K$.

To show the existence of a power vector $\boldsymbol{b}$ in (37) when $K > N$ is slightly more complicated. First, we set the power allocation to $\boldsymbol{b} = \alpha\boldsymbol{b}_0$, where $\boldsymbol{b}_0$ is taken from the interior of the $K-1$ dimensional unit simplex

$$\mathcal{S} := \left\{\boldsymbol{x} \middle| \sum_{k\in\mathcal{K}} x_k = 1 \wedge x_k \geq 0 \quad \forall k\right\} \tag{43}$$

and α will later go to infinity. Since $\boldsymbol{b}_0 \in \operatorname{int}(\mathcal{S})$, the strict inequality $\boldsymbol{b}_0 > \boldsymbol{0}$ holds. For $\alpha \to \infty$, we may omit the scaled identity $\sigma^2\mathbf{I}$ in the denominator of $f_k(\boldsymbol{b};\boldsymbol{\varepsilon}^{\text{target}})$ in (34), as the matrix $\sum_{\ell\neq k}\boldsymbol{h}_k\boldsymbol{h}_k^{\text{H}}b_{0,\ell}\alpha$ has rank N for $K > N$ and all its eigenvalues grow beyond all limits when $\alpha \to \infty$ and $\boldsymbol{b}_0 > \boldsymbol{0}$. Then, the fixed point (34) can be rewritten as

$$b_{0,k} = f_k^\infty(\boldsymbol{b}_0;\boldsymbol{\varepsilon}^{\text{target}}) := \frac{\frac{1}{\varepsilon_k^{\text{target}}}-1}{\boldsymbol{h}_k^{\text{H}}\left(\sum_{\ell\neq k}\boldsymbol{h}_\ell\boldsymbol{h}_\ell^{\text{H}}b_{0,\ell}\right)^{-1}\boldsymbol{h}_k} \tag{44}$$

which means that $\boldsymbol{b}_0 - \boldsymbol{f}^\infty(\boldsymbol{b}_0;\boldsymbol{\varepsilon}^{\text{target}}) = \boldsymbol{0}$. Obviously, the MMSE tuple obtained with power allocation $\boldsymbol{b} = \alpha\boldsymbol{b}_0$ and $\alpha \to \infty$ satisfies $\sum_{k\in\mathcal{K}}\varepsilon_k = K-N$, see (11). Since (44) and (11) evolve from (7), the target MMSE tuple $\boldsymbol{\varepsilon}^{\text{target}}$ also has to satisfy $\|\boldsymbol{\varepsilon}^{\text{target}}\|_1 = K-N$ for $\boldsymbol{f}^\infty(\boldsymbol{b}_0;\boldsymbol{\varepsilon}^{\text{target}}) = \boldsymbol{b}_0$. So far we have shown that a power allocation $\boldsymbol{b} = \alpha\boldsymbol{b}_0$ with $\boldsymbol{b}_0 \in \operatorname{int}(\mathcal{S})$ and $\alpha \to \infty$ achieves an MMSE tuple which satisfies $\boldsymbol{\varepsilon} \in \operatorname{int}(\mathcal{B})$, where $\mathcal{B}$ defines the plane which separates feasibility from infeasibility

$$\mathcal{B} := \left\{\boldsymbol{\varepsilon} \middle| \sum_{k\in\mathcal{K}}\varepsilon_k = K-N \wedge 0 \leq \varepsilon_k \leq 1 \quad \forall k \in \mathcal{K}\right\}.$$

Now, we show the converse, i.e., that there always exists a unique $\boldsymbol{b}_0 \in \operatorname{int}(\mathcal{S})$ in the power allocation $\boldsymbol{b} = \alpha\boldsymbol{b}_0$ with $\alpha \to \infty$ for *any* MMSE tuple taken from $\boldsymbol{\varepsilon} \in \operatorname{int}(\mathcal{B})$. When $\alpha \to \infty$, the SINR metric reduces to the SIR without noise component. From the various SIR-balancing papers, see, for example, [9]–[13], we know that *any* ratio between individual SIRs can be balanced since balancing is always feasible. The main objective of the SIR balancing is to find the maximum common scalar r of all SIRs

$$\underset{r,\boldsymbol{b}_0}{\text{maximize}}\ r \quad \text{s.t.}\ \text{SIR}_k = r\cdot\text{SIR}_{0,k} \quad \forall k \in \mathcal{K} \tag{45}$$

with the SIR definition

$$\text{SIR}_k := b_{0,k}\boldsymbol{h}_k^{\text{H}}\left(\sum_{\ell\neq k}\boldsymbol{h}_\ell\boldsymbol{h}_\ell^{\text{H}}b_{0,\ell}\right)^{-1}\boldsymbol{h}_k.$$

