ROBUST TRANSCEIVER OPTIMIZATION FOR FREQUENCY SELECTIVE MIMO CHANNELS

Nikola Vucic

Fraunhofer Institute
for Telecommunications HHI
Einsteinufer 37, 10587 Berlin, Germany
nikola.vucic@hhi.fraunhofer.de

Holger Boche

Technical University Berlin
Heinrich Hertz Chair for Mobile Communications
Einsteinufer 25, 10587 Berlin, Germany
holger.boche@mk.tu-berlin.de

ABSTRACT

We study the problem of joint transmit and receive filters optimization in a frequency selective, multiple-input multiple-output setup. The information about the channel at the transmitter is imperfect and belongs to a specified uncertainty set, defined by bounding the norm of the error transfer function. The framework for a robust optimization of the system, with mean-square-error (MSE) as the performance measure, is derived. Robustness is defined in the worst-case sense, and a broad range of MSE-optimization problems is supported. The algorithms are constructed in an iterative manner, where each iteration consists of two efficiently solvable semidefinite programs. The proofs of the convergence are provided, as well. Numerical examples show significant performance gains in comparison to the system which performs the optimization of the precoder only.

1. INTRODUCTION

The future wireless networks are expected to provide significant additional performance gains by utilizing antenna arrays at the transmit and receive sides, and, thereby, exploiting the spatial diversity [1]. This idea leads to the concept of multiple-input, multiple-output (MIMO) systems. Since the wireless channels are frequency selective, the task of combating the intersymbol interference (ISI) becomes more involved, as the spatial dimension of the system must be included, as well.

In such MIMO-ISI scenarios, linear transmit filters (precoders) and receive equalizers can be designed to optimize the system, with the mean-square-error (MSE) taken as the performance measure. Most of the work in the literature, assumed that the perfect channel state information (CSI) is available at the side where the optimization is performed. With an assumption of having the perfect channel knowledge, if either the receiver or the transmitter is kept fixed, the solutions for the precoder and the receive equalizer, respectively, are readily obtained as standard Wiener filters. The joint transmit and receive filter optimization is considerably more involved [2, 3]. The solutions for several scenarios of interest in block-based ISI systems, where inserted additional non-information symbols eliminate the interblock interference, are provided in [4, 5].

However, the provision of the perfect CSI is often a formidable task in wireless systems. This especially holds for the transmit side, which usually obtains the channel knowledge from the receivers using the feedback channels. Since these channels are of limited capacity, the issue of errors due to the (often rough) quantization must be taken into account [6]. In time division duplex (TDD) systems, it might be also possible for the transmitter to use the channel values that were estimated in a phase when it operated as a receiver. In this situation, there exists a danger of having the outdated estimates, due to the fast varying wireless environment. Finally, for both approaches, the errors in channel estimation itself are inevitable. Therefore, there exists a significant practical interest in providing filters which are robust with respect to the imperfect CSI.

We derive a framework that handles several robust MSE optimization problems in a system where the precoder, the channel and the receive equalizer are causal and have finite impulse responses (FIRs). Robustness is defined in accordance with the worst-case philosophy [7]. In other words, we will study either *minimax* designs, where the largest MSE with respect to the uncertainty should be minimized, or the problems that aim at minimizing the total transmit power under a strict MSE constraint that must be satisfied under the uncertainty. The errors in the channel knowledge will be bounded by putting a constraint on the error transfer function. Contrary to the block-based systems, the setup of interest in this paper is defined by standard convolutional equations, with no guard intervals needed.

In this particular setting, to the best of our knowledge, the problem of joint transmit and receive filters design is still open even with a perfect CSI assumption. An iterative algorithm, that provides a suboptimal solution in the perfect CSI case, is mentioned already in [2]. The main contribution of this paper is in showing that a similar reasoning can be applied for the robust optimization, as well. In the case of the imperfect channel knowledge, the problems are more intricate due to the fact that an infinite number of channels, contained in the uncertainty regions, must be supported with a single transmit and receive filter pair. It will be proved, however, that this issue can be resolved using numerical, semidefinite programming (SDP) methods from convex optimization theory [8].

Finally, we remark that the work in this paper presents an extension of the results from [9], where the robust optimization of a causal FIR MIMO filter is performed either at the transmit or at the receive side. The related robust designs in MIMO-ISI channels were studied recently also in [10, 11]. In [10], however, the optimization is performed only at one side, and the filters had infinite impulse responses with no causality constraint imposed. [11] considered a different model, based on block-circulant matrices, that does not correspond to the convolutional MIMO-ISI setup analyzed in our work.

