Transceiver Design for Sum-MSE Optimization in MIMO-MAC with Imperfect Channel Estimation

Patricia Layec*, Pablo Piantanida†, Raphaël Visoz*, and Antoine O. Berthet†

* Orange Labs, Issy-les-Moulineaux, France

†SUPELEC, Department of Telecommunications, Gif-sur-Yvette, France

{patricia.layec; raphael.visoz}@orange-ftgroup.com; {pablo.piantanida; antoine.berthet}@supelec.fr

***Abstract*— Linear transceiver design for multiple access channels (MACs) with spatial correlation at both transmitter and receiver is investigated in the presence of inaccurate channel state information (CSI). We consider a training-based channel estimation at the receiver while a limited-rate feedback channel conveys the transmitter information. Imperfect knowledge comes from the channel estimation errors and the quantization noise. Restricting the decoder to be linear yields to minimize of the sum-mean square error (sum-MSE) subject to individual power constraints. Although no closed-form solution is possible in a multi-user setting, an efficient iterative algorithm relying on the KKT conditions is derived. Numerical results show sum-MSE and BER performance to measure the sensitivity of a mismatched design as well as the effect of quantization noise. Furthermore, the study of channel uncertainty enables to assess the relative impact of imperfect CSI at both ends.**

I. Introduction

Multiple-input multiple-output (MIMO) systems are widely recognized to substantially increase the spectral efficiency of wireless channels. However, the benefits of multi-user MIMO highly depend on the type of channel state information (CSI) at both ends and on the level of accuracy of this information. Practical high data rates wireless systems can only have imperfect CSI at the receiver (CSIR), i.e., an estimate of the channel based on training sequence. As for the information fed back to the transmitter, either analog or digital methods, imperfection can come from partial knowledge (e.g., channel distribution information (CDI), channel quality information (CQI)) and uncertainty (e.g., quantization induced noise or noisy feedback channel). In all cases the adaptation to the channel information errors is mandatory to design reliable solutions.

Practical implementations of multi-user wireless systems often consider linear precoding to reduce interference and linear MMSE receivers, whose design is based on the sum-Mean Square Error (MSE) minimization. This choice, albeit sub-optimal, provides a good trade-off between performance and complexity. The joint transceiver design has been widely studied in the case of perfect CSI in [1] [2] for single-user, while the multi-user extension has been developed in [3]. Interestingly, the problem of finding out the optimum covariance matrices maximizing the mutual information is known to be closely related to the computing of optimum precoders/decoders minimizing the total MSE [4].

However, robust transceiver designs should incorporate the quality of the channel state information. On one hand, several works have considered imperfect CSIR owing to training-based estimation. [5] has studied the problem of optimum balanced power allocation between data and pilots. More recently, the authors in [6] have found the optimal closed-form solution of a single-user MIMO system with spatial transmit correlation. As various criteria exist to measure the system performance, [7] has proposed two general classes of cost functions, namely Schur-convex and Schur-concave functions. On the other hand, only few works have considered imperfect CSI at transmitter (CSIT). Coupled with imperfect CSIR, [8] has investigated partial knowledge CSIT through mean and covariance feedback. The proposed robust transceiver designs exhibit similar structure with the perfect CSI case but with a different noise covariance matrix. None of those works extend to quantized feedback, where the channel state information is conveyed to the transmitter through a limited-rate feedback channel.

In this paper, we evaluate the minimization of the sum-MSE in a spatially correlated uplink MIMO channel. We focus on linear transceiver design handling with imperfect CSI at both receiver and transmitter sides. In contrast to previous work, we consider a general channel estimation model [9], which leads to a composite channel averaging the channel law over all channel estimation errors. Furthermore, by imperfect CSIT, we allow for the impact of quantization errors brought by the precoders compression. The Lagrangian formulation of the sum-MSE optimization enables to derive an efficient iterative numerical algorithm, since in a multi-user scenario no closed-form solution exists. We show how the system can benefit from adaptive precoding even when subject to limited-rate feedback.

The paper is organized as follows. In section II, we describe the communication model in details. Section III formulates the sum-MSE optimization and presents an iterative algorithm. Section IV provides some simulation results before drawing some conclusions in Section V.

Notation: Lower-case and capital bold letters are used to denote vectors and matrices, respectively. Let $\mathbf{A}$ be a matrix, $[\mathbf{A}]_{i,j}$ designates the (i, j) entry of $\mathbf{A}$. The superscript † indicate Hermitian transpose. The operator $tr(.)$ stands for the trace of square matrices.

