

Transceiver Design Framework for Multiuser MIMO-OFDM Broadcast Systems with Channel Gram Matrix Feedback

Daniel Sacristán-Murga, Miquel Payaró, and Antonio Pascual-Iserte

Abstract

This work considers a multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing based multiuser broadcast system with precoding at the transmitter and feedback of channel state information. A general framework is presented for the transceiver design, and also for the design of the feedback link based on the quantization of the users' MIMO channel Gram matrices. The proposed design of the feedback link exploits the correlation of the channel response in the frequency domain due to the finite length of the channel time impulse responses to outperform other schemes based on feedback of the per carrier frequency responses. The transceiver design framework is based on a unitary linear transformation applied at the receivers which allows the computation of equivalent triangular channel response matrices at the transmitter. An analytic study of the error propagation due to the channel quantization in the feedback link and the computation of the equivalent triangular channel matrices is also performed. Based on the previous concepts, all the usual transceiver design criteria can be applied within this framework, and the particular case of a space-frequency precoder for robust mean square error minimization is derived as an example. Finally, the benefits of the proposed strategy are evaluated by means of numerical simulations and compared to other existing techniques.

Index Terms

MIMO systems, OFDM, matrix decomposition, robust designs, feedback communication, broadcast channel, multiuser communications, minimum mean square error.

D. Sacristán-Murga and M. Payaró are with the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), 08860 Castelldefels, Barcelona, Spain (email: {daniel.sacristan, miquel.payaro}@cttc.es).

A. Pascual-Iserte is with the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC), 08034 Barcelona, Spain, and with the CTTC (email: antonio.pascual@upc.edu).

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 248267 (BuNGee). It has received further support from the Catalan Government under grants 2009 SGR 891 and 2010 VALOR 198, by the Spanish Government under project TEC2011-29006-C03 (GRE3N), and by the European Cooperation in Science and Technology under project COST Action IC0902.

A limited part of this work was presented at the IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC), San Francisco, USA, June 2011.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is an effective and extensively implemented strategy that converts a frequency selective channel into a set of parallel flat fading channels. This is done by dividing the available channel bandwidth into F subchannels. When the subchannel bandwidth is sufficiently narrow, the frequency response across each subchannel can be considered approximately flat, which avoids the need of using complex equalization procedures [1].

Also, in the last years, multiple-input multiple-output (MIMO) communication systems have been in the focus of research, due to the large benefits in terms of throughput and resilience to noise and interference that they can provide [2], [3]. One of the main advantages of multi-antenna systems is that they allow to implement schemes where multiple users access the channel simultaneously in time and overlapping in frequency. In the case of the multiuser broadcast channel (BC), the optimum transmission strategy is the non-linear processing technique called dirty paper coding (DPC) [4]. However, DPC is not implemented in practice due to its high computational complexity. Instead, much simpler linear transceiver designs have been shown to achieve almost the same capacity using much lower computational resources [5].

Following the OFDM principle, it is possible to transform a MIMO frequency selective channel into a collection of F parallel flat fading MIMO channels. In such a system, the maximum achievable diversity order is the product of the number of transmit antennas, the number of receive antennas, and the number of propagation paths represented by the channel impulse response length [6], [7]. In order to achieve this full diversity the information symbols should be allowed to be spread not only over the transmitting antennas, but also over the carriers. Note, however, that conventional linear space codes are designed to exploit the spatial diversity of flat fading MIMO channels, and do not take into account the frequency diversity of an OFDM scheme.

This work presents a framework for the transceiver design in multiuser MIMO-OFDM BC systems which, instead of the feedback of the MIMO channel response matrix \mathbf{H} for each user, considers the feedback of the channel Gram matrices (i.e., $\mathbf{H}^H\mathbf{H}$). This idea is based on the previous works [8], [9], where it was proved that for point-to-point single-user MIMO systems, the minimum amount of channel state information (CSI) required at the transmitter in order to perform the optimum linear precoding corresponds to such channel Gram matrix. This conclusion is extended in this paper for the multiuser scenario and also to robust designs through a unique decomposition of the channel Gram matrix of each user, which results in a triangular equivalent propagation channel response matrix. Additionally, a feedback based on the temporal channel impulse response is proposed, as opposed to the usual quantization and feedback based on the frequency response per carrier. This enables to exploit

the frequency correlation of the CSI in order to further improve the efficiency of the quantization and feedback. The propagation of the CSI quantization error through the computation of the equivalent channel is also studied analytically and this result is later used, as an example, in the design of a robust precoding scheme.

The proposed framework is valid for any design criterion and we considered, as an illustrative example, the robust minimization of the sum of the mean square errors (MSE) of all the symbol streams for all the users, with fixed decoders. This design maps information symbols to antennas and carriers in order to exploit both spatial and frequency diversity, and requires estimates of the multiuser channel responses at the transmitter. It is a robust design in the sense that it takes into account the errors in the quantization for the feedback transmission to optimize performance. Note that there is a wide range of designs based on MSE in the literature, such as: [10], [11] which also consider fixed decoders, or [12]–[14], which present iterative designs. Note that the algorithm presented in this work could be applied in iterative designs such as the ones in [12], [13], and [14], at the step where the transmitter is computed at each iteration. Other works such as [15], [16] assume single-antenna receivers, in which case the decoder design is not an issue.

Summarizing, the main contributions of this paper can be listed as follows:

- 1) An extension of the feedback based on Gram matrices to the case of broadcast multiuser MIMO.
- 2) The computation of the equivalent triangular propagation channels from the channel Gram matrices.
- 3) The analysis of the propagation of the quantization errors through the computation of the equivalent MIMO triangular channels.
- 4) An example of robust design is implemented and evaluated numerically.

The first three topics are what constitute the general framework. On top of it any design criteria or architecture can be implemented (even non-linear designs). The fourth is an example of application of the proposed framework to a particular scenario and design criterion.

The remainder of this paper is organized as follows. The system and signal models are described in section II. Section III describes the proposed per user feedback scheme and presents the linear transformation applied to uniquely obtain the equivalent triangular channels. The model considered for the CSI quantization error and its propagation through the processing at the transmitter are presented in section IV, while section V presents an example of application consisting in a robust precoder that takes into account the errors in the CSI. Finally, section VI provides numerical simulations to evaluate the performance of the proposed strategies in a MIMO-OFDM BC system. Section VII concludes the paper.

II. SYSTEM AND SIGNAL MODELS

We consider a multiuser MIMO-OFDM BC system with F carriers and K users, denoted by the indices $f = 0, \dots, F - 1$ and $k = 1, \dots, K$, respectively. The transmitter features n_T antennas and

the k th receiver has $n_R^{(k)}$ antennas. The propagation channel of user k is characterized by its temporal impulse response, which consists of a maximum of L taps¹ and is denoted by $\bar{\mathbf{H}}_l^{(k)} \in \mathbb{C}^{n_R^{(k)} \times n_T}$, $l = 0, \dots, L - 1$ (the horizontal overline is used to denote that the variable is defined in the time domain). Accordingly, the frequency channel response at carrier f and user k is given by:

$$\mathbf{H}_f^{(k)} = \sum_{l=0}^{L-1} \bar{\mathbf{H}}_l^{(k)} e^{-j\frac{2\pi}{F}fl} \in \mathbb{C}^{n_R^{(k)} \times n_T}. \quad (1)$$

Classically, parallel linear precoding per carrier at the transmitter is considered, which is denoted by a precoding matrix $\mathbf{P}_f^{(k)} \in \mathbb{C}^{n_T \times n_{S_f}^{(k)}}$, for user k and carrier f , where $n_{S_f}^{(k)}$ is the number of streams transmitted to the k th receiver through the f th carrier. The corresponding linear processing at the k th receiver for carrier f is represented by the decoding matrix $\mathbf{D}_f^{(k)} \in \mathbb{C}^{n_{S_f}^{(k)} \times n_R^{(k)}}$. Following this model, the estimated symbols $\hat{\mathbf{x}}_f^{(k)} \in \mathbb{C}^{n_{S_f}^{(k)}}$ corresponding to the f th carrier at the k th receiver are given by:

$$\hat{\mathbf{x}}_f^{(k)} = \mathbf{D}_f^{(k)} \mathbf{H}_f^{(k)} \sum_{i=1}^K \mathbf{P}_f^{(i)} \mathbf{x}_f^{(i)} + \mathbf{D}_f^{(k)} \mathbf{w}_f^{(k)} \in \mathbb{C}^{n_{S_f}^{(k)}}, \quad \forall f, k, \quad (2)$$

where $\mathbf{x}_f^{(k)} \in \mathbb{C}^{n_{S_f}^{(k)}}$ is the vector containing the streams of symbols transmitted to user k through carrier f and $\mathbf{w}_f^{(k)} \in \mathbb{C}^{n_R^{(k)}}$ is the additive white Gaussian noise (AWGN) at receiver k .

Using a notation with block diagonal matrices to group the symbols transmitted through all carriers corresponding to each receiver, the estimated symbols at the k th receiver are given by:

$$\hat{\mathbf{x}}^{(k)} = \mathbf{D}^{(k)} \mathbf{H}^{(k)} \sum_{i=1}^K \mathbf{P}^{(i)} \mathbf{x}^{(i)} + \mathbf{D}^{(k)} \mathbf{w}^{(k)} \in \mathbb{C}^{\sum_{f=0}^{F-1} n_{S_f}^{(k)}}, \quad \forall k, \quad (3)$$

where $\mathbf{D}^{(k)} = \text{blockdiag} \left(\mathbf{D}_0^{(k)}, \dots, \mathbf{D}_{F-1}^{(k)} \right) \in \mathbb{C}^{\sum_{f=0}^{F-1} n_{S_f}^{(k)} \times F n_R^{(k)}}$, $\hat{\mathbf{x}}^{(k)} = \left[\hat{\mathbf{x}}_0^{(k)T}, \dots, \hat{\mathbf{x}}_{F-1}^{(k)T} \right]^T \in \mathbb{C}^{\sum_{f=0}^{F-1} n_{S_f}^{(k)}}$, $\mathbf{x}^{(k)} = \left[\mathbf{x}_0^{(k)T}, \dots, \mathbf{x}_{F-1}^{(k)T} \right]^T \in \mathbb{C}^{\sum_{f=0}^{F-1} n_{S_f}^{(k)}}$, $\mathbf{H}^{(k)} = \text{blockdiag} \left(\mathbf{H}_0^{(k)}, \dots, \mathbf{H}_{F-1}^{(k)} \right) \in \mathbb{C}^{F n_R^{(k)} \times F n_T}$, $\mathbf{w}^{(k)} = \left[\mathbf{w}_0^{(k)T}, \dots, \mathbf{w}_{F-1}^{(k)T} \right]^T \in \mathbb{C}^{F n_R^{(k)}}$, and $\mathbf{P}^{(k)} = \text{blockdiag} \left(\mathbf{P}_0^{(k)}, \dots, \mathbf{P}_{F-1}^{(k)} \right) \in \mathbb{C}^{F n_T \times \sum_{f=0}^{F-1} n_{S_f}^{(k)}}$ (see Fig. 1 for a complete diagram of the BC system).

Note that in expressions (2) and (3) each symbol is constrained to be transmitted over one single carrier as shown by the fact that the precoding and decoding matrices, $\mathbf{P}^{(k)}$ and $\mathbf{D}^{(k)}$, are block-diagonal. It is possible to achieve higher diversity by not forcing a block-diagonal structure at the precoding $\mathbf{P}^{(k)}$ and decoding $\mathbf{D}^{(k)}$ stages. According to this, in this work a space-frequency precoder is designed as a mean to extract both spatial and frequency diversity and, consequently, the precoding and decoding matrices are not constrained to be block-diagonal. Since the symbols are now spread among the carriers, it does not make sense to use the notation $n_{S_f}^{(k)}$ corresponding to the number of symbols per carrier. Instead we consider a total number of streams $n_S^{(k)}$ transmitted to receiver k through all carriers, a global precoder $\mathbf{P}^{(k)} \in \mathbb{C}^{F n_T \times n_S^{(k)}}$ for receiver k , and a global decoder $\mathbf{D}^{(k)} \in \mathbb{C}^{n_S^{(k)} \times F n_R^{(k)}}$

¹For the case where the channel impulse responses of the different users have different number of taps, L is defined as the maximum among the number of taps for all users.

for receiver k . Note that equation (3) is still correct in this setup (with $\mathbf{P}^{(k)}$ and $\mathbf{D}^{(k)}$ no longer being forced to be block diagonal), by simply substituting $\sum_{f=0}^{F-1} n_{S_f}^{(k)}$ by $n_S^{(k)}$. Observe that, as will be shown in section V-C, for some particular cases the precoder and decoder matrices do turn out block-diagonal as a result of the optimization and without being imposed from the beginning.

For the sake of compactness in the notation, it is convenient to further group the symbols estimated at all receivers in a single vector $\hat{\mathbf{x}}$, which can be expressed as:

$$\hat{\mathbf{x}} = \mathbf{DHP}\mathbf{x} + \mathbf{D}\mathbf{w} \in \mathbb{C}^{\sum_{k=1}^K n_S^{(k)}}, \quad (4)$$

where $\mathbf{x} = [\mathbf{x}^{(1)T}, \dots, \mathbf{x}^{(K)T}]^T \in \mathbb{C}^{\sum_{k=1}^K n_S^{(k)}}$, $\hat{\mathbf{x}} = [\hat{\mathbf{x}}^{(1)T}, \dots, \hat{\mathbf{x}}^{(K)T}]^T \in \mathbb{C}^{\sum_{k=1}^K n_S^{(k)}}$, $\mathbf{H} = [\mathbf{H}^{(1)T}, \dots, \mathbf{H}^{(K)T}]^T \in \mathbb{C}^{F \sum_{k=1}^K n_R^{(k)} \times \sum_{k=1}^K n_S^{(k)}}$, $\mathbf{P} = [\mathbf{P}^{(1)}, \dots, \mathbf{P}^{(K)}] \in \mathbb{C}^{F n_T \times \sum_{k=1}^K n_S^{(k)}}$, $\mathbf{w} = [\mathbf{w}^{(1)T}, \dots, \mathbf{w}^{(K)T}]^T \in \mathbb{C}^{F \sum_{k=1}^K n_R^{(k)}}$, and $\mathbf{D} = \text{blockdiag}(\mathbf{D}^{(1)}, \dots, \mathbf{D}^{(K)}) \in \mathbb{C}^{\sum_{k=1}^K n_S^{(k)} \times F \sum_{k=1}^K n_R^{(k)}}$. Here we define $\mathbf{R}_x^{(k)} = \mathbb{E}\{\mathbf{x}^{(k)}\mathbf{x}^{(k)H}\}$ and $\mathbf{R}_w^{(k)} = \mathbb{E}\{\mathbf{w}^{(k)}\mathbf{w}^{(k)H}\}$; therefore, $\mathbf{R}_x = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \text{blockdiag}(\mathbf{R}_x^{(1)}, \dots, \mathbf{R}_x^{(K)})$ and $\mathbf{R}_w = \mathbb{E}\{\mathbf{w}\mathbf{w}^H\} = \text{blockdiag}(\mathbf{R}_w^{(1)}, \dots, \mathbf{R}_w^{(K)})$.

