ROBUST MMSE PRECODING FOR THE MIMO COMPLEX GAUSSIAN BROADCAST CHANNEL

Jonathan Duplicy, Luc Vandendorpe

Université catholique de Louvain
Communications and remote sensing laboratory
B-1348 Louvain-la-Neuve, Belgium
{jonathan.duplicy, luc.vandendorpe}@uclouvain.be

ABSTRACT

This paper addresses the design of the linear precoders and decoders of the complex Gaussian broadcast channel in which both the base station and the remote station are equipped with arrays of multiple antennas. An imperfect channel knowledge is assumed at the base station and the Minimization of the sum of the Mean Square Errors (MMSE) of the system's substreams is chosen as optimization criterion. A stochastic approach is taken to make the design robust against the channel estimation errors. The solution is based on an iterative algorithm whose convergence is guaranteed. Simulations results emphasize the benefit of the proposed design.

***Index Terms*—** Broadcast channel, MIMO, Least mean square methods, Robustness.

1. INTRODUCTION

Since the seminal works of Alamouti and Tarokh, Multiple Input Multiple Output (MIMO) systems have attracted an increasing interest as they allow a dramatic capacity improvement in comparison with single antenna systems. Among others, the broadcast channel has recently captured a lot of attention. Although, the well-known Costa's dirty paper coding scheme has been proven to achieve the channel capacity [1], it faces serious implementation issues due to its high complexity. Linear schemes are therefore seen as a suboptimal alternative for practical systems.

In this paper, we address the design of the linear precoders and decoders of the MIMO complex Gaussian broadcast channel with the minimization of the sum of the mean square errors of the system' substreams as objective. With the assumption of a perfect Channel State Information (CSI) available at the base station, such MMSE criterion has already been optimally solved in [2] making use of the uplink-downlink duality. Furthermore, a suboptimal iterative design has been proposed in [3]. Although suboptimal, this solution iterates between pre/decoding *closed-formed* designs which makes it appealing for an implementation point of view.

However, a perfect CSI available at the transmit side is impossible to obtain in real systems and one has to account for estimation errors. Two strategies are commonly used to deal with them: worst case and stochastic. The first strategy aims at designing the beamformers such that the specifications will be fulfilled provided the estimation errors are lower than a given level [4]. Although interesting for QoS-constrained (Quality of Service) systems, this approach often leads to too much conservative designs. The second strategy exploits the knowledge of some statistics of the channels and the estimation errors to deal with the channel uncertainties. Such a stochastic approach has been used in [5] and [6] for the design of multicarrier single-user precoders. The single receive antenna MMSE stochastic design has been proposed in [7]. In this work, we take this second strategy to design the MMSE precoders and decoders robust to the channel estimation errors. As these precoders rely on the decoders and vice versa, we use an iterative algorithm whose convergence is guaranteed. Simulations emphasize the benefit of the proposed robust scheme in comparison with a naive design built by considering the estimated channels as the actual ones.

The following notations will be used. Matrices and vectors are represented with bold capital letters and bold lowercase letters, respectively. $\boldsymbol{A}^*$, $\boldsymbol{A}^T$ and $\boldsymbol{A}^\dagger$ are the conjugate, transpose and Hermitian of matrix $\boldsymbol{A}$. Moreover, its trace is written as $\mathrm{tr}\{\boldsymbol{A}\}$ whereas $[A]_{ij}$ denotes the element in row i of column j. Finally, $\boldsymbol{I}_r$ is the identity matrix of size r and $E\{.\}$ is the mathematical expectation operator.

2. SYSTEM MODEL

The system model we consider is depicted in figure 1 for an exemplary two-user system. A base station equipped with M transmit antennas broadcasts to a set of K remote stations equipped with arrays of N_k receive antennas (in which k is the user index). Linear processing is assumed at both the transmit and receive sides. Furthermore, we assume that several symbols can be simultaneously transmitted to the *same* user. Hence, e.g., the l^{th} symbol of user k ($l \in [1, \ldots, L_k]$, with

Jonathan Duplicy would like to thank the Belgian fund FRIA for its financial support. This work has been also partly achieved in the context of the CELTIC WISQUAS project, funded by BELSPO.