We will use small and large bold fonts to denote vectors and matrices, respectively. $\boldsymbol{I}$ is the identity matrix, and $\mathbf{0}$ is the zero-matrix. $\boldsymbol{A} \succeq \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ is positive semidefinite [8]. $\boldsymbol{A} \otimes \boldsymbol{B}$ is the Kronecker product of $\boldsymbol{A}$ and $\boldsymbol{B}$ [12]. $\|(\cdot)\|_2$ denotes either the spectral norm of a matrix or the H_2 norm of a matrix transfer function

This work is supported by the STREP project No. IST-026905 (MASCOT) within the sixth framework programme of the European Commission.

(MTF) [13], and the meaning will be clear from the context. $\|(\cdot)\|_F$ denotes the Frobenius norm of a matrix [14]. The transpose, the conjugate (Hermitian) transpose and the trace of a matrix are written as $(\cdot)^T$, $(\cdot)^*$ and $\mathrm{Tr}\{\cdot\}$, respectively. $\boldsymbol{A}_{(:,m)}$ ($\boldsymbol{A}_{(m,:)}$) denotes the mth column (row) of $\boldsymbol{A}$. If $\boldsymbol{A}$ is a matrix with M columns, $\mathrm{vec}(\boldsymbol{A})$ is defined as $\mathrm{vec}(\boldsymbol{A}) = \begin{bmatrix} \boldsymbol{A}_{(:,1)}^T & \cdots & \boldsymbol{A}_{(:,M)}^T \end{bmatrix}^T$. $\mathrm{E}\{\cdot\}$ is the expectation operator.

2. SYSTEM MODEL

A schematic representation of the considered frequency selective MIMO setup is given in Fig. 1. The transmitter is supposed to

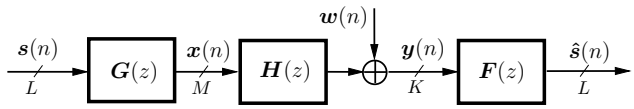

Fig. 1. Block-Scheme of the studied MIMO-ISI system.

send a sequence of vectors containing L spatial data streams (layers) $\boldsymbol{s}(n) = \begin{bmatrix} s_1(n) & \cdots & s_L(n) \end{bmatrix}^T$, where n denotes the time instant. It will be assumed that the transmit sequence is white, with

$$\mathrm{E}\left\{\boldsymbol{s}(n)\boldsymbol{s}^*(n-k)\right\} = \begin{cases} \sigma_s^2 \boldsymbol{I}_L, & k = 0 \\ \boldsymbol{0}_{L\times L}, & k \neq 0 \end{cases}. \tag{1}$$

The transmitter and the receiver are equipped with M and K antennas and with linear, causal FIR filters $\boldsymbol{G}(z)$ and $\boldsymbol{F}(z)$, respectively

$$\boldsymbol{G}(z) = \sum_{k=0}^{N_G} \boldsymbol{G}_k z^{-k}, \quad \boldsymbol{F}(z) = \sum_{k=0}^{N_F} \boldsymbol{F}_k z^{-k}. \tag{2}$$

The $K \times M$ MIMO-ISI channel is of order N_H, with the MTF

$$\boldsymbol{H}(z) = \hat{\boldsymbol{H}}(z) + \boldsymbol{\Delta}(z), \tag{3}$$

where the coefficients of

$$\hat{\boldsymbol{H}}(z) = \sum_{k=0}^{N_H} \hat{\boldsymbol{H}}_k z^{-k}, \quad \boldsymbol{\Delta}(z) = \sum_{k=0}^{N_H} \boldsymbol{\Delta}_k z^{-k} \tag{4}$$

are channel estimates and disturbances, respectively. It is assumed that the transmitter is provided only with the erroneous estimate $\hat{\boldsymbol{H}}(z)$ and a bound ε on the H_2 norm of the error MTF

$$\|\boldsymbol{\Delta}\|_2 \triangleq \sqrt{\sum_{k=0}^{N_H} \|\boldsymbol{\Delta}_k\|_F^2} \leq \varepsilon. \tag{5}$$

The receiver will be supposed to have the perfect CSI, though this assumption can be somewhat relaxed, as discussed in Section 3. The model (5) presents an ISI-suitable extension of the ellipsoidal uncertainty model that is often used for modeling errors in flat-fading channels [15]. A similar error model for broadband channels is used also in [10]. In a practical application, the value ε would be determined by the roughness of the channel quantization, or according to the desired outage probability in the system, if the CSI errors are unbounded, as often assumed in the estimation process.