III. FEEDBACK AND EQUIVALENT CHANNELS

As will be seen in section V, in general, in order to optimally design the precoding matrix \mathbf{P} at the transmitter under a generic optimization criterion, the CSI needed from each user corresponds to the channel propagation matrices for all F carriers. That is, knowledge of $\mathbf{H}_f^{(k)} \in \mathbb{C}^{n_R^{(k)} \times n_T}$, $\forall f, k$ (or, equivalently, $\mathbf{H}^{(k)}$ as defined in (3)) is used to build the optimum precoder. In [8], [17], however, the authors proved that, for the single user point-to-point MIMO scenario, the optimum linear precoding design can be calculated with just the channel Gram matrix (which is defined as $\mathbf{R}_f^{(k)} = \mathbf{H}_f^{(k)H} \mathbf{H}_f^{(k)}$), for all the usual criteria. This fact was applied in [9], [18], [19] to design more efficient feedback algorithms that exploit the differential geometry of the set of all possible channel Gram matrices, which are Hermitian and positive semidefinite by construction, and that perform better than the feedback of the channel matrix $\mathbf{H}_f^{(k)}$. The motivation for using Gram matrix feedback is that it contains less information than the channel matrix feedback, but this information is sufficient for the precoder designs. Since less information has to be quantized and sent through the feedback link, the feedback can be performed more efficiently and the system performance is better, as shown in the related works [9], [17] among others. In some specific multiuser systems, such as in the BC with block diagonalization (BD) [20], the transmitter design also depends only on the channel Gram matrix of each user and, therefore, the efficient techniques for quantization and feedback presented in [9], [18], [19] could also be applied, as shown in [21]. Note, however, that in the general multiuser scenario (i.e., without constraining a BD transmission and/or for a general quality criterion), as well as in robust precoder designs, complete knowledge of each $\mathbf{H}_f^{(k)}$ has been assumed so far by the research community. As a contribution of this paper, in section III-A, we present a linear transformation technique that still enables the use of channel Gram matrix feedback even in the general multiuser scenario by adding a unitary pre-transformation at the decoder to identify uniquely an equivalent triangular MIMO channel

for each receiver. That is, knowledge of $\mathbf{R}_f^{(k)} \equiv \mathbf{H}_f^{(k)H} \mathbf{H}_f^{(k)} \in \mathbb{C}^{n_T \times n_T}$ of each carrier and user is sufficient to design the optimum precoder.

In this work we also analyze the possibility of performing feedback of the temporal CSI instead of the frequency CSI in a per carrier basis, as is usually done in OFDM systems. Following from (1), the F Gram matrices $\mathbf{R}_f^{(k)}$ of size $n_T \times n_T$ for each user can also be computed as:

$$\mathbf{R}_f^{(k)} \equiv \mathbf{H}_f^{(k)H} \mathbf{H}_f^{(k)} = \sum_{n=0}^{L-1} \sum_{m=0}^{L-1} \bar{\mathbf{H}}_n^{(k)H} \bar{\mathbf{H}}_m^{(k)} e^{-j\frac{2\pi}{F}f(m-n)} \in \mathbb{C}^{n_T \times n_T}, \quad f = 0, \dots, F-1. \quad (5)$$

Using this, the necessary CSI corresponding to the k th user at the transmitter can be computed with knowledge of the L matrices $\bar{\mathbf{H}}_l^{(k)} \in \mathbb{C}^{n_R^{(k)} \times n_T}$ or, alternatively, using one temporal Gram matrix $\bar{\mathbf{R}}^{(k)} \in \mathbb{C}^{Ln_T \times Ln_T}$ defined as (note that the sub-blocks of the following matrix are used directly within the sum in (5)):

$$\begin{aligned} \bar{\mathbf{R}}^{(k)} &\equiv \begin{bmatrix} \bar{\mathbf{H}}_0^{(k)H} \\ \bar{\mathbf{H}}_1^{(k)H} \\ \vdots \\ \bar{\mathbf{H}}_{L-1}^{(k)H} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{H}}_0^{(k)} & \bar{\mathbf{H}}_1^{(k)} & \dots & \bar{\mathbf{H}}_{L-1}^{(k)} \end{bmatrix} \\ &= \begin{bmatrix} \bar{\mathbf{H}}_0^{(k)H} \bar{\mathbf{H}}_0^{(k)} & \bar{\mathbf{H}}_0^{(k)H} \bar{\mathbf{H}}_1^{(k)} & \dots & \bar{\mathbf{H}}_0^{(k)H} \bar{\mathbf{H}}_{L-1}^{(k)} \\ \bar{\mathbf{H}}_1^{(k)H} \bar{\mathbf{H}}_0^{(k)} & \bar{\mathbf{H}}_1^{(k)H} \bar{\mathbf{H}}_1^{(k)} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathbf{H}}_{L-1}^{(k)H} \bar{\mathbf{H}}_0^{(k)} & \bar{\mathbf{H}}_{L-1}^{(k)H} \bar{\mathbf{H}}_1^{(k)} & \dots & \bar{\mathbf{H}}_{L-1}^{(k)H} \bar{\mathbf{H}}_{L-1}^{(k)} \end{bmatrix} \in \mathbb{C}^{Ln_T \times Ln_T}. \end{aligned} \quad (6)$$

$$(7)$$

Observe that matrices $\mathbf{R}_f^{(k)}$ and $\bar{\mathbf{R}}^{(k)}$ are positive semidefinite and Hermitian by construction. Since, as will be seen in the following subsections, the precoder design depends on the propagation channel of each user through $\mathbf{R}_f^{(k)}$, the most straightforward approach would be to quantize and feed back $\mathbf{R}_f^{(k)}$ individually for each carrier and user. However, this is suboptimal because: a) it does not exploit the correlation in frequency of the propagation channel, and b) in systems with many carriers the feedback overhead would be too large. Therefore, in order to improve the performance of the CSI quantization, the scheme proposed in this paper considers the possibility of feeding back the temporal channel Gram matrix $\bar{\mathbf{R}}^{(k)}$ of each user. This allows to exploit the correlation in frequency of the channel and also the fact that the size of the matrix to be quantized grows with the number of channel impulse response taps L instead of the number of carriers F . This can help to greatly improve the performance of the CSI quantization in some situations, as will be shown in section VI.

From the knowledge of the temporal channel Gram matrix $\bar{\mathbf{R}}^{(k)}$ (and therefore, the individual Gram matrices $\mathbf{R}_f^{(k)}$ associated to each carrier), it is possible to compute (as is described in the next section), for each user and carrier, a unique equivalent channel response triangular matrix which can be used to apply any type of multiuser design on top of it (without the restrictions of BD, which spends degrees

of freedom to completely avoid inter-user interference), for any quality criterion in the same way as if knowledge of the actual channel response matrix $\mathbf{H}_f^{(k)}$ was available.

A. Equivalent channel transformation

Following the feedback model presented in the previous section, knowledge of an estimation of the channel Gram matrix of each user is assumed at the transmitter. Note that there are multiple possible channel matrices $\mathbf{H}^{(k)}$ that generate the same Gram matrix $\mathbf{R}^{(k)} \triangleq \mathbf{H}^{(k)H} \mathbf{H}^{(k)}$ (for example $\mathbf{H}^{(k)}$ and $\mathbf{U}\mathbf{H}^{(k)}$, with \mathbf{U} being a unitary matrix, generate the same Gram matrix). Since the transmitter has knowledge only of the Gram matrix, it cannot know which of the multiple channel matrices that generate that Gram matrix is the actual channel matrix. However, there is only one of the possible matrices that can generate the Gram matrix that is upper triangular and with real positive elements in the diagonal, and this matrix can be computed at the transmitter, as will be shown in this subsection. On the other hand, note also that, by applying a properly calculated unitary linear transformation $\mathbf{Q}^{(k)H}$, with $\mathbf{Q}^{(k)} \in \mathbb{C}^{F n_R^{(k)} \times F n_R^{(k)}}$ at each receiver k , it is possible to generate an equivalent channel response matrix $\mathbf{T}^{(k)} \in \mathbb{C}^{F n_R^{(k)} \times F n_T}$ from the transmitter to the output of such unitary transformation at each receiver (as shown in Fig. 2), which is upper triangular and with real elements in the diagonal. Since there is only one possible upper triangular matrix with real elements in the diagonal associated to the Gram matrix, and the unitary transformation does not change the Gram matrix, the equivalent channel matrix computed at the transmitter and the equivalent channel generated by the application of such unitary transformation $\mathbf{Q}^{(k)}$ at the receiver are the same. This scheme is depicted in Fig. 2 (note that in Fig. 2, $\tilde{\mathbf{D}}^{(k)}$ is defined as $\tilde{\mathbf{D}}^{(k)} = \mathbf{D}^{(k)} \mathbf{Q}^{(k)} \in \mathbb{C}^{n_S^{(k)} \times F n_R^{(k)}}$, so that $\mathbf{D}^{(k)} = \tilde{\mathbf{D}}^{(k)} \mathbf{Q}^{(k)H}$).

Finally, it is important to note that this matrix $\mathbf{Q}^{(k)}$ does not introduce a penalty in the performance of the system since it is a unitary linear transformation that could be reversed at the receiver by a proper processing after such unitary transformation, and it allows any arbitrary multiuser transmission design to be applied on top of the basis of these new equivalent channel matrices $\{\mathbf{T}^{(k)}\}_{k=1}^K$, where now the transmitter is required to know only the channel Gram matrices $\mathbf{R}_f^{(k)}$ or $\bar{\mathbf{R}}^{(k)}$.

According to the previous definitions, the equivalent triangular channel, $\mathbf{T}^{(k)}$, is such that satisfies:

$$\mathbf{H}^{(k)} = \mathbf{Q}^{(k)} \mathbf{T}^{(k)}, \quad \forall k = 1, \dots, K. \quad (8)$$

The computation of the equivalent channel response matrix $\mathbf{T}^{(k)}$ at the transmitter is described in subsection III-A1, as a function of the matrix $\mathbf{R}^{(k)}$. At the receiver, matrices $\mathbf{Q}^{(k)}$ and $\mathbf{T}^{(k)}$ are computed from $\mathbf{H}^{(k)}$ as shown in subsection III-A2.

Observe that, since $\mathbf{H}^{(k)}$ is a block diagonal matrix (we recall that $\mathbf{H}^{(k)} = \text{blockdiag}(\mathbf{H}_0^{(k)}, \dots, \mathbf{H}_{F-1}^{(k)})$ as defined in section II), the resulting $\mathbf{Q}^{(k)}$ and $\mathbf{T}^{(k)}$ are also block diagonal. In fact, the computation can be performed in parallel for each block of $\mathbf{H}^{(k)}$, which corresponds to the channel response matrix for each carrier, $\mathbf{H}_f^{(k)} \forall f = 0, \dots, F-1$. This greatly reduces the required computational complexity, and, for this reason, the transformation will be presented next for each individual matrix $\mathbf{H}_f^{(k)}, \mathbf{R}_f^{(k)}$.

The channel response matrix for the f th carrier and the k th receiver $\mathbf{H}_f^{(k)}$ can be written as:

$$\mathbf{H}_f^{(k)} = \mathbf{Q}_f^{(k)} \mathbf{T}_f^{(k)}, \quad \forall f = 0, \dots, F-1, \quad (9)$$

where $\mathbf{Q}_f^{(k)} \in \mathbb{C}^{n_R^{(k)} \times n_R^{(k)}}$ is unitary and $\mathbf{T}_f^{(k)} \in \mathbb{C}^{n_R^{(k)} \times n_T}$ is upper triangular [22] (consequently, the channel Gram matrix $\mathbf{R}_f^{(k)}$ can also be written as $\mathbf{R}_f^{(k)} = \mathbf{H}_f^{(k)H} \mathbf{H}_f^{(k)} = \mathbf{T}_f^{(k)H} \mathbf{T}_f^{(k)}$). The decomposition that allows to calculate $\mathbf{T}_f^{(k)}$ from $\mathbf{R}_f^{(k)}$ is unique if we force $\mathbf{T}_f^{(k)}$ to be upper triangular and with the elements $t_{i,i}$ on the main diagonal being positive and real, as will be shown next.

1) *Calculation of $\mathbf{T}_f^{(k)}$ at the transmitter:* Matrix $\mathbf{R}_f^{(k)}$ is obtained at the transmitter from the $\bar{\mathbf{R}}^{(k)}$ received through the feedback link using (5) and (7). Using $\mathbf{R}_f^{(k)}$, the transmitter can uniquely compute $\mathbf{T}_f^{(k)}$ as described next. Observe that there is only one possible $\mathbf{T}_f^{(k)} \in \mathbb{C}^{n_R^{(k)} \times n_T}$ satisfying that the elements $t_{i,i}$ are real and positive and $t_{i,j} = 0, \forall i > j$.

From $\mathbf{R}_f^{(k)} = \mathbf{T}_f^{(k)H} \mathbf{T}_f^{(k)}$, matrix $\mathbf{T}_f^{(k)}$ is computed as ($j \geq i$):²

$$t_{i,j} = \begin{cases} \sqrt{r_{1,1}}, & i = j = 1, \\ \frac{r_{1,j}}{t_{1,1}}, & i = 1, \forall j > 1, \\ \frac{r_{i,j} - \sum_{k=1}^{i-1} t_{k,i}^* t_{k,j}}{t_{i,i}}, & \forall i, j; 1 < i < j, \\ \sqrt{r_{i,i} - \sum_{k=1}^{i-1} |t_{k,i}|^2}, & \forall i, j; i = j > 1, \end{cases} \quad (10)$$

where $r_{i,j}$ and $t_{i,j}$ are the elements i, j of $\mathbf{R}_f^{(k)}$ and $\mathbf{T}_f^{(k)}$, respectively, and where, for the sake of clarity in the notation, we have dropped the dependence of $r_{i,j}$ and $t_{i,j}$ on f .

2) *Calculation of $\mathbf{T}_f^{(k)}$ and $\mathbf{Q}_f^{(k)}$ at the receiver:* The receiver knows $\mathbf{H}_f^{(k)}$ and can obtain $\mathbf{T}_f^{(k)} = \mathbf{Q}_f^{(k)H} \mathbf{H}_f^{(k)}$, as is described in the following algorithm based on the QR decomposition. From (9) we have that matrices $\mathbf{T}_f^{(k)}$ and $\mathbf{Q}_f^{(k)}$ can be computed as³:

$$\mathbf{q}_i = \begin{cases} \frac{\mathbf{h}_i}{\|\mathbf{h}_i\|}, & i = 1, \\ \frac{(\mathbf{I} - \sum_{k=1}^{i-1} \mathbf{q}_k \mathbf{q}_k^H) \mathbf{h}_i}{\|(\mathbf{I} - \sum_{k=1}^{i-1} \mathbf{q}_k \mathbf{q}_k^H) \mathbf{h}_i\|}, & \forall i; 1 < i \leq n_T, \end{cases} \quad (11)$$

$$t_{i,j} = \begin{cases} \mathbf{q}_i^H \mathbf{h}_j, & \forall i < j, \\ \|\mathbf{h}_1\|, & i = j = 1, \\ \left\| \left(\mathbf{I} - \sum_{k=1}^{i-1} \mathbf{q}_k \mathbf{q}_k^H \right) \mathbf{h}_i \right\|, & \forall i, j; i = j > 1, \end{cases} \quad (12)$$

where \mathbf{q}_i and \mathbf{h}_i correspond to the i th column of matrices $\mathbf{Q}_f^{(k)}$ and $\mathbf{H}_f^{(k)}$, respectively, and, again, the dependence on f is omitted for clarity in the notation.

In the case where $n_R^{(k)} > n_T$, the last $n_R^{(k)} - n_T$ columns of $\mathbf{Q}_f^{(k)}$ are chosen such that $\mathbf{Q}_f^{(k)H} \mathbf{Q}_f^{(k)} = \mathbf{I}$, i.e., they just have to be orthogonal with each other and with the previous columns and have a norm

²Note that the decomposition presented in (10) is not exactly the Cholesky factorization because the resulting matrix $\mathbf{T}_f^{(k)} \in \mathbb{C}^{n_R^{(k)} \times n_T}$ is not forced to be square (the equivalent channel propagation matrix $\mathbf{T}_f^{(k)}$ must have the same dimensions as the actual channel propagation matrix $\mathbf{H}_f^{(k)}$).

³Note that, since at the receiver there is knowledge of $\mathbf{H}_f^{(k)}$, it would also be possible to calculate $\mathbf{R}_f^{(k)} = \mathbf{H}_f^{(k)H} \mathbf{H}_f^{(k)}$ and compute $\mathbf{T}_f^{(k)}$ locally as explained in the previous subsection describing the decomposition to be calculated at the transmitter.

equal to 1 and they can be calculated following the Gram-Schmidt procedure. In this case, we assume that the rank of the matrix $\mathbf{R}_f^{(k)}$ is given by n_T . It is important to note that the equivalent channel $\mathbf{T}_f^{(k)}$ is a tall matrix with the last $n_R^{(k)} - n_T$ rows equal to zero. This means that, at the receiver, after the application of the transformation represented by $\mathbf{Q}_f^{(k)H}$, the last $n_R^{(k)} - n_T$ outputs contain only noise, which is uncorrelated with the data. Observe that, in the case of spatially white noise, i.e., if $\mathbf{R}_{w_f}^{(k)} = \sigma_{w_f}^2 \mathbf{I}$, the last $n_R^{(k)} - n_T$ outputs contain no useful information. Consequently, for the computation of the estimates of the transmitted symbols at the receiver, the last $n_R^{(k)} - n_T$ columns of $\mathbf{Q}_f^{(k)}$ could be ignored, reducing the complexity of computing $\mathbf{Q}_f^{(k)}$. In the case where $n_R^{(k)} \leq n_T$ we assume that the rank of the matrix $\mathbf{R}_f^{(k)}$ is given by $n_R^{(k)}$ and the transformation is performed as already described in this subsection.