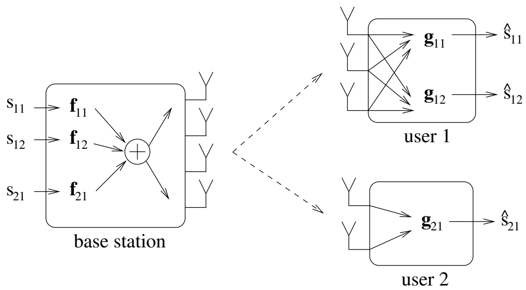

Fig. 1. Exemplary system: K=2; M=4; $N_{1,2}$=3, 2; $L_{1,2}$=2, 1.

$L_k \leq N_k$ and $\sum_u L_u \leq M$) is estimated as follows:

$$\hat{s}_{kl} = \frac{1}{\eta} \boldsymbol{g}_{kl} \left(\boldsymbol{H}_k \boldsymbol{f}_{kl}\, s_{kl} + \boldsymbol{H}_k \sum_{\substack{u=1 \\ um \neq kl}}^{K} \sum_{m=1}^{L_u} \boldsymbol{f}_{um} s_{um} + \boldsymbol{\nu}_k \right) \quad (1)$$

in which $\boldsymbol{\nu}_k$ is a zero-mean complex white Gaussian noise with correlation matrix $\boldsymbol{R}_k^{\boldsymbol{\nu}}$. The $(M \times 1)$ precoder (transmit beamformer) associated with symbol s_{kl} is denoted as $\boldsymbol{f}_{kl}$, whereas $\boldsymbol{g}_{kl}$ is the $(1 \times N_k)$ corresponding decoder (receive beamformer). Symbols are assumed to be of unitary variance, non correlated between each other and non correlated with the noise. Moreover, $\boldsymbol{H}_k$ is the $(N_k \times M)$ complex Gaussian channel matrix of user k. Finally, scalar η is a scaling factor common to all substreams. Equation (1) emphasizes that the symbol estimates suffer from the additive noise, the interference of the symbols of the other users but also from the other symbols of the *same* user (self-interference). Furthermore, the transmit power at the base station is globally constrained:

$$\sum_{k=1}^{K} \sum_{l=1}^{L_k} \operatorname{tr}\left\{ \boldsymbol{f}_{kl} \boldsymbol{f}_{kl}^{\dagger} \right\} \leq P. \quad (2)$$

Assuming a perfect (section 3) or imperfect (section 4) channel knowledge, our goal is to design the precoders and decoders so as to minimize the sum (or equivalently, the arithmetic mean) of the MSEs of the $\sum_{k=1}^{K} L_k$ substreams. The optimization is assumed to be conducted at the base station which sends (through control channels) the decoders to the remote stations whose complexity can then be reduced.

3. PERFECT CSI-BASED MMSE DESIGN

The perfect CSI-based iterative MMSE design has been derived in [8] for the MISO case and for the MIMO case in [3]. Here, we briefly recall the first method as it will be the basis of the robust design we propose in the next section. Assuming a perfect CSI available, the channels are treated as deterministic and the optimization problem is stated as:

$$\min_{\substack{\boldsymbol{f}_{um},\, \boldsymbol{g}_{um},\, \eta \\ \forall u,m}} \sum_{k=1}^{K} \sum_{l=1}^{L_k} MSE_{kl} \qquad \text{s.t. } (2), \quad (3)$$

in which MSE_{kl} is the mean square error of substream kl:

$$MSE_{kl} = E_{s_{kl}, \boldsymbol{\nu}_k} \left\{ \| \hat{s}_{kl} - s_{kl} \|^2 \right\}, \quad (4)$$

where the expectation is to be taken over the symbols and the additional noise. Thanks to the assumptions of section 2, (4) can be rewritten as:

$$\begin{aligned} MSE_{kl} \;\; &= \;\; 1 + \frac{1}{\eta^2} \boldsymbol{g}_{kl} \boldsymbol{H}_k \left(\sum_{u=1}^{K} \sum_{m=1}^{L_j} \boldsymbol{f}_{um} \boldsymbol{f}_{um}^{\dagger} \right) \boldsymbol{H}_k^{\dagger} \boldsymbol{g}_{kl}^{\dagger} \\ &+ \frac{1}{\eta^2} \boldsymbol{g}_{kl} \boldsymbol{R}_k^{\boldsymbol{\nu}} \boldsymbol{g}_{kl}^{\dagger} - \frac{1}{\eta} \left(\boldsymbol{g}_{kl} \boldsymbol{H}_k \boldsymbol{f}_{kl} + \boldsymbol{f}_{kl}^{\dagger} \boldsymbol{H}_k^{\dagger} \boldsymbol{g}_{kl}^{\dagger} \right). \quad (5) \end{aligned}$$