II. Communication model

A. Channel model

We consider a K-user MIMO-MAC communication system where the K transmitters are equipped with $N_1 \cdots N_K$ antennas and the receiver with N_R antennas. Feedback information is conveyed to the transmitter through a noiseless feedback link. Transmission occurs over a Rayleigh flat fading channel. Each user spreads its data over $M_k \leq N_k$ streams. The received signal vector $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$ is

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{F}_k \mathbf{x}_k + \mathbf{n} \tag{1}$$

where $\mathbf{x}_k \in \mathbb{C}^{M_k \times 1}$, $\mathbf{F}_k \in \mathbb{C}^{N_k \times M_k}$ and $\mathbf{H}_k \in \mathbb{C}^{N_R \times N_k}$ denote, respectively, the transmit signal vector, the linear precoder and the channel matrix of user $k = 1, \ldots, K$. The additive noise vector $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ is independent and identically distributed (i.i.d.) zero-mean circularly symmetric complex Gaussian with covariance matrix $\mathbf{\Sigma}_0 = \sigma_n^2 \mathbf{I}_{N_R}$. Each user is subject to a power constraint P_k such that $\mathrm{tr}(\mathbf{F}_k \mathbb{E}[\mathbf{x}_k \mathbf{x}_k^\dagger] \mathbf{F}_k^\dagger) \leq P_k$ with $\mathbb{E}[\mathbf{x}_k \mathbf{x}_k^\dagger] = \mathbf{I}_{M_k}$. The total transmit power from all users is $\sum_{k=1}^{K} P_k = P$. We further assume that the elements of $\mathbf{H}_k$ are correlated. The channel matrix of user k can thus be factorized as

$$\mathbf{H}_k = \mathbf{R}_{N_R}^{1/2} \mathbf{H}_{W,k} \mathbf{R}_{N_k}^{1/2} \tag{2}$$

where $\mathbf{H}_{W,k}$ is the spatially white channel matrix. The pdf of $\mathbf{H}_k$ is $\psi_{\mathbf{H}_k} = \mathcal{CN}(0, \mathbf{R}_{N_k} \otimes \mathbf{\Sigma}_{H_k})$, where $\mathbf{\Sigma}_{H_k} = \sigma_{h,k}^2 \mathbf{R}_{N_R}$. Note that the diagonal elements of the transmit and receive correlation matrices are equal to 1.

At the receiver side, the received signal of user k is decoded with the use of a linear receiver $\mathbf{G}_k \in \mathbb{C}^{M_k \times N_R}$. The signal estimate vector $\mathbf{r}_k$ is given by

$$\mathbf{r}_k = \mathbf{G}_k \mathbf{y}. \tag{3}$$

To study the robust design of linear transceivers when inaccurate channel estimation is considered, the performance metric used is the sum-MSE. Hence, the MSE matrix of user k is given by

$$MSE_k = \mathbb{E}\big[(\mathbf{r}_k - \mathbf{x}_k)(\mathbf{r}_k - \mathbf{x}_k)^\dagger\big] \tag{4}$$

Next, the channel estimation model is presented so as to have an exact definition of the sum-MSE in function of the channel estimation errors.

B. Estimation model

Each channel matrix estimate $\hat{\mathbf{H}}_k$ may be obtained by the use of a training sequence sent from user k to the receiver, before transmitting the data. The training sequence is constituted of L_k vectors $\mathbf{X}_{T,k} = (\mathbf{x}_{T,k,1}, \cdots, \mathbf{x}_{T,k,L_k})$. We assume that the receiver can capture the statistics of the channel accurately, notably the transmit correlation. Therefore, we choose a training sequence $\mathbf{X}'_{T,k}$ that decorrelates the transmit correlation, i.e., $\mathbf{X}'_{T,k} = \mathbf{R}_{N_k}^{-1/2} \mathbf{X}_{T,k}$. The average energy of the training symbols may be expressed as $P_{T,k} = \frac{1}{L_k N_k} \mathrm{tr}(\mathbf{X}_{T,k} \mathbf{X}_{T,k}^\dagger)$. The corresponding received signal $\mathbf{Y}_{T,k} = \mathbf{R}_{N_R}^{1/2} \mathbf{H}_{W,k} \mathbf{X}_{T,k} + \mathbf{N}_{T,k}$ allows the receiver to perform ML estimation. Since to estimate $\mathbf{R}_{N_R}^{1/2} \mathbf{H}_{W,k}$ we need at least $N_R N_k$ measurements, and each symbol time yields N_R samples, we must have $L_k \geq N_k$ provided that $\mathbf{X}_{T,k}$ is full rank. Let us then consider that $\hat{\mathbf{H}}_k = \left(\mathbf{R}_{N_R}^{1/2} \mathbf{H}_{W,k} + \mathbf{E}_{k,w}\right) \mathbf{R}_{N_k}^{1/2}$, where $\mathbf{E}_{W,k}$ is a white estimation error matrix if we assume that $\mathbf{X}_{T,k}$ is orthogonal. The estimation variance is $\sigma_{e,k}^2 = \mathrm{SNR}_{T,k}^{-1}$ with $\mathrm{SNR}_{T,k} = \frac{L_k P_{T,k}}{\sigma_n^2}$. As a result, the conditional pdf of $\hat{\mathbf{H}}_k$ given $\mathbf{H}_k$ is