At the reception, additive white noise $\boldsymbol{w}(n)$ is present, with

$$\mathrm{E}\left\{\boldsymbol{w}(n)\boldsymbol{w}^*(n-k)\right\} = \begin{cases} \sigma_w^2 \boldsymbol{I}_K, & k = 0 \\ \boldsymbol{0}_{K\times K}, & k \neq 0 \end{cases}. \tag{6}$$

The transmission of data in the system can be described with convolutional equations

$$\boldsymbol{x}(n) = \sum_{k=0}^{N_G} \boldsymbol{G}_k \boldsymbol{s}(n-k), \tag{7}$$

$$\boldsymbol{y}(n) = \sum_{k=0}^{N_H} \boldsymbol{H}_k \boldsymbol{x}(n-k) + \boldsymbol{w}(n), \;\; \hat{\boldsymbol{s}}(n) = \sum_{k=0}^{N_F} \boldsymbol{F}_k \boldsymbol{y}(n-k).$$

Finally, as the performance measure, the MSE between the sent and the estimated signal

$$\mathrm{MSE} \triangleq \mathrm{E}\left\{\|\boldsymbol{s}(n) - \hat{\boldsymbol{s}}(n)\|_2^2\right\}, \tag{8}$$

is used. In the following two sections, we analyze robust counterparts of several MSE-optimization problems in the described setting.

3. MINIMAX TRANSCEIVER OPTIMIZATION

In this section, we study the problem of *minimax* robust design

$$\min_{\boldsymbol{G}(z),\,\boldsymbol{F}(z)} \; \max_{\|\boldsymbol{\Delta}\|_2 \leq \varepsilon} \mathrm{MSE}, \quad \text{s.t. } P \leq P_{\max}, \tag{9}$$

where MSE is given by (8), and P is the total transmit power that can be calculated as

$$P = \mathrm{E}\left\{\mathrm{Tr}\left\{\boldsymbol{x}(n)\boldsymbol{x}^*(n)\right\}\right\} = \sigma_s^2 \|\boldsymbol{G}\|_2^2. \tag{10}$$

In order to understand the structure of the problem (9), we start with transforming the expression for the MSE (8)

$$\mathrm{MSE} = \sigma_s^2 \left\|\tilde{\boldsymbol{F}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{G}} - \tilde{\boldsymbol{I}}\right\|_F^2 + \sigma_w^2 \|\boldsymbol{f}\|_2^2, \tag{11}$$

where $\tilde{\boldsymbol{F}}$ and $\tilde{\boldsymbol{H}}$ are $(N_F + N_H + N_G + 1)L \times (N_H + N_G + 1)K$ and $(N_H + N_G + 1)K \times (N_G + 1)M$ matrices, respectively, with

$$\tilde{\boldsymbol{F}} = \begin{bmatrix} \boldsymbol{F}_0 & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{F}_{N_F} & \ddots & \boldsymbol{F}_0 \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{F}_{N_F} \end{bmatrix}, \; \tilde{\boldsymbol{H}} = \begin{bmatrix} \boldsymbol{H}_0 & \cdots & \boldsymbol{0} \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{N_H} & \ddots & \boldsymbol{H}_0 \\ \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{H}_{N_H} \end{bmatrix} \tag{12}$$

and where

$$\tilde{\boldsymbol{G}} = \begin{bmatrix} \boldsymbol{G}_0 \\ \vdots \\ \boldsymbol{G}_{N_G} \end{bmatrix}, \quad \tilde{\boldsymbol{I}} = \begin{bmatrix} \boldsymbol{I}_L \\ \boldsymbol{0} \end{bmatrix}, \quad \boldsymbol{f} = \mathrm{vec}\left(\begin{bmatrix} \boldsymbol{F}_0 \\ \vdots \\ \boldsymbol{F}_{N_F} \end{bmatrix}\right). \tag{13}$$

We proceed by noticing that (11) can be rewritten as

$$\mathrm{MSE} = \sigma_s^2 \left\|\mathrm{vec}\left(\tilde{\boldsymbol{F}}\tilde{\boldsymbol{H}}\tilde{\boldsymbol{G}} - \tilde{\boldsymbol{I}}\right)\right\|_2^2 + \sigma_w^2 \|\boldsymbol{f}\|_2^2 \tag{14}$$

$$= \sigma_s^2 \Big\| \left(\tilde{\boldsymbol{G}}^T \otimes \tilde{\boldsymbol{F}}\right) \mathrm{vec}\left(\tilde{\hat{\boldsymbol{H}}}\right) - \mathrm{vec}\left(\tilde{\boldsymbol{I}}\right) + \left(\tilde{\boldsymbol{G}}^T \otimes \tilde{\boldsymbol{F}}\right) \mathrm{vec}\left(\tilde{\boldsymbol{\Delta}}\right) \Big\|_2^2 + \sigma_w^2 \|\boldsymbol{f}\|_2^2, \tag{15}$$

where in (15), the relation $\mathrm{vec}\left(\boldsymbol{ABC}\right) = \left(\boldsymbol{C}^T \otimes \boldsymbol{A}\right) \mathrm{vec}\left(\boldsymbol{B}\right)$ is used [12], and $\tilde{\hat{\boldsymbol{H}}}$ and $\tilde{\boldsymbol{\Delta}}$ are defined analogously to $\tilde{\boldsymbol{H}}$.

