Robust Precoding with Limited Feedback Design based on Precoding MSE for MU-MISO Systems

Michael Joham, *Member, IEEE,* Paula M. Castro, *Member, IEEE,* Wolfgang Utschick, *Senior Member, IEEE,*and Luis Castedo, *Member, IEEE*

Abstract

For the separation of the signals in the vector *Broadcast Channel* (BC), some information about the channel state is necessary at the transmitter. In many cases, this *Channel State Information* (CSI) must be fed back from the receivers to the transmitter. We jointly design the channel estimators and the quantizers at the receivers together with the precoder at the transmitter based on a precoder-centric criterion, i.e., the minimization of a *Mean Square Error* (MSE) metric appropriate for the precoder design. This is in contrast to our previous works, where the quantizer design was based on a CSI MSE metric, i.e., based on the minimization of the MSE between the true channel and the channel recovered by the transmitter using a feedback channel. Interestingly, the estimators resulting from this joint formulation are independent of the used codebook. The codebook entries are the employed precoders. Therefore, each receiver feeds back the index of a set of precoders and the intersection of the sets gives the appropriate precoder. Since the quantizers of the different receivers have to work separately, the metric for the computation of the partition cells cannot be expressed as a simple squared error depending on the quantizer output. The proposed system based on a joint optimization clearly outperforms previous designs with separate optimization of feedback and precoding.

Index Terms

Feedback channel, Bayesian approach, imperfect CSI, robust precoding, precoding MSE metric.

I. Introduction

A *Multi User Multiple Input Single Output* (MU-MISO) system is an appropriate model for the downlink of a cellular system where it is reasonable to assume that the transmitter (base station) is equipped with multiple antennas whereas the receivers (mobile stations) only support a single antenna in order to reduce size, power consumption,

Michael Joham and Wolfgang Utschick are with the Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany, e-mail:{joham,utschick}@tum.de.

Paula M. Castro and Luis Castedo are with the Department of Electronics and Systems, University of A Coruña, A Coruña, 15071 Spain, e-mail: {pcastro,luis}@udc.es.

and cost. As the receivers have no interference suppressing capabilities, the transmitter is in charge of all tasks related to eliminating the inter-user interference.

The availability of CSI at the transmitter is crucial for the signal separation in the considered vector BC. In cellular systems that use *Frequency Division Duplexing* (FDD), the utilization of a finite-rate feedback channel is common to send the CSI estimated at the receiver to the transmitter. The standard assumption for the design of these limited feedback channels is to assume that the receivers have a perfect CSI knowledge (see [1]–[5]). In practice, however, the information about the channels obtained by the transmitter via limited rate feedback is always erroneous. Thus, perfect interference suppression with precoding is impossible. Additionally, an information theoretic approach to the design of limited feedback channels with imperfect CSI is difficult due to the fact that the computation of the mutual information cannot be found in closed form and is costly to be estimated via simulations (see [6], [7]). For this reason, in this work we have resorted to precoding and limited feedback channel designs based on the minimum MSE criterion. More specifically, we propose to jointly design the CSI estimator and quantizer at the receiver together with the precoder at the transmitter based on a precoder-centric criterion, i.e., the minimization of an MSE metric appropriate for the precoder design [8].

The utilization of such a precoding MSE for the design of both the precoders and the feedback is motivated as follows. In [9], it has been demonstrated that a function of the MSE is a lower bound to the mutual information for Gaussian signaling and for perfect CSI at receiver. This result has been generalized in [10], i.e., a lower bound for the mutual information can be found that is a function of the MSE and that is applicable irrespective of the quality of CSI and the modulation format. Thus, the minimization of the MSE considered in this paper corresponds to the maximization of a lower bound to the mutual information. Additionally, functions of the MSE constitute upper bounds for the symbol error rate of QAM symbols (e.g., [11]) and for the bit error rate of QPSK symbols (e.g., [12]). Thus, the minimization of the MSE can also be interpreted as the minimization of an upper bound of error probability.

The proposed limited feedback channel design procedure works as follows. First, the channel estimator is designed to minimize the MSE between the transmitted symbols and the symbols recovered by the users (including the precoder) averaged over all possible channel realizations, assuming a given quantizer (see Section IV). Interestingly, the estimators resulting from this joint optimization are independent of the used quantizer codebook and are equal to the estimators obtained from CSI MSE metrics.

Next, we design the codebook entries in Subsection V-A that consist of the precoders to be employed. These precoders are found by minimizing the precoding MSE conditioned on the fed-back index. The utilization of white estimates (by dropping the coloring with the square root of the respective covariance matrix) and the restriction to rectangular regions leads to a simple computation of the conditional means necessary for the precoding design step. The most difficult part of the proposed scheme is the design of the partition cells. The cell boundaries are designed by minimizing the precoding MSE conditioned on the quantizer input (see Subsection V-B). We also focus on how to implement bit allocation in Subsection V-D, and on how we can solve the problems related to its computational complexity by means of a heuristic strategy. Finally, we present the results of some computer simulations in Section VI that were carried out to illustrate the performance of the proposed limited feedback channel design in terms of uncoded BER.

Note that each user feeds back the index of a set of precoders and the intersection of the sets performed at the transmitter gives the appropriate precoder to be used during the transmission. Since the quantizers of the different receivers have to work separately, the metric for the computation of the partition cells cannot be expressed as a simple squared error depending on the quantizer output and its computation is quite complex as shown in this work.

All derivations are based on the assumption of perfect knowledge of the second-order statistics of the noise, the symbols, and the channels. However, these parameters have to be estimated and reported to the transmitter in practice, although we will not deal with this problem in this work. We assume that all random variables are zero-mean and stationary.

Vectors and matrices are denoted by lower case bold and capital bold letters, respectively. The $K \times K$ identity matrix is denoted by $\mathbf{I}_K$ and $\mathbf{0}_K$ is a K-dimensional zero vector. We use $\mathrm{E}[\bullet]$, $\Re(\bullet)$, $\Im(\bullet)$, $\mathrm{tr}(\bullet)$, $(\bullet)^*$, $(\bullet)^{\mathrm{T}}$, $(\bullet)^{\mathrm{H}}$, $\det(\bullet)$, and $\|\bullet\|_2$ for expectation, real and imaginary part of the argument, trace of a matrix, complex conjugation, transposition, conjugate transposition, determinant of a matrix, and Euclidean norm, respectively. The i-th element of a vector $\boldsymbol{x}$ is x_i. With $f_{\mathrm{G}}(\boldsymbol{x}, \boldsymbol{\mu}_x, \boldsymbol{C}_x)$, we refer to a circularly symmetric complex Gaussian *Probability Density Function* (PDF) of $\boldsymbol{x} \in \mathbb{C}^m$ with the mean $\boldsymbol{\mu}_{\boldsymbol{x}} \in \mathbb{C}^m$ and the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}} \in \mathbb{C}^{m \times m}$, i.e., $\boldsymbol{x} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{C}_{\boldsymbol{x}})$ and

$$f_{\mathrm{G}}\left(\boldsymbol{x}, \boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{C}_{\boldsymbol{x}}\right) = \frac{\exp\left(-\left(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}}\right)^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{x}}^{-1} \left(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{x}}\right)\right)}{\pi^m \det(\boldsymbol{C}_{\boldsymbol{x}})}.$$

II. System Model

Fig. 1 depicts the block diagram of a MU–MISO system with linear precoding. We assume a transmitter equipped with N antennas and K single-antenna receivers. Let us denote the information symbols by $\boldsymbol{u} \in \mathbb{C}^K$, a vector of zero-mean complex-valued modulated signals with unit covariance matrix, i.e., $\boldsymbol{C}_{\boldsymbol{u}} = \mathrm{E}[\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}}] = \mathbf{I}$. This vector is linearly transformed by the precoder $\boldsymbol{P} \in \mathbb{C}^{N \times K}$ to obtain the transmit signal $\boldsymbol{x} \in \mathbb{C}^N$. This signal propagates over the channel $\boldsymbol{h}_k \in \mathbb{C}^N$ to the k-th receiver to produce the received signal

$$y_k = \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{x} + \eta_k \qquad k = 1, ..., K \tag{1}$$

where η_k is the *Additive White Gaussian Noise* (AWGN). The channel $\boldsymbol{h}_k \in \mathbb{C}^N$ is assumed to be time-varying and modeled by means of a vector of zero-mean complex-valued Gaussian random variables, i.e., $\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{h},k})$ with the channel covariance matrix for the k-th user $\boldsymbol{C}_{\boldsymbol{h},k} = \mathrm{E}[\boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}}] \in \mathbb{C}^{N \times N}$. The receiver applies the common receive weight $g \in \mathbb{C}$ to get the estimate $\hat{u}_k = g y_k$. Note that the common weight g is only assumed in the precoder design to allow for a closed form solution of the precoder $\boldsymbol{P}$ (see also the discussion in [13]) and to simplify the presentation. In contrast, every receiver applies an MMSE optimal receiver weight in the final system (see Subsection V-A). As shown in Fig. 1, combining the signals at the output of the different receivers yields

$$\hat{\boldsymbol{u}} = g\boldsymbol{H}\boldsymbol{P}\boldsymbol{u} + g\boldsymbol{\eta} \tag{2}$$

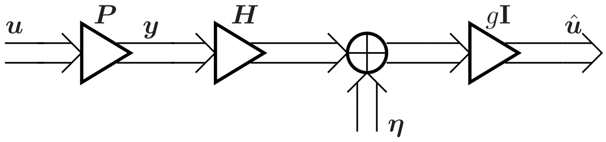

Fig. 1. System model for MU–MISO linear precoding combining signals from all users.

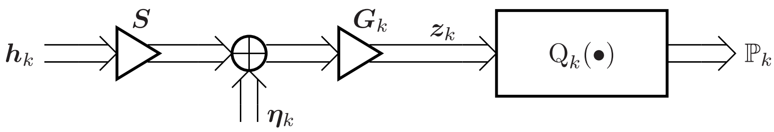

Fig. 2. System model for feedback.

where $\hat{\boldsymbol{u}} = [\hat{u}_1, \ldots, \hat{u}_K]^{\mathrm{T}} \in \mathbb{C}^K$, $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_K] \in \mathbb{C}^K$ with $\boldsymbol{\eta} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{\eta}})$, and $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]^{\mathrm{T}} \in \mathbb{C}^{K \times N}$.

We impose the constraint that the average total transmit energy is upper bounded by E_{tx}, i.e.,

$$\mathrm{E}\left[\|\boldsymbol{P}\boldsymbol{u}\|_2^2\right] \leq E_{\mathrm{tx}}.$$

Fig. 2 depicts the block diagram of the estimation and quantization of the CSI performed at the receivers. The resulting index representing the set $\mathbb{P}_k$ is fed back to the transmitter. We assume that the centralized transmitter sends a sequence of N_{tr} pilot symbols from all transmit antennas. The received noisy pilot symbols are passed through the linear estimator $\boldsymbol{G}_k \in \mathbb{C}^{N \times N_{\mathrm{tr}}}$ to obtain the channel estimate

$$\boldsymbol{z}_k = \boldsymbol{G}_k \left(\boldsymbol{S}\boldsymbol{h}_k + \boldsymbol{\eta}_k\right) \in \mathbb{C}^N. \tag{3}$$

This channel estimate will be the input to the quantizer $\mathrm{Q}_k(\bullet)$ of user k. The matrix $\boldsymbol{S} \in \mathbb{C}^{N_{\mathrm{tr}} \times N}$ contains the pilot symbols and $\boldsymbol{\eta}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{\eta},k})$ is the noise of the pilot channel to the k-th receiver. For simplicity reasons, the feedback channel is assumed to be error-free and without delay. The delay effect is relatively easy to correct (see [14], [15]) but at the cost of unnecessarily complicating our notation.

