On the Convexity of the MSE Region of Single-Antenna Users

Raphael Hunger and Michael Joham

Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany

Telephone: +49 89 289-28508, Fax: +49 89 289-28504, Email: {hunger,joham}@tum.de

Abstract—We prove convexity of the sum-power constrained *mean square error* (MSE) region in case of two single-antenna users communicating with a multi-antenna base station. Due to the MSE duality this holds both for the vector broadcast channel and the dual multiple access channel. Increasing the number of users to more than two, we show by means of a simple counter-example that the resulting MSE region is not necessarily convex any longer, even under the assumption of single-antenna users. In conjunction with our former observation that the two user MSE region is not necessarily convex for two multi-antenna users, this extends and corrects the hitherto existing notion of the MSE region geometry.

I. Introduction

Up to now, only few contributions on the geometrical structure of the mean square error region exist. In [1], the authors show that the multi-user MIMO MSE region is convex under *fixed* transmit and receive beamforming vectors both for linear and nonlinear preprocessing. Obviously, a larger set of MSE tuples can be achieved by means of *adaptive* transmit and receive beamformers. For this extended setup only the two user case has been investigated so far. Utilizing matrix inequalities of matrix-convex functions, the authors in [2] prove that the two user multi-antenna MSE region cannot exhibit a nonconvex dent between two feasible MSE points connected by a line segment with $-45°$ slope. From this observation, they claim that the MSE region is convex. For convexity, however, all possible slopes would have to be checked. As a matter of fact, a channel realization exhibiting a nonconvex MSE region with two multi-antenna users has been observed in [3] disproving the convexity theorem in [2]. A multi-carrier system where several single-antenna users communicate with a single-antenna base station has been investigated in [4]. There, the complementary MSE region of parallel broadcast channels is shown to be not necessarily convex. Since the system under consideration in [4] can be recast into a *block diagonal* MIMO broadcast channel, the authors of [4] conclude that the two user multi-antenna MSE region cannot be convex in general which again contradicts the theorem in [2]. So far, no distinct statements on convexity of the MSE region depending on the number of users and antennas per user are available in case of adaptive transmit and receive beamformers.

Some applications for which the geometry of the MSE region is of interest are for example the stream priorization according to buffer states or queue states by means of the weighted sum-MSE minimization, cf. [5]. Here, suboptimum transmit and receive filters are derived by repeatedly switching between the downlink and the dual uplink in combination with a geometric program solver for a reasonable power allocation. Balancing is considered in [6] where the weights of a weighted sum-MSE minimization are adapted until certain MSE ratios hold. Exploiting the relationship between the derivative of the the mutual information and the minimum mean square error, Christensen et al. tackle the weighted sum-rate maximization utilizing results from a weighted sum-MSE minimization, see [7]. However, convexity of the MSE region is the crucial point for the proper functionality of above applications since nonconvexity may for example prevent convergence of iterative algorithms. Finally, the MSE ε achieved with MMSE receivers is tightly related to the maximum SINR via

$$\text{SINR} = \frac{1}{\varepsilon} - 1, \tag{1}$$

and hence, also to the data rate R via the simple relation

$$R = -\log_2 \varepsilon. \tag{2}$$

Summing up, all this clearly motivates a detailed investigation.

In this paper, we extend the hitherto existing notion of the MSE region geometry. The single antenna case with two users is covered in Section II whereas statements on the convexity of the MSE region for three or more single-antenna users are presented in Section III. Finally, a conjecture on the convexity of the multi-antenna two user case is given in Section IV, and detailed proofs for the theorems and corollaries are attached in Appendices A–C for the sake of readability.

II. Convexity for Two Single-Antenna Users

In this section we present statements on the geometry of the MSE region of two single-antenna users. For this setup, convexity can always be shown:

Theorem II.1: *The MSE region of two single-antenna users is convex both in the multiple-access channel and in the vector broadcast channel.*

Proof: See Appendix A. ∎

For the most important part of the boundary of the two user MSE region (see Fig. 1) there is a functional relationship

$$\varepsilon_2 = g(\varepsilon_1) \tag{3}$$

between the two users' MSEs ε_1 and ε_2. If the channel vectors $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ describing the transmission from both users to the base station in the dual MAC are not colinear, the function g is *strictly* convex, otherwise, it is affine:

Corollary II.1. *The function* $g : \varepsilon_1 \mapsto \varepsilon_2 = g(\varepsilon_1)$ *describing the efficient set of the MSE region is strictly convex if* $\boldsymbol{h}_1$ *and* $\boldsymbol{h}_2$ *are not colinear. Otherwise,* g *is affine.*

Proof: See Appendix B. ∎

III. NONCONVEXITY EXAMPLE FOR MORE THAN TWO SINGLE-ANTENNA USERS

Although the MSE region is convex for two single-antenna users, this property may get lost when adding an additional user, even if he is equipped with only a single antenna:

Theorem III.1: *The three user MSE region of both the vector broadcast channel and the multiple-access channel is not necessarily convex.*

Proof: For the proof, we present a simple example in Appendix C where the line segment connecting two feasible MSE triples lies *outside* the MSE region. A further confirmation of Theorem III.1 results from the observation, that the weighted sum-MSE minimization has more than one local minimum, see Appendix C. ∎

Nonconvexity implies for example that not every point of the MSE efficient set can be achieved by means of the weighted sum-MSE minimization technique. Balancing approaches based on the weighted sum-MSE minimization algorithm hence may fail to achieve the desired MSE ratios, cf. [3]. Instead they are prone to oscillations. The following theorem covers the case when (three or) more than three users are present in the system:

Theorem III.2: *The MSE region of more than two users may be nonconvex both in the vector broadcast channel and in the multiple-access channel.*

Proof: The three user case has already been shown in Theorem III.1. For more than three users, the MSE region is a multi-dimensional manifold. However, setting the powers of those users to $p_4 = \ldots = p_K = 0$, the intersection of this manifold with the $K - 3$ hyperplane(s) $p_i = 0$, $i \in \{4, \ldots, K\}$, is again a three-dimensional manifold which may have the same geometry as the manifold of the three user case. Hence, the MSE region may be nonconvex for more than three users as well. ∎

IV. CONJECTURE ON THE CONVEXITY OF THE MULTI-ANTENNA TWO USER CASE

A counter-example to convexity of the MSE region when multi-antenna users are involved has been shown in [3], where two users each equipped with two antennas communicate
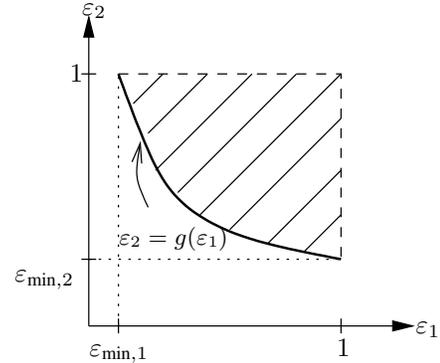
Fig. 1.   MSE ε_2 of user 2 depending on MSE ε_1 of user 1.

with a multi-antenna base-station. Similarly, the multi-carrier single-antenna system in [4] can be recast into a multi-antenna MIMO broadcast channel system where again nonconvexity was observed. Following the idea in the proof of Theorem III.2, the MSE region of two or more than two users may be nonconvex as soon as two multi-antenna users are present. Proving convexity for the case of one single-antenna user and one multi-antenna user turns out to be difficult since a parametrization of the lower left boundary of the feasible MSE region is not known, points on this boundary are obtained by limits of iterative algorithms. Nonetheless, extensive simulation results bring us to the *conjecture* that the MSE region of one single antenna user and one multi-antenna user is convex.

APPENDIX A
PROOF OF THEOREM II.1

Because of the MSE duality between the vector BC [8] and the MAC in [9], [6], [10], it suffices to prove convexity in the MAC which is easier to handle. Fig. 1 shows the basic characteristics of the two user MSE region for single-antenna users. Here, the MSEs ε_1 and ε_2 of both users are upper bounded by 1 since MMSE receivers are assumed. Allowing for other receiver types does not bring any reasonable gain since only MSE-pairs where at least one entry may lie above 1 would arise. Under the assumption of MMSE receivers, the right part of the boundary of the MSE region is obtained when user one does not transmit any data to the base station at all and user two varies its transmit power from zero to P_{Tx}. Similarly, the upper part of the boundary is reached when user two does not transmit at all whereas user one varies its transmit power from zero to P_{Tx}. Evidently, the most interesting part of the boundary is the lower left one, where the sum of both transmit powers equals the maximum available power P_{Tx}. MSE pairs lying on this boundary feature the functional relationship $\varepsilon_2 = g(\varepsilon_1)$, where the domain and the image of g are the sets $[\varepsilon_{\min,1}, 1]$ and $[\varepsilon_{\min,2}, 1]$, respectively. When less than the total transmit power P_{Tx} is consumed, points are achieved that are element of the interior of the MSE region. As a conclusion, convexity of the set of feasible MSE points corresponds to convexity of the function g relating the MSE