$$\psi_{\hat{\mathbf{H}}_k | \mathbf{H}_k} = \mathcal{CN}(\mathbf{H}_k, \mathbf{R}_{N_k} \otimes \mathbf{\Sigma}_{E_k}) \tag{5}$$

with $\mathbf{\Sigma}_{E_k} = \sigma_{e,k}^2 \mathbf{I}_{N_R}$. The *a posteriori* pdf $\psi_{\mathbf{H}_k | \hat{\mathbf{H}}_k}$ can be derived from (5) and $\psi_{\mathbf{H}_k}$, and is expressed as $\psi_{\mathbf{H}_k | \hat{\mathbf{H}}_k} = \mathcal{CN}(\mathbf{\Sigma}_{\Delta_k} \hat{\mathbf{H}}_k, \mathbf{R}_{N_k} \otimes \mathbf{\Sigma}_{\Delta_k} \mathbf{\Sigma}_{E_k})$ where $\mathbf{\Sigma}_{\Delta_k} = \mathbf{\Sigma}_{H_k} (\mathbf{\Sigma}_{H_k} + \mathbf{\Sigma}_{E_k})^{-1}$. Then by averaging the unknown channel with conditional pdf $W(\mathbf{y} \,|\mathbf{x}_1, \ldots, \mathbf{x}_K, \mathbf{H}_1, \ldots, \mathbf{H}_K)$ over all channel estimation errors and after some algebra, we obtain the composite channel, with $\hat{\mathcal{H}} = \big(\hat{\mathbf{H}}_1, \ldots, \hat{\mathbf{H}}_K\big)$,

$$\begin{aligned} \tilde{W}\big(\mathbf{y} | \mathbf{x}_1, \cdots, \mathbf{x}_K, \hat{\mathcal{H}}\big) &= \mathcal{CN}\Big(\sum_{k=1}^{K} \mathbf{\Sigma}_{\Delta_k} \hat{\mathbf{H}}_k \mathbf{F}_k \mathbf{x}_k, \tilde{\mathbf{\Sigma}}_0\Big) \\ \tilde{\mathbf{\Sigma}}_0 &= \mathbf{\Sigma}_0 + \textstyle\sum_{k=1}^{K} \mathbf{\Sigma}_{\Delta_k} \mathbf{\Sigma}_{E_k} tr(\mathbf{R}_{N_k} \mathbf{F}_k \mathbf{x}_k \mathbf{x}_k^\dagger \mathbf{F}_k^\dagger). \end{aligned} \tag{6}$$

From this expression, the MSE matrix of user k can now be expressed as

$$\begin{aligned} MSE_k = \mathbf{G}_k \Bigg[& \sum_{i=1}^{K} \mathbf{\Sigma}_{\Delta_i} \hat{\mathbf{H}}_i \mathbf{F}_i \mathbf{F}_i^\dagger \hat{\mathbf{H}}_i^\dagger \mathbf{\Sigma}_{\Delta_i} + \sigma_n^2 \mathbf{I}_{N_R} \\ & + \sum_{i=1}^{K} \mathbf{\Sigma}_{\Delta_i} \mathbf{\Sigma}_{E_i} tr(\mathbf{R}_{N_i} \mathbf{F}_i \mathbf{F}_i^\dagger) \Bigg] \mathbf{G}_k^\dagger - \mathbf{G}_k \mathbf{\Sigma}_{\Delta_k} \hat{\mathbf{H}}_k \mathbf{F}_k \\ & + \mathbf{I}_{M_k} - \big(\mathbf{G}_k \mathbf{\Sigma}_{\Delta_k} \hat{\mathbf{H}}_k \mathbf{F}_k\big)^\dagger \end{aligned} \tag{7}$$