After estimation, it is necessary to implement some type of quantization in order to compress all the information sent through the finite-rate feedback channel. Contrary to the quantizers used in [14], [15], where the codebook entries were white channel coefficients, the codebook entries of the quantizers proposed in this work are the precoders of Fig. 1, i.e., the quantized information eventually represents a precoder and not a CSI.

A. Model for Quantizers

Let us initially assume a genie-aided MU–MISO system where all the users work in a cooperative way. In this case, it is possible to carry out a joint quantization:

$$\mathrm{Q}\left(\boldsymbol{z}\right) = \sum_{i=1}^{M} \boldsymbol{P}_i \, \mathrm{S}_i\left(\boldsymbol{z}\right) \tag{4}$$

where M is the codebook size. Here, $\boldsymbol{z} = [\boldsymbol{z}_1^{\mathrm{T}}, \ldots, \boldsymbol{z}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{KN}$ represents the estimated CSI of all users. The selector function $\mathrm{S}_i(\bullet)$ is 1 if the argument lies in the partition cell $\mathcal{R}_i \subseteq \mathbb{C}^{KN}$, and 0 elsewhere. Each of the M codebook entries $\boldsymbol{P}_i \in \mathbb{C}^{N\times K}$ is a precoder and $\boldsymbol{P}_i$ is chosen if $\boldsymbol{z} \in \mathcal{R}_i$.

In practice, however, a joint quantization of the estimated CSI is impossible because receivers do not cooperate and each receiver has access only to its own CSI $\boldsymbol{z}_k$. Therefore, the partition cell $\mathcal{R}_i$ must be decomposed into K subregions $\mathcal{R}_{k,i} \subseteq \mathbb{C}^N$, i.e., $\mathcal{R}_i = \mathcal{R}_{1,i} \times \cdots \times \mathcal{R}_{K,i}$, where $\times$ denotes the cartesian product defined as

$$\mathcal{R}_i = \mathcal{R}_{1,i} \times \cdots \times \mathcal{R}_{K,i} = \{(\boldsymbol{x}_{1,i}, \ldots, \boldsymbol{x}_{K,i}) \,|\, \boldsymbol{x}_{1,i} \in \mathcal{R}_{1,i}, \ldots, \boldsymbol{x}_{K,i} \in \mathcal{R}_{K,i}\}. \tag{5}$$

Here, $\mathcal{R}_i$ denotes the total partition cell corresponding to the i–th codebook entry $\boldsymbol{P}_i$ and $\mathcal{R}_{k,i}$, with $k = 1, \ldots, K$, represents the partition cell of the i-th codebook entry corresponding to user k. The aim of the k-th user's quantizer $\mathrm{Q}_k(\bullet)$ is to identify the region $\mathcal{R}_{k,i}$ in which the CSI $\boldsymbol{z}_k$ lies. The resulting fed-back information of user k, i.e., the output of its quantizer $\mathrm{Q}_k(\boldsymbol{z}_k)$, is equivalent to a set of indices $\mathbb{P}_k$ referring to the precoder representation points that best fit to its current channel state. When collecting the fed-back information from all users, the transmitter finds the index of the final precoder representation point by intersecting the sets of indices from all users. Therefore, the selector function of the overall quantizer in Eq. (4) is finally defined as

$$\mathrm{S}_i(\boldsymbol{z}) = \begin{cases} 1 & \text{for } i \in \bigcap_{k=1}^{K} \mathrm{Q}_k(\boldsymbol{z}_k) \\ 0 & \text{else.} \end{cases}$$

Note that the above intersection gives a set with cardinality one due to the properties of the cartesian product used to split $\mathcal{R}_i$ into $\mathcal{R}_{1,i}, \ldots, \mathcal{R}_{K,i}$ [see Eq. (5)]. This complicated representation is inevitable since the users are not cooperative and, therefore, no single user has information about the others. Remember that the codebook entries are the precoder representation points and the receive weights and not the CSI.

When restricting to scalar quantization, we can further decompose $\mathcal{R}_{k,i}$ as

$$\mathcal{R}_{k,i} = \mathcal{R}_{k,i}^{(1)} \times \cdots \times \mathcal{R}_{k,i}^{(N)}$$

i.e., the cartesian product of the N rectangular regions $\mathcal{R}_{k,i}^{(n)} \subseteq \mathbb{C}$, with $n = 1, ..., N$. Remember that N is the number of transmit antennas and is thus the maximum number of scalar coefficients sent from user k to the transmitter. Let us define each (complex) rectangular region $\mathcal{R}_{k,i}^{(n)}$ by means of its corner coordinates $\alpha_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$, $\beta_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$, $\alpha_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$, and $\beta_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$. In other words, the scalar quantizer for the complex-valued $z_{k,n}$ is split into two real-valued quantizers with the two quantizer indices $j_k^{(\mathrm{Re},n)}$ and $j_k^{(\mathrm{Im},n)}$. Thus, when the real and imaginary part of the n-th entry $z_{k,n}$ of $\boldsymbol{z}_k$ corresponding to the k-th user's quantizer $\mathrm{Q}_k(\bullet)$ lies in the cells $\mathcal{C}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$ and/or $\mathcal{C}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$, respectively, the conditions $\alpha_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)} \leq \Re(z_{k,n}) < \beta_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$ and/or $\alpha_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)} \leq \Im(z_{k,n}) < \beta_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$ are respectively fulfilled. In that case, a set $\mathbb{P}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)}$ or $\mathbb{P}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)}$ of indices is implicitly chosen, for which it holds that

$$\mathbb{P}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)} = \left\{ i = 1, \ldots, M \,\middle|\, \mathrm{Re}\left(\mathcal{R}_{k,i}^{(n)}\right) = \mathcal{C}_{k,j_k^{(\mathrm{Re},n)}}^{(\mathrm{Re},n)} \right\} \tag{6}$$

or

$$\mathbb{P}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)} = \left\{ i = 1, \ldots, M \,\middle|\, \mathrm{Im}\left(\mathcal{R}_{k,i}^{(n)}\right) = \mathcal{C}_{k,j_k^{(\mathrm{Im},n)}}^{(\mathrm{Im},n)} \right\} \tag{7}$$

respectively. The information that user k feeds back are the indices $j_k^{(\mathrm{Re},n)}$ and $j_k^{(\mathrm{Im},n)}$ with $n = 1,\ldots,N$. To obtain the quantizer output $\mathrm{Q}_k(\bullet)$, the quantized results for the different real and imaginary parts of the entries $z_{k,n}$, $n = 1,\ldots,N$, i.e., $j_k^{(\mathrm{Re},n)}$ and $j_k^{(\mathrm{Im},n)}$, should be combined by simply intersecting the sets $\mathbb{P}^{(1)}_{k,j_k^{(1)}},\ldots,\mathbb{P}^{(N)}_{k,j_k^{(N)}}$, where $\mathbb{P}^{(n)}_{k,j_k^{(n)}} = \mathbb{P}^{(\mathrm{Re},n)}_{k,j_k^{(\mathrm{Re},n)}} \cap \mathbb{P}^{(\mathrm{Im},n)}_{k,j_k^{(\mathrm{Im},n)}}$:

$$\mathrm{Q}_k(\boldsymbol{z}_k) = \mathbb{P}_k = \bigcap_{n=1}^{N} \mathbb{P}^{(n)}_{k,j_k^{(n)}}.$$

III. Proposed MMSE Optimization

In this section, we focus on the optimization of the following elements pertaining to the limited feedback channel: the channel estimators $\{\boldsymbol{G}_k\}_{k=1}^K$ and the quantizers $\{\mathrm{Q}_k(\bullet)\}_{k=1}^K$, i.e., the partition cells $\{\mathcal{R}_i\}_{i=1}^M$ and the precoders representation points $\{\boldsymbol{P}_i\}_{i=1}^M$. We choose as a feasible designing criterion the minimization of the MSE between the transmitted and received symbols, that is,

$$\mathrm{MSE} = \mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2\right] = \sum_{i=1}^{M} p_i\, \mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2 \middle| \boldsymbol{z}\in\mathcal{R}_i\right] \tag{8}$$

where p_i denotes the probability that $\boldsymbol{z}\in\mathcal{R}_i$. Taking into account that the output signals at the receivers are given by $\hat{\boldsymbol{u}} = g(\boldsymbol{H}\boldsymbol{P}\boldsymbol{u}+\boldsymbol{\eta})$ [see Eq. (2)], where $\boldsymbol{P}$ is the precoder obtained from the overall quantizer, i.e., $\boldsymbol{P} = \mathrm{Q}(\boldsymbol{z}) = \sum_{i=1}^M \boldsymbol{P}_i\,\mathrm{S}_i(\boldsymbol{z})$ [cf. Eq. (4)] and $g = \sum_{i=1}^M g_i\,\mathrm{S}_i(\boldsymbol{z})$, we can further elaborate the MSE cost function as follows

$$\mathrm{MSE} = \sum_{i=1}^{M} p_i \left(K - 2g_i\Re\left(\mathrm{tr}\left(\mathrm{E}\left[\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right]\boldsymbol{P}_i\right)\right) + g_i^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\ + g_i^2\,\mathrm{tr}\left(\mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right]\boldsymbol{P}_i\boldsymbol{P}_i^{\mathrm{H}}\right)\right) \tag{9}$$

due to $\mathrm{E}[\boldsymbol{u}\boldsymbol{\eta}^{\mathrm{H}}] = \boldsymbol{0}_K$ and $\mathrm{E}[\boldsymbol{u}\boldsymbol{u}^{\mathrm{H}}] = \mathbf{I}_K$. Note again that we neglect the delay of the feedback in our system model for the sake of brevity.

The optimization problem that we have to solve is

$$\{\{\boldsymbol{G}_k\}_{k=1}^K, \{\boldsymbol{P}_i\}_{i=1}^M, \{\mathcal{R}_i\}_{i=1}^M\}_{\text{opt}} = \underset{\{\{\boldsymbol{G}_k\}_{k=1}^K, \{\boldsymbol{P}_i\}_{i=1}^M, \{\mathcal{R}_i\}_{i=1}^M\}}{\mathrm{argmin}} \mathrm{MSE} = \mathrm{E}\left[\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2\right]$$
$$\text{subject to: } \mathrm{E}\left[\|\boldsymbol{P}\boldsymbol{u}\|_2^2\right] \le E_{\text{tx}}. \tag{10}$$

Unfortunately, no closed form expressions can be obtained for both the estimators and the quantizers of the feedback systems. Instead, we will follow an alternating optimization approach to minimize the MSE, because it is possible to obtain closed form expressions for the minimization of certain quantities while the other quantities are kept fixed. Indeed, let us start by fixing the partition regions $\mathcal{R}_i$ and the precoder representation points $\boldsymbol{P}_i$. It is possible to obtain a closed-form expression for the optimum estimator $\boldsymbol{G}_k$ and afterwards use the Lloyd algorithm to iteratively optimize the partition cells and codebook representation points of the quantizers of each user.