ε_1 of user one to the MSE ε_2 of user two on the lower left boundary of the MSE region. In the following, we show that

$$\frac{\partial^2 \varepsilon_2}{\partial \varepsilon_1^2} = \frac{\partial^2 g(\varepsilon_1)}{\partial \varepsilon_1^2} \geq 0 \tag{4}$$

holds which immediately implies convexity of g.

Unfortunately, a direct functional relationship between ε_2 and ε_1 is not available. Instead, the two MSEs ε_1 and ε_2 are parametrized by the transmit power of one of them, for example by the transmit power $p \in \mathbb{D} = [0, P_{\mathrm{Tx}}]$ of user one:

$$\varepsilon_1 = f_1(p), \quad \varepsilon_2 = f_2(p).$$

We can conclude that user two has to transmit with power $P_{\mathrm{Tx}} - p$ in order to utilize the complete power budget. In conjunction with MMSE receivers, the mean square error of user one reads as

$$\varepsilon_1 = f_1(p) = 1 - p\boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{X}^{-1}(p)\boldsymbol{h}_1 > 0, \tag{5}$$

with the positive definite covariance matrix of the received signal

$$\boldsymbol{X}(p) = \sigma_\eta^2 \mathbf{I}_N + p\boldsymbol{h}_1\boldsymbol{h}_1^{\mathrm{H}} + (P_{\mathrm{Tx}} - p)\boldsymbol{h}_2\boldsymbol{h}_2^{\mathrm{H}} \tag{6}$$

and $\sigma_\eta^2 > 0$ represents the variance of the noise at every antenna element. Similarly, the MSE of user two is denoted by

$$\varepsilon_2 = f_2(p) = 1 - (P_{\mathrm{Tx}} - p)\boldsymbol{h}_2^{\mathrm{H}} \boldsymbol{X}^{-1}(p)\boldsymbol{h}_2 > 0. \tag{7}$$

Combining (5), (7), and (6), the function f_1 turns out to be strictly monotonically decreasing in p, i.e.,

$$\dot{\varepsilon}_1 := \frac{\partial f_1(p)}{\partial p} < 0 \quad \forall p \in \mathbb{D}, \tag{8}$$

whereas f_2 is strictly monotonically increasing in p:

$$\dot{\varepsilon}_2 := \frac{\partial f_2(p)}{\partial p} > 0 \quad \forall p \in \mathbb{D}. \tag{9}$$

From (8) and (9), pseudo-convexity of g already follows. Before validating (4), we compute the first derivative:

$$\frac{\partial g(\varepsilon_1)}{\partial \varepsilon_1} = \left. \frac{\frac{\partial f_2(p)}{\partial p}}{\frac{\partial f_1(p)}{\partial p}} \right|_{p = f_1^{-1}(\varepsilon_1)}. \tag{10}$$

Note that $f_1^{-1}(\varepsilon_1)$ denotes the inverse function of f_1 which exists due to (8). Differentiating (10) again with respect to ε_1 yields

$$\begin{aligned}
\frac{\partial^2 g(\varepsilon_2)}{\partial \varepsilon_1^2} &= \frac{\partial}{\partial \varepsilon_1} \left(\left. \frac{\frac{\partial f_2(p)}{\partial p}}{\frac{\partial f_1(p)}{\partial p}} \right|_{p = f_1^{-1}(\varepsilon_1)} \right) \\
&= \left. \left(\frac{\partial}{\partial p} \frac{\frac{\partial f_2(p)}{\partial p}}{\frac{\partial f_1(p)}{\partial p}} \right) \right|_{p = f_1^{-1}(\varepsilon_1)} \cdot \frac{\partial f_1^{-1}(\varepsilon_1)}{\partial \varepsilon_1} \\
&= \left. \frac{\ddot{\varepsilon}_2 \dot{\varepsilon}_1 - \ddot{\varepsilon}_1 \dot{\varepsilon}_2}{(\dot{\varepsilon}_1)^2} \right|_{p = f_1^{-1}(\varepsilon_1)} \cdot \frac{1}{\dot{\varepsilon}_1 |_{p = f_1^{-1}(\varepsilon_1)}} \\
&= \left. \frac{\ddot{\varepsilon}_2 \dot{\varepsilon}_1 - \ddot{\varepsilon}_1 \dot{\varepsilon}_2}{(\dot{\varepsilon}_1)^3} \right|_{p = f_1^{-1}(\varepsilon_1)}.
\end{aligned} \tag{11}$$