Similarly to the last equation in (13), we group the unknown precoder coefficients into a vector $\boldsymbol{g}$

$$\boldsymbol{g} = \mathrm{vec}\left(\begin{bmatrix} \boldsymbol{G}_0^T & \cdots & \boldsymbol{G}_{N_G}^T \end{bmatrix}^T\right), \tag{16}$$

and, for notational convenience, define a function $\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f})$

$$\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) = \left(\tilde{\boldsymbol{G}}^T \otimes \tilde{\boldsymbol{F}}\right) \mathrm{vec}\left(\tilde{\hat{\boldsymbol{H}}}\right) - \mathrm{vec}\left(\tilde{\boldsymbol{I}}\right). \tag{17}$$

Finally, after some calculations we obtain the following equivalent expression for the MSE, that will be convenient for the robust analysis

$$\mathrm{MSE} = \sigma_s^2 \left\|\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) + \boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f})\boldsymbol{\delta}\right\|_2^2 + \sigma_w^2 \left\|\boldsymbol{f}\right\|_2^2, \tag{18}$$

where

$$\boldsymbol{\delta} = \mathrm{vec}\left(\begin{bmatrix} \boldsymbol{\Delta}_0^T & \cdots & \boldsymbol{\Delta}_{N_H}^T \end{bmatrix}^T\right), \tag{19}$$

contains the whole uncertainty, and

$$\boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f}) = \left(\tilde{\boldsymbol{G}}^T \otimes \tilde{\boldsymbol{F}}\right) \begin{bmatrix} \bar{\boldsymbol{I}} \\ \boldsymbol{0}_{(N_G+1)K \times (N_H+1)MK} \\ \bar{\boldsymbol{I}} \\ \boldsymbol{0}_{(N_G+1)K \times (N_H+1)MK} \\ \vdots \\ \bar{\boldsymbol{I}} \end{bmatrix}, \tag{20}$$

with $((N_H+1)M + N_G(M-1))K \times (N_H+1)MK$ matrix $\bar{\boldsymbol{I}}$

$$\bar{\boldsymbol{I}} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0}^{(1)} & \cdots & \boldsymbol{0}^{(1)} \\ \boldsymbol{0}^{(2)} & \boldsymbol{0}^{(2)} & \cdots & \boldsymbol{0}^{(2)} \\ \boldsymbol{0}^{(1)} & \boldsymbol{I} & \cdots & \boldsymbol{0}^{(1)} \\ \boldsymbol{0}^{(2)} & \boldsymbol{0}^{(2)} & \cdots & \boldsymbol{0}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0}^{(1)} & \boldsymbol{0}^{(1)} & \cdots & \boldsymbol{I} \end{bmatrix}, \tag{21}$$

where $\boldsymbol{0}^{(1)}$ and $\boldsymbol{0}^{(2)}$ have dimensions $(N_H+1)K \times (N_H+1)K$ and $N_G K \times (N_H+1)K$, respectively. Consider the formula (18) with no uncertainty, i.e., assuming $\boldsymbol{\delta} = \boldsymbol{0}$. By analyzing (17), it can be noticed that (18) is not convex in the unknown vectors $\boldsymbol{g}$ and $\boldsymbol{f}$ jointly because of the multiplication via the Kronecker product. However, due to the affinity of (17) in $\boldsymbol{g}$ and $\boldsymbol{f}$ separately, it is easily seen that (18) is convex in $\boldsymbol{g}$ if $\boldsymbol{f}$ is kept fixed and vice versa. This leads to an idea that the problem can be approached using an iterative procedure, if the perfect CSI is provided. We will show now that the same principle is applicable for the robust case $\boldsymbol{\delta} \neq \boldsymbol{0}$, as well. We summarize the iterative algorithm for the problem (9) in Table 1.