IV. Channel Estimators

In this subsection, the channel estimator $\boldsymbol{G}_k$ is optimized for a given codebook (precoder and receiver weights) and partition cells. It is apparent from Eq. (3) that

$$\boldsymbol{C}_{\boldsymbol{z},k} = \mathrm{E}\left[\boldsymbol{z}_k\boldsymbol{z}_k^{\mathrm{H}}\right] = \boldsymbol{G}_k\left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)\boldsymbol{G}_k^{\mathrm{H}}.$$

Thus, we can write the following alternative parameterization of the channel estimator

$$\boldsymbol{G}_k = \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{X}_k^{\mathrm{H}} \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2} \tag{11}$$

where the unknown $\boldsymbol{X}_k \in \mathbb{C}^{N_{\mathrm{tr}} \times N}$ has orthonormal columns, i.e., $\boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{X}_k = \mathbf{I}_N$. It is very easy to verify that this expression for $\boldsymbol{G}_k$ leads to $\boldsymbol{C}_{\boldsymbol{z},k}$ when we substitute it into $\boldsymbol{G}_k(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})\boldsymbol{G}_k^{\mathrm{H}}$. Note that the transformation of $\boldsymbol{S}\boldsymbol{h}_k + \boldsymbol{\eta}_k$ with $(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})^{-1/2}$ leads to an uncorrelated signal with unit covariance matrix and the additional transformation with $\boldsymbol{X}_k^{\mathrm{H}}$ again gives an uncorrelated signal with unit covariance matrix no matter the choice for $\boldsymbol{C}_{\boldsymbol{z},k}$. Therefore, the optimization with respect to $\boldsymbol{G}_k$ can be split into an optimization with respect to $\boldsymbol{X}_k$ and a subsequent optimization with respect to $\boldsymbol{C}_{\boldsymbol{z},k}$.

Before carrying out the minimization of the MSE $\mathrm{E}[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2]$ with respect to $\boldsymbol{X}_k$, let us rewrite the MSE in terms of an auxiliary matrix $\boldsymbol{A}_k$ defined as

$$\boldsymbol{A}_k = \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k} \right)^{-1/2} \in \mathbb{C}^{N \times N_{\mathrm{tr}}}. \tag{12}$$

To this end, let us obtain the conditional moments $\mathrm{E}[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i]$ and $\mathrm{E}[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i]$. Taking into account that $\boldsymbol{h}_k$ and $\boldsymbol{z}_k$ are jointly Gaussian, we have

$$\begin{bmatrix} \boldsymbol{h}_k \\ \boldsymbol{z}_k \end{bmatrix} \sim \mathcal{N}_{\mathbb{C}} \left(\boldsymbol{0}, \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{h},k} & \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}^{\mathrm{H}} \\ \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k} & \boldsymbol{C}_{\boldsymbol{z},k} \end{bmatrix} \right)$$

where $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}$ is given by [see Eqs. (3), (11), and (12)]

$$\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k} = \mathrm{E}\left[\boldsymbol{z}_k \boldsymbol{h}_k^{\mathrm{H}}\right] = \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}}. \tag{13}$$

Thus, the conditional moments are (e.g., [16])

$$\begin{aligned}
\mathrm{E}[\boldsymbol{h}_k|\boldsymbol{z}_k] &= \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{z},k}^{-1} \boldsymbol{z}_k = \boldsymbol{A}_k \boldsymbol{X}_k \boldsymbol{C}_{\boldsymbol{z},k}^{-1/2} \boldsymbol{z}_k \\
\mathrm{E}[\boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}}|\boldsymbol{z}_k] &= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{z},k}^{-1} \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k} + \mathrm{E}[\boldsymbol{h}_k|\boldsymbol{z}_k]\,\mathrm{E}[\boldsymbol{h}_k|\boldsymbol{z}_k]^{\mathrm{H}} \\
&= \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k \boldsymbol{X}_k \boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}} + \boldsymbol{A}_k \boldsymbol{X}_k \boldsymbol{C}_{\boldsymbol{z}_k}^{-1/2} \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{z},k}^{-1/2,\mathrm{H}} \boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}}.
\end{aligned}$$

Clearly, it holds that $\mathrm{E}[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i]) = \mathrm{E}[\mathrm{E}[\boldsymbol{H}|\boldsymbol{z}]|\boldsymbol{z} \in \mathcal{R}_i]$. Therefore, taking into account that $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]^{\mathrm{T}}$, we have

$$\begin{aligned}
\mathrm{E}[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i] &= [\boldsymbol{A}_1 \boldsymbol{X}_1 \boldsymbol{\mu}_{1,i}, \ldots, \boldsymbol{A}_K \boldsymbol{X}_K \boldsymbol{\mu}_{K,i}]^{\mathrm{T}} \qquad (14) \\
\mathrm{E}[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i] &= \sum_{k=1}^{K} \left(\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k \boldsymbol{X}_k \left(\mathbf{I} - \boldsymbol{R}_{k,i} \right) \boldsymbol{X}_k^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}} \right)^{\mathrm{T}}
\end{aligned}$$

with [cf. Eq. (5)]

$$\begin{aligned}
\boldsymbol{\mu}_{k,i} &= \mathrm{E}\left[\boldsymbol{C}_{\boldsymbol{z},k}^{-1/2} \boldsymbol{z}_k \middle| \boldsymbol{z}_k \in \mathcal{R}_{k,i} \right] \\
\boldsymbol{R}_{k,i} &= \mathrm{E}\left[\boldsymbol{C}_{\boldsymbol{z},k}^{-1/2} \boldsymbol{z}_k \boldsymbol{z}_k^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{z},k}^{-1/2,\mathrm{H}} \middle| \boldsymbol{z}_k \in \mathcal{R}_{k,i} \right].
\end{aligned}$$

Notice that $\boldsymbol{\mu}_{k,i}$ and $\boldsymbol{R}_{k,i}$ only depend on the choice of the partition regions $\mathcal{R}_{k,i}$ which are assumed to be given in this section.

The obtained results for $\mathrm{E}[\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i]$ and $\mathrm{E}[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i]$ can be substituted into Eq. (9). Thus, the MSE for the given codebook entries $\{\boldsymbol{P}_i, g_i\}_{i=1}^M$ and partition cells $\{\mathcal{R}_i\}_{i=1}^M$ is expressed as

$$\mathrm{MSE} = \sum_{i=1}^M p_i \left(K - 2g_i \Re\left(\mathrm{tr}\left(\left[\boldsymbol{A}_1\boldsymbol{X}_1\boldsymbol{\mu}_{1,i},\ldots,\boldsymbol{A}_K\boldsymbol{X}_K\boldsymbol{\mu}_{K,i}\right]^{\mathrm{T}}\boldsymbol{P}_i\right)\right) + g_i^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\right.$$
$$\left. + g_i^2\sum_{k=1}^K \mathrm{tr}\left(\left(\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k\boldsymbol{X}_k\left(\mathbf{I}-\boldsymbol{R}_{k,i}\right)\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}\right)^{\mathrm{T}}\boldsymbol{P}_i\boldsymbol{P}_i^{\mathrm{H}}\right)\right). \tag{15}$$

As mentioned before, thanks to introducing the alternative representation of the channel estimator $\boldsymbol{G}_k$ in Eq. (11), we can obtain the optimum channel estimator by finding the basis $\boldsymbol{X}_k$ that minimizes the above MSE expression for a fixed $\boldsymbol{C}_{\boldsymbol{z},k}$, i.e.,

$$\boldsymbol{X}_{\mathrm{opt},k} = \underset{\boldsymbol{X}_k}{\mathrm{argmin}}\,\mathrm{MSE} \qquad \text{subject to} \quad \boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{X}_k = \mathbf{I}_N$$

where the constraint has been introduced to ensure the sub-unitarity of $\boldsymbol{X}_k\in\mathbb{C}^{N_{\mathrm{tr}}\times N}$. Let us solve this optimization problem using the Lagrangian multipliers method. The corresponding Lagrangian function reads as

$$L(\boldsymbol{X}_k,\boldsymbol{\Lambda}_k) = \mathrm{MSE} + \mathrm{tr}\left(\boldsymbol{\Lambda}_k\left(\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{X}_k - \mathbf{I}\right)\right)$$

where $\boldsymbol{\Lambda}_k\in\mathbb{C}^{N\times N}$ is the Lagrangian multiplier which is Hermitian by definition since the constraint is Hermitian. A necessary condition for optimality is that

$$\frac{\partial L(\boldsymbol{X}_k,\boldsymbol{\Lambda}_k)}{\partial \boldsymbol{X}_k^{\mathrm{T}}} = \frac{\partial \mathrm{MSE}}{\partial \boldsymbol{X}_k^{\mathrm{T}}} + \boldsymbol{\Lambda}_k\boldsymbol{X}_k^{\mathrm{H}} = \mathbf{0}.$$

From this Karush-Kuhn-Tucker (KKT) condition we obtain that [cf. Eq. (15)]

$$\sum_{i=1}^M -p_i\boldsymbol{\mu}_{k,i}\boldsymbol{e}_k^{\mathrm{T}}\boldsymbol{P}_i^{\mathrm{T}}g_i\boldsymbol{A}_k - p_i\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}g_i^2\boldsymbol{P}_i^*\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{A}_k + p_i\boldsymbol{R}_{k,i}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}g_i^2\boldsymbol{P}_i^*\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{A}_k + \boldsymbol{\Lambda}_k\boldsymbol{X}_k^{\mathrm{H}} = \mathbf{0}.$$

Since the range of the first three summands reachable for row vectors multiplied from the left is the span of the rows of $\boldsymbol{A}_k$, the space spanned by the rows of $\boldsymbol{X}_k^{\mathrm{H}}$ must be the same to fulfill the above condition and thus

$$\mathrm{range}\left(\boldsymbol{X}_k\right) = \mathrm{range}\left(\boldsymbol{A}_k^{\mathrm{H}}\right). \tag{16}$$

By considering the *Singular Value Decomposition* (SVD) of a matrix $\boldsymbol{B}=\boldsymbol{M}\boldsymbol{D}\boldsymbol{N}^{\mathrm{H}}$, where $\boldsymbol{D}$ is a square diagonal matrix and $\boldsymbol{M}$ and $\boldsymbol{N}$ are unitary or sub-unitary, it is satisfied that the range of $\boldsymbol{B}$ is equal to the range of $\boldsymbol{M}$ [17]. Having in mind this property and the SVD decomposition of $\boldsymbol{A}_k$ given by

$$\boldsymbol{A}_k = \boldsymbol{V}_k\boldsymbol{\Phi}_k\boldsymbol{W}_k^{\mathrm{H}}$$

with unitary $\boldsymbol{V}_k\in\mathbb{C}^{N\times N}$, diagonal $\boldsymbol{\Phi}_k = \mathrm{diag}(\phi_{k,1},\cdots,\phi_{k,N})\in\mathbb{R}^{N\times N}$ whose diagonal elements $\phi_{k,i}$ are positive, and sub-unitary $\boldsymbol{W}_k\in\mathbb{C}^{N_{\mathrm{tr}}\times N}$, we have that $\mathrm{range}(\boldsymbol{A}_k^{\mathrm{H}}) = \mathrm{range}(\boldsymbol{W}_k)$. Thus, we can conclude that the optimal basis is given by

$$\boldsymbol{X}_{\mathrm{opt},k} = \boldsymbol{W}_k\boldsymbol{U}_k^{\mathrm{H}}\in\mathbb{C}^{N_{\mathrm{tr}}\times N} \tag{17}$$

to fulfill the condition in Eq. (16).