Since f_1^{-1} maps from $[\varepsilon_{\min,1}, 1]$ to $\mathbb{D}$, and since $\dot{\varepsilon}_1 < 0$ holds $\forall p \in \mathbb{D}$, the function g is convex iff [see (11) and cond. (4)]

$$\ddot{\varepsilon}_2 \dot{\varepsilon}_1 - \ddot{\varepsilon}_1 \dot{\varepsilon}_2 \leq 0 \Leftrightarrow g \text{ is convex.} \tag{12}$$

For notational brevity, we introduce the two substitutions

$$a_{i,j} = \boldsymbol{h}_i^{\mathrm{H}} \boldsymbol{X}^{-1}(p)\boldsymbol{h}_j \quad \text{and} \quad b_{i,j} = \boldsymbol{h}_i^{\mathrm{H}} \boldsymbol{X}^{-2}(p)\boldsymbol{h}_j, \tag{13}$$

which satisfy $a_{i,j} = a_{j,i}^*$ and $b_{i,j} = b_{j,i}^*$. Making use of

$$\frac{\partial \boldsymbol{X}^{-1}(p)}{\partial p} = -\boldsymbol{X}^{-1}(p) \frac{\partial \boldsymbol{X}(p)}{\partial p} \boldsymbol{X}^{-1}(p),$$

the first derivatives with respect to p in (8) and (9) can be shown to equal

$$\begin{aligned}
\dot{\varepsilon}_1 &= -\sigma_\eta^2 b_{1,1} - P_{\mathrm{Tx}} |a_{1,2}|^2, \\
\dot{\varepsilon}_2 &= +\sigma_\eta^2 b_{2,2} + P_{\mathrm{Tx}} |a_{1,2}|^2,
\end{aligned} \tag{14}$$

respectively. Differentiating (14) again w.r.t. p, we obtain

$$\ddot{\varepsilon}_1 = 2\sigma_\eta^2 [a_{1,1} b_{1,1} - \Re\{a_{1,2} b_{2,1}\}] + 2P_{\mathrm{Tx}} |a_{1,2}|^2 (a_{1,1} - a_{2,2}),$$
$$\ddot{\varepsilon}_2 = 2\sigma_\eta^2 [a_{2,2} b_{2,2} - \Re\{a_{2,1} b_{1,2}\}] + 2P_{\mathrm{Tx}} |a_{1,2}|^2 (a_{2,2} - a_{1,1}).$$

Inserting (14) and the last two equations into (12) and applying $\Re\{a_{2,1} b_{1,2}\} = \Re\{a_{1,2} b_{2,1}\}$ results in

$$\begin{aligned}
\ddot{\varepsilon}_2 \dot{\varepsilon}_1 &- \ddot{\varepsilon}_1 \dot{\varepsilon}_2 = \\
&2\sigma_\eta^2 P_{\mathrm{Tx}} |a_{1,2}|^2 \left[2\Re\{a_{1,2} b_{2,1}\} - a_{2,2} b_{1,1} - a_{1,1} b_{2,2} \right] \\
&+ 2\sigma_\eta^4 b_{1,1} \left(\Re\{a_{1,2} b_{2,1}\} - a_{1,1} b_{2,2} \right) \\
&+ 2\sigma_\eta^4 b_{2,2} \left(\Re\{a_{2,1} b_{1,2}\} - a_{2,2} b_{1,1} \right).
\end{aligned} \tag{15}$$

In order to prove that (15) is not positive to fulfill the convexity requirement in (12), we will reveal that all three summands in (15) are not positive.