In the sequel, it will be proved that the problems $(\mathcal{P}1)$ and $(\mathcal{P}2)$ are convex optimization problems with efficient numerical solutions, and that the algorithm converges. Notice that $\boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f})$ in (18) is also affine in $\boldsymbol{g}$ and $\boldsymbol{f}$ individually, and that the bounded uncertainty set (5) has an equivalent representation

$$\left\|\boldsymbol{\Delta}\right\|_2 \leq \varepsilon \;\Leftrightarrow\; \left\|\boldsymbol{\delta}\right\|_2 \leq \varepsilon. \tag{22}$$

For the further robust analysis, we use the following lemma, whose proof can be found in [16]

Lemma 1 *Let $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ be given matrices, with $\boldsymbol{A} = \boldsymbol{A}^*$. The relation*

$$\boldsymbol{A} \succeq \boldsymbol{B}^*\boldsymbol{D}\boldsymbol{C} + \boldsymbol{C}^*\boldsymbol{D}^*\boldsymbol{B}, \quad \forall \boldsymbol{D}:\ \left\|\boldsymbol{D}\right\|_2 \leq \varepsilon, \tag{23}$$

is valid, if and only if

$$\exists \lambda \geq 0, \quad \begin{bmatrix} \boldsymbol{A} - \lambda \boldsymbol{C}^*\boldsymbol{C} & -\varepsilon \boldsymbol{B}^* \\ -\varepsilon \boldsymbol{B} & \lambda \boldsymbol{I} \end{bmatrix} \succeq \boldsymbol{0}. \tag{24}$$

Table 1 Iterative solution for the problem (9).

1: *initialization:* Set $t = 0$ (step number), the maximal number of iterations $t_{\max}$, the desired accuracy $\epsilon > 0$, the initial $\mathrm{MSE}_{\mathrm{new}} \gg 0$ and the initial receive filter $\boldsymbol{F}(z)$.
2: **repeat**
3: $t \leftarrow t+1$, $\mathrm{MSE}_{\mathrm{old}} \leftarrow \mathrm{MSE}_{\mathrm{new}}$
4: Solve the optimization problem

$$\min_{\boldsymbol{G}(z)} \max_{\|\boldsymbol{\Delta}\|_2 \leq \varepsilon} \mathrm{MSE}, \qquad \text{s.t.} \quad P \leq P_{\max}, \tag{$\mathcal{P}1$}$$

with the current values determining $\boldsymbol{F}(z)$ fixed.
5: Update the receive filter by solving the following problem with the fixed transmit filter $\boldsymbol{G}(z)$ obtained from $(\mathcal{P}1)$:

$$\mathrm{MSE}_{\mathrm{new}} = \min_{\boldsymbol{F}(z)} \max_{\|\boldsymbol{\Delta}\|_2 \leq \varepsilon} \mathrm{MSE} \tag{$\mathcal{P}2$}$$

6: **until** : $t > t_{\max}$ or $|\mathrm{MSE}_{\mathrm{new}} - \mathrm{MSE}_{\mathrm{old}}| \leq \epsilon$

Now we are equipped with the tolls which enable us to transform the uncertain problems $(\mathcal{P}1)$ and $(\mathcal{P}2)$ into tractable forms.

Theorem 1 *The uncertain problem $(\mathcal{P}1)$ is equivalent to a convex optimization problem*

$$\begin{aligned} &\min_{\boldsymbol{g},\tau,\lambda} \ \tau^2 + \sigma_w^2 \left\|\boldsymbol{f}\right\|_2^2, \quad \text{s.t.} \quad \left\|\boldsymbol{g}\right\|_2 \leq P_{\max}, \\ &\begin{bmatrix} \frac{\tau}{\sigma_s} - \lambda & \boldsymbol{\psi}^*(\boldsymbol{g}, \boldsymbol{f}) & \boldsymbol{0} \\ \boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) & \frac{\tau}{\sigma_s}\boldsymbol{I} & -\varepsilon \boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f}) \\ \boldsymbol{0} & -\varepsilon \boldsymbol{\Psi}^*(\boldsymbol{g}, \boldsymbol{f}) & \lambda \boldsymbol{I} \end{bmatrix} \succeq \boldsymbol{0}, \end{aligned} \tag{25}$$

where τ and λ are slack variables.

Proof: The objective function in (25) is obtained from (18), by introducing a slack variable τ, so that $\tau \geq \sigma_s \left\|\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) + \boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f})\boldsymbol{\delta}\right\|_2$. This constraint can be equivalently reformulated using the Schur complements lemma [14]

$$\begin{aligned} &\begin{bmatrix} \frac{\tau}{\sigma_s} & \boldsymbol{\psi}^*(\boldsymbol{g}, \boldsymbol{f}) \\ \boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) & \frac{\tau}{\sigma_s}\boldsymbol{I} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f}) \end{bmatrix} \boldsymbol{\delta} \begin{bmatrix} 1 & \boldsymbol{0} \end{bmatrix} \\ &+ \begin{bmatrix} 1 \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{\delta}^* \begin{bmatrix} \boldsymbol{0} & \boldsymbol{\Psi}^*(\boldsymbol{g}, \boldsymbol{f}) \end{bmatrix} \succeq \boldsymbol{0}. \end{aligned} \tag{26}$$