The so far undefined unitary $\boldsymbol{U}_k \in \mathbb{C}^{N\times N}$ must be chosen to minimize the precoding MSE in Eq. (15). Since $\boldsymbol{\Phi}_k \boldsymbol{W}_k^{\mathrm{H}} = \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{A}_k$, the optimal estimator must have the form [cf. Eq. (11)]

$$\begin{aligned}\boldsymbol{G}_{\text{opt},k} &= \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{U}_k \boldsymbol{W}_k^{\mathrm{H}} \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)^{-1/2} = \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{U}_k \boldsymbol{\Phi}_k^{-1} \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{A}_k \left(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k}\right)^{-1/2} \\ &= \boldsymbol{C}_{\boldsymbol{z},k}^{1/2} \boldsymbol{U}_k \boldsymbol{\Phi}_k^{-1} \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{G}_{\text{MMSE-estim},k} \end{aligned} \tag{18}$$

where the conventional linear *Minimum Mean Square Error* (MMSE) estimator is given by

$$\boldsymbol{G}_{\text{MMSE-estim},k} = \boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}}(\boldsymbol{S}\boldsymbol{C}_{\boldsymbol{h},k}\boldsymbol{S}^{\mathrm{H}} + \boldsymbol{C}_{\boldsymbol{\eta},k})^{-1}.$$

Let us examine in more detail the expression for the optimal estimator given by Eq. (18). Notice that $\boldsymbol{V}_k^{\mathrm{H}}$ decorrelates the output of the linear MMSE estimator and $\boldsymbol{\Phi}_k^{-1}$ forces that its variance be the identity matrix. Then, some rotation with $\boldsymbol{U}_k$ is applied that does not change the property of unit covariance and, finally, the estimate is colored with $\boldsymbol{C}_{\boldsymbol{z},k}^{1/2}$. This result is quite surprising and is a consequence of not optimizing the mean squared error between the true channel and the channel recovered at the transmitter but the precoding MSE $\mathrm{E}[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2]$ [see Eq. (10)].

We also see from Eq. (18) that the optimal estimator $\boldsymbol{G}_{\text{opt},k}$ can be written in closed form except for the covariance matrix $\boldsymbol{C}_{\boldsymbol{z},k}$ and the unitary matrix $\boldsymbol{U}_k$. The optimization of these two parts of the estimator is difficult and cannot be done analytically. However, they can be moved into the quantizer $\mathrm{Q}_k(\bullet)$ as in [14] by a proper redefinition of the partition cells $\mathcal{R}_{k,i}$. Therefore, we can set without loss of optimality that

$$\boldsymbol{G}_{\text{opt},k} = \boldsymbol{\Phi}_k^{-1} \boldsymbol{V}_k^{\mathrm{H}} \boldsymbol{G}_{\text{MMSE-estim},k} \in \mathbb{C}^{N\times N_{\text{tr}}} \tag{19}$$

and proceed to the quantization of this estimator's output instead of quantizing the output of the estimator given in Eq. (18). Accordingly, the optimal $\boldsymbol{X}_k$ in the parameterization of Eq. (11) is

$$\boldsymbol{X}_{\text{opt},k} = \boldsymbol{W}_k \tag{20}$$

with the SVD $\boldsymbol{A}_k = \boldsymbol{V}_k \boldsymbol{\Phi}_k \boldsymbol{W}_k^{\mathrm{H}}$. Additionally, $\boldsymbol{C}_{\boldsymbol{z},k} = \mathbf{I}$.

A. MSE with Optimal Estimators

The advantage of the approach described above is that now the optimal estimator is independent of the codebook and the other estimators. Additionally, notice that the estimator's output $\boldsymbol{z}_k$ is Gaussian distributed with unit covariance matrix. Thus, we rename the estimator output as $\boldsymbol{w}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$. Due to the relationship between $\boldsymbol{X}_{\text{opt},k}$ and $\boldsymbol{A}_k$ [see Eq. (20)], we have [cf. Eq. (15)]

$$\boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{A}_k \boldsymbol{X}_{\text{opt},k} \boldsymbol{X}_{\text{opt},k}^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}} = \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_k \boldsymbol{\Phi}_k^2 \boldsymbol{V}_k^{\mathrm{H}}$$

and

$$\boldsymbol{A}_k \boldsymbol{X}_{\text{opt},k} \boldsymbol{R}_{k,i} \boldsymbol{X}_{\text{opt},k}^{\mathrm{H}} \boldsymbol{A}_k^{\mathrm{H}} = \boldsymbol{V}_k \boldsymbol{\Phi}_k \boldsymbol{R}_{k,i} \boldsymbol{\Phi}_k \boldsymbol{V}_k^{\mathrm{H}}.$$

Having in mind the above results, the conditional moments from Eq. (14) can be rewritten as

$$\begin{aligned}\mathrm{E}\left[\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right] &= \left[\boldsymbol{\mu}_{1,i},\ldots,\boldsymbol{\mu}_{K,i}\right]^{\mathrm{T}}\\ \mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right] &= \sum_{k=1}^{K}\left(\boldsymbol{C}_{\boldsymbol{h},k}-\boldsymbol{V}_k\boldsymbol{\Phi}_k^2\boldsymbol{V}_k^{\mathrm{H}}+\boldsymbol{R}_{k,i}\right)^{\mathrm{T}}\end{aligned} \tag{21}$$

where $\boldsymbol{\mu}_{k,i}$ and $\boldsymbol{R}_{k,i}$ are redefined as

$$\begin{aligned}\boldsymbol{\mu}_{k,i} &= \boldsymbol{V}_k\boldsymbol{\Phi}_k\,\mathrm{E}\left[\boldsymbol{w}_k\,|\,\boldsymbol{w}_k\in\mathcal{R}_{k,i}\right]\\ \boldsymbol{R}_{k,i} &= \boldsymbol{V}_k\boldsymbol{\Phi}_k\,\mathrm{E}\left[\boldsymbol{w}_k\boldsymbol{w}_k^{\mathrm{H}}\,|\,\boldsymbol{w}_k\in\mathcal{R}_{k,i}\right]\boldsymbol{\Phi}_k\boldsymbol{V}_k^{\mathrm{H}}\end{aligned} \tag{22}$$

with $\boldsymbol{w}_k\sim\mathcal{N}_{\mathbb{C}}(\boldsymbol{0},\mathbf{I})\ \forall k$. Applying $\mathrm{E}[\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}}|\boldsymbol{x}]=\mathrm{E}[(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}})(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}})^{\mathrm{H}}|\boldsymbol{x}]+\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}}\boldsymbol{\mu}_{\boldsymbol{y}|\boldsymbol{x}}^{\mathrm{H}}$ to $\boldsymbol{R}_{k,i}$ leads to

$$\begin{aligned}\mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z}\in\mathcal{R}_i\right] &= \sum_{k=1}^{K}\Big(\boldsymbol{C}_{\boldsymbol{h},k}-\boldsymbol{V}_k\boldsymbol{\Phi}_k^2\boldsymbol{V}_k^{\mathrm{H}}+\boldsymbol{\mu}_{k,i}\boldsymbol{\mu}_{k,i}^{\mathrm{H}}\\ &\quad+\boldsymbol{V}_k\boldsymbol{\Phi}_k\,\mathrm{E}\left[\left(\boldsymbol{w}_k-\boldsymbol{\Phi}_k^{-1}\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\mu}_{k,i}\right)\left(\boldsymbol{w}_k-\boldsymbol{\Phi}_k^{-1}\boldsymbol{V}_k^{\mathrm{H}}\boldsymbol{\mu}_{k,i}\right)^{\mathrm{H}}|\,\boldsymbol{w}_k\in\mathcal{R}_{k,i}\right]\boldsymbol{\Phi}_k\boldsymbol{V}_k^{\mathrm{H}}\Big)^{\mathrm{T}}\\ &= \sum_{k=1}^{K}(\underbrace{\boldsymbol{C}_{\boldsymbol{h},k}-\boldsymbol{V}_k\boldsymbol{\Phi}_k^2\boldsymbol{V}_k^{\mathrm{H}}}_{\boldsymbol{C}_{\text{estim},k}}+\boldsymbol{\mu}_{k,i}\boldsymbol{\mu}_{k,i}^{\mathrm{H}}+\underbrace{\boldsymbol{V}_k\boldsymbol{\Phi}_k\boldsymbol{C}_{\mathrm{Q},k,i}\boldsymbol{\Phi}_k\boldsymbol{V}_k^{\mathrm{H}}}_{\boldsymbol{C}_{\text{quantize},k,i}})^{\mathrm{T}} \qquad (23)\\ &= \sum_{k=1}^{K}\left(\boldsymbol{C}_{\boldsymbol{h},k}+\boldsymbol{\mu}_{k,i}\boldsymbol{\mu}_{k,i}^{\mathrm{H}}-\boldsymbol{V}_k\boldsymbol{\Phi}_k\boldsymbol{\Gamma}_{k,i}\boldsymbol{\Phi}_k\boldsymbol{V}_k^{\mathrm{H}}\right)^{\mathrm{T}}. \qquad (24)\end{aligned}$$

Notice that $\boldsymbol{C}_{\text{estim},k}$ is the MSE error matrix due to the estimation with $\boldsymbol{G}_{\text{MMSE-estim},k}$ and $\boldsymbol{C}_{\text{quantize},k,i}$ is the error covariance matrix due to the quantization error. The matrix $\boldsymbol{\Gamma}_{k,i}=\mathbf{I}-\boldsymbol{C}_{\mathrm{Q},k,i}\in\mathbb{R}^{0,+}$ depends only on the quantizer parameters. Notice that when we assume perfect channel knowledge at the receiver, i.e., when there are no errors caused by estimation, $\boldsymbol{C}_{\text{estim},k}=\boldsymbol{0}$, and when there is no limited rate for the feedback, i.e., no quantization errors, we have that $\boldsymbol{C}_{\text{quantize},k,i}=\boldsymbol{0}$. Therefore, the regularization that is introduced due to imperfect CSI at the transmitter is given by $\boldsymbol{C}_{\text{estim},k}+\boldsymbol{C}_{\text{quantize},k,i}$.

Remember that the effect of feedback delay was omitted when deriving Eqs. (23) and (24). If we assume a simple Jakes model, we would have that the correlation between the channel $\boldsymbol{h}_k[q]$ at slot q and $\boldsymbol{h}_k[\nu]$, the channel delayed by $D=q-\nu$ slots, is given by

$$\mathrm{E}\left[\boldsymbol{h}_k[q]\boldsymbol{h}_k^{\mathrm{H}}[\nu]\right]=\mathrm{J}_0(2\pi f_{\mathrm{D,max},k}D/f_{\mathrm{slot}})\boldsymbol{C}_{\boldsymbol{h},k}=r_k\boldsymbol{C}_{\boldsymbol{h},k}$$

where $f_{\mathrm{D,max},k}$ is the maximum Doppler frequency of the k-th user, f_{slot} is the slot rate, and $\mathrm{J}_0(\bullet)$ is the zero-th order Bessel function of the first kind [18]. The factor r_k in the last equality is implicitly defined. Notice that the delay can be neglected when considering a speed value of $v=0$ km/h ($r_k=1$). However, the only impact on the previous derivations is that this term r_k must be included into the expression of $\boldsymbol{A}_k$ in Eq. (12) since the input of the quantizer $\boldsymbol{z}_k$ given by Eq. (3) is obtained from outdated channel vectors and, therefore, $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{h},k}=r_k\boldsymbol{C}_{\boldsymbol{z},k}^{1/2}\boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{A}_k^{\mathrm{H}}$ [cf. Eq. (13)]. Consequently, also $\boldsymbol{\Phi}_k$ is weighted with r_k.