For the first summand, this turns out to be very easy: Noticing that $a_{i,i} > 0$ and $b_{i,i} > 0$, the first summand in (15) is nonpositive if

$$4\Re^2\{a_{2,1} b_{1,2}\} \leq (a_{2,2} b_{1,1} + a_{1,1} b_{2,2})^2. \tag{16}$$

Clearly, we can upper bound the real part by the magnitude and apply the *Cauchy-Schwarz*-inequality with (13) to bound the magnitude:

$$4\Re^2\{a_{2,1} b_{1,2}\} \leq 4|a_{2,1} b_{1,2}|^2 \leq 4a_{2,2} a_{1,1} b_{1,1} b_{2,2}. \tag{17}$$

Validating the inequality

$$(a_{2,2} b_{1,1} + a_{1,1} b_{2,2})^2 \geq 4a_{2,2} a_{1,1} b_{1,1} b_{2,2}$$
$$\Leftrightarrow (a_{2,2} b_{1,1} - a_{1,1} b_{2,2})^2 \geq 0$$

leads in conjunction with (17) to the conclusion that (16) is fulfilled, i.e., the first summand in (15) is nonpositive.

Nonpositivity of the second summand in (15) is resembled by the inequality

$$\Re \left\{ \frac{a_{1,2}}{a_{1,1}} \frac{b_{2,1}}{b_{2,2}} \right\} \leq 1. \tag{18}$$

To prove (18) we explicitly have to exploit the structure of $\boldsymbol{X}(p)$ in (6) which makes the proof longer than the one for

the first summand. Interestingly, the real part operator in (18) is redundant as its argument turns out to be real-valued. Applying the matrix inversion lemma several times, we get

$$\frac{a_{1,2}}{a_{1,1}} = \frac{\sigma_\eta^2 \boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{h}_2}{\sigma_\eta^2 \|\boldsymbol{h}_1\|_2^2 + d(P_{\mathrm{Tx}} - p)}, \tag{19}$$

with the substitution

$$d = \|\boldsymbol{h}_1\|_2^2 \|\boldsymbol{h}_2\|_2^2 - |\boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{h}_2|^2 \geq 0. \tag{20}$$

Applying several times the matrix inversion lemma as for the first fraction, the second fraction in (18) can be expressed as

$$\frac{b_{2,1}}{b_{2,2}} = \frac{\boldsymbol{h}_2^{\mathrm{H}} \boldsymbol{h}_1 \left[\sigma_\eta^4 - p(P_{\mathrm{Tx}} - p)d\right]}{\sigma_\eta^4 \|\boldsymbol{h}_2\|_2^2 + dp(2\sigma_\eta^2 + p\|\boldsymbol{h}_1\|_2^2)}. \tag{21}$$

Multiplying (19) by (21) yields the real-valued expression

$$\frac{b_{2,1} a_{1,2}}{a_{1,1} b_{2,2}} = \frac{\sigma_\eta^6 |\boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{h}_2|^2 - c_1}{\sigma_\eta^6 \|\boldsymbol{h}_1\|_2^2 \|\boldsymbol{h}_2\|_2^2 + c_2} \ \in \mathbb{R}$$

with the two substitutions

$$c_1 = \sigma_\eta^2 |\boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{h}_2|^2 pd(P_{\mathrm{Tx}} - p) \geq 0,$$
$$c_2 = \left[\sigma_\eta^2 \|\boldsymbol{h}_1\|_2^2 + d(P_{\mathrm{Tx}} - p)\right] dp \left(2\sigma_\eta^2 + p\|\boldsymbol{h}_1\|_2^2\right)$$
$$+ \sigma_\eta^4 \|\boldsymbol{h}_2\|_2^2 d(P_{\mathrm{Tx}} - p) \geq 0.$$

Since both c_1 and c_2 are nonnegative, we find

$$\frac{b_{2,1} a_{1,2}}{a_{1,1} b_{2,2}} \leq \frac{\sigma_\eta^6 |\boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{h}_2|^2}{\sigma_\eta^6 \|\boldsymbol{h}_1\|_2^2 \|\boldsymbol{h}_2\|_2^2}$$

as an upper bound from which (18) directly follows due to the *Cauchy-Schwarz*-inequality. Thus, the nonpositivity of the second summand in (15) is proven.