Now, Lemma 1 can be applied with

$$\begin{aligned} &\boldsymbol{A} = \begin{bmatrix} \frac{\tau}{\sigma_s} & \boldsymbol{\psi}^*(\boldsymbol{g}, \boldsymbol{f}) \\ \boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f}) & \frac{\tau}{\sigma_s}\boldsymbol{I} \end{bmatrix}, \quad \boldsymbol{D} = \boldsymbol{\delta}, \\ &\boldsymbol{B} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{\Psi}^*(\boldsymbol{g}, \boldsymbol{f}) \end{bmatrix}, \quad \boldsymbol{C} = -\begin{bmatrix} 1 & \boldsymbol{0} \end{bmatrix}. \end{aligned} \tag{27}$$

Remember that in $(\mathcal{P}1)$, the receiver, uniquely represented by the vector $\boldsymbol{f}$ is fixed. Since $\boldsymbol{\psi}(\boldsymbol{g}, \boldsymbol{f})$ and $\boldsymbol{\Psi}(\boldsymbol{g}, \boldsymbol{f})$ are affine in the unknown transmit filter coefficients, grouped in the vector $\boldsymbol{g}$, the second constraint in (25) is a convex constraint in the form of a linear matrix inequality (LMI) [8]. The objective function and the total transmit power constraint are clearly convex, which concludes the proof. ■

Beside being convex, it can be noticed that the problem (26) belongs to a class of semidefinite programming (SDP) problems, which are known to be solvable very efficiently using interior point methods [8, 17].

The problem $(\mathcal{P}2)$ can be solved using the same methodology. For an equivalent formulation of $(\mathcal{P}2)$, (25) can be used directly

with only two minor differences. Firstly, $\boldsymbol{g}$ will be fixed, and $\boldsymbol{f}$ should be optimized. Secondly, the total transmit power constraint can be omitted, since this condition will always remain satisfied after solving $(\mathcal{P}2)$, because $(\mathcal{P}2)$ updates the receiver only. These modifications clearly do not change the conclusions about the problem's convexity and solvability. Finally, we prove that the algorithm in Table 1 *always* converges.

Theorem 2 *The convergence of the algorithm listed in Table 1 is guaranteed for any choice of the initial receive equalizer.*

Proof: It can be noticed that $(\mathcal{P}1)$ is always feasible. Let $\mathrm{MSE}^{(1a)}$ be the achieved minimax MSE under the uncertainty after the employment of $(\mathcal{P}1)$ in the first step of the first iteration. The minimax problem $(\mathcal{P}2)$ is also solvable for any input parameters, and it will yield a new achievable MSE under the uncertainty, $\mathrm{MSE}^{(1)}_{\mathrm{new}}$. By the definition of $(\mathcal{P}2)$, $\mathrm{MSE}^{(1)}_{\mathrm{new}} \leq \mathrm{MSE}^{(1a)}$ must hold. In the second iteration, the transmit filter $\boldsymbol{G}(z)$ is recalculated with the updated receiver $\boldsymbol{F}(z)$. However, solving $(\mathcal{P}1)$ in the second iteration can only give the MSE which is smaller or equal to $\mathrm{MSE}^{(1)}_{\mathrm{new}}$. Therefore, the procedure can be continued in an iterative manner, and a monotonically decreasing sequence of MSEs (and the corresponding transceiver coefficients) is obtained. This sequence is bounded below (e.g., with $\mathrm{MSE} = 0$), which concludes the proof of the theorem. ■

Although the convergence to the global optimum cannot be asserted, it will be shown in Section 5 that the described iterative procedure performs very well in practice and enables significant gains comparing to the optimization of the precoder only. Finally, we remark that the algorithm from Table 1 is meant to be implemented by the transmitter that possesses only the imperfect CSI. The receiver either obtains the coefficients of $\boldsymbol{F}(z)$ from the transmitter in an initial phase of communication over one channel using the presumably error-free control channels, or, after the transmitter determines its precoder, estimates (perfectly) the equivalent channel $\boldsymbol{H}(z)\boldsymbol{G}(z)$ and adopts the standard Wiener solution for $\boldsymbol{F}(z)$, which can only additionally reduce the MSE.

4. RELATED MSE-OPTIMIZATION PROBLEMS

In this section, we focus firstly on the problem of transmit power minimization with respect to a fixed MSE target μ, that must be satisfied despite the uncertainty

$$\min_{\boldsymbol{G}(z),\,\boldsymbol{F}(z)} P, \qquad \text{s.t.} \quad \mathrm{MSE} \leq \mu, \quad \forall \left\|\boldsymbol{\Delta}\right\|_2 \leq \varepsilon. \tag{28}$$

The problem (28) can be approached with a slight modification of the iterative algorithm proposed in Section 3.