C. Limited feedback link

With a limited feedback channel, the choice of the precoder for each user is contained on a pre-defined codebook whose size depends on the feedback rate constraint. The need for compression of precoders induces some distortion, which is modelled as an additive quantization noise. This yields to the following expression

$$\mathbf{F}_k = \hat{\mathbf{F}}_k + \mathbf{D}_k \tag{8}$$

where $\mathbf{D}_k$ is the quantization error matrix, with i.i.d. zero-mean circularly symmetric complex Gaussian entries of variance $\sigma_{d,k}^2$. Let B be the total number of feedback bits allowed in the low-rate feedback channel, shared by the K users. It results that $\sigma_{d,k}^2$ is a function of B and K. Moreover, its expression mainly depends on the quantizer design including both dimension (e.g., scalar, vector...) and method (e.g., random, Lloyd-Max...). Consequently, the robust transceiver design should incorporate all sources of CSI imperfection.

Based on the proposed distortion model (8), the MSE matrix of user k is finally given by

$$\begin{aligned} MSE_k &= \mathbf{G}_k \Big[\sum_{i=1}^{K} \mathbf{\Sigma}_{\Delta_i} \hat{\mathbf{H}}_i \hat{\mathbf{F}}_i \hat{\mathbf{F}}_i^{\dagger} \hat{\mathbf{H}}_i^{\dagger} \mathbf{\Sigma}_{\Delta_i} + \sigma_n^2 \mathbf{I}_{N_R} \\ &+ \sum_{i=1}^{K} \mathbf{\Sigma}_{\Delta_i} \hat{\mathbf{H}}_i \sigma_{d,i}^2 \hat{\mathbf{H}}_i^{\dagger} \mathbf{\Sigma}_{\Delta_i} + \sum_{i=1}^{K} \mathbf{\Sigma}_{\Delta_i} \mathbf{\Sigma}_{E_i} tr(\mathbf{R}_{N_i} \hat{\mathbf{F}}_i \hat{\mathbf{F}}_i^{\dagger}) \\ &+ \sum_{i=1}^{K} \mathbf{\Sigma}_{\Delta_i} \mathbf{\Sigma}_{E_i} \sigma_{d,i}^2 tr(\mathbf{R}_{N_i}) \Big] \mathbf{G}_k^{\dagger} - \mathbf{G}_k \mathbf{\Sigma}_{\Delta_k} \hat{\mathbf{H}}_k \hat{\mathbf{F}}_k \\ &+ \mathbf{I}_{M_k} - \left(\mathbf{G}_k \mathbf{\Sigma}_{\Delta_k} \hat{\mathbf{H}}_k \hat{\mathbf{F}}_k \right)^{\dagger} \end{aligned} \quad (9)$$

To avoid a large feedback overhead, the base station computes the optimized precoders and receivers in a centralized fashion. Then, the resulting precoders are quantized and then broadcasted to all users. Next section details the optimization process.

III. Sum-MSE with imperfect CSIR and quantized CSIT

Usually the multi-user MAC transceiver optimization solves the minimization of the sum-MSE under individual power constraints. Solution of this optimization leads to a set of receive matrices $\{\mathbf{G}_k\}_{k=1}^{K}$ and a set of quantized precoders $\{\hat{\mathbf{F}}_k\}_{k=1}^{K}$.

A. Joint optimization of precoding and receive filters

The objective function considers that a maximum transmit power P_k is allowed for each user k. These individual power constraints ensure that none of the power budgets are exceeded. Therefore the sum-MSE minimization can be formulated as

$$\min_{\{\mathbf{G}_k, \mathbf{F}_k\}} \quad \sum_{k=1}^{K} tr\big(MSE_k\big) \quad (10)$$

$$\text{subject to} \quad tr\big(\hat{\mathbf{F}}_k \hat{\mathbf{F}}_k^{\dagger}\big) \leq P_k, \ k = 1, \ldots, K \quad (11)$$

Although this problem has no closed-form solution, an efficient algorithm based on the Lagrangian can be derived. Associated with the minimization problem (10), the Lagrangian function $\mathcal{L}(\{\hat{\mathbf{F}}_k\}, \{\mathbf{G}_k\}, \{\mu_k\})$ is given by

$$\mathcal{L} = \sum_{k=1}^{K} tr\big(MSE_k\big) + \sum_{k=1}^{K} \mu_k \big(tr(\hat{\mathbf{F}}_k \hat{\mathbf{F}}_k^{\dagger}) - P_k \big) \quad (12)$$

where $\{\mu_k \geq 0\}_{k=1}^{K}$ is the Lagrangian multiplier with respect to the power constraints. The optimal transceiver design is then obtained with the KKT conditions. By taking the partial derivatives of expression (12) with respect to $\mathbf{G}_k$ and $\hat{\mathbf{F}}_k$, we obtain the expressions (13) and (14) respectively. Each Lagrange multiplier is calculated so as to satisfy its corresponding transmit power constraint. Then, the computation of the Lagrange multiplier can be solved for

$$\mu_k \Big[tr\big(\hat{\mathbf{F}}_k \hat{\mathbf{F}}_k^{\dagger}\big) - P_k \Big] = 0.$$

To find $\{\mu_k\}_{k=1}^{K}$, a closed-from solution may be obtained using similar derivation to that employed in [3]. It is explicitly detailed in the Appendix.