Finally, the nonpositivity of the third summand in (15) is shown by the same reasoning as for the second summand:

$$\Re\left\{\frac{a_{1,2}}{a_{2,2}} \frac{b_{2,1}}{b_{1,1}}\right\} \leq 1 \tag{22}$$

is deduced from

$$\frac{b_{2,1} a_{1,2}}{a_{2,2} b_{1,1}} = \frac{\sigma_\eta^6 |\boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{h}_2|^2 - d_1}{\sigma_\eta^6 \|\boldsymbol{h}_1\|_2^2 \|\boldsymbol{h}_2\|_2^2 + d_2} \ \in \mathbb{R},$$

where d_i follows from c_i by interchanging indices and powers:

$$d_1 = c_1,$$
$$d_2 = \left(\sigma_\eta^2 \|\boldsymbol{h}_2\|_2^2 + dp\right) d(P_{\mathrm{Tx}} - p) \left[2\sigma_\eta^2 + (P_{\mathrm{Tx}} - p)\|\boldsymbol{h}_2\|_2^2\right]$$
$$+ \sigma_\eta^4 \|\boldsymbol{h}_1\|_2^2 dp \geq 0.$$

As all three summands in (15) are nonpositive, the inequality in (12) is satisfied and the proof for the convexity of the MSE region is complete.

APPENDIX B
PROOF OF COROLLARY II.1

If the inequality in (12) is strict for all $p \in \mathbb{D}$, g is strictly convex. Excluding equality in (12) therefore ensures that g is strictly convex. The difference in (15) is zero if and only if all three summands are zero since each summand is nonpositive.

In order to let the first summand vanish, the *Cauchy-Schwarz*-inequality in (17) has to be fulfilled with equality. To this end, $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ have to be colinear which also fulfills (16) with equality. If both channel vectors are colinear, $d = 0$ results from (20) and the variables c_1, c_2, d_1, and d_2 are zero as well. Obviously, (18) holds with equality and the last two summands in (15) vanish. Thus, we have shown that if the two channel vectors $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are not colinear, then the function g is strictly convex. Additionally, if both vectors are colinear, g has curvature zero for all powers $p \in \mathbb{D}$. As a consequence, g is affine. In the latter case, we have the relationship

$$g(\varepsilon_1) = -\varepsilon_1 \frac{|\alpha|^2 + |\alpha|^2 \gamma \|\boldsymbol{h}_1\|_2^2}{1 + |\alpha|^2 \gamma \|\boldsymbol{h}_1\|_2^2} + 1 + \frac{|\alpha|^2}{1 + |\alpha|^2 \gamma \|\boldsymbol{h}_1\|_2^2}, \tag{23}$$

where $\gamma = P_{\mathrm{Tx}}/\sigma_\eta^2$ denotes the transmit SNR, $\boldsymbol{h}_2 = \alpha \boldsymbol{h}_1$, and $\varepsilon_1 \in [\varepsilon_{\min,1}, 1]$ with

$$\varepsilon_{\min,1} = \frac{1}{1 + \gamma \|\boldsymbol{h}_1\|_2^2}. \tag{24}$$

APPENDIX C
PROOF OF THEOREM III.1

A nonconvex three user MSE region can for example be obtained by the channel matrix

$$\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \tag{25}$$

and a transmit power $P_{\mathrm{Tx}} = 10$. In this case, the base station is equipped with $N = 2$ antennas, and the channel vector $\boldsymbol{h}_3$ is the sum of $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$. Note that the base station has fewer antennas than users are present in the system in this special case. Nonconvexity of the MSE region can also be and has been observed when the channel vectors of all users are linearly independent ($N \geq K$ must hold then). If the MSE region was convex, the line segment between every two feasible MSE triples would have to be a subset of the region. Moreover, the weighted sum-MSE minimization with arbitrary nonnegative weights $\boldsymbol{w} = [w_1, \ldots, w_K]^{\mathrm{T}} \geq \boldsymbol{0}_K$, $\boldsymbol{w} \neq \boldsymbol{0}_K$ may have stationary points fulfilling the KKT conditions *with only one common value* of the weighted minimization. In the following, we show that these conditions are violated for the channel in (25).