Theorem 3 *The problem (28) can be solved using the algorithm from Table 1 with the following changes:*

- *Except in the problem $(\mathcal{P}2)$, P_{old} and P_{new} should take places of $\mathrm{MSE}_{\mathrm{old}}$ and $\mathrm{MSE}_{\mathrm{new}}$, respectively.*
- *Instead of $(\mathcal{P}1)$, the optimization problem*

$$\mathrm{P}_{\mathrm{new}} = \min_{\boldsymbol{G}(z)} P, \qquad \text{s.t.} \quad \mathrm{MSE} \leq \mu, \quad \forall \left\|\boldsymbol{\Delta}\right\|_2 \leq \varepsilon. \tag{29}$$

should be solved with a fixed receive equalizer $\boldsymbol{F}(z)$.

The problem (29) has an efficiently solvable SDP representation. If the problem (29) is feasible in the first step of the first iteration, the convergence of the algorithm is guaranteed.

Proof: An equivalent SDP problem for (29): $\min_{\boldsymbol{g},\tau,\lambda} \left\|\boldsymbol{g}\right\|_2^2$, subject to a convex constraint $\tau^2 + \sigma_w^2 \left\|\boldsymbol{f}\right\|_2^2 \leq \mu$ and the same LMI as in (25), is derived in a similar manner as in Theorem 1. Under a technical condition that the first optimization problem is feasible (otherwise the algorithm stops immediately), the convergence of the algorithm can be proved by noticing that it yields a monotonically decreasing sequence of transmit powers (and the corresponding transceiver coefficients), which is clearly bounded below. Notice that the update of the receivers in $(\mathcal{P}2)$ does not change the transmit power, but it has beneficial influence since it reduces the MSE, having in mind the uncertainty. After one complete iteration, (29) operates with an updated receive filter, and the new resulting transmit power must be smaller or equal to the power obtained after the previous solution of (29). ■

The balancing problems, where the layers are treated as individual data streams and the maximum over the layers' MSEs, $\mathrm{MSE}_l = \mathrm{E}\left\{|s_l(n) - \hat{s}_l(n)|^2\right\}$, should be minimized under the uncertainty

$$\min_{\boldsymbol{F}(z),\,\boldsymbol{G}(z)} \max_{\substack{\|\boldsymbol{\Delta}\|_2 \leq \varepsilon,\\ 1 \leq l \leq L}} \mathrm{MSE}_l, \quad \text{s.t. } P \leq P_{\max}, \tag{30}$$

are also readily supported with small modifications of the presented framework. Furthermore, per-antenna power constraints are easily included in all problem formulations, as they are second-order cone constraints $\sigma_s \left\| \left[\begin{array}{ccc} \boldsymbol{G}_{0_{(m,:)}} & \cdots & \boldsymbol{G}_{N_{G(m,:)}} \end{array} \right]^T \right\|_2 \leq \sqrt{P_m}$, where P_m is the power limit for antenna m. Due to the lack of space, we leave the detailed analysis of these issues for a future, more comprehensive work on this topic.

5. NUMERICAL EXAMPLES

The numerical examples in this section are obtained using SeDuMi [18]. System parameters $\sigma_s^2 = 1$, $M = 3$, $K = 2$, $N_G = 3$, $N_H = 2$, $N_F = 1$ and $\epsilon = 10^{-3}$, are kept fixed in all simulations. The initial receive equalizer had always $\boldsymbol{F}_0 = \boldsymbol{I}$, and the other taps were zero. We remark that any, e.g., non-robustly designed transceiver can also be selected as a starting point and leave the issue of smart initialization as an interesting open topic for the future work.

In Fig. 2, we plot the resulting minimax MSE versus the bound on the uncertainty ε in (5), as a solution of the algorithm from Table 1. The simulations present an average over 10^2 channel realizations, where real and imaginary parts of each channel coefficient had distributions $\mathcal{N}\left(0, \frac{1}{2}\right)$. Signal-to-Noise (SNR) scenarios 1 and 2, with $P_{\max} = 5$, $\sigma_w^2 = 0.1$ and $P_{\max} = 10$, $\sigma_w^2 = 0.01$, respectively, were considered. It can be seen the introduction of only one additional tap at the receive side, with the optimization of the whole system, renders significant gains comparing to the precoder optimization alone.

The necessary number of iterations for the termination of the algorithm ranged from 13.74 to 15.52 for the scenario 1, and from 3.54 to 7.32 for the scenario 2, depending on ε. From Fig. 2, it can be seen that already after the second iteration (or, in other words, with a fixed complexity), the performance improvement is conspicuous. We note that the complexity of the complete iterative algorithm is practically of the same order as the optimization of the transmitter only (one application of $(\mathcal{P}1)$), however, with a significantly larger multiplicative constant.