B. Iterative algorithm

In the multi-user case, the resulting (colored) noise covariance matrix is also a function of the covariance matrices of all other users. Thus, it is not possible to find a closed form solution simultaneously diagonalizing (13) and (14). Therefore, we iterate successively between $\{\mathbf{G}_k\}_{k=1}^{K}$ and $\{\hat{\mathbf{F}}_k\}_{k=1}^{K}$ to obtain the optimized set of receivers and precoders. A general algorithm follows.

Proposed algorithm - minimum sum-MSE

- **Initialize** $\{\hat{\mathbf{F}}_k\}_{k=1}^{K}$ to a diagonal matrix, where its (m, n)-th entry is equal to 0 if $m \neq n$ or equal to $\sqrt{\frac{P_k}{M_k}}$ elsewhere.
- **Step 1:** for $k = 1, \ldots, K$ compute $\mathbf{G}_k$ using the partial derivative of the Lagrangian with respect to $\mathbf{G}_k$ (13). The linear receive filters correspond to the LMMSE receiver.
- **Step 2:** Satisfy the power constraint: update μ_k for each $k = 1, \ldots, K$.
- **Step 3:** for $k = 1, \ldots, K$ compute $\hat{\mathbf{F}}_k$ using partial derivative of the Lagrangian function with respect to $\hat{\mathbf{F}}_k$ (14).
- **Repeat** until convergence.

The proposed algorithm is guaranteed to converge. Indeed, the iterative process between users is a monotonically decreasing function of the sum-MSE. Besides the objective function is clearly lower bounded by zero, as a result the algorithm will always converge to a local optimum. However, we cannot guarantee to reach the global optimum as the sum-MSE function is not jointly convex over all $\{\hat{\mathbf{F}}_k, \mathbf{G}_k\}_{k=1}^{K}$. When $K = 1$ a unique solution exists using similar arguments as in [10, chapter 4], which can be derived in closed-form, where some constants need to be computed numerically.

IV. Numerical results

In this section, we present some numerical results when we assume that the total number of feedback bits is equally shared between all users. Furthermore, since the focus of this paper is the additional impact of quantization errors (and not the optimized design of quantizer), we consider a scalar quantization. Based on the rate-distortion theory [11], the variance induced by quantization noise of any user k can thus be formulated as

$$\sigma_{d,k}^2 = 2^{-\frac{B}{K N_k M_k}} \quad (15)$$

The (i, j)-th entry of the correlation matrix is modelled as $\big[\mathbf{R}\big]_{i,j} = \rho^{|i-j|}$, where $\rho \in [0, 1]$ is the correlation parameter (with subscript R or k to denote receive or transmit correlation). We assume that each user have the same parameters $\sigma_{h,k}^2 = 1$, $\sigma_{d,k}^2 = \sigma_d^2$, $N_k = N_T$, $L_k = L$, $\rho_k = \rho_T$, and the same individual power constraints $P_k = P/K$. The

$$\mathbf{G}_k = \hat{\mathbf{F}}_k^\dagger \hat{\mathbf{H}}_k^\dagger \boldsymbol{\Sigma}_{\Delta_k} \Big[\sum_{i=1}^K \boldsymbol{\Sigma}_{\Delta_i} \hat{\mathbf{H}}_i \hat{\mathbf{F}}_i \hat{\mathbf{F}}_i^\dagger \hat{\mathbf{H}}_i^\dagger \boldsymbol{\Sigma}_{\Delta_i} + \sum_{i=1}^K \boldsymbol{\Sigma}_{\Delta_i} \hat{\mathbf{H}}_i \sigma_{d,i}^2 \hat{\mathbf{H}}_i^\dagger \boldsymbol{\Sigma}_{\Delta_i} + \sigma_n^2 \mathbf{I}_{N_R} + \sum_{i=1}^K \boldsymbol{\Sigma}_{\Delta_i} \boldsymbol{\Sigma}_{E_i} tr(\mathbf{R}_{N_i} \hat{\mathbf{F}}_i \hat{\mathbf{F}}_i^\dagger) + \sum_{i=1}^K \boldsymbol{\Sigma}_{\Delta_i} \boldsymbol{\Sigma}_{E_i} \sigma_{d,i}^2 tr(\mathbf{R}_{N_i}) \Big]^{-1} \quad (13)$$