IV. ERROR ANALYSIS

In general, robust transceiver designs require the characterization of the CSI error in order to minimize its effect. In the presented framework, this characterization has been performed in terms of the second-order statistics of the resulting error at the equivalent triangular channel matrix, as will be shown in this section. This will be exploited in the robust design presented in section V as an illustrative example of application.

The error in the equivalent triangular channel is the result of the propagation of the initial error generated in the CSI quantization. In general, only the second-order statistics of the initial quantization error are known. This section presents the analytic study of the linear relation, for small errors, between the initial quantization error and the final error in the equivalent triangular channel matrix. From this derivation it will be possible to compute the second order statistics of the error in the equivalent triangular channel, which is often required in statistical robust designs.

A. Error model

There are three different sources of inaccuracies in the CSI sent through the feedback link from the receiver to the transmitter: estimation errors at the receiver, quantization errors which are inherent to the feedback process, and errors due to noise in the feedback link. In practical situations, where the part of the transmission dedicated to feedback is greatly constrained, the quantization error is the dominant factor of the error. Consequently, in this paper we will consider only the errors resulting from the quantization process. The CSI at the transmitter $\bar{\mathbf{R}}_{\text{quant}}^{(k)} \equiv \text{quantiz}(\bar{\mathbf{R}}^{(k)})$ is then modeled with the error matrix $\bar{\mathbf{R}}_{\text{err}}^{(k)}$:

$$\bar{\mathbf{R}}^{(k)} = \bar{\mathbf{R}}_{\text{quant}}^{(k)} + \bar{\mathbf{R}}_{\text{err}}^{(k)} \in \mathbb{C}^{Ln_T \times Ln_T}. \quad (13)$$

Note that, although the rank of matrix $\bar{\mathbf{R}}^{(k)} \in \mathbb{C}^{Ln_T \times Ln_T}$ before the quantization is $\min(n_R^{(k)}, Ln_T)$, depending on the quantization strategy that is applied and if $n_R^{(k)} < Ln_T$ it is possible that after the quantization the rank of the resulting matrix is increased up to Ln_T . If the rank is increased, matrix

$\bar{\mathbf{R}}_{\text{quant}}^{(k)}$ should be projected on the space of the matrices with a rank equal to $\min(n_R^{(k)}, Ln_T)$ at the transmitter in order to maintain the rank.⁴ The error after the projection, is then propagated at the transmitter through the Fourier transformation used to compute the channel Gram matrix associated to carrier f and user k based on the Fourier transform and through the computation of the equivalent triangular channel $\mathbf{T}_f^{(k)}$. The final error in the resulting knowledge of $\mathbf{T}_f^{(k)}$ is denoted in the following by $\mathbf{T}_{\text{err}_f}^{(k)}$:

$$\mathbf{T}_f^{(k)} = \mathbf{T}_{\text{quant}_f}^{(k)} + \mathbf{T}_{\text{err}_f}^{(k)}, \quad (14)$$

where $\mathbf{T}_{\text{quant}_f}^{(k)}$ is the estimated value of the actual equivalent triangular channel $\mathbf{T}_f^{(k)}$ calculated using the quantized and projected channel Gram matrix sent through the feedback link. The expression of $\mathbf{T}_{\text{err}_f}^{(k)}$ as a result of the error propagation is derived in subsection IV-B.

Following this notation, (4) can be rewritten reflecting the errors in the CSI at the transmitter and incorporating the notation corresponding to the equivalent channels as:

$$\hat{\mathbf{x}} = \tilde{\mathbf{D}}\mathbf{T}_{\text{quant}}\mathbf{P}\mathbf{x} + \tilde{\mathbf{D}}\mathbf{T}_{\text{err}}\mathbf{P}\mathbf{x} + \tilde{\mathbf{D}}\mathbf{Q}^H\mathbf{w} \in \mathbb{C}^{\sum_{k=1}^K n_S^{(k)}}, \quad (15)$$

where $\mathbf{T}_{\text{quant}}^{(k)} = \text{blockdiag}(\mathbf{T}_{\text{quant}_0}^{(k)}, \dots, \mathbf{T}_{\text{quant}_{F-1}}^{(k)}) \in \mathbb{C}^{Fn_R^{(k)} \times Fn_T}$, $\mathbf{T}_{\text{err}}^{(k)} = \text{blockdiag}(\mathbf{T}_{\text{err}_0}^{(k)}, \dots, \mathbf{T}_{\text{err}_{F-1}}^{(k)}) \in \mathbb{C}^{Fn_R^{(k)} \times Fn_T}$, $\tilde{\mathbf{D}} = \text{blockdiag}(\tilde{\mathbf{D}}^{(1)}, \dots, \tilde{\mathbf{D}}^{(K)}) \in \mathbb{C}^{\sum_{k=1}^K n_S^{(k)} \times F \sum_{k=1}^K n_R^{(k)}}$, $\mathbf{Q} = \text{blockdiag}(\mathbf{Q}^{(1)}, \dots, \mathbf{Q}^{(K)}) \in \mathbb{C}^{F \sum_{k=1}^K n_R^{(k)} \times F \sum_{k=1}^K n_R^{(k)}}$, $\mathbf{T}_{\text{quant}} = [\mathbf{T}_{\text{quant}}^{(1)T}, \dots, \mathbf{T}_{\text{quant}}^{(K)T}]^T \in \mathbb{C}^{F \sum_{k=1}^K n_R^{(k)} \times Fn_T}$, and $\mathbf{T}_{\text{err}} = [\mathbf{T}_{\text{err}}^{(1)T}, \dots, \mathbf{T}_{\text{err}}^{(K)T}]^T \in \mathbb{C}^{F \sum_{k=1}^K n_R^{(k)} \times Fn_T}$.

B. Error propagation

As shown in the previous subsection, the CSI inaccuracies are defined in the matrix $\bar{\mathbf{R}}^{(k)}$, which is quantized and then sent through the feedback link. At the transmitter, the received matrix $\bar{\mathbf{R}}_{\text{quant}}^{(k)}$ is projected, if needed, to guarantee that its rank is $\min(n_R^{(k)}, Ln_T)$ and an estimation of the channel Gram matrix $\mathbf{R}_{\text{quant}_f}^{(k)}$ is computed using the projected matrix $\bar{\mathbf{R}}_{\text{quant},p}^{(k)}$, following (5). In a third step, the estimated equivalent channel $\mathbf{T}_{\text{quant}}^{(k)}$ is computed from $\mathbf{R}_{\text{quant}_f}^{(k)}$ following (10).

First, the propagation of the error through the semidefinite projection is presented. Next, the transformation of the resulting error through the computation of $\mathbf{R}_f^{(k)}$ at the transmitter is studied. Finally, the result is further propagated through the equivalent channel computation, in order to obtain an expression of $\mathbf{T}_{\text{err}}^{(k)}$. This result will then be used for the design of the robust precoder in section V.

1) *Propagation through the semidefinite projection:* As it has been pointed out above, the first step where the quantization error propagates through is the positive semidefinite projection. Note that if $n_R^{(k)} \geq Ln_T$ this step is not necessary, since $\bar{\mathbf{R}}_{\text{quant}}^{(k)}$ already maintains the same rank as $\bar{\mathbf{R}}^{(k)}$, so, in

⁴In the case where $n_R^{(k)} \geq Ln_T$ the quantization process maintains the rank of the matrix, but in the case that $n_R^{(k)} < Ln_T$ the projection presented here is required in order to keep the same rank.

the following of this subsection we will assume that $n_R^{(k)} < Ln_T$. The positive semidefinite projection operator, \mathcal{P} , is defined as:

$$\mathcal{P}(\mathbf{X}, N) = \sum_{i=1}^N \lambda_i(\mathbf{X}) \mathbf{u}_i(\mathbf{X}) \mathbf{u}_i(\mathbf{X})^H, \quad (16)$$

where \mathbf{X} represents any Hermitian matrix, $\lambda_i(\mathbf{X})$ is the i -th eigenvalue of \mathbf{X} (sorted in decreasing order), and $\mathbf{u}_i(\mathbf{X})$ is its associated unitary eigenvector.

Consequently, this first step can be formally written as $\bar{\mathbf{R}}_{\text{quant,p}}^{(k)} = \mathcal{P}(\bar{\mathbf{R}}_{\text{quant}}^{(k)}, n_R^{(k)})$ and, from (13), we have that $\bar{\mathbf{R}}_{\text{quant,p}}^{(k)} = \mathcal{P}(\bar{\mathbf{R}}^{(k)} - \bar{\mathbf{R}}_{\text{err}}^{(k)}, n_R^{(k)})$. This can also be written as:

$$\bar{\mathbf{R}}_{\text{quant,p}}^{(k)} = \mathcal{P}(\bar{\mathbf{R}}^{(k)} - \bar{\mathbf{R}}_{\text{err}}^{(k)}, n_R^{(k)}) = \mathcal{P}(\bar{\mathbf{R}}^{(k)}, n_R^{(k)}) - \bar{\mathbf{R}}_{\text{err,p}}^{(k)} = \bar{\mathbf{R}}^{(k)} - \bar{\mathbf{R}}_{\text{err,p}}^{(k)}, \quad (17)$$

where the error after the projection is defined as $\bar{\mathbf{R}}_{\text{err,p}}^{(k)} \triangleq \mathcal{P}(\bar{\mathbf{R}}^{(k)}, n_R^{(k)}) - \mathcal{P}(\bar{\mathbf{R}}^{(k)} - \bar{\mathbf{R}}_{\text{err}}^{(k)}, n_R^{(k)}) = \bar{\mathbf{R}}^{(k)} - \mathcal{P}(\bar{\mathbf{R}}^{(k)} - \bar{\mathbf{R}}_{\text{err}}^{(k)}, n_R^{(k)})$. Now, it remains to linearly relate the real and imaginary parts of the elements in $\bar{\mathbf{R}}_{\text{err,p}}^{(k)}$ with those in $\bar{\mathbf{R}}_{\text{err}}^{(k)}$, which can be done using a first order approximation and the results in [23] as:

$$\widetilde{\text{vec}}(\bar{\mathbf{R}}_{\text{err,p}}^{(k)}) \approx \left(D_{\bar{\mathbf{r}}^{(k)}} \mathcal{P}(\bar{\mathbf{R}}^{(k)}, n_R^{(k)}) \right) \widetilde{\text{vec}}(\bar{\mathbf{R}}_{\text{err}}^{(k)}), \quad (18)$$

where we have used the operator $\widetilde{\text{vec}}(\cdot) = [\text{vech}(\Re(\cdot))^T \text{veci}(\Im(\cdot))^T]^T$ ⁵ and $\bar{\mathbf{r}}^{(k)} = \widetilde{\text{vec}}(\bar{\mathbf{R}}^{(k)})$. Finally, with a modicum of algebra and using the results in [24, Lem. A.4 and B.7], the expression of the Jacobian matrix $D_{\bar{\mathbf{r}}^{(k)}} \mathcal{P}(\bar{\mathbf{R}}^{(k)}, n_R^{(k)})$ can be computed as:

$$\begin{aligned} D_{\bar{\mathbf{r}}^{(k)}} \mathcal{P}(\bar{\mathbf{R}}^{(k)}, n_R^{(k)}) &= \sum_{i=1}^{n_R^{(k)}} \text{vec}(\mathbf{u}_i(\bar{\mathbf{R}}^{(k)}) \mathbf{u}_i(\bar{\mathbf{R}}^{(k)})^H) (D_{\bar{\mathbf{r}}^{(k)}} \lambda_i(\bar{\mathbf{R}}^{(k)})) \\ &+ \sum_{i=1}^{n_R^{(k)}} \lambda_i(\bar{\mathbf{R}}^{(k)}) (\mathbf{u}_i(\bar{\mathbf{R}}^{(k)})^* \otimes \mathbf{I}_{n_T}) (D_{\bar{\mathbf{r}}^{(k)}} \mathbf{u}_i(\bar{\mathbf{R}}^{(k)})) + \sum_{i=1}^{n_R^{(k)}} \lambda_i(\bar{\mathbf{R}}^{(k)}) (\mathbf{I}_{n_T} \otimes \mathbf{u}_i(\bar{\mathbf{R}}^{(k)})) (D_{\bar{\mathbf{r}}^{(k)}} \mathbf{u}_i(\bar{\mathbf{R}}^{(k)}))^*, \end{aligned}$$

where the explicit expressions for $D_{\bar{\mathbf{r}}^{(k)}} \lambda_i(\bar{\mathbf{R}}^{(k)})$ and $D_{\bar{\mathbf{r}}^{(k)}} \mathbf{u}_i(\bar{\mathbf{R}}^{(k)})$ can be straightforwardly found from the results in [23, Ch. 9] and are not reproduced here for the sake of space.

2) *Propagation through the Fourier transformation:* Equation (5) describes the computation of $\mathbf{R}_f^{(k)}$ from $\bar{\mathbf{R}}^{(k)}$, which corresponds to the Fourier transformation of the projected time domain Gram matrix. The error propagation through the computation of $\mathbf{R}_f^{(k)}$ is studied next. From (5) and (17) it follows that:

$$\mathbf{R}_f^{(k)} = \mathbf{F}_f^H \bar{\mathbf{R}}^{(k)} \mathbf{F}_f = \mathbf{F}_f^H \left(\bar{\mathbf{R}}_{\text{quant,p}}^{(k)} + \bar{\mathbf{R}}_{\text{err,p}}^{(k)} \right) \mathbf{F}_f, \quad (19)$$

where \mathbf{F}_f is the extended Fourier matrix defined as

$$\mathbf{F}_f = \left[e^{-j\frac{2\pi}{F}(f-0)}, e^{-j\frac{2\pi}{F}(f-1)}, \dots, e^{-j\frac{2\pi}{F}(f-(L-1))} \right]^T \otimes \mathbf{I}_{n_T} \in \mathbb{C}^{Ln_T \times n_T}.$$

⁵The operators $\text{vec}(\cdot)$, $\text{vech}(\cdot)$ and $\text{veci}(\cdot)$ act upon matrices and transform them into vectors as described next. First, $\text{vec}(\cdot)$ represents the vector obtained by stacking the columns from left to right. Next $\text{vech}(\cdot)$ transforms its matrix argument into a vector, by stacking only the elements of each column that lie on or below the main diagonal. Similarly, $\text{veci}(\cdot)$ represents the result of stacking only the elements of each column that lie strictly below the main diagonal.