Using a classical Lagrangian-based optimization, the optimal pre/decoders are easily derived and given by [3][8]:

$$\begin{cases} \boldsymbol{f}_{kl}^{\text{opt}} = \eta \left[\mathbf{A} + \alpha \boldsymbol{I}_M \right]^{-1} \boldsymbol{H}_k^{\dagger} \boldsymbol{g}_{kl}^{\dagger} \\ \boldsymbol{g}_{kl}^{\text{opt}} = \eta \boldsymbol{f}_{kl}^{\dagger} \boldsymbol{H}_k^{\dagger} \left[\boldsymbol{H}_k \left(\sum_{u=1}^{K} \sum_{m=1}^{L_u} \boldsymbol{f}_{um} \boldsymbol{f}_{um}^{\dagger} \right) \boldsymbol{H}_k^{\dagger} + \boldsymbol{R}_k^{\boldsymbol{\nu}} \right]^{-1} \end{cases} \quad (6)$$

with

$$\begin{cases} \boldsymbol{A} = \sum_{u=1}^{K} \boldsymbol{H}_u^{\dagger} \left(\sum_{m=1}^{L_u} \boldsymbol{g}_{um}^{\dagger} \boldsymbol{g}_{um} \right) \boldsymbol{H}_u \\ \alpha = \frac{1}{P} \left(\sum_{u=1}^{K} \sum_{m=1}^{L_u} \boldsymbol{g}_{um} \boldsymbol{R}_u^{\boldsymbol{\nu}} \boldsymbol{g}_{um}^{\dagger} \right) \\ \eta = \sqrt{\frac{P}{\operatorname{tr}\{[\mathbf{A} + \alpha \boldsymbol{I}_M]^{-1} \mathbf{A} [\mathbf{A} + \alpha \boldsymbol{I}_M]^{-1}\}}} \end{cases} \quad (7)$$

The algorithm consists in iterating between the transmit and receive designs until convergence. As each step of the algorithm decreases the lower bounded objective function ($\sum_k \sum_l MSE_{kl} \geq 0$), the convergence is guaranteed. However, as local optima can be reached, this iterative strategy is not optimal. Nevertheless, as it relies on closed-formed designs and as the convergence speed is high, this scheme can be interesting for practical implementations.

4. STOCHASTIC ROBUST DESIGN

Although the above perfect CSI assumption is theoretically appealing, quantization errors, imperfect estimation at the receive side, feedback delay,... make it unreachable for practical systems. Of course, one could treat the channel estimates as perfect and use the pre/decoders (6). However, this approach is not optimal if one has access to some statistics of the channels. In this work, we assume that the means ($\boldsymbol{\mu}$) and covariances ($\boldsymbol{C}$) of the channels and the estimation noises are perfectly available at the base station. This assumption makes sense since these parameters vary slowly and are therefore easier to estimate than the instantaneous channel. Hence, the MMSE design turns into the minimization of the *expected* (over the channels) mean square error of the system[1]:

$$\min_{\substack{\boldsymbol{f}_{um},\, \boldsymbol{g}_{um},\, \eta \\ \forall u,m}} \sum_{k=1}^{K} \sum_{l=1}^{L_k} EMSE_{kl} \qquad \text{s.t. } (2), \quad (8)$$

where $EMSE_{kl}$ is the conditional expectation of MSE_{kl}:

$$EMSE_{kl} = E_{\boldsymbol{H}_k | \hat{\boldsymbol{H}}_k, \boldsymbol{\mu}_{\boldsymbol{H}_k}, \boldsymbol{C}_{\boldsymbol{H}_k}} \left\{ MSE_{kl} \right\} \quad (9)$$

The estimated channel is assumed to be the addition of the actual channel and an estimation noise:

$$\hat{\boldsymbol{H}}_k = \boldsymbol{H}_k + \boldsymbol{W}_k. \quad (10)$$

[1] Note that we assume the statistics of the additive white noise perfectly known as it is a slow varying parameter. However, the extension to an imperfect knowledge is easy.