Finally, for the sake of notational brevity, we introduce

$$\begin{aligned}\boldsymbol{M}_i &= [\boldsymbol{\mu}_{1,i},\ldots,\boldsymbol{\mu}_{K,i}]^{\mathrm{T}} \in \mathbb{C}^{K\times N}\\ \boldsymbol{C}_{\text{estim}} &= \sum_{k=1}^{K} \boldsymbol{C}_{\boldsymbol{h},k} - \boldsymbol{V}_k \boldsymbol{\Phi}_k^2 \boldsymbol{V}_k^{\mathrm{H}} \in \mathbb{C}^{N\times N}\\ \boldsymbol{C}_{\text{quantize},i} &= \sum_{k=1}^{K} \boldsymbol{V}_k \boldsymbol{\Phi}_k \boldsymbol{C}_{\mathrm{Q},k,i} \boldsymbol{\Phi}_k \boldsymbol{V}_k^{\mathrm{H}} \in \mathbb{C}^{N\times N}.\end{aligned} \tag{25}$$

This way, the precoding MSE when using the optimal estimators can be concisely written as

$$\mathrm{MSE} = \sum_{i=1}^{M} p_i \Big(K - 2\Re\left(\mathrm{tr}\left(\boldsymbol{M}_i g_i \boldsymbol{P}_i\right)\right) + g_i^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) + g_i^2\,\mathrm{tr}\left(\left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right)\boldsymbol{P}_i\boldsymbol{P}_i^{\mathrm{H}}\right)\Big). \tag{26}$$

In the ensuing section, we assume that the optimal estimators $\boldsymbol{G}_{\text{opt},k}$, $k = 1,\ldots,K$, are employed, i.e., the precoding MSE given by Eq. (26) has to be minimized when designing the quantizers. It is interesting to note that the conditional moments provided by this scheme are equal to the conditional moments obtained for the joint optimization based on a CSI-metric (see [14], [15], [19]).

V. Codebook Entries

A. *Codebook entries: precoder representation points*

In this section, we proceed with solving Eq. (10) by designing the codebook entries (precoder representation points) $\boldsymbol{P}_i$ and the respective receive weights g_i in order to minimize the precoding MSE of Eq. (26) under a transmit power constraint for a given set of partition cells $\mathcal{R}_i$, $i = 1,\ldots,M$:

$$\{\boldsymbol{P}_{\text{opt},i}, g_{\text{opt},i}\} = \underset{\{\boldsymbol{P}_i,g_i\}}{\operatorname{argmin}}\,\mathrm{MSE} \qquad \text{subject to: } \mathrm{E}\left[\|\boldsymbol{P}_i\boldsymbol{u}\|_2^2\right] \le E_{\text{tx}}. \tag{27}$$

Again, this constrained optimization problem will be solved using the method of Lagrangian multipliers.

Without destroying optimality, we make a change of variables and set $\boldsymbol{P}_i = g_i^{-1}\boldsymbol{F}_i$. Consequently, the Lagrangian function reads as

$$\begin{aligned}L\left(\boldsymbol{F}_i, g_i, \lambda\right) = \sum_{i=1}^{M} p_i \Big(& K - 2\Re\left(\mathrm{tr}\left(\boldsymbol{M}_i\boldsymbol{F}_i\right)\right) + g_i^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right)\\ & + \mathrm{tr}\left(\left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right)\boldsymbol{F}_i\boldsymbol{F}_i^{\mathrm{H}}\right) + \lambda\left(g_i^{-2}\,\|\boldsymbol{F}_i\|_{\mathrm{F}}^2 - E_{\text{tx}}\right)\Big)\end{aligned} \tag{28}$$

with the Lagrangian multiplier $\lambda \in \mathbb{R}^{0,+}$.

One KKT condition is obtained by deriving with respect to g_i, which is assumed to be real. Equating this derivative to zero yields

$$\frac{\partial L\left(\bullet\right)}{\partial g_i} = 2g_i\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) - 2\lambda g_i^{-3}\,\|\boldsymbol{F}_i\|_{\mathrm{F}}^2 = 0$$

which leads to $\lambda = g_i^2 \frac{\mathrm{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{g_i^{-2}\|\boldsymbol{F}_i\|_{\mathrm{F}}^2} > 0$. As it is apparent that the transmit energy constraint is active, that is, $g_i^{-2}\|\boldsymbol{F}_i\|_{\mathrm{F}}^2 = E_{\text{tx}}$, we have $\lambda = g_i^2\frac{\mathrm{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})}{E_{\text{tx}}}$.

When we set the derivative with respect to $\boldsymbol{F}_i^*$ to zero, we obtain the following KKT condition

$$\frac{\partial L(\bullet)}{\partial \boldsymbol{F}_i^*} = -\boldsymbol{M}_i^{\mathrm{H}} + \left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right)\boldsymbol{F}_i + \frac{\lambda}{g_i^2}\boldsymbol{F}_i = \boldsymbol{0}. \tag{29}$$

This result, together with the above result for λ and the transmit power constraint, leads to the optimal precoder representation point (codebook entry) corresponding to the i-th partition cell $\mathcal{R}_i$ given by

$$\boxed{\begin{aligned} \boldsymbol{F}_{\text{opt},i} &= \left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}} + \xi\mathbf{I}\right)^{-1}\boldsymbol{M}_i^{\mathrm{H}} \\ g_{\text{opt},i} &= \sqrt{\frac{1}{E_{\text{tx}}}\operatorname{tr}\left(\left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right)^{-2}\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i\right)} \end{aligned}} \tag{30}$$

where $\xi = \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}})/E_{\text{tx}}$. Interestingly, this result can be interpreted as the *centroid condition*. Note that we use MMSE optimal receiver weights (different for different receivers) although the optimization of Eq. (27) gives $g_{\text{opt},i}$. The MMSE optimal receiver weights correct the phase and lead to an approximately coherent detection (see [15] for more details).

Note also that the solution for the precoder representation points is inherently robust against errors, since the respective error covariance matrices regularize the pseudo inversion in the definition of $\boldsymbol{F}_{\text{opt},i} = g_{\text{opt},i}\boldsymbol{P}_{\text{opt},i}$.

Due to the expectations $\mathrm{E}\left[\boldsymbol{w}_k \mid \boldsymbol{w}_k \in \mathcal{R}_{k,i}\right]$ for $k = 1, \ldots, K$ [see Eqs. (22) and (25)], the computation of the precoder $\boldsymbol{F}_i$ is difficult for a general set of partition cells $\mathcal{R}_{1,i}, \ldots, \mathcal{R}_{K,i}$ such as those obtained when using *vector quantization*. However, by restricting ourselves to *scalar quantization*, the integration over the rectangular regions $\mathcal{R}_{k,i}^{(n)}$ can be solved in closed form (see [14], [15]). Note that this precoder is basically the same precoder as that based on the CSI MSE metric although the design considered in this paper is based on the precoding MSE only (see [14], [15]). Both linear precoders are robust against errors in CSI by means of regularization terms. Contrary to the CSI MSE metric, however, where the precoder is based on already optimized and fixed partition cells that are independent of the channel statistics, [1] the joint design according to the precoding MSE metric shown in this work optimizes the precoder and the partition cells using the Lloyd algorithm. The Lloyd algorithm switches between the precoder design and the partition cell computation and converges to locally optimum precoders and regions since every step reduces the MSE, and the MSE is lower bounded. Note that both, precoders and partition cells, must be recomputed as soon as the channel statistics change. Additionally, note that the obtained estimators in Eq. (18) are optimal for any codebook and the codebook entries in Eq. (30) are optimal for given partition cells. In the next subsection, the optimal partition cells for given codebook entries are derived. This motivates the alternating optimization of the Lloyd algorithm.

B. Partition Cells

In this subsection, we explain how to optimize the quantizer partition cells. Since the receivers do not cooperate, the estimates of other users are unknown to the quantizer of user ℓ. Thus, we will design the regions $\mathcal{R}_{\ell,i}$ of the ℓ-th quantizer in order to minimize the distortion $\mathrm{d}_\ell = \mathrm{E}[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2|\boldsymbol{z}_\ell]$ for given codebook entries $\boldsymbol{P}_i$, $i = 1, \ldots, M$,

[1]Neglecting the effect of bit allocation.

and according receive weights g_i, $i = 1, \dots, M$. Motivated by the fact that $\boldsymbol{z}_\ell \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$, i.e., the quantizer's inputs are Gaussian and uncorrelated, and that the computation of the precoders is difficult for vector quantization, we restrict ourselves to scalar quantization which implies that the entries of $\boldsymbol{z}_\ell$ are quantized separately. In this case, the partition cells $\mathcal{C}^{(\mathrm{Re},n)}_{\ell,j_k^{(\mathrm{Re},n)}}$ and $\mathcal{C}^{(\mathrm{Im},n)}_{\ell,j_k^{(\mathrm{Im},n)}}$ [see Eqs. (6) and (7)], that is, their corner coordinates $\alpha^{(\mathrm{Re},n)}_{\ell,j_k^{(\mathrm{Re},n)}}$, $\beta^{(\mathrm{Re},n)}_{\ell,j_k^{(\mathrm{Re},n)}}$, $\alpha^{(\mathrm{Im},n)}_{\ell,j_k^{(\mathrm{Im},n)}}$, and $\beta^{(\mathrm{Im},n)}_{\ell,j_k^{(\mathrm{Im},n)}}$ of the scalar quantizers for, respectively, real and imaginary parts of the n-th entry $z_{\ell,n}$ of $\boldsymbol{z}_\ell$, minimize the distortions

$$\mathrm{d}^{(\mathrm{Re},n)}_\ell(\Re[z_{\ell,n}]) = \mathrm{E}\left[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2 \middle| \Re[z_{\ell,n}]\right] = \sum_{j=1}^{M_\ell^{(n)}} \mathrm{S}^{(\mathrm{Re},n)}_{\ell,j}(\Re[z_{\ell,n}])\, \mathrm{d}^{(\mathrm{Re},n)}_{\ell,j}(\Re[z_{\ell,n}]) \tag{31}$$

and

$$\mathrm{d}^{(\mathrm{Im},n)}_\ell(\Im[z_{\ell,n}]) = \mathrm{E}\left[\|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2 \middle| \Im[z_{\ell,n}]\right] = \sum_{j=1}^{M_\ell^{(n)}} \mathrm{S}^{(\mathrm{Im},n)}_{\ell,j}(\Im[z_{\ell,n}])\, \mathrm{d}^{(\mathrm{Im},n)}_{\ell,j}(\Im[z_{\ell,n}]) \tag{32}$$

respectively. Here, $M_\ell^{(n)}$ is the number of codebook entries for the quantizers of $\Re[z_{\ell,n}]$ and $\Im[z_{\ell,n}]$. As a result of computing these expressions for each $z_{\ell,n}$, we can obtain the indices $j_\ell^{(\mathrm{Re},n)}$ and $j_\ell^{(\mathrm{Im},n)}$ that minimize these distortions, i.e., the respective partition cells $\mathcal{C}^{(\mathrm{Re},n)}_{\ell,j_k^{(\mathrm{Re},n)}}$ and $\mathcal{C}^{(\mathrm{Im},n)}_{\ell,j_k^{(\mathrm{Im},n)}}$ are optimized. Note that, given the n–th quantizer input of user ℓ, $z_{\ell,n}$, we assume that the other quantizer inputs $z_{k,n}$, with $k \neq \ell$, are unknown and, therefore, it is necessary to average over all the possible $z_{k,n}$. Although the other entries $z_{\ell,\nu}$ with $\nu \neq n$ are known to receiver ℓ, also over these quantities is averaged, since scalar quantizers are used. However, the corresponding cells are given since the codebook design is centralized at the transmitter and stored at both the transmitter and all the receivers.