Following from (19), the error in the computation of $\mathbf{R}_f^{(k)}$ at the transmitter is given by

$$\mathbf{R}_{\text{err}_f}^{(k)} \equiv \mathbf{F}_f^H \bar{\mathbf{R}}_{\text{err,p}}^{(k)} \mathbf{F}_f, \quad (20)$$

where we have that $\mathbf{R}_f^{(k)} = \mathbf{R}_{\text{quant}_f}^{(k)} + \mathbf{R}_{\text{err}_f}^{(k)}$ and $\mathbf{R}_{\text{quant}_f}^{(k)}$ is the estimated Gram matrix associated to receiver k and carrier f , i.e., $\mathbf{R}_{\text{quant}_f}^{(k)} = \mathbf{F}_f^H \bar{\mathbf{R}}_{\text{quant,p}}^{(k)} \mathbf{F}_f$. Following the structure used in the previous subsection, the propagation of the error through the Fourier transformation, (20), can be expressed as:

$$\widetilde{\text{vec}} \left(\mathbf{R}_{\text{err}_f}^{(k)} \right) = \widetilde{\mathbf{F}}_f \widetilde{\text{vec}} \left(\bar{\mathbf{R}}_{\text{err,p}}^{(k)} \right), \quad (21)$$

where

$$\widetilde{\mathbf{F}} = \begin{bmatrix} \mathbf{D}_{n_T}^+ \left(\mathbf{F}_{f_r}^T \otimes \mathbf{F}_{f_r}^T + \mathbf{F}_{f_i}^T \otimes \mathbf{F}_{f_i}^T \right) \mathbf{D}_{Ln_T} & \mathbf{D}_{n_T}^+ \left(\mathbf{F}_{f_r}^T \otimes \mathbf{F}_{f_i}^T - \mathbf{F}_{f_i}^T \otimes \mathbf{F}_{f_r}^T \right) \mathbf{D}_{Ln_T} \\ \mathbf{C}_{n_T}^+ \left(\mathbf{F}_{f_i}^T \otimes \mathbf{F}_{f_r}^T - \mathbf{F}_{f_r}^T \otimes \mathbf{F}_{f_i}^T \right) \mathbf{C}_{Ln_T} & \mathbf{C}_{n_T}^+ \left(\mathbf{F}_{f_r}^T \otimes \mathbf{F}_{f_r}^T + \mathbf{F}_{f_i}^T \otimes \mathbf{F}_{f_i}^T \right) \mathbf{C}_{Ln_T} \end{bmatrix}, \quad (22)$$

and the following notation was used: $\mathbf{F}_{f_r} = \Re(\mathbf{F}_f)$ and $\mathbf{F}_{f_i} = \Im(\mathbf{F}_f)$. Also, $(\cdot)^+$ stands for the pseudo-inverse operation and \mathbf{D}_n corresponds to the duplication matrix, whose definition is given in [23]. Similarly, the antiduplication matrix \mathbf{C}_n is defined as the unique matrix such that, for all $\mathbf{X} \in \mathbb{R}^{n \times n}$, $\text{vec}(\mathbf{X} - \mathbf{X}^T) = \mathbf{C}_n \text{veci}(\mathbf{X} - \mathbf{X}^T)$.

3) *Propagation through the equivalent channel computation:* After the computation of $\mathbf{R}_{\text{quant}_f}^{(k)}$, the error $\mathbf{R}_{\text{err}_f}^{(k)}$ is propagated through the matrix factorization at the transmitter described in section III-A1. The objective now is to obtain the expression of the error in the equivalent triangular channel response matrix $\mathbf{T}_{\text{err}_f}^{(k)}$ as a function of $\mathbf{R}_{\text{err}_f}^{(k)}$. A first order approximation of the error propagation is considered, which is valid for small errors. From (19) we have that

$$\mathbf{R}_f^{(k)} = \mathbf{T}_f^{(k)H} \mathbf{T}_f^{(k)}, \quad (23)$$

$$\mathbf{R}_{\text{quant}_f}^{(k)} + \mathbf{R}_{\text{err}_f}^{(k)} = \left(\mathbf{T}_{\text{quant}_f}^{(k)} + \mathbf{T}_{\text{err}_f}^{(k)} \right)^H \left(\mathbf{T}_{\text{quant}_f}^{(k)} + \mathbf{T}_{\text{err}_f}^{(k)} \right). \quad (24)$$

After some manipulations described in appendix A, the error in the equivalent channel response matrix $\mathbf{T}_{\text{err}_f}^{(k)}$ can be expressed as a function of $\mathbf{R}_{\text{err}_f}^{(k)}$ as:

$$\widetilde{\text{vec}} \left(\mathbf{T}_{\text{err}_f}^{(k)T} \right) \approx \left(\mathbf{D}_{\mathbf{r}_f^{(k)}} \mathbf{t}_f^{(k)} \right) \widetilde{\text{vec}} \left(\mathbf{R}_{\text{err}_f}^{(k)} \right), \quad (25)$$

where $\mathbf{D}_{\mathbf{r}_f^{(k)}} \mathbf{t}_f^{(k)}$ is the Jacobian matrix of $\mathbf{t}_f^{(k)}$ as defined in (49), and is derived in appendix A.

4) *Summary:* The complete CSI error propagation process described in the previous subsections is summarized in the diagram in Fig. 3, which reflects also the notation used through the computation. Mathematically, the complete error propagation process can be expressed from (18), (21) and (25) as:

$$\widetilde{\text{vec}} \left(\mathbf{T}_{\text{err}_f}^{(k)T} \right) \approx \mathbf{X}_f^{(k)} \bar{\mathbf{r}}_{\text{err}}^{(k)}, \quad (26)$$

where $\bar{\mathbf{r}}_{\text{err}}^{(k)} = \widetilde{\text{vec}} \left(\bar{\mathbf{R}}_{\text{err}}^{(k)} \right) \in \mathbb{R}^{L^2 n_T^2 \times 1}$ and $\mathbf{X}_f^{(k)}$ is the linear transformation that results from the error propagation through all the steps:

$$\mathbf{X}_f^{(k)} = \begin{cases} \left(\mathbf{D}_{\mathbf{r}_f^{(k)}} \mathbf{t}_f^{(k)} \right) \widetilde{\mathbf{F}}_f \left(\mathbf{D}_{\bar{\mathbf{r}}^{(k)}} \mathcal{P} \left(\bar{\mathbf{R}}^{(k)}, n_R^{(k)} \right) \right), & \text{if } n_R^{(k)} < Ln_T \\ \left(\mathbf{D}_{\mathbf{r}_f^{(k)}} \mathbf{t}_f^{(k)} \right) \widetilde{\mathbf{F}}_f, & \text{if } n_R^{(k)} \geq Ln_T \end{cases} \quad (27)$$

Finally, appendix A-A presents a notation that relates the subindices of the error in the triangular matrix $\mathbf{T}_{\text{err}_f}^{(k)}$ with the corresponding row index of matrix $\mathbf{X}_f^{(k)}$. This notation will be used in the following section.

V. APPLICATION TO ROBUST PRECODER DESIGN

In this section we present an example of robust design of the precoder matrix taking into account the error in the available CSI due to the quantization for the feedback transmission. As explained before, with the help of the transformation described in section III-A, the transmitter is able to compute an equivalent propagation channel using the channel Gram matrices sent through the feedback links. This will allow to apply a robust MSE precoding strategy, which takes into account explicitly the statistics of the inaccuracies in the CSI at the transmitter defined by $\bar{\mathbf{R}}_{\text{err}}^{(k)}$ in (13). The advantage of the robust design is that it is less sensitive to such errors.

A. Optimization of the MSE

It is important to emphasize that the equivalent channel transformation from section III-A can be used to apply any arbitrary design criterion and system architecture (also for joint precoder and decoder design) on top of it. In this section, and for illustrative purposes, the specific design criterion of minimization of the MSE with fixed decoders is considered as an example of application and because it is analytically tractable. First, the expression of the MSE is presented and then the robust precoder design is derived.

In order to adjust the dynamic range of $\hat{\mathbf{x}}$ before computing the MSE, the factor β is included, as in [25], [26], which could be understood as a gain control at the receivers. The MSE is then given by:

$$\begin{aligned} \text{MSE}(\mathbf{P}, \beta) &= \sum_{k=1}^K \mathbb{E}_{\mathbf{T}_{\text{err}}^{(k)}, \mathbf{x}^{(k)}, \mathbf{w}^{(k)}} \left\{ \|\mathbf{x}^{(k)} - \beta^{-1} \hat{\mathbf{x}}^{(k)}\|_2^2 \right\} = \mathbb{E}_{\mathbf{T}_{\text{err}}, \mathbf{x}, \mathbf{w}} \left\{ \|\mathbf{x} - \beta^{-1} \hat{\mathbf{x}}\|_2^2 \right\} \\ &= \frac{1}{\beta^2} \text{tr} \left(\tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \mathbf{P} \mathbf{R}_x \mathbf{P}^H \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H - \beta \tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \mathbf{P} \mathbf{R}_x - \beta \mathbf{R}_x \mathbf{P}^H \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H + \beta^2 \mathbf{R}_x \right) \\ &\quad + \frac{1}{\beta^2} \text{tr} \left(\mathbf{P} \mathbf{R}_x \mathbf{P}^H \Delta \right) + \frac{1}{\beta^2} \text{tr} \left(\tilde{\mathbf{D}} \mathbf{Q}^H \mathbf{R}_w \mathbf{Q} \tilde{\mathbf{D}}^H \right), \end{aligned} \quad (28)$$

with $\Delta = \mathbb{E}_{\mathbf{T}_{\text{err}}} \left\{ \mathbf{T}_{\text{err}}^H \tilde{\mathbf{D}}^H \tilde{\mathbf{D}} \mathbf{T}_{\text{err}} \right\} = \sum_{k=1}^K \mathbb{E}_{\mathbf{T}_{\text{err}}^{(k)}} \left\{ \mathbf{T}_{\text{err}}^{(k)H} \tilde{\mathbf{D}}^{(k)H} \tilde{\mathbf{D}}^{(k)} \mathbf{T}_{\text{err}}^{(k)} \right\}$. Note that matrix Δ depends on the second-order statistics of the error in the equivalent channel matrix \mathbf{T}_{err} . Such statistics can be computed assuming that the second-order statistics of the original quantization error $\bar{\mathbf{R}}_{\text{err}}^{(k)}$ are known and using the analytic study of the error propagation presented in section IV. Appendix B presents how matrix Δ can be computed for the particular case where the error matrix $\bar{\mathbf{R}}_{\text{err}}^{(k)}$ is composed of i.i.d. elements and using explicitly the derivation presented in Section IV. Note that the extension for any correlation of the elements in $\bar{\mathbf{R}}_{\text{err}}^{(k)}$ could be addressed following similar steps as those presented in Appendix B.

B. Robust precoder design

The robust system design can be expressed as the following optimization problem based on the MSE criterion (28) and including a constraint on the maximum power P_t available at the transmitter:

$$[\mathbf{P}_{rob}^*, \beta_{rob}^*] = \arg \min_{\{\mathbf{P}, \beta\}} \text{MSE}(\mathbf{P}, \beta) \quad (29)$$

$$\text{s.t.}: \quad \mathbb{E} \{ \|\mathbf{P}\mathbf{x}\|_2^2 \} = P_t. \quad (30)$$

Note that the MSE is not jointly convex in \mathbf{P} and β . However, two necessary conditions arise from the fact that the optimum solution must fulfill that the optimum \mathbf{P} minimizes the MSE for the optimum β subject to the power constraint and at the same time, the optimum value of β must minimize the MSE for \mathbf{P} equal to its optimum value:

$$\mathbf{P}_{rob}^* = \arg \min_{\mathbf{P}} \text{MSE}(\mathbf{P}, \beta_{rob}^*) \quad (31)$$

$$\text{s.t.}: \quad \mathbb{E} \{ \|\mathbf{P}\mathbf{x}\|_2^2 \} = P_t. \quad (32)$$

and

$$\beta_{rob}^* = \arg \min_{\beta} \text{MSE}(\mathbf{P}_{rob}^*, \beta). \quad (33)$$

The problem (31)-(32) is convex and therefore the optimum solution must satisfy the expression obtained by constructing the Lagrangian function $L(\mathbf{P}; \lambda)$ with Lagrange multiplier $\lambda \in \mathbb{R}^+$ and setting its derivatives equal to zero [27]:

$$L(\mathbf{P}; \lambda) = \mathbb{E} \left\{ \|\mathbf{x} - \beta_{rob}^{*-1} \hat{\mathbf{x}}\|_2^2 \right\} + \lambda (\text{tr}(\mathbf{P}\mathbf{R}_x\mathbf{P}^H) - P_t). \quad (34)$$

$$\nabla_{\mathbf{P}^H} L = \frac{1}{\beta_{rob}^{*2}} \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \mathbf{P} \mathbf{R}_x - \frac{1}{\beta_{rob}^*} \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \mathbf{R}_x + \frac{1}{\beta_{rob}^{*2}} \Delta \mathbf{P} \mathbf{R}_x + \lambda \mathbf{P} \mathbf{R}_x = 0. \quad (35)$$

Similarly, the condition (33) is convex in β and deriving the MSE with respect to β results in:

$$\begin{aligned} \frac{\partial \text{MSE}(\mathbf{P}_{rob}^*, \beta)}{\partial \beta} &= -\frac{2}{\beta^3} \text{tr} \left(\tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \mathbf{P}_{rob}^* \mathbf{R}_x \mathbf{P}_{rob}^{*H} \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \right) - \frac{2}{\beta^3} \text{tr} \left(\mathbf{P}_{rob}^* \mathbf{R}_x \mathbf{P}_{rob}^{*H} \Delta \right) \\ &\quad - \frac{2}{\beta^3} \text{tr} \left(\tilde{\mathbf{D}} \mathbf{Q}^H \mathbf{R}_w \mathbf{Q} \tilde{\mathbf{D}}^H \right) + \frac{1}{\beta^2} \text{tr} \left(\tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \mathbf{P}_{rob}^* \mathbf{R}_x \right) + \frac{1}{\beta^2} \text{tr} \left(\mathbf{R}_x \mathbf{P}_{rob}^{*H} \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \right) = 0, \end{aligned} \quad (36)$$

which, using the fact that $\text{tr} \left(\tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \mathbf{P}_{rob}^* \mathbf{R}_x \right) = \text{tr} \left(\mathbf{R}_x \mathbf{P}_{rob}^{*H} \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \right)$, results in:

$$-\text{tr} \left(\tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \mathbf{P}_{rob}^* \mathbf{R}_x \mathbf{P}_{rob}^{*H} \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \right) - \text{tr} \left(\mathbf{P}_{rob}^* \mathbf{R}_x \mathbf{P}_{rob}^{*H} \Delta \right) - \text{tr} \left(\tilde{\mathbf{D}} \mathbf{Q}^H \mathbf{R}_w \mathbf{Q} \tilde{\mathbf{D}}^H \right) + \beta \text{tr} \left(\tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \mathbf{P}_{rob}^* \mathbf{R}_x \right) = 0. \quad (37)$$

We now introduce the following change of variables $\xi = \lambda \beta^2$ and $\mathbf{P} = \beta \tilde{\mathbf{P}}$. Then (37) results in⁶:

$$\xi \beta^2 \text{tr} \left(\tilde{\mathbf{P}}_{rob}^* \mathbf{R}_x \tilde{\mathbf{P}}_{rob}^{*H} \right) - \text{tr} \left(\tilde{\mathbf{D}} \mathbf{Q}^H \mathbf{R}_w \mathbf{Q} \tilde{\mathbf{D}}^H \right) = 0 \Rightarrow \xi_{rob}^* = \frac{\text{tr} \left(\tilde{\mathbf{D}} \mathbf{Q}^H \mathbf{R}_w \mathbf{Q} \tilde{\mathbf{D}}^H \right)}{P_t}. \quad (38)$$

From (35), it follows that:

$$\mathbf{P}_{rob}^* = \beta_{rob}^* \left(\mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} + \Delta + \xi_{rob}^* \mathbf{I} \right)^{-1} \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H, \quad (39)$$

⁶The fact that $\text{tr} \left(\tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} \tilde{\mathbf{P}} \mathbf{R}_x \right) = \text{tr} \left(\left(\mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} + \xi \mathbf{I} \right) \tilde{\mathbf{P}} \mathbf{R}_x \tilde{\mathbf{P}}^H \right)$ was used in the derivations.

and from the power constraint $\text{tr}(\mathbf{P}\mathbf{R}_x\mathbf{P}^H) = P_t$ and (39), it follows directly that:

$$\beta_{rob}^* = \sqrt{\frac{P_t}{\text{tr}\left[\left(\mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \tilde{\mathbf{D}} \mathbf{T}_{\text{quant}} + \mathbf{\Delta} + \xi_{rob}^* \mathbf{I}\right)^{-2} \mathbf{T}_{\text{quant}}^H \tilde{\mathbf{D}}^H \mathbf{R}_x \tilde{\mathbf{D}} \mathbf{T}_{\text{quant}}\right]}}. \quad (40)$$

Since there is only one solution (up to a phase change) that satisfies the two necessary conditions, (31)-(32) and (33), this solution will be the optimum one.

Note that, for the computation of (38), an additional parameter associated to the noise power of each user, $\xi^{(k)} = \text{tr}\left(\tilde{\mathbf{D}}^{(k)} \mathbf{Q}^{(k)H} \mathbf{R}_w^{(k)} \mathbf{Q}^{(k)} \tilde{\mathbf{D}}^{(k)H}\right)$, has to be fed back to the transmitter. However, this scalar parameter varies very slowly over time and does not imply a relevant increase in the feedback load. From (38), it follows that $\xi_{rob}^* = \frac{\sum_{k=1}^K \xi^{(k)}}{P_t}$. Observe that, in the case that there is no knowledge of the CSI error at the transmitter, a naive design would assume $\mathbf{\Delta} = 0$, and in this case (39)-(40) results in a non-robust design which coincides with the optimum non-robust design derived in [25].