We row-wise vectorize these matrices into $(MN_k \times 1)$vectors:

$$\hat{\boldsymbol{h}}_k = \boldsymbol{h}_k + \boldsymbol{w}_k. \tag{11}$$

The actual channel as well as the estimation noise are assumed to be jointly complex Gaussian: $\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}_{\boldsymbol{h}_k}, \boldsymbol{C}_{\boldsymbol{h}_k})$ and $\boldsymbol{w}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}_{\boldsymbol{w}_k}, \boldsymbol{C}_{\boldsymbol{w}_k})$ [9]. Furthermore, we assume the independence between $\boldsymbol{h}_k$ and $\boldsymbol{w}_k$. Hence, $\hat{\boldsymbol{h}}_k$ is complex Gaussian distributed: $\hat{\boldsymbol{h}}_k \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}_{\boldsymbol{h}_k} + \boldsymbol{\mu}_{\boldsymbol{w}_k}, \boldsymbol{C}_{\boldsymbol{h}_k} + \boldsymbol{C}_{\boldsymbol{w}_k})$. Furthermore, from [9], the conditional channel $(\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k, \boldsymbol{\mu}_{\boldsymbol{h}_k}, \boldsymbol{C}_{\boldsymbol{h}_k}, \boldsymbol{\mu}_{\boldsymbol{w}_k}, \boldsymbol{C}_{\boldsymbol{w}_k})$, is also Gaussian distributed with mean and covariance written like[2]:

$$\left\{ \begin{array}{l} \boldsymbol{\mu}_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} = \boldsymbol{\mu}_{\boldsymbol{h}_k} + \boldsymbol{C}_{\boldsymbol{h}_k\hat{\boldsymbol{h}}_k} \boldsymbol{C}^{-1}_{\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k} \left(\hat{\boldsymbol{h}}_k - \boldsymbol{\mu}_{\hat{\boldsymbol{h}}_k}\right) \\ \boldsymbol{C}_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} = \boldsymbol{C}_{\boldsymbol{h}_k\boldsymbol{h}_k} - \boldsymbol{C}_{\boldsymbol{h}_k\hat{\boldsymbol{h}}_k} \boldsymbol{C}^{-1}_{\hat{\boldsymbol{h}}_k\hat{\boldsymbol{h}}_k} \boldsymbol{C}_{\hat{\boldsymbol{h}}_k\boldsymbol{h}_k} \end{array} \right. \tag{12}$$

As in [5], and thanks to the matrix inversion lemma, we rewrite (12) in the following smarter way (with $\boldsymbol{C}_{\boldsymbol{x}} = \boldsymbol{C}_{\boldsymbol{x}\boldsymbol{x}}$):

$$\left\{ \begin{array}{l} \boldsymbol{\mu}_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} = \boldsymbol{\mu}_{\boldsymbol{h}_k} + \boldsymbol{C}_{\boldsymbol{h}_k} \left(\boldsymbol{C}_{\boldsymbol{h}_k} + \boldsymbol{C}_{\boldsymbol{w}_k}\right)^{-1} \left(\hat{\boldsymbol{h}}_k - \boldsymbol{\mu}_{\hat{\boldsymbol{h}}_k}\right) \\ \boldsymbol{C}_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} = \left(\boldsymbol{C}^{-1}_{\boldsymbol{h}_k} + \boldsymbol{C}^{-1}_{\boldsymbol{w}_k}\right)^{-1} \end{array} \right. \tag{13}$$

In addition, if as in [5], the estimation noise is zero-mean and small ($\boldsymbol{C}_{\boldsymbol{w}_k} \ll \boldsymbol{C}_{\boldsymbol{h}_k}$), (13) can be approximated as:

$$\left\{ \begin{array}{l} \boldsymbol{\mu}_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \simeq \hat{\boldsymbol{h}}_k \\ \boldsymbol{C}_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \simeq \boldsymbol{C}_{\boldsymbol{w}_k} \end{array} \right. \tag{14}$$