The distortions due to the j-th codebook entry for both real and imaginary entries of the input $z_{\ell,n}$ read respectively as [cf. Eq. (9)]

$$\begin{aligned}\mathrm{d}^{(\mathrm{Re},n)}_{\ell,j}(\Re[z_{\ell,n}]) = \sum_{i\in\mathbb{P}^{(\mathrm{Re},n)}_{\ell,j}} \frac{p_i}{p^{(\mathrm{Re},n)}_{\ell,j}} \Bigg(& K + g_i^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}}) - \sum_{k=1,k\neq\ell}^{K} 2\Re\left(\boldsymbol{\mu}_{k,i}^{\mathrm{T}} \boldsymbol{F}_i \boldsymbol{e}_k\right) \\ & - 2\Re\left(\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Re},n),\mathrm{T}} \boldsymbol{F}_i \boldsymbol{e}_\ell\right) + \operatorname{tr}\left(\boldsymbol{C}_{\text{estim}}^{\mathrm{T}} \boldsymbol{F}_i \boldsymbol{F}_i^{\mathrm{H}}\right) + \sum_{k=1,k\neq\ell}^{K} \operatorname{tr}\left(\boldsymbol{R}_{k,i}^{\mathrm{T}} \boldsymbol{F}_i \boldsymbol{F}_i^{\mathrm{H}}\right) + \operatorname{tr}\left(\boldsymbol{R}_{\ell,i}^{(\mathrm{Re},n),\mathrm{T}} \boldsymbol{F}_i \boldsymbol{F}_i^{\mathrm{H}}\right)\Bigg)\end{aligned} \tag{33}$$

and

$$\begin{aligned}\mathrm{d}^{(\mathrm{Im},n)}_{\ell,j}(\Im[z_{\ell,n}]) = \sum_{i\in\mathbb{P}^{(\mathrm{Im},n)}_{\ell,j}} \frac{p_i}{p^{(\mathrm{Im},n)}_{\ell,j}} \Bigg(& K + g_i^2 \operatorname{tr}(\boldsymbol{C}_{\boldsymbol{\eta}}) - \sum_{k=1,k\neq\ell}^{K} 2\Re\left(\boldsymbol{\mu}_{k,i}^{\mathrm{T}} \boldsymbol{F}_i \boldsymbol{e}_k\right) \\ & - 2\Re\left(\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Im},n),\mathrm{T}} \boldsymbol{F}_i \boldsymbol{e}_\ell\right) + \operatorname{tr}\left(\boldsymbol{C}_{\text{estim}}^{\mathrm{T}} \boldsymbol{F}_i \boldsymbol{F}_i^{\mathrm{H}}\right) + \sum_{k=1,k\neq\ell}^{K} \operatorname{tr}\left(\boldsymbol{R}_{k,i}^{\mathrm{T}} \boldsymbol{F}_i \boldsymbol{F}_i^{\mathrm{H}}\right) + \operatorname{tr}\left(\boldsymbol{R}_{\ell,i}^{(\mathrm{Im},n),\mathrm{T}} \boldsymbol{F}_i \boldsymbol{F}_i^{\mathrm{H}}\right)\Bigg)\end{aligned} \tag{34}$$

where $\boldsymbol{F}_i = g_i \boldsymbol{P}_i$ and $\boldsymbol{e}_k$ denotes the k-th column of the $K \times K$ identity matrix. For $\boldsymbol{\mu}_{k,i}$ and $\boldsymbol{R}_{k,i}$, see Eq. (22). $p^{(\mathrm{Re},n)}_{\ell,j} = \sum_{i\in\mathbb{P}^{(\mathrm{Re},n)}_{l,j}} p_i$ and $p^{(\mathrm{Im},n)}_{\ell,j} = \sum_{i\in\mathbb{P}^{(\mathrm{Im},n)}_{l,j}} p_i$ are the probabilities of $\Re[z_{\ell,n}] \in \mathcal{C}^{(\mathrm{Re},n)}_{\ell,j}$ and $\Im[z_{\ell,n}] \in \mathcal{C}^{(\mathrm{Im},n)}_{\ell,j}$ [see Eqs. (6) and (7)], respectively. Additionally, the conditional moments $\boldsymbol{\mu}_{\ell,i}$ and $\boldsymbol{R}_{\ell,i}$ under the conditions $\Re[z_{\ell,n}]$

and $\Im[z_{\ell,n}]$, denoted by $\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Re},n)}$, $\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Im},n)}$, $\boldsymbol{R}_{\ell,i}^{(\mathrm{Re},n)}$, and $\boldsymbol{R}_{\ell,i}^{(\mathrm{Im},n)}$, can be found as follows [cf. Eq. (22)]:

$$\begin{aligned}\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Re},n)} &= \boldsymbol{V}_\ell \boldsymbol{\Phi}_\ell \,\mathrm{E}\left[\boldsymbol{z}_\ell \,|\, \boldsymbol{z}_\ell \in \mathcal{R}_{\ell,i}, \Re[z_{\ell,n}]\right] \\ \boldsymbol{\mu}_{\ell,i}^{(\mathrm{Im},n)} &= \boldsymbol{V}_\ell \boldsymbol{\Phi}_\ell \,\mathrm{E}\left[\boldsymbol{z}_\ell \,|\, \boldsymbol{z}_\ell \in \mathcal{R}_{\ell,i}, \Im[z_{\ell,n}]\right]\end{aligned} \tag{35}$$

and

$$\begin{aligned}\boldsymbol{R}_{\ell,i}^{(\mathrm{Re},n)} &= \boldsymbol{V}_\ell \boldsymbol{\Phi}_\ell \,\mathrm{E}\left[\boldsymbol{z}_\ell \boldsymbol{z}_\ell^{\mathrm{H}} \,|\, \boldsymbol{z}_\ell \in \mathcal{R}_{\ell,i}, \Re[z_{\ell,n}]\right] \boldsymbol{\Phi}_\ell \boldsymbol{V}_\ell^{\mathrm{H}} \\ \boldsymbol{R}_{\ell,i}^{(\mathrm{Im},n)} &= \boldsymbol{V}_\ell \boldsymbol{\Phi}_\ell \,\mathrm{E}\left[\boldsymbol{z}_\ell \boldsymbol{z}_\ell^{\mathrm{H}} \,|\, \boldsymbol{z}_\ell \in \mathcal{R}_{\ell,i}, \Im[z_{\ell,n}]\right] \boldsymbol{\Phi}_\ell \boldsymbol{V}_\ell^{\mathrm{H}}.\end{aligned} \tag{36}$$

Following the *nearest neighbor condition*, the partition cells $\mathcal{C}_{\ell,j}^{(\mathrm{Re},n)}$ must be chosen such that for any input $\Re[z_{\ell,n}]$ the minimum distortion $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$ is picked by the quantizer. Equivalently, for the imaginary part, the partition cells $\mathcal{C}_{\ell,j}^{(\mathrm{Im},n)}$ are chosen such that for any input $\Im[z_{\ell,n}]$ the quantizer uses the minimum distortion $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$. Since $\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Re},n)}$ and $\boldsymbol{\mu}_{\ell,i}^{(\mathrm{Im},n)}$ are linear, and $\boldsymbol{R}_{\ell,i}^{(\mathrm{Re},n)}$ and $\boldsymbol{R}_{\ell,i}^{(\mathrm{Im},n)}$ are quadratic functions of $\Re[z_{\ell,n}]$ and $\Im[z_{\ell,n}]$, respectively, the distortions $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$ are also quadratic functions. Thus, for the real part of $z_{\ell,n}$ the optimal cell borders $\alpha_{\ell,j}^{(\mathrm{Re},n)}$ and $\beta_{\ell,j}^{(\mathrm{Re},n)}$ are simply the roots of the quadratic polynomial equations $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) - \mathrm{d}_{\ell,j-1}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) - \mathrm{d}_{\ell,j+1}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$, respectively. The two roots that determine both cell borders, $\alpha_{\ell,j}^{(\mathrm{Re},n)}$ and $\beta_{\ell,j}^{(\mathrm{Re},n)}$, must verify $\alpha_{\ell,j-1}^{(\mathrm{Re},n)} < \alpha_{\ell,j}^{(\mathrm{Re},n)} < \beta_{\ell,j}^{(\mathrm{Re},n)}$. Again, similarly for the imaginary part of $z_{\ell,n}$, the region boundaries are given by the roots of the quadratic polynomials $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) - \mathrm{d}_{\ell,j-1}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) - \mathrm{d}_{\ell,j+1}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$.

C. Codebook Computation

Although the estimators and the quantizers are jointly optimized by minimizing the precoding MSE in Eq. (8), the codebook parameters have to be computed only once since the channel estimators are independent of the codebook choice [see Eq. (19)]. For the computation of the codebook parameters, we use the Lloyd algorithm (see [20], [21]), i.e., we alternately optimize the precoders by using the centroid condition in Eq. (30) and optimize the partition cells following the nearest neighbor condition as discussed in the previous subsection. Since the MSE in Eq. (26) is reduced in every step and the MSE is non-negative, this iterative procedure converges.

The Lloyd algorithm is initialized with the quantizers based on codebooks appropriate for unit variance complex Gaussian inputs [14]. Therefore, the parameters of these scalar quantizers can be stored and do not have to be recomputed for varying channel statistics. As a consequence, the initialization of the proposed feedback scheme based on the precoding MSE of Eq. (26) is very cheap.

Table I summarizes the overall design procedure for computing the codebook, which is basically a modified version of the Lloyd algorithm. Note that this new codebook has to be recomputed each time that the channel statistics change.