Assuming $\mathrm{SIR}_{0,k} > 0\ \forall k$ and using (12), the SIR ratio constraints can be converted into MMSE constraints:

$$\mathrm{SIR}_k = r \cdot \mathrm{SIR}_{0,k} \Leftrightarrow \varepsilon_k = \frac{1}{r\left(\frac{1}{\varepsilon_{0,k}} - 1\right) + 1}, \tag{46}$$

where $\varepsilon_{0,k} = 1/(1 + \mathrm{SIR}_{0,k})$. Exploiting the fact that in case of infinite power allocation for every user, the sum MMSE is given by $K - N$ [see (11)], we find the radius r in the SIR domain. According to (46), it is easy to see that $\sum_{k\in\mathcal{K}} \varepsilon_k$ is decreasing in r and, therefore, $\sum_{k\in\mathcal{K}} \varepsilon_k = K - N$ has a unique solution for r. In particular, $r = 1$ if all $\mathrm{SIR}_{0,\cdot}$ are chosen such that $\sum_{k\in\mathcal{K}} 1/(1 + \mathrm{SIR}_{0,k}) = K - N$. This proves that for any target MMSE tuple $\boldsymbol{\varepsilon}^{\mathrm{target}}$ taken from $\mathrm{int}(\mathcal{B})$ and thus satisfying $\|\boldsymbol{\varepsilon}^{\mathrm{target}}\|_1 = K - N$, there exists an asymptotic power allocation $\boldsymbol{b} = \alpha \boldsymbol{b}_0$ with $\alpha \to \infty$ and $\boldsymbol{b}_0 \in \mathrm{int}(\mathcal{S})$ for which $\boldsymbol{f}(\boldsymbol{b}; \boldsymbol{\varepsilon}^{\mathrm{target}}) = \boldsymbol{b}$. A slight relaxation of the target MMSE tuple $\boldsymbol{\varepsilon}^{\mathrm{target}}$ with $\|\boldsymbol{\varepsilon}^{\mathrm{target}}\|_1 = K - N$ to $\boldsymbol{\varepsilon}^{\prime,\mathrm{target}} = \beta \boldsymbol{\varepsilon}^{\mathrm{target}} > \boldsymbol{\varepsilon}^{\mathrm{target}}$ with $\beta > 1$ yields $\|\boldsymbol{\varepsilon}^{\prime,\mathrm{target}}\|_1 > K - N$. Clearly, the scaling β must be small enough such that $\boldsymbol{\varepsilon}^{\prime,\mathrm{target}}$ fulfills (8). Since $\boldsymbol{f}(\boldsymbol{b}; \boldsymbol{\varepsilon})$ in (34) is strictly *decreasing* in $\boldsymbol{\varepsilon}$, we have $\boldsymbol{f}(\boldsymbol{b}; \boldsymbol{\varepsilon}^{\mathrm{target}}) > \boldsymbol{f}(\boldsymbol{b}; \boldsymbol{\varepsilon}^{\prime,\mathrm{target}})$, and therefore

$$\begin{aligned} \boldsymbol{b} &= \boldsymbol{f}(\boldsymbol{b}; \boldsymbol{\varepsilon}^{\mathrm{target}}) \quad \text{if} \quad \|\boldsymbol{\varepsilon}^{\mathrm{target}}\|_1 = K - N, \\ \boldsymbol{b} &> \boldsymbol{f}(\boldsymbol{b}; \boldsymbol{\varepsilon}^{\mathrm{target}}) \quad \text{if} \quad \|\boldsymbol{\varepsilon}^{\mathrm{target}}\|_1 > K - N. \end{aligned}$$

Hence, the third requirement (37) for the existence of a unique fixed point is satisfied for any $\boldsymbol{\varepsilon}^{\mathrm{target}}$ whose sum MMSE is larger than $K - N$. Since $\alpha \to \infty$ and $\boldsymbol{b}_0 > \mathbf{0}$, we have $\boldsymbol{b} = \alpha \boldsymbol{b}_0 > \boldsymbol{a}$. This completes the proof for Theorem III.1.

References

[1] C. Farsakh and J. A. Nossek, "Spatial covariance based downlink beamforming in an SDMA mobile radio system," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1497–1506, Nov. 1998.

[2] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. 37th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 1999, pp. 987–996.

[3] M. Schubert and H. Boche, "Solution of the multi-user downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[4] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Processing*, vol. 54, pp. 161–176, January 2006.