C. Particular case: independent processing per carrier

In the particular case where the decoder matrix of each user $\mathbf{D}^{(k)}$ is constrained to be block diagonal, which is the case for example when joint-processing of the signals from different carriers is not possible at the receiver, the optimum solution given by (38)-(40) is also block diagonal. This means that if the decoder is not capable of processing the signals of different carriers jointly, the optimum precoder does not spread the information symbols across carriers. Consequently, in this particular case, the solution from (38)-(40) is also valid for the MIMO-OFDM scheme as described in (2).

VI. NUMERICAL RESULTS

This section numerically evaluates the performance of the proposed design framework and precoding scheme. For the simulations we consider a scenario featuring a transmitter with $n_T = 4$ antennas and $K = 2$ receivers with $n_R^{(k)} = 2, k = 1, 2$, antennas each. The l th tap of the channel impulse response is generated as $\bar{\mathbf{H}}_l^{(k)} = \sigma_l \mathbf{N}_l^{(k)}$, where σ_l characterizes the power delay profile and $\mathbf{N}_l^{(k)}$ is composed of i.i.d. zero-mean circularly symmetric complex Gaussian entries with unit variance. For the simulations, we considered an exponential decaying power delay profile given by $\sigma_l^2 = a e^{-\frac{l}{\tau}}$, (where $a = (\sum_{n=0}^{L-1} e^{-\frac{n}{\tau}})^{-1}$) with a normalized delay spread of $\tau = 3$. The simulations are averaged over a sufficiently large number of realizations. Since the joint optimal design of \mathbf{P} and \mathbf{D} is still an open problem, a decoder matrix \mathbf{D} has to be fixed for the simulations. A simple choice is to set $\mathbf{D}^{(k)} = \mathbf{I}$, as in [11]. Note that this implies that the number of streams is chosen as $n_S^{(k)} = F n_R^{(k)}$.

A. Evaluation of the robust precoder

In this subsection the performance of the proposed robust algorithm, implemented within the presented feedback framework, is numerically compared with that of the non-robust algorithm from [25]. To show the applicability of the presented framework, both the naive (i.e., non-robust) and the

robust versions of BD [20] are also implemented and compared, in a setup with $L = 16$ taps and $F = 128$ carriers.

Using this setup, Fig. 4 shows the MSE versus the transmit power P_t for different values of the variance of each element of the error matrix $\bar{\mathbf{R}}_{\text{err}}^{(k)}$, represented by $\bar{\sigma}_e^2$. These simulations show that the improvement in terms of MSE of the robust design with respect to the non-robust solution is higher as the error in the quantization and feedback increases. The same conclusion applies to the case when a SER cost function is used, as shown in Fig. 5 for a scenario featuring a QPSK constellation.

The framework allows also other design implementations and, as an example, a design based on BD [20] is now considered⁷. Using this, the robust and non-robust BD designs are applied, and the results in the considered scenario are shown in Fig. 6. Note that this is shown as an example of the applicability of the framework, but, if the designs based on BD were to be compared to the non-BD designs, the performance in terms of MSE would be worse for the BD schemes since they spend degrees of freedom to force interference nulling among users.

The performance as a function of the amount of error in the CSI is considered next. Fig. 7 shows the achievable MSE versus the SNR in the estimation of $\bar{\mathbf{R}}^{(k)}$, defined as $\text{SNR}_e = \frac{1}{\bar{\sigma}_e^2}$, for a fixed value of the transmit power $P_t = 60$ dB. The curves show that the robust designs outperform the other precoding techniques when the estimation of $\bar{\mathbf{R}}^{(k)}$ is not very good, i.e., when the SNR_e is low and consequently the error is high, while at high SNR_e the error in the CSI is very small and the curves corresponding to the non-robust techniques converge to the curves corresponding to the robust designs as it is to be expected. The designs based on BD show a small performance loss due to the fact that some degrees of freedom are used to guarantee an interference-free transmission.

B. Comparison of feedback strategies

A numerical analysis of the performance of the feedback scheme based on the quantization and feedback of one temporal channel Gram matrix $\bar{\mathbf{R}}^{(k)}$ per user (as described in section III), instead of the usual feedback per carrier per user and the traditional feedback of the complete channel propagation matrix is presented in this subsection. This performance comparison is numerically characterized by constraining the same number of quantization bits for the different approaches in order to obtain a fair evaluation. There are multiple quantization and feedback algorithms that can be used to quantize either $\bar{\mathbf{R}}^{(k)} \in \mathbb{C}^{Ln_T \times Ln_T}$ or the F matrices $\mathbf{R}_f^{(k)} \in \mathbb{C}^{n_T \times n_T}$. Since the focus of this work is on the objective of the quantization and not on the algorithm itself, the algorithm from [18] will be taken as a reference for the comparison of due to its simplicity.⁸ The performance of a system using both the

⁷The transceiver design is implemented within the proposed framework by applying the BD scheme on top of the equivalent triangular channels presented in section III in this paper.

⁸The algorithm from [18] used as a reference is based on an individual quantization of the real and imaginary non-repeated elements of the matrix, i.e., in the scheme based on quantization of temporal CSI, $L^2 n_T^2$ real scalar elements have to be quantized, while in the scheme based on quantization of frequency CSI, $F n_T^2$ real scalar elements have to be quantized.

quantization of $\bar{\mathbf{R}}^{(k)}$ and the quantization of the F matrices $\mathbf{R}_f^{(k)}$ and featuring the same number of feedback bits for both cases will be shown next.

First, a comparison of the feedback based on the channel Gram matrix versus the feedback of the complete channel response matrix is presented. A scenario with $L = 16$ taps and $F = 16$ carriers is considered, and the results are shown in Fig. 8. It can be seen that the feedback of the Gram matrix provides a lower MSE than the technique based on direct quantization of the channel response matrix, for the cases of $B = 1536$ and $B = 2560$ total feedback bits.

Next, Fig. 9 shows a comparison of the performance using feedback of the time domain CSI versus feedback of the frequency domain CSI in a scenario with $L = 8$ taps and $F = 128$ carriers, for different values of the feedback load. First, we considered $B = 12288$ bits of feedback. This means that each of the 1024 real and scalar parameters that have to be fed back in the time domain CSI feedback is quantized using 12 bits, while each of the 2048 parameters corresponding to the frequency domain CSI feedback case is quantized using 6 bits. The figure also shows the results of simulations featuring 14336 bits per feedback update (which corresponds to 14 bits for the quantization of each element in the scheme based on time domain CSI and 7 bits for each element in the scheme based on frequency domain CSI). These curves show that, for this specific setup, the quantization and feedback of matrix $\bar{\mathbf{R}}^{(k)}$ (which is based on the time domain CSI) provides a lower SER than the quantization and feedback of $\mathbf{R}_f^{(k)}$ (which is based on frequency domain CSI) when using the same feedback algorithm. This is due to the fact that in the case of time domain CSI feedback the number of parameters to be quantized is half the number of parameters to be quantized using the same number of bits in the case of frequency domain CSI feedback. In a scenario such as the one evaluated in Fig. 10, with $F = 64$ carriers instead of the 128 carriers considered in Fig. 9, the number of elements to be quantized is higher in the time domain CSI feedback than in the frequency domain CSI feedback, and the later shows better performance.

From this we conclude that the choice of the most adequate feedback scheme (feedback of the time domain CSI or feedback of the frequency domain CSI) depends on the number of taps of the temporal channel response and the number of carriers, and this should be taken into consideration at the system design stage. Note that the trend in wireless communication systems is to increase the number of carriers (WiMAX for example supports up to 1728 usable carriers [28]), in which case the feedback of the time domain CSI provides better performance.

VII. CONCLUSIONS

This work presents a framework for the design of multiuser MIMO-OFDM BC systems with CSI feedback. The proposed framework is based on the computation of an equivalent triangular channel response matrix, and enables the use of efficient CSI feedback techniques based on the quantization of the Gram matrix of the temporal response of the channels. This scheme is valid for and can be applied to any given design quality criterion. An analytical study of the propagation of CSI quantization

error through the channel Gram matrix computation and the posterior equivalent channel response matrix computation is also presented. As an illustrative example of the potential of this framework for transceiver designs, the case of MSE minimization has been considered and a closed form expression for a robust space-frequency linear precoding design has been derived. Numerical simulations reveal the advantages of the proposed feedback scheme and also of the MSE minimization precoding technique compared to other feedback techniques and to the non-robust counterpart precoding techniques.

APPENDIX A

CSI ERROR PROPAGATION THROUGH THE COMPUTATION OF THE EQUIVALENT CHANNEL

In the neighborhood of $\mathbf{R}_f^{(k)}$ the errors $\mathbf{R}_{\text{err}_f}^{(k)}$ and $\mathbf{T}_{\text{err}_f}^{(k)}$ can be approximated by the differentials, $d\mathbf{R}_f^{(k)}$ and $d\mathbf{T}_f^{(k)}$, respectively. From (23), we readily obtain:

$$d\mathbf{R}_f^{(k)} = d\mathbf{T}_f^{(k)H} \mathbf{T}_f^{(k)} + \mathbf{T}_f^{(k)H} d\mathbf{T}_f^{(k)}. \quad (41)$$

By separating the real and imaginary parts of the individual matrices, $\mathbf{R}_{f_r}^{(k)} = \Re(\mathbf{R}_f^{(k)})$, $\mathbf{R}_{f_i}^{(k)} = \Im(\mathbf{R}_f^{(k)})$, $\mathbf{T}_{f_r}^{(k)} = \Re(\mathbf{T}_f^{(k)})$, $\mathbf{T}_{f_i}^{(k)} = \Im(\mathbf{T}_f^{(k)})$, and defining $d\mathbf{R}_f^{(k)} = d\mathbf{R}_{f_r}^{(k)} + j d\mathbf{R}_{f_i}^{(k)}$ and $d\mathbf{T}_f^{(k)} = d\mathbf{T}_{f_r}^{(k)} + j d\mathbf{T}_{f_i}^{(k)}$, we obtain:

$$d\mathbf{R}_{f_r}^{(k)} = d\mathbf{T}_{f_r}^{(k)T} \mathbf{T}_{f_r}^{(k)} + d\mathbf{T}_{f_i}^{(k)T} \mathbf{T}_{f_i}^{(k)} + \mathbf{T}_{f_r}^{(k)T} d\mathbf{T}_{f_r}^{(k)} + \mathbf{T}_{f_i}^{(k)T} d\mathbf{T}_{f_i}^{(k)}, \quad (42)$$

$$d\mathbf{R}_{f_i}^{(k)} = d\mathbf{T}_{f_r}^{(k)T} \mathbf{T}_{f_i}^{(k)} - d\mathbf{T}_{f_i}^{(k)T} \mathbf{T}_{f_r}^{(k)} + \mathbf{T}_{f_r}^{(k)T} d\mathbf{T}_{f_i}^{(k)} - \mathbf{T}_{f_i}^{(k)T} d\mathbf{T}_{f_r}^{(k)}. \quad (43)$$

The following facts are considered in the derivations:

- 1) From the fact that $\mathbf{T}_f^{(k)}$ is upper triangular with real and positive elements in the main diagonal, it follows that $d\mathbf{T}_{f_i}^{(k)}$ is strictly upper triangular. Thus, the only non-zero elements of $d\mathbf{T}_{f_r}^{(k)}$ and $d\mathbf{T}_{f_i}^{(k)}$ are contained in the vectors $\text{vech}(d\mathbf{T}_{f_r}^{(k)T})$ and $\text{veci}(d\mathbf{T}_{f_i}^{(k)T})$.
- 2) From (42), it follows that $d\mathbf{R}_{f_r}^{(k)}$ is symmetric. Thus its non-repeated elements are contained in the vector $\text{vech}(d\mathbf{R}_{f_r}^{(k)})$.
- 3) From (43), it follows that $d\mathbf{R}_{f_i}^{(k)}$ is anti-symmetric. Thus, it has zeros in the main diagonal and all its non-repeated elements (up to a change of sign) are contained in the vector $\text{veci}(d\mathbf{R}_{f_i}^{(k)})$.

Consequently, from all that has been said above, in order to compute the derivative of $\mathbf{T}_f^{(k)}$ with respect to $\mathbf{R}_f^{(k)}$, our objective is to linearly relate the elements of $\text{vech}(d\mathbf{T}_{f_r}^{(k)T})$ and $\text{veci}(d\mathbf{T}_{f_i}^{(k)T})$ to those of $\text{vech}(d\mathbf{R}_{f_r}^{(k)})$ and $\text{veci}(d\mathbf{R}_{f_i}^{(k)})$ and apply the first identification theorem [23].

We start by applying vech at both sides in (42), and it follows that⁹:

$$\text{vech}(d\mathbf{R}_{f_r}^{(k)}) = \text{vech} \left(d\mathbf{T}_{f_r}^{(k)T} \mathbf{T}_{f_r}^{(k)} + d\mathbf{T}_{f_i}^{(k)T} \mathbf{T}_{f_i}^{(k)} + \mathbf{T}_{f_r}^{(k)T} d\mathbf{T}_{f_r}^{(k)} + \mathbf{T}_{f_i}^{(k)T} d\mathbf{T}_{f_i}^{(k)} \right) \quad (44)$$

$$= \mathbf{D}_{n_T}^+ \text{vec} \left(d\mathbf{T}_{f_r}^{(k)T} \mathbf{T}_{f_r}^{(k)} + d\mathbf{T}_{f_i}^{(k)T} \mathbf{T}_{f_i}^{(k)} + \mathbf{T}_{f_r}^{(k)T} d\mathbf{T}_{f_r}^{(k)} + \mathbf{T}_{f_i}^{(k)T} d\mathbf{T}_{f_i}^{(k)} \right) \quad (45)$$

$$= 2\mathbf{D}_{n_T}^+ \left(\left(\mathbf{T}_{f_r}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R} \text{vech} \left(d\mathbf{T}_{f_r}^{(k)T} \right) + \left(\mathbf{T}_{f_i}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R}^S \text{veci} \left(d\mathbf{T}_{f_i}^{(k)T} \right) \right). \quad (46)$$

⁹In the developments, we make use of the triangularization matrix $\mathbf{V}_{n,m}$, which is the unique matrix of the appropriate dimensions such that, for all lower triangular $\mathbf{X} \in \mathbb{R}^{n \times m}$, we have $\text{vec}(\mathbf{X}) = \mathbf{V}_{n,m} \text{vech}(\mathbf{X})$, and the strict triangularization matrix $\mathbf{V}_{n,m}^S$, which is the unique matrix of the appropriate dimensions such that, for all strictly lower triangular $\mathbf{X} \in \mathbb{R}^{n \times m}$, we have $\text{vec}(\mathbf{X}) = \mathbf{V}_{n,m}^S \text{veci}(\mathbf{X})$.