D. Bit Allocation

When using scalar quantization (transform coding, [21]) instead of vector quantization, the available bits have to be allocated to the different scalar coefficients. Contrary to the case of CSI MSE based feedback, the distortion

1. Set $m = 1$
2. Initial codebook $\mathcal{C}_1$ and regions $\{\mathcal{R}_i\}_{i=1}^{M}$
3. Set the threshold to stop the iterations ϵ_{min} and set $\epsilon = \infty$

while $\epsilon > \epsilon_{\text{min}}$ **do**

4. obtain the quadratic functions:
   $\forall l, j : \mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}])$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}])$
5. (*Nearest Neighbor Condition*) solve the quadratic equations:
   $\forall l, j : \mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) - \mathrm{d}_{\ell,j-1}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) = 0$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) - \mathrm{d}_{\ell,j+1}^{(\mathrm{Re},n)}(\Re[z_{\ell,n}]) = 0$
   $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) - \mathrm{d}_{\ell,j-1}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) = 0$ and $\mathrm{d}_{\ell,j}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) - \mathrm{d}_{\ell,j+1}^{(\mathrm{Im},n)}(\Im[z_{\ell,n}]) = 0$
   to get the new partition regions $\{\mathcal{R}_i\}_{i=1}^{M}$
6. compute the new conditional channel moments:
   $\mathrm{E}\left[\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i\right]$ and $\mathrm{E}\left[\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}|\boldsymbol{z} \in \mathcal{R}_i\right]$
7. (*Centroid condition*) compute the new precoders $\{\boldsymbol{P}_i\}_{i=1}^{M}$
8. compute the precoding MSE metric for the new codebook (precoders) $\{\boldsymbol{P}_i\}_{i=1}^{M}$ and the new partition regions $\{\mathcal{R}_i\}_{i=1}^{M}$
9. $m \leftarrow m + 1$

end while

TABLE I

CODEBOOK OPTIMIZATION.

function obtained for the case that the precoders are included in the optimization given by

$$\mathrm{MSE} = \sum_{i=1}^{M} p_i \Big(K - 2\Re\left(\mathrm{tr}\left(\boldsymbol{M}_i g_i \boldsymbol{P}_i\right)\right) + g_i^2\,\mathrm{tr}\left(\boldsymbol{C}_{\boldsymbol{\eta}}\right) + g_i^2\,\mathrm{tr}\left(\left(\boldsymbol{M}_i^{\mathrm{H}}\boldsymbol{M}_i + \boldsymbol{C}_{\text{estim}}^{\mathrm{T}} + \boldsymbol{C}_{\text{quantize},i}^{\mathrm{T}}\right)\boldsymbol{P}_i\boldsymbol{P}_i^{\mathrm{H}}\right)\Big) \tag{37}$$

has a very complicated structure since all the parameters are mixed together. Thus, it is impossible to separate the influence relative to each user and each scalar quantizer which makes it very difficult to find an efficient optimum bit allocation. We can therefore decide the optimum bit allocation by trying out all the possible bit allocation combinations and taking as a result the best one in terms of minimizing the MSE in Eq. (37).

The bit allocation optimization is expressed as

$$\boldsymbol{B}_{\text{opt}} = \underset{\boldsymbol{B}}{\operatorname{argmin}}\,\mathrm{MSE}\left(\boldsymbol{B}\right) \quad \text{subject to:} \quad \boldsymbol{B} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_K] \in \mathbb{B}^{N \times K}, \boldsymbol{b}_k = [b_{k,1}, \ldots, b_{k,N}]^{\mathrm{T}}$$
$$\text{with} \quad \mathbb{B} = 0, 2, 4, \ldots \quad \text{and} \quad \sum_{n=1}^{N} b_{k,n} = N_{\text{bit}} \tag{38}$$

where $\boldsymbol{B}$ is the matrix that determines the bit allocation corresponding to the coefficients of each user and N_{bit} is the number of bits available for each user. Notice that only an even number of bits is used to quantize each coefficient, since both real and imaginary parts of each coefficient make use of the same number of bits. Initially, we use the scalar quantizers (codebook entries and partition cells) obtained from the CSI metric for a unit-variance input as in [14].

Bits per user	No bit allocation	Rank reduction	Heur. bit allocation
$N_{\text{bit}} = 6$ 3 for real part 3 for imaginary part	$[2,2,2,0]^{\mathrm{T}}$	$[4,2,0,0]^{\mathrm{T}}$	Select the best from: $[2,2,2,0]^{\mathrm{T}}$, $[4,2,0,0]^{\mathrm{T}}$ $[6,0,0,0]^{\mathrm{T}}$
$N_{\text{bit}} = 8$ 4 for real part 4 for imaginary part	$[2,2,2,2]^{\mathrm{T}}$	$[4,4,0,0]^{\mathrm{T}}$	Select the best from: $[2,2,2,2]^{\mathrm{T}}$, $[4,2,2,0]^{\mathrm{T}}$ $[4,4,0,0]^{\mathrm{T}}$, $[6,2,0,0]^{\mathrm{T}}$ $[8,0,0,0]^{\mathrm{T}}$
$N_{\text{bit}} = 10$ 5 for real part 5 for imaginary part	$[4,2,2,2]^{\mathrm{T}}$	$[4,4,2,0]^{\mathrm{T}}$	Select the best from: $[4,2,2,2]^{\mathrm{T}}$, $[4,4,2,0]^{\mathrm{T}}$ $[6,4,0,0]^{\mathrm{T}}$, $[8,2,0,0]^{\mathrm{T}}$ $[10,0,0,0]^{\mathrm{T}}$

TABLE II

NUMBER OF BITS ASSIGNED PER USER'S COEFFICIENT FOR PRECODING MSE METRIC.

When the number of bits is low, there are no serious problems arising from the computational complexity, but the search for optimum bit allocation becomes infeasible as the number of bits increases. Therefore, we propose a heuristic solution to the problem by reducing the number of combinations to be tested on the MSE. It seems that an uniform distribution over all the coefficients without implementing rank reduction is the most likely allocation in the sense of minimizing the MSE. Thus, a first trial consists of distributing the bits over all the coefficients as uniformly as possible. On the other hand, it is obvious that the coefficients with more energy, i.e., the coefficients whose eigenvalues are larger, have more impact on the final MSE performance and, therefore, we must tend to allocate more bits to the first coefficients in order to minimize the MSE. Bearing this fact in mind, successive combinations will move the bits from the initial bit allocation to the coefficients with larger eigenvalues. Therefore, the MSE of Eq. (37) is sequentially computed by following this ordering for bit allocation so the process is stopped when, given a certain bit allocation, the MSE is greater than the previous one in the list. This will be termed *heuristic bit allocation*.

To illustrate this idea, let us assume that we have to distribute 8 bits for each user (see Table II). According to the heuristic bit allocation described above, the chain of possible bit allocations is given by $[2,2,2,2]^{\mathrm{T}} \rightarrow [4,2,2,0]^{\mathrm{T}} \rightarrow [4,4,0,0]^{\mathrm{T}} \rightarrow [6,2,0,0]^{\mathrm{T}} \rightarrow [8,0,0,0]^{\mathrm{T}}$. Imagine the combination given by $[4,2,2,0]^{\mathrm{T}}$ gives us less MSE than $[2,2,2,2]^{\mathrm{T}}$. In that case, we have to test the result when $[4,4,0,0]^{\mathrm{T}}$ is considered. As long as the new MSE obtained is less than the previous one, we have to continue with the search until the last possibility embodied by $[8,0,0,0]^{\mathrm{T}}$. If not, we choose $[4,2,2,0]^{\mathrm{T}}$ as the optimum bit allocation for our joint approach based on precoding MSE metric. This heuristic solution significantly reduces the computational complexity of the search with negligible loss in performance.

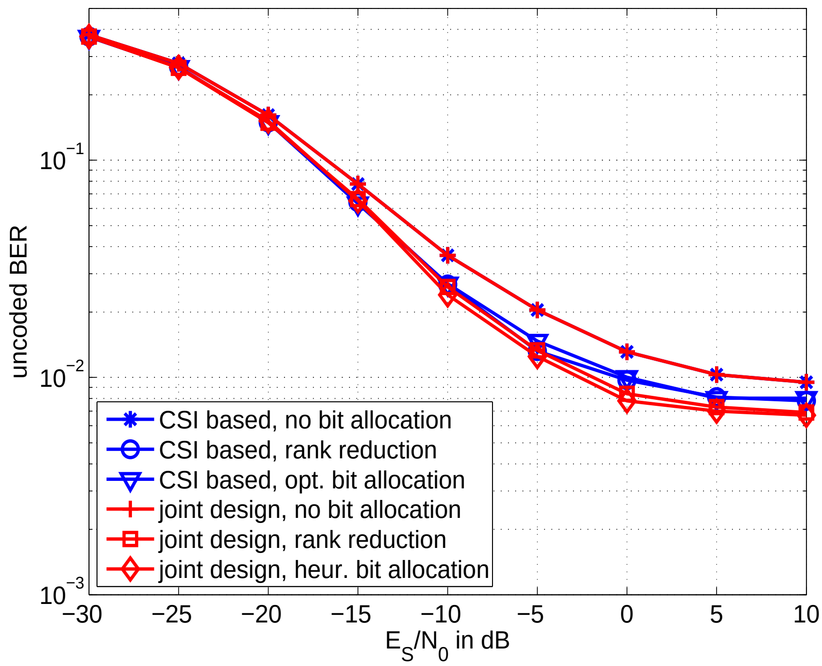

Fig. 3. MU–MISO system with robust linear precoding, $N = 4$ antennas, $K = 2$ users, and 8 bits per user.

VI. Simulation Results

Given the enormous computational complexity due to the calculation of the distortions in Subsection V-B, we consider a system with a transmitter equipped with $N = 4$ antennas that serves $K = 2$ users using QPSK modulation. We use the urban micro *Spatial Channel Model* (SCM) described in [22], which is the most difficult for precoding, out of the three spatial channel models introduced in [22], because the second and the third channel eigenvalues have a non-negligible magnitude. The results for the CSI metric are the mean of 100 channel realizations with $1\,000$ symbols being transmitted per channel realization. The number of averaged channel settings or channel covariance matrices is 10. The training sequence has $N_{\text{tr}} = 16$ symbols. In the figures, the number of bits per user is given in the legends. Although the optimization of Eq. (27) gives the weight $g_{\text{opt},i}$, we use MMSE receive weights instead those weights arising from the optimization to correct the phase caused by imperfect CSI at the transmitter and get an approximately coherent detection [14], [15].

We implemented three different types of bit allocation. First, *no bit allocation*, which tries to spread the bits as uniformly as possible (in the case that any bits are left over, e.g. with 10 bits for 4 dimensions, the dimensions corresponding to the largest $\phi_{k,i}$ get additional bits). Second, *rank reduction*, which allocates as evenly as possible the bits to the first d dimensions. And third, the *heuristic bit allocation*, which tries out different bit allocations and takes the result of the best one. Remember that we do not try all the possible combinations but the heuristic search explained in Subsection V-D is performed instead. To illustrate the different strategies, Table II summarizes the bit allocation strategies for different number of bits per user.

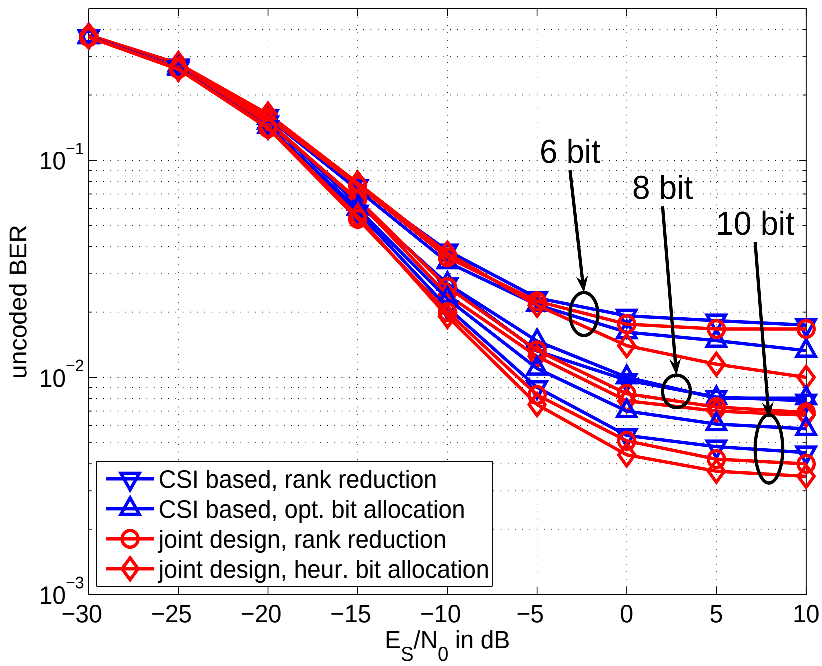

Fig. 4. MU-MISO system with robust linear precoding, $N = 4$ antennas, $K = 2$ users with different number of bits per user.