Finally, we plot in Fig. 3 the benefits in reducing the minimum transmit power as a function of the MSE target μ for the problem (28), with $\varepsilon = 0.1$ and $\sigma_w^2 = 0.01$. The averaging is performed over 10^2 randomly-chosen channels which yielded feasible problems for the observed range of μ.

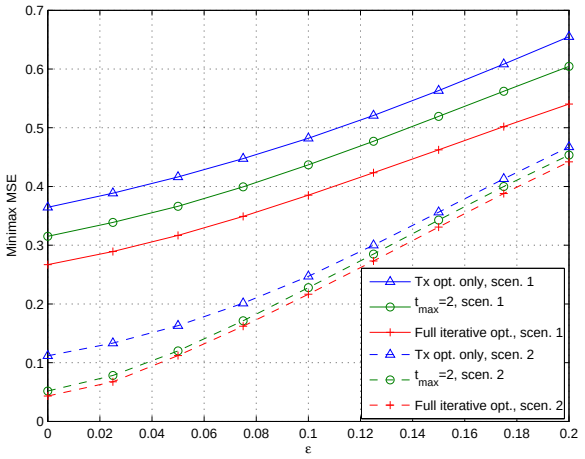

Fig. 2. Minimax MSE vs. the bound on the uncertainty ε.

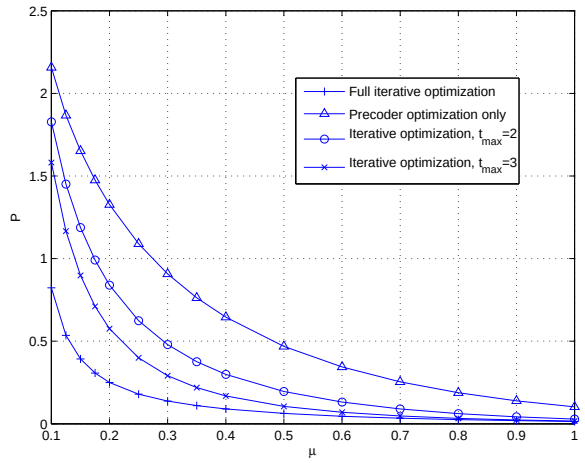

Fig. 3. Minimal transmit power vs. the MSE constraint μ.

6. CONCLUSION

An iterative framework, that is able to support a broad range of robust transceiver MSE-optimization problems in a convolutional, frequency selective, FIR MIMO setting, is proposed. Due to the intricacy of the problems, the convergence to the global optimum is unfortunately not guaranteed. However, a significantly better performance in practice is observed, with a moderate increase in computational complexity, comparing to the optimization only on one side. The gains seem to be important even with an imposed low maximum number of iterations.

7. REFERENCES

[1] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, "Great expectations: The value of spatial diversity in wireless networks," *Proc. of the IEEE*, vol. 92, no. 2, pp. 219–270, Feb. 2004.

[2] M. L. Honig, P. Crespo, and K. Steiglitz, "Suppression of near- and far-end crosstalk by linear pre- and post-filtering," *IEEE J. Select. Areas Commun.*, vol. 10, no. 3, pp. 614–629, Apr. 1992.

[3] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[4] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1051–1064, May 2002.

[5] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 1037–1052, Apr. 2004.

[6] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?," *IEEE Comm. Mag.*, vol. 42, pp. 54–59, Oct. 2004.

[7] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. of the IEEE*, vol. 73, no. 3, pp. 433–482, Mar. 1985.

[8] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, pp. 49–95, Mar. 1996.

[9] N. Vucic and H. Boche, "Robust minimax equalization of imperfectly known frequency selective MIMO channels," in *Proc. Asilomar 2007*, Pacific Grove, CA, USA, Nov. 2007.

[10] Y. Guo and B. C. Levy, "Robust MSE equalizer design for MIMO communication systems in the presence of model uncertainties," *IEEE Trans. Signal Processing*, vol. 54, no. 5, pp. 1840–1852, May 2006.

[11] A. Beck, Y. C. Eldar, and A. Ben-Tal, "Mean-squared error estimation for linear systems with block circular uncertainty," *SIAM J. Matrix Anal.*, vol. 29, no. 3, pp. 712–730, 2007.

[12] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, NY, USA, 1991.

[13] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15 of *Studies in Applied Mathematics*, SIAM, PA, USA, June 1994.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, NY, USA, 1985.

[15] A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Processing*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[16] Y. C. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Trans. Signal Processing*, vol. 52, no. 7, pp. 1931–1946, July 2004.

[17] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: Analysis, Algorithms, and Engineering Applications*, MPS-SIAM Series on Optimization, PA, USA, 2001.

[18] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB tool for optimization over symmetric cones," in *Optimization Methods and Software*, 1999, vol. 11-12, pp. 625–653.