Now, applying veci at both sides in (43), and operating similarly as before, we get:

$$\text{veci}(d\mathbf{R}_{f_i}^{(k)}) = \text{veci}\left(d\mathbf{T}_{f_r}^{(k)T} \mathbf{T}_{f_i}^{(k)} + d\mathbf{T}_{f_i}^{(k)T} \mathbf{T}_{f_r}^{(k)} + \mathbf{T}_{f_r}^{(k)T} d\mathbf{T}_{f_i}^{(k)} + \mathbf{T}_{f_i}^{(k)T} d\mathbf{T}_{f_r}^{(k)}\right) \quad (47)$$

$$= 2\mathbf{C}_{n_T}^+ \left(\left(\mathbf{T}_{f_i}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}} \text{vech}\left(d\mathbf{T}_{f_r}^{(k)T}\right) - \left(\mathbf{T}_{f_r}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}}^S \text{veci}\left(d\mathbf{T}_{f_i}^{(k)T}\right) \right). \quad (48)$$

Now, defining

$$\mathbf{t}_f^{(k)} \equiv \widetilde{\text{vec}}\left(\mathbf{T}_f^{(k)T}\right), \quad d\mathbf{t}_f^{(k)} \equiv \widetilde{\text{vec}}\left(d\mathbf{T}_f^{(k)T}\right), \quad \mathbf{r}_f^{(k)} \equiv \widetilde{\text{vec}}\left(\mathbf{R}_f^{(k)}\right), \quad d\mathbf{r}_f^{(k)} \equiv \widetilde{\text{vec}}\left(d\mathbf{R}_f^{(k)}\right), \quad (49)$$

from (46) and (48), it follows that

$$d\mathbf{r}_f^{(k)} = 2 \begin{bmatrix} \mathbf{D}_{n_T}^+ \left(\mathbf{T}_{f_r}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}} & \mathbf{D}_{n_T}^+ \left(\mathbf{T}_{f_i}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}}^S \\ \mathbf{C}_{n_T}^+ \left(\mathbf{T}_{f_i}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}} & -\mathbf{C}_{n_T}^+ \left(\mathbf{T}_{f_r}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}}^S \end{bmatrix} d\mathbf{t}_f^{(k)}. \quad (50)$$

It only remains to take the pseudo-inverse in the last equation to obtain the desired Jacobian matrix:

$$\mathbf{D}_{\mathbf{r}_f^{(k)}} \mathbf{t}_f^{(k)} = \frac{1}{2} \begin{bmatrix} \mathbf{D}_{n_T}^+ \left(\mathbf{T}_{f_r}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}} & \mathbf{D}_{n_T}^+ \left(\mathbf{T}_{f_i}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}}^S \\ \mathbf{C}_{n_T}^+ \left(\mathbf{T}_{f_i}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}} & -\mathbf{C}_{n_T}^+ \left(\mathbf{T}_{f_r}^{(k)T} \otimes \mathbf{I}_{n_T} \right) \mathbf{V}_{n_T, n_R^{(k)}}^S \end{bmatrix}^+. \quad (51)$$

Consequently, we have that $\mathbf{T}_{\text{err}_f}^{(k)}$ as a function of $\mathbf{R}_{\text{err}_f}^{(k)}$ can be computed as:

$$\widetilde{\text{vec}}\left(\mathbf{T}_{\text{err}_f}^{(k)T}\right) \approx \mathbf{D}_{\mathbf{r}_f^{(k)}} \mathbf{t}_f^{(k)} \widetilde{\text{vec}}\left(\mathbf{R}_{\text{err}_f}^{(k)}\right). \quad (52)$$

A. On the element-wise propagation error vector

In this subsection we will express the total error propagation from (26) using element-wise notation.

We will denote the i th row of matrix $\mathbf{X}_f^{(k)}$ as $\mathbf{x}_{f(i)}^{(k)}$. The element n, r of the error matrix $\mathbf{T}_{\text{err}_f}^{(k)} \in \mathbb{C}^{n_R^{(k)} \times n_T}$ will be denoted by $t_{\text{err}_f(n,r)}^{(k)}$ and can be computed as:

$$\Re(t_{\text{err}_f(n,r)}^{(k)}) \approx \begin{cases} \mathbf{a}_{f(n,r)}^{(k)} \bar{\mathbf{r}}_{err}^{(k)}; & \forall n \leq r, \\ 0 & \forall n > r \end{cases} \quad (53)$$

$$\Im(t_{\text{err}_f(n,r)}^{(k)}) \approx \begin{cases} \mathbf{b}_{f(n,r)}^{(k)} \bar{\mathbf{r}}_{err}^{(k)}; & \forall n < r, \\ 0 & \forall n \geq r. \end{cases} \quad (54)$$

Based on the antenna topology, two cases have to be considered:

1) $n_R^{(k)} \geq n_T$: In this case, $\widetilde{\text{vec}}\left(\mathbf{T}_{\text{err}_f}^{(k)T}\right) \in \mathbb{R}^{n_T^2 \times 1}$ and $\mathbf{X}_f^{(k)} \in \mathbb{C}^{n_T^2 \times L^2 n_T^2}$. Because of this matrix structure, we have that $\mathbf{a}_{f(n,r)}^{(k)} = \mathbf{x}_f^{(k)} \begin{pmatrix} r+n_T(n-1)-\sum_{x=1}^n x \\ x \neq n \\ x \neq n-1 \end{pmatrix}$ and $\mathbf{b}_{f(n,r)}^{(k)} = \mathbf{x}_f^{(k)} \begin{pmatrix} \frac{n_T(n_T+1)}{2} + n_T(n-1) - \frac{n(n-1)}{2} + r \end{pmatrix}$.

2) $n_R^{(k)} < n_T$: In this case, $\widetilde{\text{vec}}\left(\mathbf{T}_{\text{err}_f}^{(k)T}\right) \in \mathbb{R}^{n_R^{(k)}(2n_T - n_R^{(k)}) \times 1}$ and $\mathbf{X}_f^{(k)} \in \mathbb{C}^{n_R^{(k)}(2n_T - n_R^{(k)}) \times L^2 n_T^2}$. This results in $\mathbf{a}_{f(n,r)}^{(k)} = \mathbf{x}_f^{(k)} \begin{pmatrix} r+n_T(n-1)-\sum_{x=1}^n x \\ x \neq n \\ x \neq n-1 \end{pmatrix}$ and $\mathbf{b}_{f(n,r)}^{(k)} = \mathbf{x}_f^{(k)} \begin{pmatrix} \frac{n_R^{(k)}(n_R^{(k)}+1)}{2} + (n_T - n_R^{(k)})n_R^{(k)} + n_T(n-1) - \frac{n(n-1)}{2} + r \end{pmatrix}$.

APPENDIX B

COMPUTATION OF MATRIX Δ

From (28), the element i, j of matrix $\Delta \in \mathbb{C}^{Fn_T \times Fn_T}$ can be computed as:

$$\Delta_{(i,j)} = \sum_{k=1}^K \mathbb{E}_{\mathbf{T}_{\text{err}}^{(k)}} \left\{ \mathbf{t}_{\text{err}_i}^{(k)H} \tilde{\mathbf{D}}^{(k)H} \tilde{\mathbf{D}}^{(k)} \mathbf{t}_{\text{err}_j}^{(k)} \right\} = \sum_{k=1}^K \mathbb{E}_{\mathbf{T}_{\text{err}}^{(k)}} \left\{ \text{tr} \left(\mathbf{t}_{\text{err}_i}^{(k)H} \tilde{\mathbf{D}}^{(k)H} \tilde{\mathbf{D}}^{(k)} \mathbf{t}_{\text{err}_j}^{(k)} \right) \right\} \quad (55)$$

$$= \sum_{k=1}^K \text{tr} \left(\tilde{\mathbf{D}}^{(k)H} \tilde{\mathbf{D}}^{(k)} \mathbb{E}_{\mathbf{T}_{\text{err}}^{(k)}} \left\{ \mathbf{t}_{\text{err}_j}^{(k)} \mathbf{t}_{\text{err}_i}^{(k)H} \right\} \right), \quad (56)$$

where $\mathbf{t}_{\text{err}_j}^{(k)}$ is the j th column of matrix $\mathbf{T}_{\text{err}}^{(k)}$.

For simplicity with the notation, the elements of $\Delta \in \mathbb{C}^{Fn_T \times Fn_T}$ will be denoted as $\Delta_{(n_T f + n, n_T g + m)}$, where $f, g \in \{0, \dots, F-1\}$ and $n, m \in \{1, \dots, n_T\}$. Using this notation, (56) can be expressed as:

$$\Delta_{(n_T f + n, n_T g + m)} = \sum_{k=1}^K \text{tr} \left(\tilde{\mathbf{D}}^{(k)H} \tilde{\mathbf{D}}^{(k)} \mathbb{E} \left\{ \mathbf{t}_{\text{err}_{n_T g + m}}^{(k)} \mathbf{t}_{\text{err}_{n_T f + n}}^{(k)H} \right\} \right) \quad (57)$$

$$= \sum_{k=1}^K \sum_{x=1}^{n_R^{(k)}} \sum_{r=1}^{n_R^{(k)}} \bar{d}_{\binom{(k)}{(n_R^{(k)} f + x, n_R^{(k)} g + r)}} \mathbb{E} \left\{ t_{\text{err}_{g(r,m)}^{(k)}} t_{\text{err}_{f(x,n)}^{(k)*}} \right\}, \quad (58)$$

where $t_{\text{err}_{f(i,j)}^{(k)}}$ is the element i, j of matrix $\mathbf{T}_{\text{err}_f}^{(k)}$ (see appendices A-A1, and A-A2 for the expression of $t_{\text{err}_{f(i,j)}^{(k)}}$ in the cases of $n_R^{(k)} \geq n_T$ and $n_R^{(k)} < n_T$, respectively) and $\bar{d}_{(i,j)}^{(k)}$ is the element i, j of matrix $\tilde{\mathbf{D}}^{(k)H} \tilde{\mathbf{D}}^{(k)}$. Note that some of the elements of the summation in (58) are zero due to the fact that matrix $\mathbf{T}_{\text{err}_f}^{(k)}$ is upper triangular. Expression (58) can be manipulated further to write it as a function of $\mathbb{E} \left\{ \bar{\mathbf{r}}_{\text{err}}^{(k)} \bar{\mathbf{r}}_{\text{err}}^{(k)T} \right\}$, the variance of the error introduced in the quantization of the temporal Gram matrix:

$$\Delta_{(n_T f + n, n_T g + m)} \approx \sum_{k=1}^K \sum_{x=1}^{n_R^{(k)}} \sum_{r=1}^{n_R^{(k)}} \bar{d}_{\binom{(k)}{(n_R^{(k)} f + x, n_R^{(k)} g + r)}} \left(\mathbf{a}_{g(r,m)}^{(k)} + j\mathbf{b}_{g(r,m)}^{(k)} \right) \mathbb{E} \left\{ \bar{\mathbf{r}}_{\text{err}}^{(k)} \bar{\mathbf{r}}_{\text{err}}^{(k)T} \right\} \left(\mathbf{a}_{f(x,n)}^{(k)T} - j\mathbf{b}_{f(x,n)}^{(k)T} \right). \quad (59)$$

In the particular case when the error in the CSI at the transmitter, $\bar{\mathbf{R}}_{\text{err}}^{(k)}$, is composed of i.i.d. elements with zero mean and variance $\bar{\sigma}^{(k)2}$, it follows that $\mathbb{E} \left\{ \bar{\mathbf{r}}_{\text{err}}^{(k)} \bar{\mathbf{r}}_{\text{err}}^{(k)T} \right\} = \bar{\sigma}^{(k)2} \mathbf{I}$ and (59) results in:

$$\Delta_{(n_T f + n, n_T g + m)} \approx \sum_{k=1}^K \sum_{x=1}^{n_R^{(k)}} \sum_{r=1}^{n_R^{(k)}} \bar{\sigma}^{(k)2} \bar{d}_{\binom{(k)}{(n_R^{(k)} f + x, n_R^{(k)} g + r)}} \left(\mathbf{a}_{g(r,m)}^{(k)} + j\mathbf{b}_{g(r,m)}^{(k)} \right) \left(\mathbf{a}_{f(x,n)}^{(k)T} - j\mathbf{b}_{f(x,n)}^{(k)T} \right). \quad (60)$$

REFERENCES

- [1] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, May 2000.
- [2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
- [3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
- [4] M. Costa, "Writing on dirty paper," *IEEE Trans. on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [5] J. Lee and N. Jindal, "Dirty paper coding vs. linear precoding for MIMO broadcast channels," in *Fortieth Asilomar Conference on Signals, Systems and Computers*, Oct. 2006, pp. 779–783.
- [6] H. Bölcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC) 2000*, vol. 1, Sep. 2000, pp. 1–6.

- [7] B. Lu and X. Wang, "Space-time code design in OFDM systems," in *Proc. IEEE Global Communications Conference*, vol. 2, 2000, pp. 1000–1004.
- [8] M. Payaró and D. P. Palomar, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution," in *Proc. IEEE International Symposium on Information Theory (ISIT'09)*, Jul. 2009, pp. 1085 – 1089.
- [9] D. Sacristán-Murga and A. Pascual-Iserte, "Differential feedback of MIMO channel Gram matrices based on geodesic curves," *IEEE Trans. on Wireless Communications*, vol. 9, no. 12, pp. 3714–3727, Dec. 2010.
- [10] J. Wang and D. P. Palomar, "Robust MMSE precoding in MIMO channels with pre-fixed receivers," *IEEE Trans. on Signal Processing*, vol. 58, no. 11, pp. 5802–5818, Nov. 2010.
- [11] H. Sung, S. Lee, and I. Lee, "Generalized channel inversion methods for multiuser MIMO systems," *IEEE Trans. on Communications*, vol. 57, no. 11, pp. 3489–3499, Nov. 2009.
- [12] A. Tenenbaum and R. S. Adve, "Joint multiuser transmit-receive optimization using linear processing," in *Proc. IEEE Intl. Conf. on Communications*, Jul. 2004, pp. 588–592.
- [13] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: duality and sum-MSE minimization," *IEEE Trans. on Signal Processing*, vol. 55, no. 11, pp. 5436–5446, Nov. 2007.
- [14] N. Vučić, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser MIMO systems," *IEEE Trans. on Signal Processing*, vol. 57, no. 9, pp. 3576–3587, Sep. 2009.
- [15] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. on Information Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.
- [16] B. K. Chalise, S. Shahbazpanahi, A. Czylik, and A. B. Gershman, "Robust downlink beamforming based on outage probability specifications," *IEEE Trans. on Wireless Communications*, vol. 6, no. 10, pp. 3498–3503, Oct. 2007.
- [17] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.
- [18] C.-B. Chae, D. Mazzarese, N. Jindal, and R. W. Heath, "Coordinated beamforming with limited feedback in the MIMO broadcast channel," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1505–1515, Oct. 2008.
- [19] D. Sacristán-Murga and A. Pascual-Iserte, "Differential feedback of MIMO channel correlation matrices based on geodesic curves," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2009, pp. 2553–2556.
- [20] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. on Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
- [21] D. Sacristán-Murga and A. Pascual-Iserte, "Differential feedback of channel Gram matrices for block diagonalized multiuser MIMO systems," in *Proc. IEEE Intl. Conf. on Communications*, May 2010, pp. 1–5.
- [22] R. A. Horn and C. R. Johnson, Eds., *Matrix Analysis*. Cambridge University Press, 1985.
- [23] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 3rd ed. Wiley, 2002.
- [24] M. Payaró and D. P. Palomar, "Hessian and concavity of mutual information, differential entropy, and entropy power in linear vector Gaussian channels," *IEEE Trans. on Information Theory*, vol. 55, no. 8, pp. 3613–3628, Aug. 2009.
- [25] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener filter for the downlink of TDD-CDMA systems," in *Proc. IEEE International Symposium on Spread-Spectrum Tech. & Appl.*, Sep. 2002.
- [26] R. L. Choi and R. D. Murch, "Transmit MMSE pre-rake pre-processing with simplified receivers for the downlink of MISO TDD-CDMA systems," in *Proc. IEEE Global Communications Conference*, Nov. 2002, pp. 429–433.
- [27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [28] *IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems*, IEEE Std. 802.16, 2005.