Given the above statistics, let's compute $EMSE_{kl}$, the expected mean square error of the l^{th} substream of user k:

$$\begin{aligned} EMSE_{kl} &= 1 + \frac{1}{\eta^2} \boldsymbol{g}_{kl} \boldsymbol{R}^{\boldsymbol{\nu}}_k \boldsymbol{g}^\dagger_{kl} \\ &- \frac{1}{\eta} E_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \left\{ \boldsymbol{g}_{kl} \boldsymbol{H}_k \boldsymbol{f}_{kl} + \boldsymbol{f}^\dagger_{kl} \boldsymbol{H}^\dagger_k \boldsymbol{g}^\dagger_{kl} \right\} \\ &+ \frac{1}{\eta^2} E_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \left\{ \boldsymbol{g}_{kl} \boldsymbol{H}_k \left(\sum_{u=1}^{K} \sum_{m=1}^{L_j} \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \right) \boldsymbol{H}^\dagger_k \boldsymbol{g}^\dagger_{kl} \right\} \end{aligned} \tag{15}$$

The first expectation is obvious to compute and is written as:

$$E_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \left\{ \boldsymbol{g}_{kl} \boldsymbol{H}_k \boldsymbol{f}_{kl} + \boldsymbol{f}^\dagger_{kl} \boldsymbol{H}^\dagger_k \boldsymbol{g}^\dagger_{kl} \right\} = \boldsymbol{g}_{kl} \boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \boldsymbol{f}_{kl} + \boldsymbol{f}^\dagger_{kl} \boldsymbol{\mu}^\dagger_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \boldsymbol{g}^\dagger_{kl}, \tag{16}$$

where $\boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k}$ is the non vectorized version of (12). Let's now focus on the second expectation in (15) and start by turning the matrix products into sums:

$$\begin{aligned} a_{kl} &= E_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \left\{ \boldsymbol{g}_{kl} \boldsymbol{H}_k \left(\sum_{u=1}^{K} \sum_{m=1}^{L_u} \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \right) \boldsymbol{H}^\dagger_k \boldsymbol{g}^\dagger_{kl} \right\} \\ &= \boldsymbol{g}_{kl} E_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \left\{ \boldsymbol{H}_k \left(\sum_{u=1}^{K} \sum_{m=1}^{L_u} \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \right) \boldsymbol{H}^\dagger_k \right\} \boldsymbol{g}^\dagger_{kl} \\ &= \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} [\boldsymbol{g}_{kl}]_i \left(\sum_{u=1}^{K} \sum_{m=1}^{L_u} [\boldsymbol{B}_{kum}]_{ij} \right) [\boldsymbol{g}^*_{kl}]_j \, , \end{aligned} \tag{17}$$

with $\boldsymbol{B}_{kum} = E_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \left\{ \boldsymbol{H}_k \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \boldsymbol{H}^\dagger_k \right\}$. We now develop the ij^{th} element. Rewriting it as sums and using classical stochastic algebra, the following identities can be obtained:

$$\begin{aligned} [\boldsymbol{B}_{kum}]_{ij} &= E_{\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k} \left\{ \sum_{q=1}^{M} \sum_{n=1}^{M} [\boldsymbol{H}_k]_{iq} \left[\boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um}\right]_{qn} \left[\boldsymbol{H}^\dagger_k\right]_{nj} \right\} \\ &= \int p(\boldsymbol{h}_k|\hat{\boldsymbol{h}}_k) \sum_{q=1}^{M} \sum_{n=1}^{M} [\boldsymbol{H}_k]_{iq} \left[\boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um}\right]_{qn} \left[\boldsymbol{H}^\dagger_k\right]_{nj} d\boldsymbol{h}_k \\ &= \sum_{q=1}^{M} \sum_{n=1}^{M} \left[\boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um}\right]_{qn} \int p\left([\boldsymbol{H}_k]_{iq}; [\boldsymbol{H}_k]_{jn} \left| \hat{\boldsymbol{h}}_k \right. \right) [\boldsymbol{H}_k]_{iq} \, [\boldsymbol{H}^*_k]_{jn} d\boldsymbol{h}_k \\ &= \sum_{q=1}^{M} \sum_{n=1}^{M} \left[\boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um}\right]_{qn} \left(\mu_{[\boldsymbol{H}_k]_{iq}|\hat{\boldsymbol{h}}_k} \cdot \mu^*_{[\boldsymbol{H}_k]_{jn}|\hat{\boldsymbol{h}}_k} + \boldsymbol{C}_{[\boldsymbol{H}_k]_{iq}[\boldsymbol{H}_k]_{jn}|\hat{\boldsymbol{h}}_k} \right) \\ &= \left[\boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \boldsymbol{\mu}^\dagger_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \right]_{ij} + \mathrm{tr} \left\{ \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \boldsymbol{C}^T_{[\boldsymbol{H}_k]_{i:}[\boldsymbol{H}_k]_{j:}|\hat{\boldsymbol{h}}_k} \right\} \end{aligned} \tag{18}$$

in which $\boldsymbol{C}_{[\boldsymbol{H}_k]_{i:}[\boldsymbol{H}_k]_{j:}|\hat{\boldsymbol{h}}_k}$ is the conditional covariance matrix between the i^{th} and j^{th} lines of $\boldsymbol{H}_k$. Hence, (17) can be rewritten as:

$$a_{kl} = \boldsymbol{g}_{kl} \left(\boldsymbol{D}_k + \boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \left(\sum_{u=1}^{K} \sum_{m=1}^{L_u} \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \right) \boldsymbol{\mu}^\dagger_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \right) \boldsymbol{g}^\dagger_{kl} \tag{19}$$

where $\boldsymbol{D}_k$ is a $(N_k \times N_k)$ matrix with ij^{th} element given by:

$$[\boldsymbol{D}_k]_{ij} = \mathrm{tr} \left\{ \left(\sum_{u=1}^{K} \sum_{m=1}^{L_u} \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \right) \boldsymbol{C}^T_{[\boldsymbol{H}_k]_{i:}[\boldsymbol{H}_k]_{j:}|\hat{\boldsymbol{h}}_k} \right\} \tag{20}$$

Therefore, (15) can be eventually rewritten as follows:

$$\begin{aligned} EMSE_{kl} &= 1 + \frac{1}{\eta^2} \boldsymbol{g}_{kl} \boldsymbol{R}^{\boldsymbol{\nu}}_k \boldsymbol{g}^\dagger_{kl} \\ &- \frac{1}{\eta} \left(\boldsymbol{g}_{kl} \boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \boldsymbol{f}_{kl} + \boldsymbol{f}^\dagger_{kl} \boldsymbol{\mu}^\dagger_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \boldsymbol{g}^\dagger_{kl} \right) \\ &+ \frac{1}{\eta^2} \boldsymbol{g}_{kl} \left(\boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} \left(\sum_{u=1}^{K} \sum_{m=1}^{L_u} \boldsymbol{f}_{um} \boldsymbol{f}^\dagger_{um} \right) \boldsymbol{\mu}^\dagger_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k} + \boldsymbol{D}_k \right) \boldsymbol{g}^\dagger_{kl} \end{aligned} \tag{21}$$

Note the similarities between (21) and (5). The robust case is obtained by replacing the actual channel by its conditional mean and by adding an extra term ($\boldsymbol{D}_k$) which takes into account the channel correlations.

[2]To lighten the notations, the next conditioned quantities will be written like "$.|\boldsymbol{h}_k$" but note that their also implicitly conditioned on the means and covariances of the channels and the estimation noises.