In Fig. 3, the feedback design based on the CSI MSE discussed in [14], [15] is compared to the scheme proposed in this paper, that minimizes the precoding MSE, for 8 bits fed back per user. As expected, bit allocation has a considerable impact on the BER performance and the feedback design based on the precoding MSE outperforms the CSI MSE feedback.

Though the result that the uncoded BER saturates for high SNR is disappointing, it cannot be avoided in a system with limited rate feedback (e.g., eight bits per user in Fig. 3). The saturation of the BER results from the residual interference caused by the errors in the channel state information delivered to the transmitter via limited rate feedback. To circumvent this saturation, a feedback data rate increasing with the SNR would be necessary (see e.g., [23]). However, such a setup is impractical.

Similar results were obtained for a higher and lower number of bits per user, as shown in Fig. 4. Not surprisingly, a higher number of bits per user improves the BER performance of all schemes. Additionally, it seems that the advantage of the precoding MSE based design compared to the CSI MSE based design becomes more pronounced for a higher number of bits as the degrees of freedom increase.

Notice that, independently from the number of bits fed back per user, rank reduction always shows a loss in performance with respect to heuristic bit allocation since the information contained on some coefficients is dropped.

VII. Conclusions

In this work, we have shown how to obtain the robust precoder parameters, the channel estimators, and the quantizer parameters in a joint optimization by minimizing the MSE between the transmitted symbols and the estimated symbols. Interestingly, the channel estimators and precoders obtained with the metric oriented to the precoder are equal to the estimators and precoders resulting from the joint optimization based on minimizing the MSE between the true and estimated channel presented in [14], [15]. However, the crucial part of the scheme proposed in this work is the design of the partition cells corresponding to each user, which are designed by minimizing its own distortion but averaging over the quantizer inputs for the other users, since there is no cooperation between users in the downlink of a multiuser MISO system.

As a result, we get better BER performance with a no increase of the overhead in the feedback channel. The transmitter performs the intersection of the precoder sets corresponding to the indices received from all the users to find out the optimal precoder to be used during the transmission. It is important to note that the codebook entries are now the precoders rather than the white channel coefficients. Therefore, it is obvious that the design of the quantizer parameters (i.e., the codebook entries and the partition cells) becomes the hardest part of this new precoding approach, with the advantage of minimizing the MSE by including the precoder in the optimization. This improvement is even more significant when the number of fed-back bits per user is increased, albeit at the cost of higher computational complexity.

Acknowledgment

This work was supported by Xunta de Galicia, Ministerio de Educación y Ciencia, Ministerio de Ciencia e Innovación of Spain, and FEDER funds of the European Union under grants number 09TIC008105PR, TEC2010-19545-C04-01, and CSD2008-00010.

References

[1] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient Use of Side Information in Multiple-Antenna Data Transmission over Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, October 1998.

[2] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback Gain in Multiple Antenna Systems," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 785–798, May 2002.

[3] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian Beamforming for Multiple-Input Multiple-Output Wireless Systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2735–2747, October 2003.

[4] V. Lau, Y. Liu, and T. Chen, "On the Design of MIMO Block-Fading Channels With Feedback-Link Capacity Constraint," *IEEE Transactions on Communications*, vol. 52, no. 1, pp. 62–70, January 2004.

[5] J. C. Roh and B. D. Rao, "Transmit Beamforming in Multiple-Antenna Systems With Finite Rate Feedback: A VQ-Based Approach," *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1101–1112, March 2006.

[6] V. Lau, Y. Liu, and T. Chen, "Capacity of Memoryless Channels and Block-Fading Channels With Designable Cardinality-Constrained Channel State Feedback," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 2038–2049, September 2004.

[7] T. Yoo and A. Goldsmith, "Capacity and Power Allocation for Fading MIMO Channels With Channel Estimation Error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[8] M. Joham, P. M. Castro, J. Zhen, L. Castedo, and W. Utschick, "Joint Design of Limited Feedback and Multiuser Precoding Based on a Precoding MSE Metric," in *Proc. 42nd Asilomar Conference on Signals, Systems, and Computers*, October 2008, pp. 1305–1309.

[9] P. Stoica, Y. Jiang, and J. Li, "On MIMO Channel Capacity: An Intuitive Discussion," *IEEE Signal Processing Magazine*, vol. 22, no. 3, pp. 83–84, May 2005.

[10] A. Vosoughi and A. Scaglione, "On the Effect of Receiver Estimation Error Upon Channel Mutual Information," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 459–472, February 2006.

[11] G. J. Foschini and J. Salz, "Digital Communications Over Fading Radio Channels," *The Bell System Technical Journal*, vol. 62, no. 2, pp. 429–456, February 1983.

[12] J. H. Winters and J. Salz, "Upper Bounds on the Bit-Error Rate of Optimum Combining in Wireless Systems," *IEEE Transactions on Communications*, vol. 46, no. 12, pp. 1619–1624, December 1998.

[13] R. Hunger, M. Joham, and W. Utschick, "Extension of linear and nonlinear transmit filters for decentralized receivers," in *European Wireless 2005*, April 2005, pp. 40–46, vol. 1.

[14] M. Joham, P. M. Castro, L. Castedo, and W. Utschick, "MMSE Optimal Feedback of Correlated CSI for Multi-User Precoding," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008)*, April 2008, pp. 3129–3132.

[15] ——, "Robust Precoding with Bayesian Error Modeling for Limited Feedback MU-MISO Systems," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4954–4960, September 2010.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing – Estimation Theory*. Prentice Hall, 1993.

[17] L. L. Scharf, *Statistical Signal Processing – Detection, Estimation, and Time Series Analysis*. Addison-Wesley, 1991.

[18] P. Dent, G. E. Bottomley, and T. Croft, "Jakes Fading Model Revisited," *Electronics Letters*, vol. 29, no. 13, pp. 1162–1163, June 1993.

[19] P. M. Castro, M. Joham, W. Utschick, and L. Castedo, "Optimized CSI Feedback for Robust THP Design," in *Proc. 41st Asilomar Conference on Signals, Systems and Computers*, October 2007, pp. 1956–1960.

[20] S. P. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, vol. IT-28, no. 2, pp. 129–137, March 1982.

[21] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Kluwer, 1993.

[22] 3rd Generation Partnership Project, Technical Specification Group Radio Access Network, "Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations (Release 6)," September 2003, 3GPP TR 25.996, v6.1.0.

[23] N. Jindal, "MIMO Broadcast Channels With Finite-Rate Feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, November 2006.

Michael Joham was born in Kufstein, Austria, 1974. He received the Dipl.-Ing. and Dr.-Ing. degrees (both summa cum laude) in electrical engineering from the Technische Universität München (TUM), Germany, in 1999 and 2004, respectively.

Dr. Joham was with the Institute for Circuit Theory and Signal Processing at the TUM from 1999 to 2004. Since 2004, he has been with the Associate Institute for Signal Processing at the TUM, where he is currently a senior researcher. In the summers of 1998 and 2000, he visited the Purdue University, IN. In spring 2007, he was a guest lecturer at the University of A Coruña, Spain. In spring 2008, he was a guest lecturer at the University of the German Federal Armed Forces in Munich, Germany, and a guest professor at the University of A Coruña, Spain. In Winter 2009, he was a guest lecturer at the University of Hanover, Germany. In Fall 2010 and 2011, he was a guest lecturer at the Singapore Institute of Technology. His current research interests are precoding in mobile and satellite communications, limited rate feedback, MIMO communications, and robust signal processing.

Dr. Joham received the VDE Preis for his diploma thesis in 1999 and the Texas-Instruments-Preis for his dissertation in 2004. In 2007, he was a co-recipient of the best paper award at the International ITG/IEEE Workshop on Smart Antennas in Vienna. In 2011, he received the ITG Award 2011 of the German Information Technology Society (ITG).

Paula M. Castro was born in Orense, Spain, 1975. She received her MS and PhD degrees in Electrical Engineering (2001) from the University of Vigo, Spain, and in Computer Engineering (2009) from the University of A Coruña, Spain, respectively. She got a Master degree on “Managing Abilities and Working Techniques” at the Complutense University of Madrid in 2000. She was with some private companies until 2001. Since 2002 she is with the Group of Electronic Technology and Communications (GTEC) in the Department of Electronics and Systems at the University of A Coruña. From 2002 to 2006 she had the benefit of a FPI scholarship granted by the Spanish Government. She became Assistant Professor in 2006 and has been an Associate Professor since July 2011.

In 2006, she was with the Associate Institute for Signal Processing at the Technische Universität München (TUM), Germany, during four months. She has published five papers in high-impact journals, more than twenty-five papers in other journals and international conferences and seven papers in national conferences. She has also participated as a research member in more than thirty research projects and contracts with the regional and national governments. She was a co-recipient of the Best Paper Award at the International ITG/IEEE Workshop on Smart Antennas, Vienna, 2007. She has also supervised two Master Thesis. Her research is currently devoted to the design of limited feedback together with robust precoding techniques in multiuser wireless communications systems when the transmitter has no full knowledge about the channel.

Wolfgang Utschick (SM’06) was born on May 6, 1964 in Germany. He completed several industrial education programs before he received the high-school degree as well as the diploma and doctoral degrees, both with honors, in electrical engineering from the Technische Universität München (TUM) in 1993 and 1998. In this period he held a scholarship of the Bavarian Ministry of Education for exceptional students. From 1998 to 2002, he codirected the Signal Processing Group of the Institute of Circuit Theory and Signal Processing, TUM. Since 2000, he has been consulting in 3GPP standardization in the field of multi-element antenna systems. In 2002, he has been appointed Professor at the TUM where he is Head of the Fachgebiet Methoden der Signalverarbeitung. Dr. Utschick was awarded in 2006 for his excellent teaching records at TUM, and in 2007 received the ITG Award of the German Society for Information Technology (ITG). He is a senior member of the German VDE/ITG and appointed member in the ITG Expert Committee for Information and System Theory in 2009. He is currently also serving as a Chairman of the national DFG Focus Program ’Communications in Interference Limited Networks’ (COIN). Wolfgang Utschick is an Associate Editor for the IEEE Transactions on Signal Processing and an editor and founder of the Springer Book Series ’Foundations in Signal Processing, Communications and Networking’. Since 2010 he is appointed member in the Technical Committee of Signal Processing for Communications and Networking in the IEEE Signal Processing Society.

Luis Castedo was born in Santiago de Compostela, Spain, in 1966. He received the Ingeniero de Telecomunicación and Doctor Ingeniero de Telecomunicación degrees, both from Universidad Politécnica de Madrid (UPM), Spain, in 1990 and 1993, respectively. Between November 1994 and October 2001, he has been an Associate professor at the Departamento de Electrónica y Sistemas at Universidad de A Coruña, Spain, where he is currently Full Professor. He has been chairman of the Department between 2003 and 2009.

Luis Castedo is coauthor of more than one hundred and fifty papers in peer-reviewed international journals and conferences. He has also been principal researcher in more than thirty research projects funded by public organisms and private companies. His research interests are signal processing and digital communications with special emphasis on blind adaptive filtering, estimation/equalization of MIMO channels, space-time coding and prototyping of digital communication equipments.