The weighted sum-MSE minimization reads as

$$\underset{p_1, \ldots, p_K}{\mathrm{minimize}} \sum_{k=1}^{K} w_k \varepsilon_k \quad \text{s.t.:} \ \sum_{k=1}^{K} p_k \leq P_{\mathrm{Tx}}, \ \ p_k \geq 0 \ \forall k, \tag{26}$$

where p_k is the power with which user k transmits in the uplink and the MSE of user k reads as

$$\varepsilon_k = 1 - p_k \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{X}^{-1} \boldsymbol{h}_k \tag{27}$$

with the received signal covariance matrix

$$\boldsymbol{X} = \sigma_\eta^2 \mathbf{I}_N + \sum_{\ell=1}^{K} p_\ell \boldsymbol{h}_\ell \boldsymbol{h}_\ell^{\mathrm{H}}. \tag{28}$$

The *Lagrangian* function associated to (26) reads as

$$L = \sum_{k=1}^{K} w_k \varepsilon_k + \lambda \Big(\sum_{k=1}^{K} p_k - P_{\text{Tx}} \Big) - \sum_{k=1}^{K} \mu_k p_k. \qquad (29)$$

Note that the *Lagrangian* multipliers λ and $\mu_1, \ldots, \mu_K$ have to be nonnegative real. If above *Lagrangian* L has stationary points with different values for L, the underlying MSE region is not convex since more than one hyperplane with normal vector $[w_1, \ldots, w_K]^{\text{T}}$ locally supporting the MSE region exists. The KKT conditions read as

$$\boldsymbol{h}_k^{\text{H}} \check{\boldsymbol{X}}^{-1} (w_k \check{\boldsymbol{X}} - \check{\boldsymbol{S}}) \check{\boldsymbol{X}}^{-1} \boldsymbol{h}_k = \check{\lambda} - \check{\mu}_k \ \forall k, \qquad (30)$$

$$\check{p}_k \geq 0 \ \forall k, \qquad (31)$$

$$\check{p}_k \check{\mu}_k = 0 \ \forall k, \qquad (32)$$

$$\check{\mu}_k \geq 0 \ \forall k, \qquad (33)$$

$$\sum_{k=1}^{K} \check{p}_k \leq P_{\text{Tx}}, \qquad (34)$$

$$\check{\lambda} \Big(\sum_{k=1}^{K} \check{p}_k - P_{\text{Tx}} \Big) = 0, \qquad (35)$$

$$\check{\lambda} \geq 0, \qquad (36)$$

with the substitution

$$\boldsymbol{S} = \sum_{\ell=1}^{K} w_\ell p_\ell \boldsymbol{h}_\ell \boldsymbol{h}_\ell^{\text{H}}. \qquad (37)$$

Note that checked variables $(\check{\cdot})$ are those which fulfill the KKT conditions. Assuming a weight vector

$$\boldsymbol{w} = [0.22, 0.54, 0.24]^{\text{T}}, \qquad (38)$$

the weighted sum-MSE minimization (26) features *two* stationary points satisfying the KKT conditions (30)–(36) for the channel vectors (25) and a transmit power $P_{\text{Tx}} = 10$. The first set of primal and dual variables fulfilling the KKTs reads as

$$\check{\boldsymbol{p}}^{(1)} = [3.6753, \ 6.3247, \ 0]^{\text{T}}, \quad \check{\lambda}^{(1)} = 0.0101,$$
$$\check{\mu}_1^{(1)} = \check{\mu}_2^{(1)} = 0, \quad \check{\mu}_3^{(1)} = 0.0266, \qquad (39)$$

and achieves a weighted sum-MSE $\sum_{k=1}^{3} w_k \check{\varepsilon}_k^{(1)} = 0.36078$. The second set of variables reads as

$$\check{\boldsymbol{p}}^{(2)} = [0, \ 7.0794, \ 2.9206]^{\text{T}}, \quad \check{\lambda}^{(2)} = 0.0115,$$
$$\check{\mu}_1^{(2)} = 0.007, \quad \check{\mu}_2^{(2)} = \check{\mu}_3^{(2)} = 0, \qquad (40)$$

and obtains a slightly larger metric $\sum_{k=1}^{3} w_k \check{\varepsilon}_k^{(2)} = 0.3828$. The existence of two KKT points with different values algebraically proves the nonconvexity of the MSE region.