$$\hat{\mathbf{F}}_k = \Big[\sum_{i=1}^K \hat{\mathbf{H}}_k^\dagger \boldsymbol{\Sigma}_{\Delta_k} \mathbf{G}_i^\dagger \mathbf{G}_i \boldsymbol{\Sigma}_{\Delta_k} \hat{\mathbf{H}}_k + \mu_k \mathbf{I}_{N_k} + \sum_{i=1}^K tr(\mathbf{G}_i \boldsymbol{\Sigma}_{\Delta_k} \boldsymbol{\Sigma}_{E_k} \mathbf{G}_i^\dagger) \mathbf{R}_{N_k} \Big]^{-1} \hat{\mathbf{H}}_k^\dagger \boldsymbol{\Sigma}_{\Delta_k} \mathbf{G}_k^\dagger \quad (14)$$

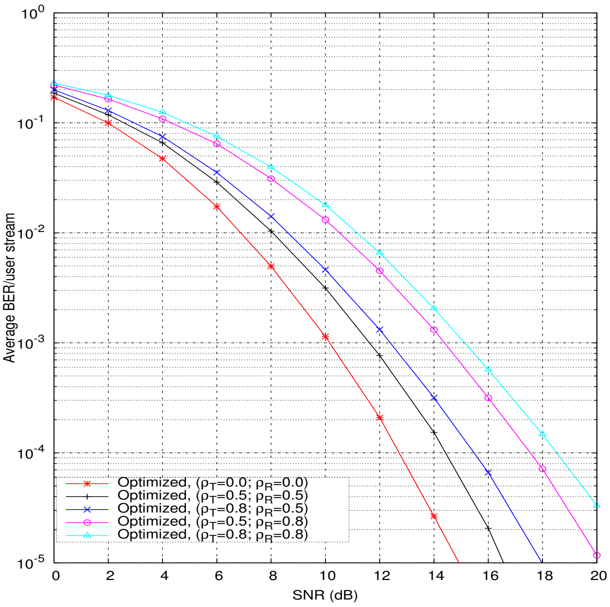

Fig. 1. Rayleigh channel. Impact of the channel correlation factor with 1 stream/user, $K = 2, L = 4, \sigma_d^2 = 0.01$.

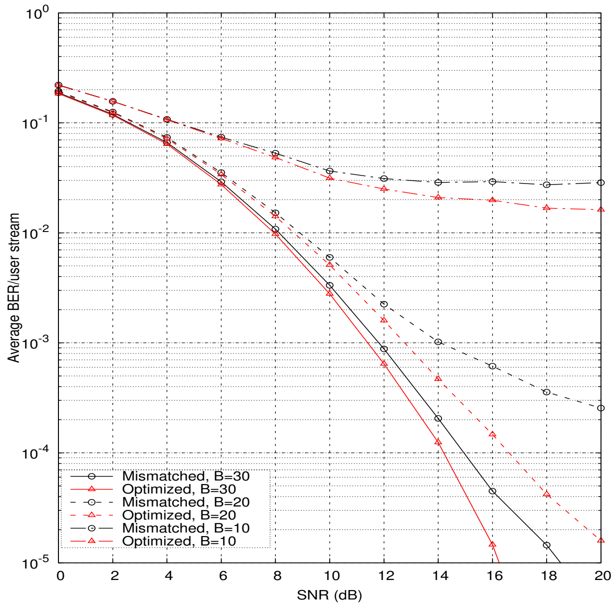

Fig. 2. Rayleigh channel. Impact of quantization errors with 1 stream/user, $K = 2, L = 4, \rho_T = \rho_R = 0.5$.

average energy of the training symbols is set equal to the power constraint, i.e., $P_{T,k} = P_k$. To obtain the Bit Error Rate (BER) performance, we use uncoded 4-QAM on each user data stream. Our results are plotted for 2 users equipped with 2 antennas and we assume that $N_R = 4$ receive antennas.

Fig. 1 shows the impact of the channel correlation. For the optimized design (10), we compare the BER for different correlation factors. Quite naturally, a high correlated channel exhibits worse performance. It appears that the receive correlation has a more negative impact than transmit correlation. Indeed, when $\rho_T = \{0.5; 0.8\}$, the increase of ρ_R from 0.5 to 0.8 results in a 3 dB-gap to achieve a target BER of 10^{-3}.