As for the perfect CSI case, problem (8) is solved with the help of a Lagrangian optimization. This gives rise to the hereafter optimal pre/decoders:

$$\begin{cases} \boldsymbol{f}_{kl}^{\text{opt}} = \eta\left[\mathbf{A}+\mathbf{B}+\alpha\boldsymbol{I}_M\right]^{-1}\boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k}^{\dagger}\boldsymbol{g}_{kl}^{\dagger} \\ \boldsymbol{g}_{kl}^{\text{opt}^{\dagger}} = \eta\boldsymbol{f}_{kl}^{\dagger}\boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k}^{\dagger}\left[\boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k}\left(\sum_{u=1}^{K}\sum_{m=1}^{L_u}\boldsymbol{f}_{um}\boldsymbol{f}_{um}^{\dagger}\right)\boldsymbol{\mu}_{\boldsymbol{H}_k|\hat{\boldsymbol{h}}_k}^{\dagger}\right. \\ \qquad\qquad \left.+\boldsymbol{R}_k^{\boldsymbol{\nu}}+\text{real}\{\boldsymbol{D}_k\}\right]^{-1} \end{cases} \tag{22}$$

with

$$\begin{cases} \boldsymbol{A} = \sum_{u=1}^{K}\boldsymbol{\mu}_{\boldsymbol{H}_u|\hat{\boldsymbol{h}}_u}^{\dagger}\left(\sum_{m=1}^{L_u}\boldsymbol{g}_{um}^{\dagger}\boldsymbol{g}_{um}\right)\boldsymbol{\mu}_{\boldsymbol{H}_u|\hat{\boldsymbol{h}}_u} \\ \boldsymbol{B} = \sum_{u=1}^{K}\sum_{m=1}^{L_u}\left(\boldsymbol{g}_{um}\otimes\boldsymbol{I}_M\right)\boldsymbol{Q}_u\left(\boldsymbol{g}_{um}^{\dagger}\otimes\boldsymbol{I}_M\right) \\ \alpha = \frac{1}{P}\left(\sum_{u=1}^{K}\sum_{m=1}^{L_u}\boldsymbol{g}_{um}\boldsymbol{R}_u^{\boldsymbol{\nu}}\boldsymbol{g}_{um}^{\dagger}\right) \\ \eta = \sqrt{\frac{P}{\text{tr}\{[\mathbf{A}+\mathbf{B}+\alpha\boldsymbol{I}_M]^{-1}\mathbf{A}[\mathbf{A}+\mathbf{B}+\alpha\boldsymbol{I}_M]^{-1}\}}} \end{cases} \tag{23}$$

and where $\boldsymbol{Q}_u$ is a $(MN_u\times MN_u)$ block matrix with block ij given by $\boldsymbol{C}^T_{[\boldsymbol{H}_u]_{i:}[\boldsymbol{H}_u]_{j:}|\hat{\boldsymbol{h}}_u}+\boldsymbol{C}^*_{[\boldsymbol{H}_u]_{i:}[\boldsymbol{H}_u]_{j:}|\hat{\boldsymbol{h}}_u}$ $(i,j=1...N_u)$ whereas $\otimes$ is the Kronecker product. As the optimal precoders rely on the optimal decoders and vice versa, we resort to a two-step loop iterative algorithm. Each step optimizes one set of variables (i.e. the precoders or the decoders) while considering the other variables as fixed values. As each step of the algorithm monotonously decreases the lower bounded objective function ($\sum_k\sum_l EMSE_{kl}\geq 0$), the algorithm is guaranteed to converge. Nevertheless, note that although each step of the algorithm is convex, the stationary point reached is not assured to be *globally* optimal. Finally, note that for our simulations, we have initialized the algorithm with random decoders.

5. SIMULATION RESULTS

Figure 2 illustrates a Monte-Carlo simulation for a two-user system in which the base station is equipped with four antennas and sends two streams to each remote station (both making use of two receive antennas). The two users undergo the same additive noise variance (σ_ν^2) and zero-mean unitary non correlated channels: $\boldsymbol{h}_k\sim\mathcal{N}_{\mathbb{C}}(\mathbf{0},\boldsymbol{I}_{MN_k})\ \forall k$. The channel estimation noise statistics are: $\boldsymbol{w}_k\sim\mathcal{N}_{\mathbb{C}}(\mathbf{0},\sigma_{\boldsymbol{w}}^2\boldsymbol{I}_{N_kM})\ \ \forall k$, with two simulated levels of variance: $\sigma_{\boldsymbol{w}}^2=0.03, 0.07$. In addition to the perfect CSI curve, three curves are plotted for each level: the "naive" curve uses precoders (6) considering the estimated channels as the actual ones, the "robust_app" and "robust" curves use the robust designs (22) with the approximated (14) and actual (13) conditional means and covariances, respectively. The figure clearly shows the gain obtained with our robust design in comparison with a naive design, especially at high P/σ_ν^2 values[3]. As expected, the approximated robust design tends to the true robust design as the estimation noise becomes smaller. Finally, note that the sum-MSE of the naive design *increases* at high P/σ_ν^2 values. This comes from that, as the noise variance decreases, the diagonal loading term bringing robustness decreases.