A geometrical proof is shown in Fig. 2, where the three-user MSE region for the channel in (25) is plotted with $P_{\text{Tx}} = 10$. The two KKT points in (39) and (40) achieve individual MSE triples

$$\check{\boldsymbol{\varepsilon}}^{(1)} = \Big[\check{\varepsilon}_1^{(1)}, \check{\varepsilon}_2^{(1)}, \check{\varepsilon}_3^{(1)} \Big]^{\text{T}} = [0.2139, \ 0.1365, \ 1]^{\text{T}},$$
$$\check{\boldsymbol{\varepsilon}}^{(2)} = \Big[\check{\varepsilon}_1^{(2)}, \check{\varepsilon}_2^{(2)}, \check{\varepsilon}_3^{(2)} \Big]^{\text{T}} = [1, \ 0.1977, \ 0.2335]^{\text{T}}. \qquad (41)$$
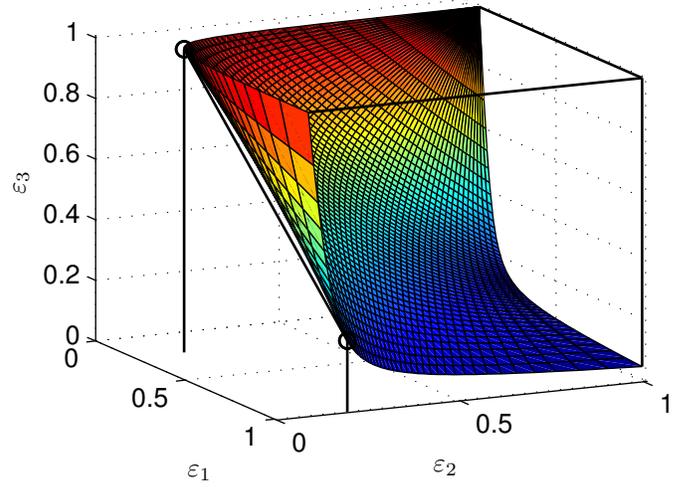
Fig. 2. Example of a nonconvex MSE region for $K = 3$ users. The line segment connecting two feasible points lies *outside* the region.

However, the line segment connecting $\check{\boldsymbol{\varepsilon}}^{(1)}$ and $\check{\boldsymbol{\varepsilon}}^{(2)}$ does not completely belong to the MSE region, it lies outside the region and touches the boundary of the MSE region at $\check{\boldsymbol{\varepsilon}}^{(1)}$ and $\check{\boldsymbol{\varepsilon}}^{(2)}$. Evidently, the MSE region cannot be convex.

REFERENCES

[1] S. Shi and M. Schubert, "Convexity Analysis of the Feasible MSE Region of Sum-Power Constrained Multiuser MIMO Systems," in *Proc. IEEE Internat. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC), Berlin, Germany*, Sept. 2005.

[2] E. Jorswieck and H. Boche, "Transmission Strategies for the MIMO MAC with MMSE Receiver: Average MSE Optimization and Achievable Individual MSE Region," *IEEE Trans. on Signal Processing*, vol. 51, no. 11, pp. 2872–2881, Nov. 2003, special issue on MIMO wireless communications.

[3] R. Hunger, M. Joham, and W. Utschick, "Efficient MSE Balancing for the Multi-User MIMO Downlink," in *Proc. 41st Asilomar Conference on Signals, Systems, and Computers*, November 2007.

[4] G. Wunder, I. Blau, and T. Michel, "Utility Optimization based on MSE for Parallel Broadcast Channels: The Square Utility Optimization based on MSE for Parallel Broadcast Channels: The Square Root Law," in *45th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, Illinois, September 2007.

[5] M. Codreanu, A. Tolli, M. Juntti, and M. Latva-aho, "Weighted Sum MSE Minimization for MIMO Broadcast Channel," in *17th International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, September 2006.

[6] A. Mezghani, M. Joham, R. Hunger, and W. Utschick, "Transceiver Design for Multi-User MIMO Systems," in *Proc. ITG/IEEE WSA 2006*, March 2006.

[7] S. Christensen, R. Agarwal, E. d. Carvalho, and J. M. Cioffi, "Beamforming Design for Weighted Sum-Rate Maximization in MIMO Broadcast Channels," *Submitted to IEEE Transactions on Wireless Communications*.

[8] W. Yu and J. M. Cioffi, "Sum Capacity of Gaussian Vector Broadcast Channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1875–1892, September 2004.

[9] R. Hunger, M. Joham, and W. Utschick, "On the MSE-Duality of the Broadcast Channel and the Multiple Access Channel," *Submitted to IEEE Transactions on Signal Processing*.

[10] M. Schubert, S. Shi, E. A. Jorswieck, and H. Boche, "Downlink Sum-MSE Transceiver Optimization for Linear Multi-User MIMO Systems," in *Proc. Asilomar Conf. on Signals, Systems and Computers, Monterey, CA*, Sept. 2005.