Next, we illustrate the influence of imperfect CSIT on Fig. 2. As expected, decreasing the number of feedback bits can critically damage the error probability. With high value of distortion (e.g., $B = 10$), an error floor occurs. Note that employing vector or matrix quantizer enables to lower the average distortion, or similarly permits to reduce the number of feedback bits needed. Additionally, we compare the optimized multi-user system design with the mismatched design. By mismatched, we mean that precoders and receivers are designed by considering the set of $\{\hat{\mathbf{H}}_k\}_{k=1}^K$ and $\{\hat{\mathbf{F}}_k\}_{k=1}^K$ as the true values. For any SNR > 10 dB, the mismatched design leads to non-negligible performance loss with respect to the optimized design. For instance when $B = 20$, 1.2 dB gap is observed at a target BER of 10^{-3}. Fig. 3 is similar to Fig. 2 except that we plot the average sum-MSE performance instead of the BER. Analogous performance trends are observed. However, with this metric, the gap between the optimized and mismatched designs is nearly constant at all SNR. To conclude, quantization errors have a major impact specifically when the SNR > 10 dB or when the variance induced by distortion is too high (here, when $B = 10$).

Finally, Fig. 4 compares the optimized for different CSI assumptions. This comparison reveals that at low SNR the impact of imperfect CSIT is small, while in contrast, the channel estimation errors significantly contribute to the corruption of data. In addition, we also plot a non-precoded system (e.g., precoders are set to a scaled identity matrix) but only with imperfect CSIR. No quantization is needed since precoders cannot be adapted. The non-precoding curve shows the worse performance suggesting that even with quantization errors, it is still useful to have an adaptive precoded system design.

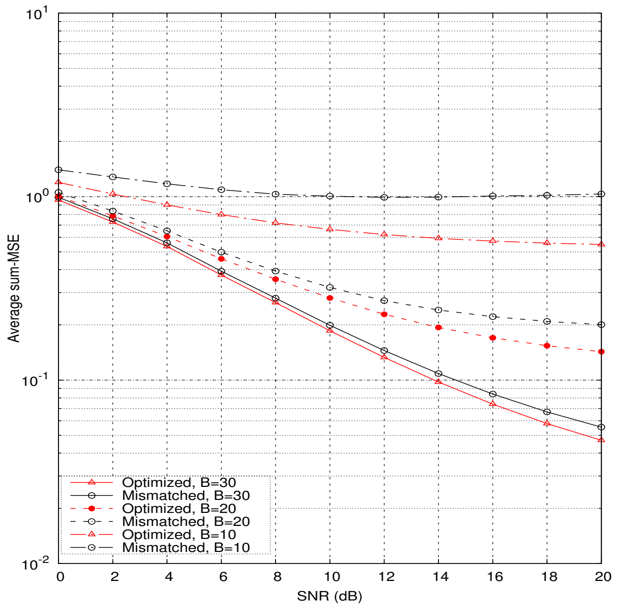

Fig. 3. Rayleigh channel. Impact of quantization errors on the sum-MSE performance with 1 stream/user, $K = 2, \rho_T = \rho_R = 0.5, L = 4$.

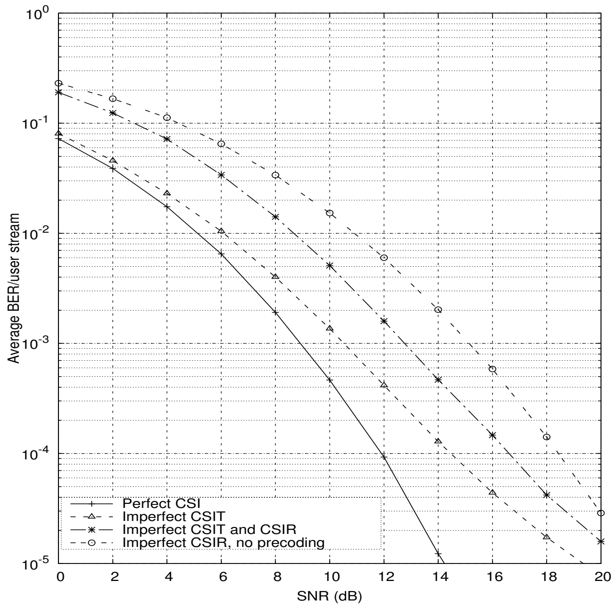

Fig. 4. Rayleigh channel. Impact of channel uncertainty on BER performance with 1 stream/user, $K = 2, \rho_T = \rho_R = 0.5, (L = 4, B = 20)$.