[3]Notice that we have considered the same estimation error level for the whole range of P/σ_ν^2 values. Nevertheless, in practical systems, the estimation noise variance should decrease as P/σ_ν^2 becomes higher.

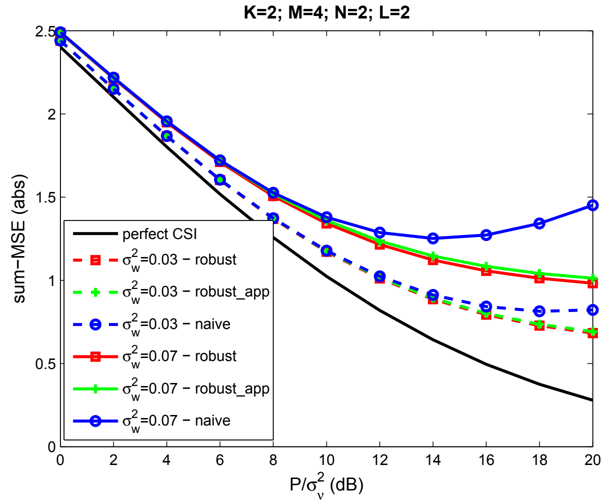

Fig. 2. Exemplary simulation results.

6. CONCLUSIONS

In this paper, we have proposed a robust MMSE design for the linear pre/decoders of the MIMO complex Gaussian broadcast channel. The scheme makes use of the knowledge of the first and second order statistics of the channels and their estimation noise to make the design more robust against the estimation errors. The proposed algorithm iterates between the precoder and decoder designs and is guaranteed to converge. Simulations illustrate the benefit of this enhanced approach in comparison with the so-called naive method which consists in considering the estimated channels as the actual ones.

7. REFERENCES

[1] H. Weingarten, Y. Steinberg, and S. Shamai, "The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel," *IEEE Transactions on Information Theory*, Vol. 52, No. 9, pp. 3936-3964, September 2006.

[2] M. Schubert, S. Shi, E.A. Jorswieck and H. Boche, "Downlink sum-MSE transceiver optimization for linear multi-user MIMO systems," *Proceedings Thirty-Ninth Asilomar Conference on Signals, Systems and Computers,* 2005.

[3] J. Duplicy, J. Louveaux and L. Vandendorpe, "Optimization of linear pre/decoders for multi-user closed-loop MIMO-OFDM," *Proceedings of IEEE International Conference on Communications*, Seoul, South Korea, May 2005.

[4] A. Abdel-Samad, T.N. Davidson and A.B. Gershman, "Robust Transmit Eigen Beamforming Based on Imperfect Channel State Information,", *IEEE Transactions on signal processing*, Vol. 54, No. 5, pp. 1596-1609, May 2006.

[5] D.P. Palomar, "A Unified Framework for Communications through MIMO Channels", Ph.D. dissertation, Universitat Politcnica de Catalunya, Barcelona, May 2003.

[6] F. Rey, "Feedback-channel and adaptive MIMO coded modulations," Ph.D. dissertation, Universitat Politcnica de Catalunya, Barcelona, November 2005.

[7] F.A. Dietrich, W. Utschick, and P. Breun, "Linear precoding based on a stochastic MSE criterion," *Proceedings Eusipco 2005*, Sept. 2005.

[8] M. Joham, K. Kusume, M.H. Gzara, W. Utschick, and J.A. Nossek, "Transmit Wiener filter for the downlink of TDD DS-CDMA systems," *Proceedings of ISSTA 2002*, Sept. 2002.

[9] S.M. Kay, "Fundamentals of statistical signal processing, estimation theory," volume 1, Prentice Hall, 1993.