[5] A. M. Khachan, A. J. Tenenbaum, and R. S. Adve, "Linear processing for the downlink in multiuser MIMO systems with multiple data streams," in *Proc. Int. Conf. Commun. (ICC)*, Jun. 2006, pp. 4113–4118.

[6] 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.

[7] A. J. Fehske, F. Richter, and G. P. Fettweis, "SINR balancing for the multi-user downlink under general power constraints," in *Proc. IEEE Global Commun. Conf. (Globecom)*, New Orleans, LA, Nov. 2008, pp. 1–6.

[8] G. Montalbano and D. T. M. Slock, "Matched filter bound optimization for multiuser downlink transmit beamforming," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Oct. 1998, vol. 1, pp. 677–681.

[9] H. Boche and M. Schubert, "SIR balancing for multiuser downlink beamforming—A convergence analysis," in *Proc. Int. Conf. Commun. (ICC)*, New York, Apr. 2002, vol. 2, pp. 841–845.

[10] I. Koutsopoulos, T. Ren, and L. Tassiulas, "The impact of space division multiplexing on resource allocation: A unified approach," in *Proc. INFOCOM 2003*, Mar. 2003, vol. 1, pp. 533–543.

[11] H. Boche and M. Schubert, "A general theory for SIR balancing," *EURASIP J. Wireless Commun. Networking*, vol. 2006, no. 2, pp. 1–18, Apr. 2006.

[12] H. Boche and M. Schubert, "Multiuser interference balancing for general interference functions—A convergence analysis," in *Proc. Int. Conf. Commun. (ICC)*, Glasgow, Scotland, Jun. 2007, pp. 4664–4669.

[13] H. Boche and M. Schubert, "On the structure of the multiuser QoS region," *IEEE Trans. Signal Process.*, vol. 55, no. 7-1, pp. 3484–3495, Jul. 2007.

[14] S. Stanczak and H. Bochec, "The infeasible SIR region is not a convex set," *IEEE Trans. Commun.*, vol. 54, no. 11, pp. 1905–1907, Nov. 2006.

[15] M. Joham and R. Hunger, "Feasible rate region of the MIMO broadcast channel with linear transceivers," in *Proc. Int. ITG Workshop on Smart Antennas (WSA)*, Feb. 2010, pp. 342–349.

[16] M. Mohseni, R. Zhang, and J. M. Cioffi, "Optimized transmission for fading multiple-access and broadcast channels with multiple antennas," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1627–1639, Aug. 2006.

[17] 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.

[18] 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.

[19] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1968–1983, Sep. 1999.

[20] J. Kennan, "Uniqueness of positive fixed points for increasing concave functions on $\mathbb{R}^n$: An elementary result," *Rev. Economic Dynamics*, vol. 4, no. 4, pp. 893–899, October 2001.

[21] 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, Sep. 1995.

[22] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York: Academic, 1979.

Raphael Hunger (S'06) was born in Illertissen, Germany, 1979. He studied electrical engineering at the Technische Universität München (TUM), Munich, Germany, from 1999 until 2005 and, in 2004, at the RWTH Aachen, Germany. He received the B.Sc., the Dipl.-Ing., and M.S. degrees in electrical engineering from the Technische Universität München in 2003, 2004, and 2005, respectively. Since 2005, he has been working towards his Doctorate degree at the Associate Institute for Signal Processing, TUM.

His research interests focus on the joint optimization of transmitters and receivers in multiuser MIMO communications.

Michael Joham (S'99–M'05) was born in Kufstein, Austria, in 1974. He received the Dipl.-Ing. and Dr.-Ing. degrees (both *summa cum laude*) in electrical engineering from the Technical University Munich (TUM), Munich, Germany, in 1999 and 2004, respectively.

He was with the Institute for Circuit Theory and Signal Processing, TUM, from 1999 to 2004. Since 2004, he has been with the Associate Institute for Signal Processing, TUM, where he is currently a Senior Researcher. In the summers of 1998 and 2000, he visited Purdue University, IN. In the Spring of 2008, he was a Guest Lecturer at the University of the German Federal Armed Forces, Munich, and a Guest Professor at the University of A Coruna, Spain. In the Winter of 2009, he was a Guest Lecturer at the University of Hanover, Germany. His current research interests are precoding in mobile and satellite communications, limited rate feedback, MIMO communications, and robust signal processing.

Dr. Joham received the VDE Preis for his diploma thesis in 1999 and the Texas-Instruments-Preis for his dissertation in 2004. In 2007, he was a corecipient of the Best Paper Award at the International ITG/IEEE Workshop on Smart Antennas in Vienna.