V. Conclusions

In this work we have studied the sum-MSE minimization over spatially correlated channels. The multi-user uplink transceiver design with individual power constraints is addressed under the assumption of imperfect CSI arising from inaccurate channel estimation and quantization errors. Although no closed-form solution is possible, the optimized design can be obtained by means of efficient numerical algorithms based on the KKT conditions. Simulation results show that channel estimation errors impact the performance at all SNR. In contrast, imperfect CSIT is more pronounced at higher SNR. Significant gains can thus be obtained with an adaptive precoding design if the noise variance induced by quantization is acceptable. The sensitivity of the mismatched transceiver design has also been assessed.

References

[1] H. Sampath, P. Stoica, A. Paulraj, "Generalized Linear Precoder and Decoder Design for MIMO Channels using the weighted MMSE Criterion," *IEEE Trans. Commun.*, vol 49., no. 12, pp. 2198-2206, Dec. 2001.

[2] D.P. Palomar, J.M. Cioffi, M.A. Lagunas, "Joint Tx-Rx Beamforming Design for Multicarrier MIMO Channels: A Unified Framework for Convex Optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381-2401, Sept. 2003.

[3] S. Serbetli, A. Yener, "Transceiver Optimization for Multiuser MIMO Systems," *IEEE Trans. Signal Process.*, vol. 52, no. 1, Jan. 2004.

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

[5] S. Serbetli, A. Yener, "MMSE Transmitter Design for Correlated MIMO Systems with Imperfect Channel Estimates: Power Allocation Trade-offs," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2295-2304, Aug. 2006.

[6] M. Ding, S.D. Blostein, "MIMO LMMSE Transceiver Design with Imperfect CSI at Both Ends," *IEEE ICC'07*, Jun. 2007.

[7] X. Zhang, D.P. Palomar, B. Ottersten, "Robust Design of Linear MIMO Transceivers under Channel Uncertainty," *IEEE ICASSP'06*, vol. 4, May 2006.

[8] X. Zhang, D.P. Palomar, B. Ottersten , "Statistically Robust Design of Linear MIMO Transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678-3689, Aug. 2008.

[9] P. Piantanida, and P. Duhamel," Dirty-paper Coding without Channel Information at the Transmitter and Imperfect Estimation at the Receiver," *IEEE ICC'07*, Jun. 2007.

[10] M. Ding, "Multiple-input Multiple-output Wireless Systems Designs with Imperfect Channel State Knowledge," *Ph.D. Thesis* Queen's University, July 2008.

[11] T.M. Cover, J.A. Thomas, "Elements of Information Theory," *Wiley Series in Telecommunication*, 1991.

Appendix

The computation of Langrange multipliers relies on the optimized precoder $\hat{\mathbf{F}}_k$. It is easily seen from (14) that $\hat{\mathbf{F}}_k$ has the form

$$\hat{\mathbf{F}}_k = \left(\mu_k \mathbf{I}_{M_k} + \mathbf{C}_k\right)^{-1} \mathbf{B}_k \tag{16}$$

As μ_k should satisfy the individual power constraint, we have $tr\left(\hat{\mathbf{F}}_k \hat{\mathbf{F}}_k^\dagger\right) = P_k$. Since $\mathbf{C}_k$ and $\mathbf{B}_k$ are known we can eigendecompose $\mathbf{C}_k = \mathbf{U}_k \boldsymbol{\Sigma}_k \mathbf{U}_k^\dagger$. It leads to

$$tr\left[\left(\mu_k \mathbf{I}_{M_k} + \boldsymbol{\Sigma}_k\right)^{-2} \mathbf{U}_k^\dagger \mathbf{B}_k \mathbf{B}_k^\dagger \mathbf{U}_k\right] = P_k \tag{17}$$

Let $\mathbf{J}_k = \mathbf{U}_k^\dagger \mathbf{B}_k \mathbf{B}_k^\dagger \mathbf{U}_k$, we can thus expressed the previous equation as

$$\sum_{i=1}^{N_T} \frac{j_{i,i}}{(\mu_k + \sigma_{i,i})^2} = P_k \tag{18}$$

where $j_{i,i}$ and $\sigma_{i,i}$ are the ith diagonal coefficients of $\mathbf{J}_k$ and $\boldsymbol{\Sigma}_k$ respectively. The LHS of (18) is a monotically decreasing funtion of μ_k. Therefore, it exists only one non-negative real value of μ_k satisfying (18), otherwise, it is 0.