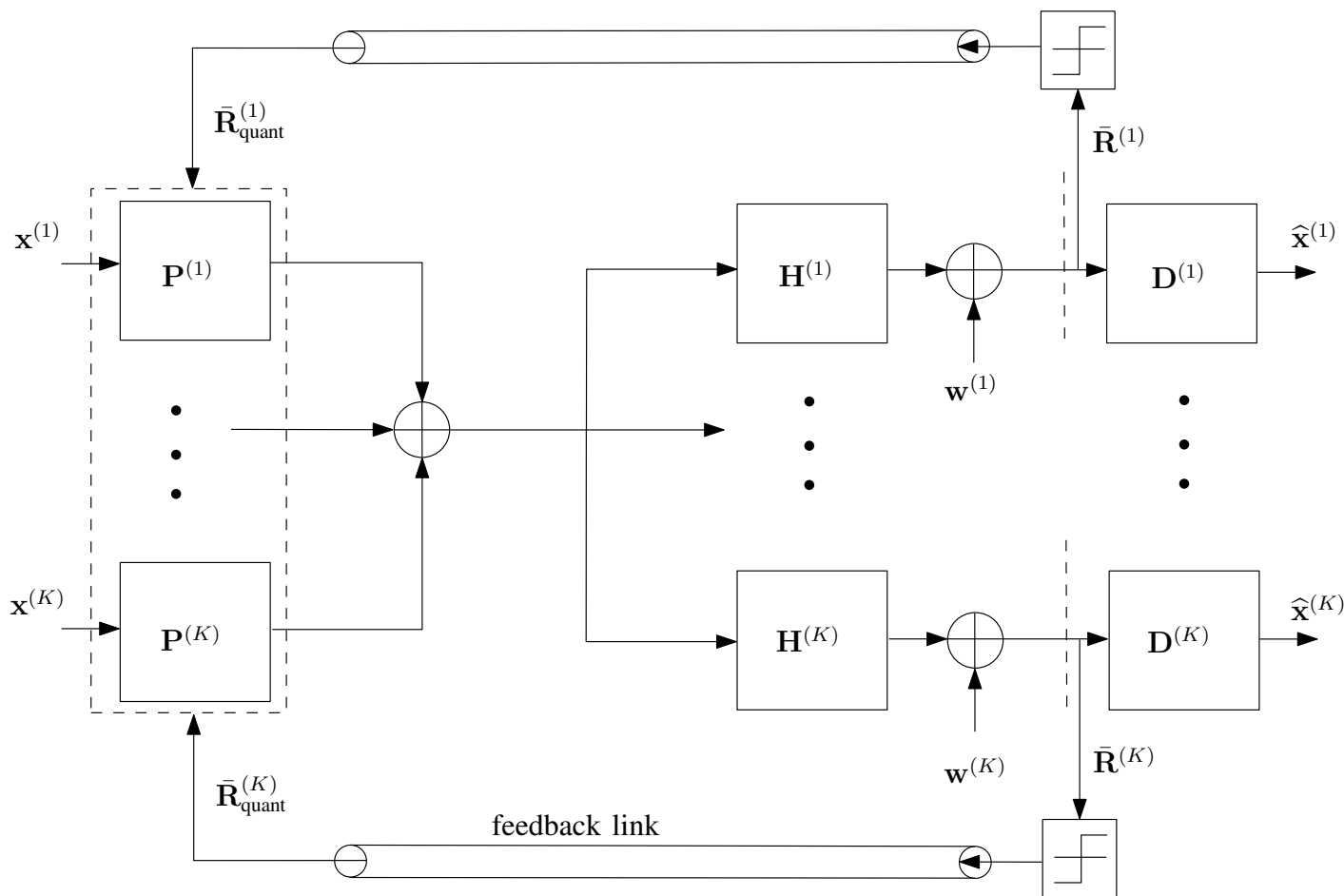


Fig. 1. MIMO-OFDM broadcast system model with feedback.

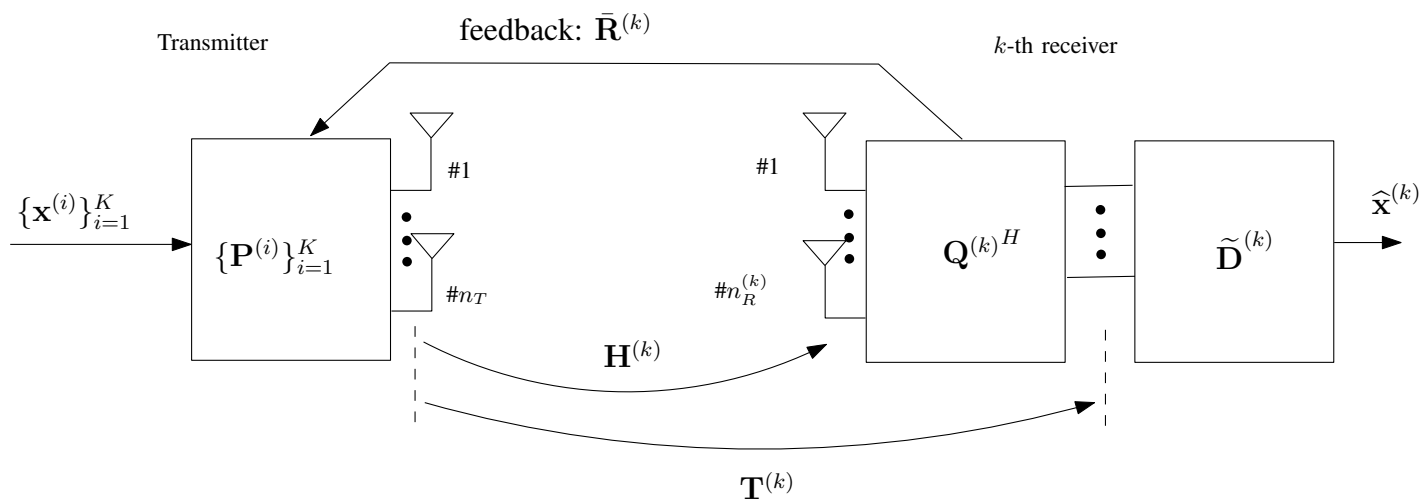


Fig. 2. Equivalent channel model.

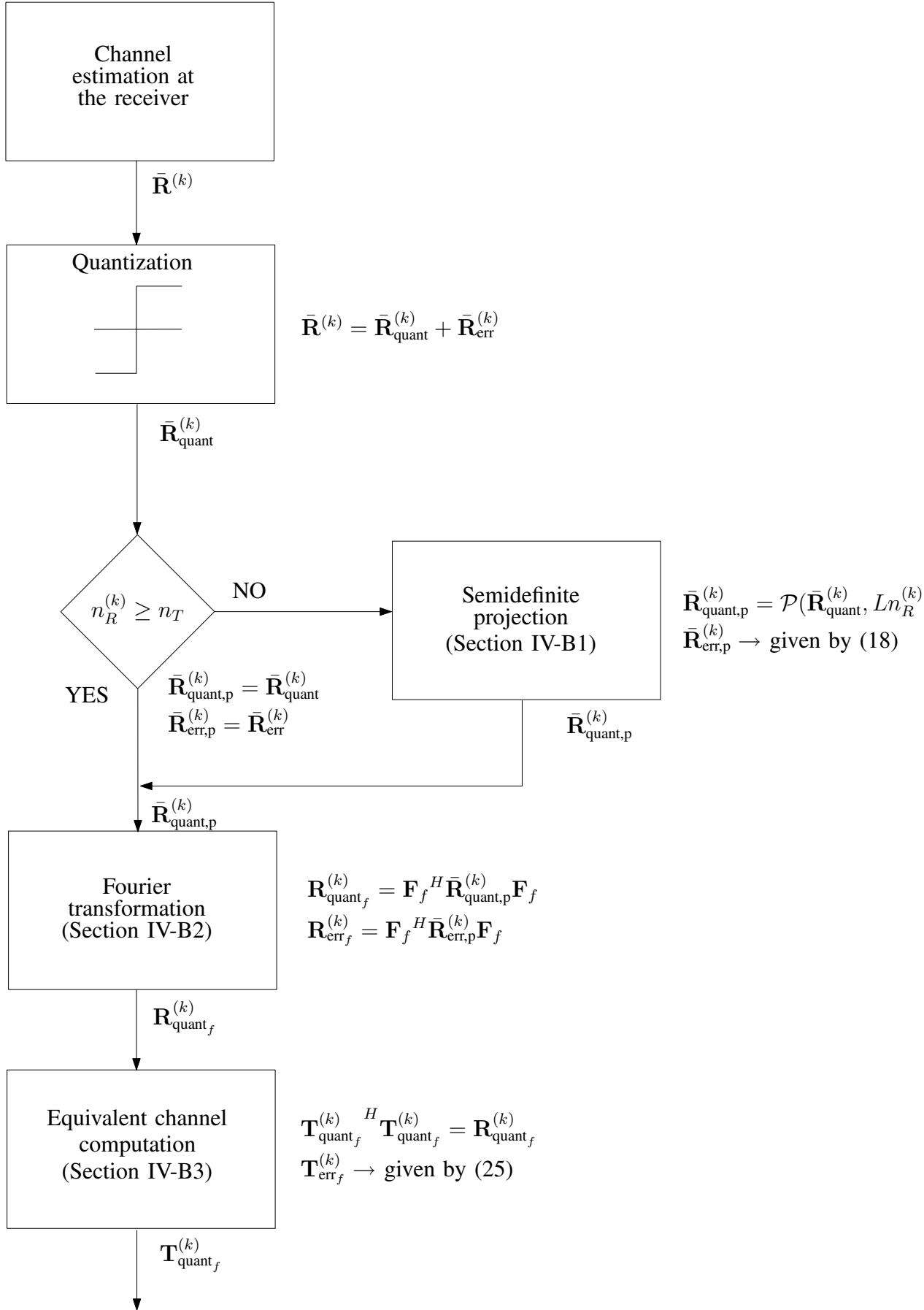


Fig. 3. Diagram of the complete CSI processing and the error propagation through the stages of such processing.

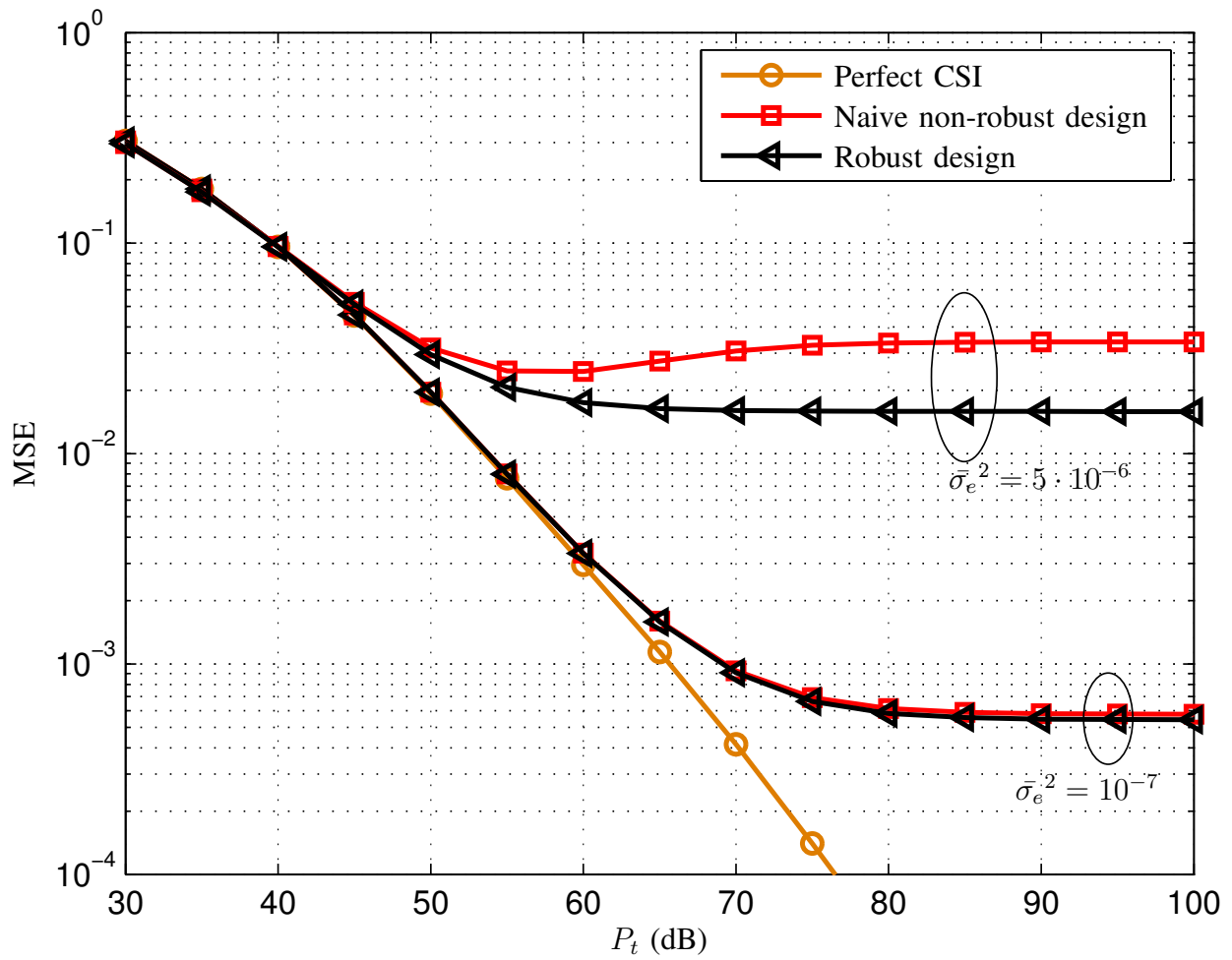


Fig. 4. MSE versus the total transmission power allocated among all the 128 carriers in a $4 \times \{2,2\}$ system with $\bar{\mathbf{R}}_{\text{err}}^{(k)}$ generated using i.i.d. elements following a Gaussian distribution with zero mean and variance $\bar{\sigma}_e^2$.

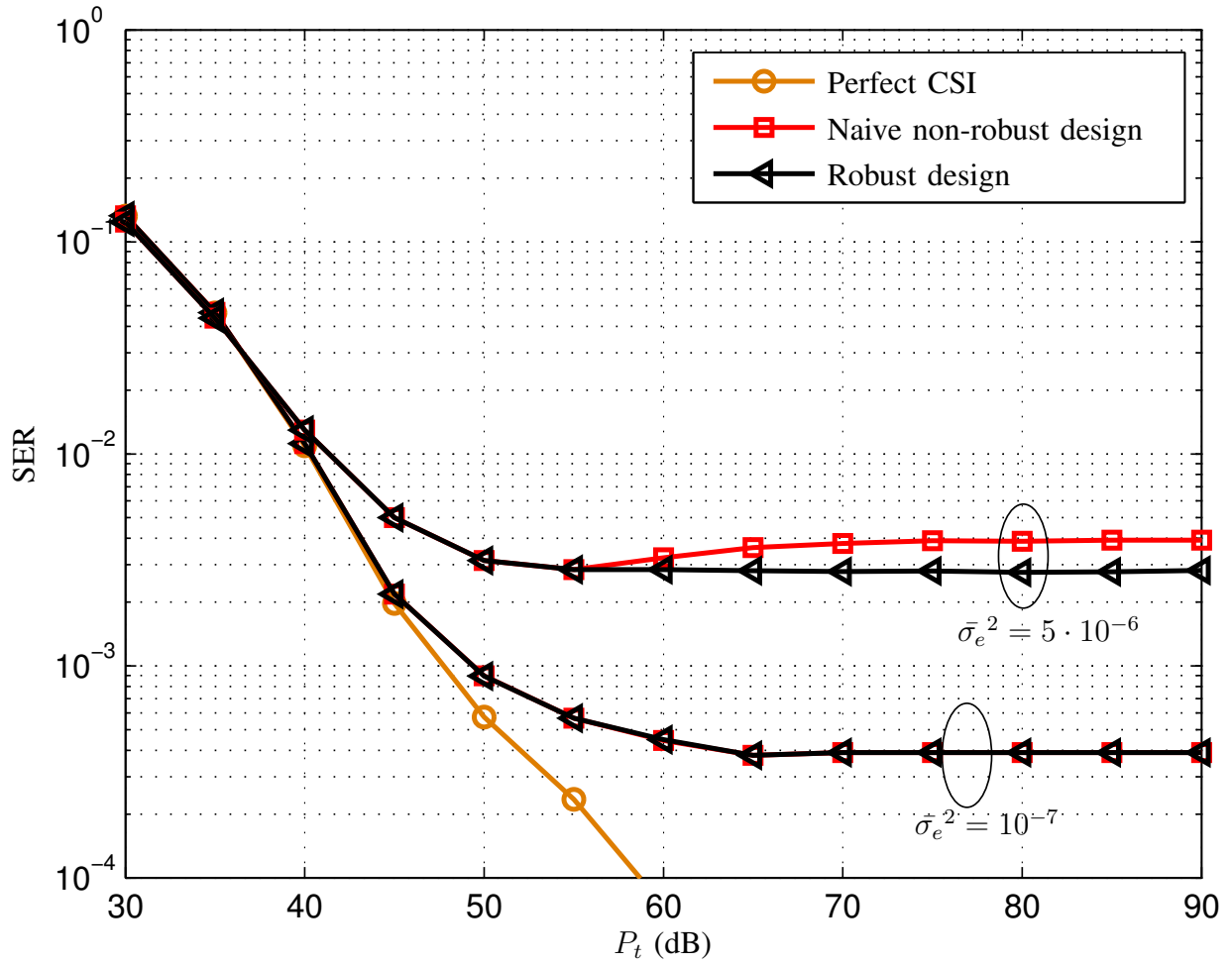


Fig. 5. SER versus the total transmission power allocated among all the 128 carriers in a $4 \times \{2,2\}$ system with QPSK constellation and with 512 symbols transmitted simultaneously. $\bar{\mathbf{R}}_{\text{err}}^{(k)}$ consists of i.i.d. elements following a Gaussian distribution with zero mean and variance $\bar{\sigma}_e^2$.

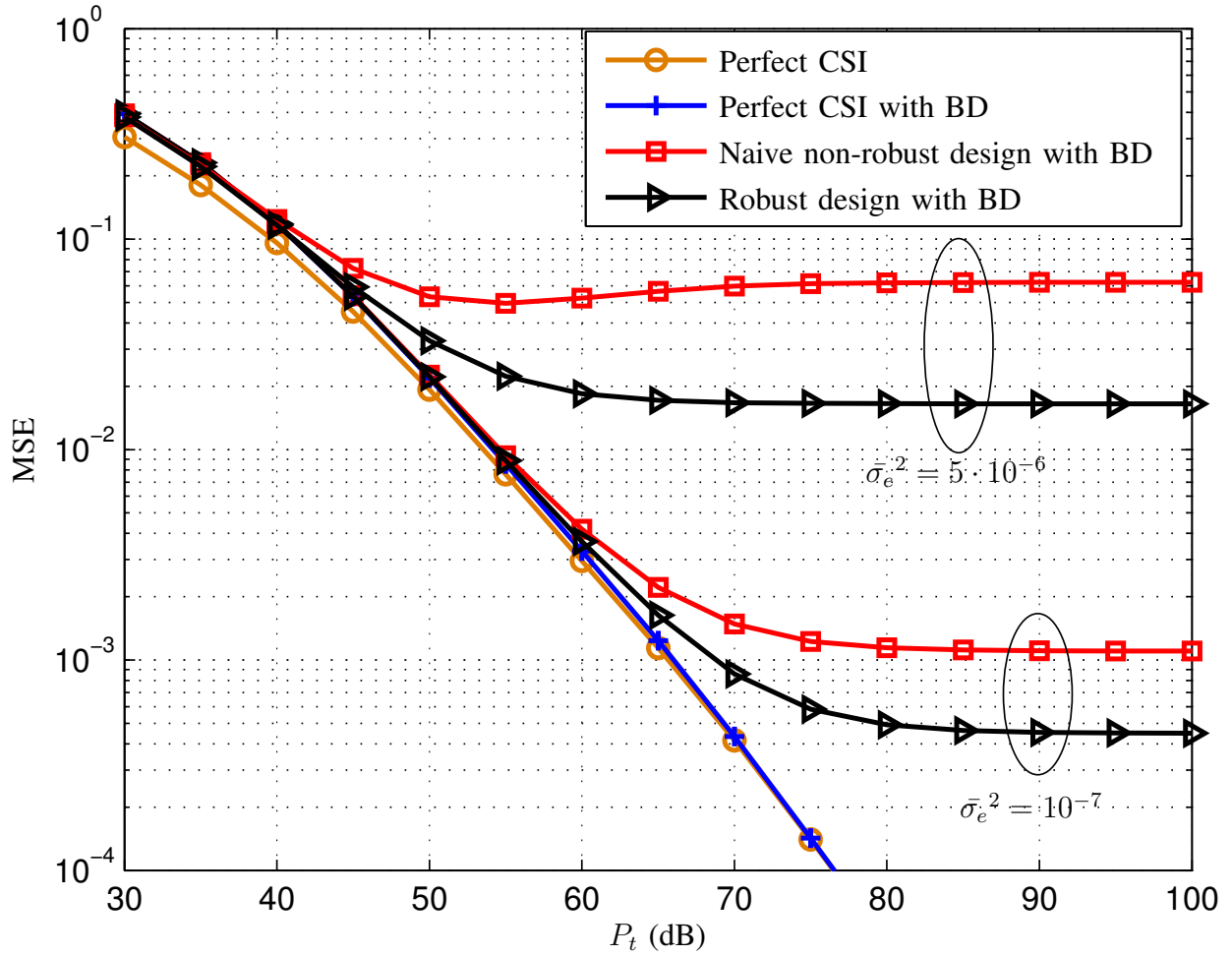


Fig. 6. MSE versus the total transmission power allocated among all the 128 carriers in a $4 \times \{2,2\}$ system with $\bar{\mathbf{R}}_{\text{err}}^{(k)}$ generated using i.i.d. elements following a Gaussian distribution with zero mean and variance σ_e^2 , and an implementation of a BD design.

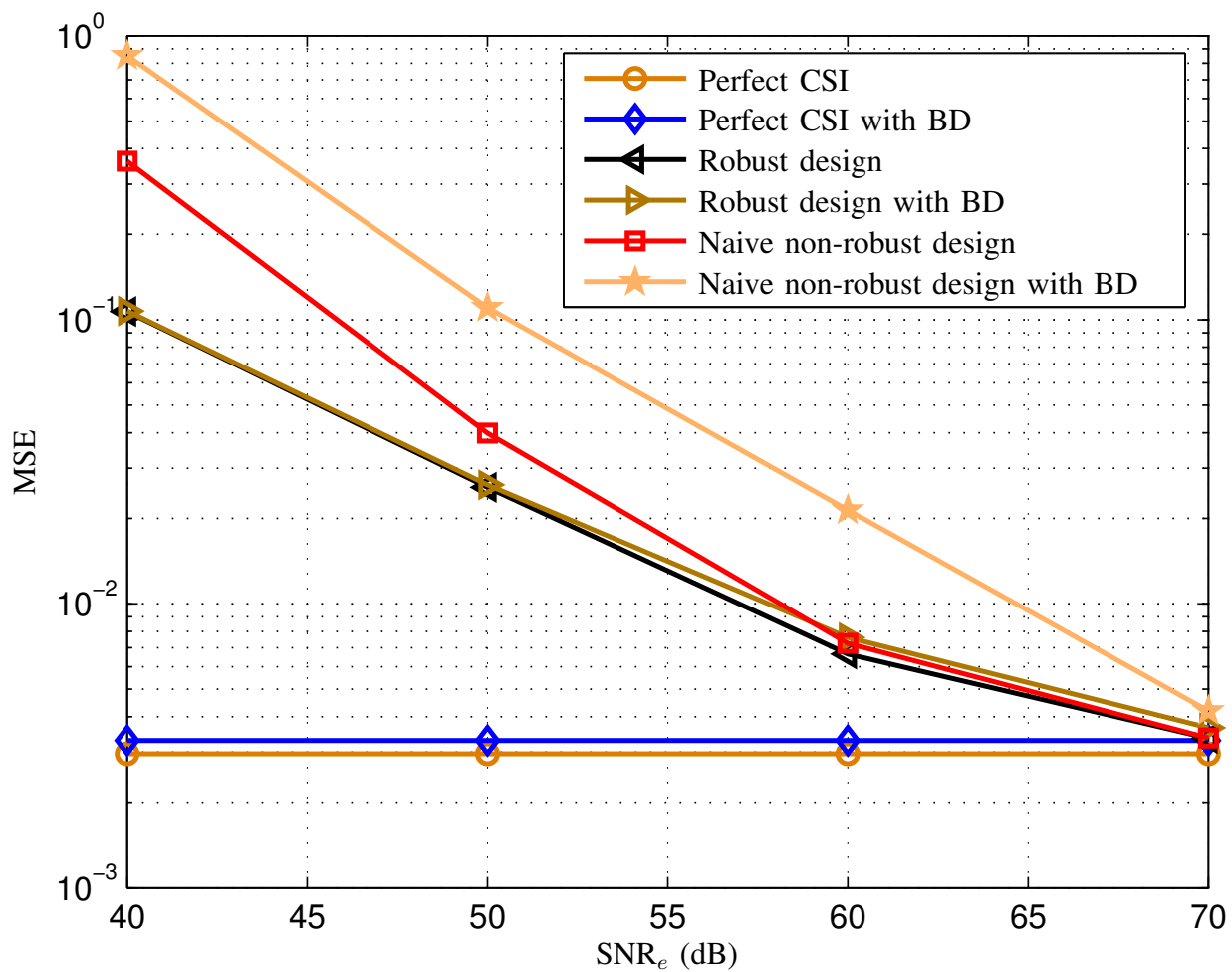


Fig. 7. MSE versus SNR_e in a 4x{2,2} system with a total transmission power allocated among all the 128 carriers of P_t = 60 dB above the noise level.

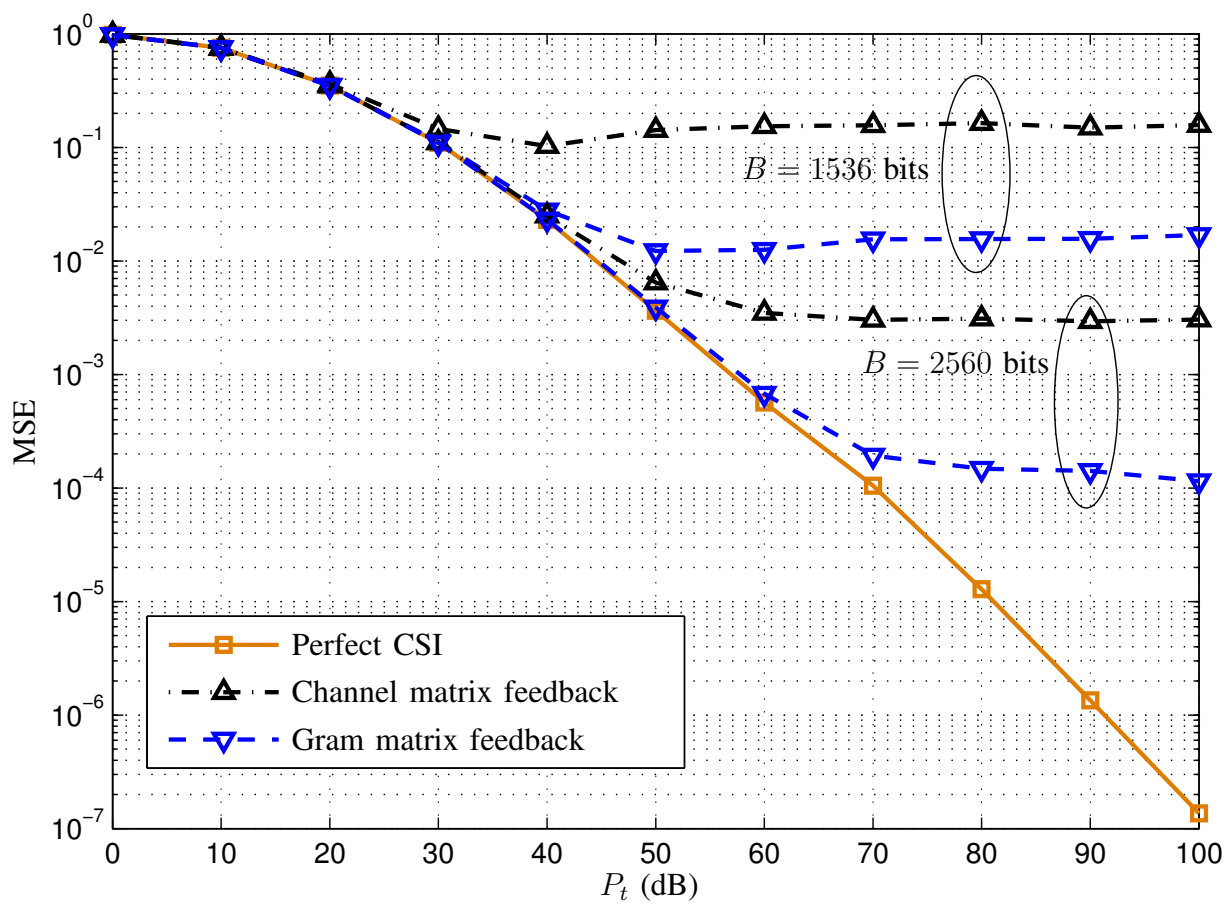


Fig. 8. Feedback based on the channel Gram matrix versus feedback based on the complete channel response matrix, for different values of the feedback overhead in number of bits.

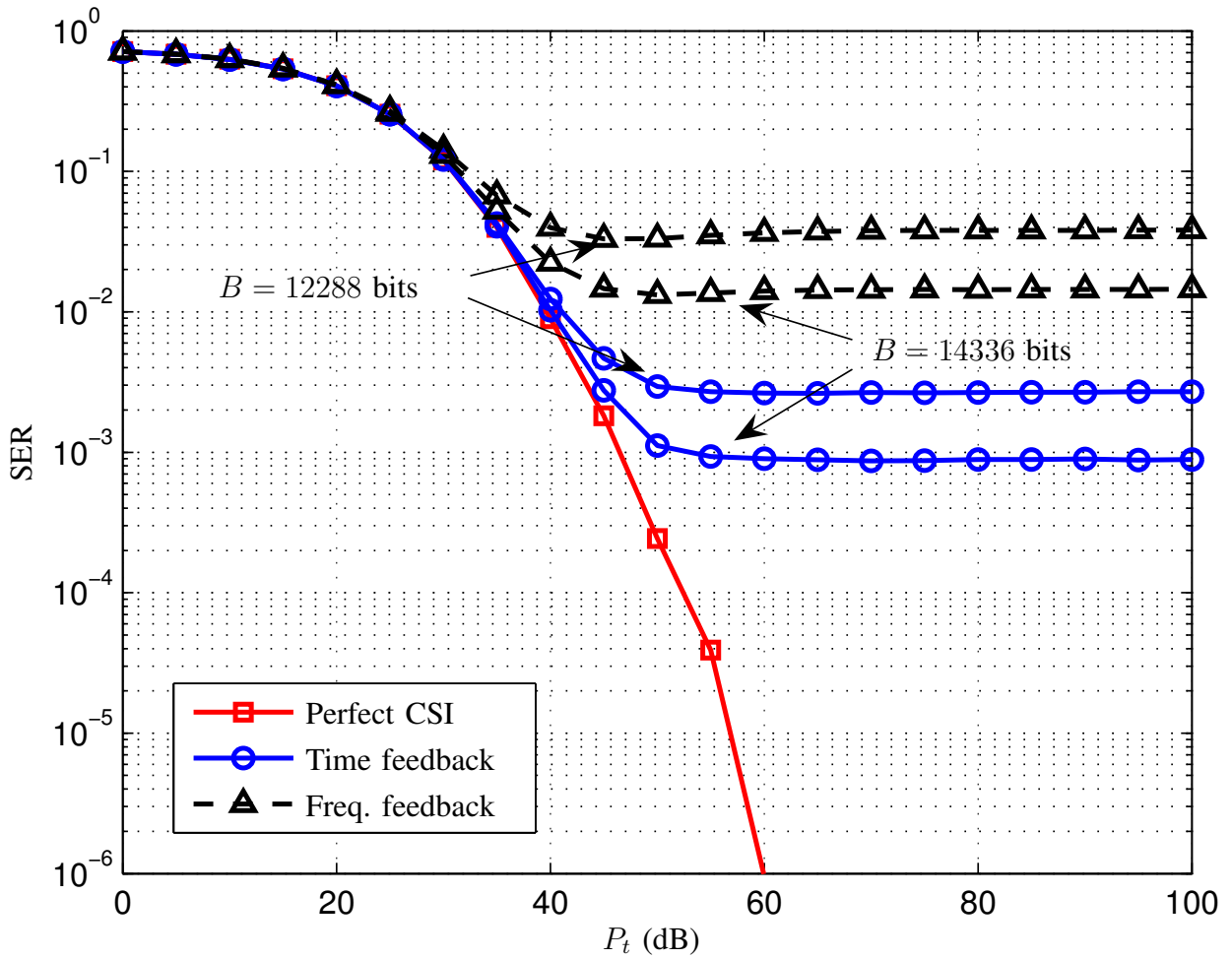


Fig. 9. Feedback of the frequency domain CSI versus feedback of the time domain CSI. Scenario with $L = 8$ taps, $F = 128$ carriers, and different values of the feedback overhead in number of bits. The transmit power P_t is spread over all 128 carriers and all 4 antennas, and 512 QPSK symbols are transmitted simultaneously.

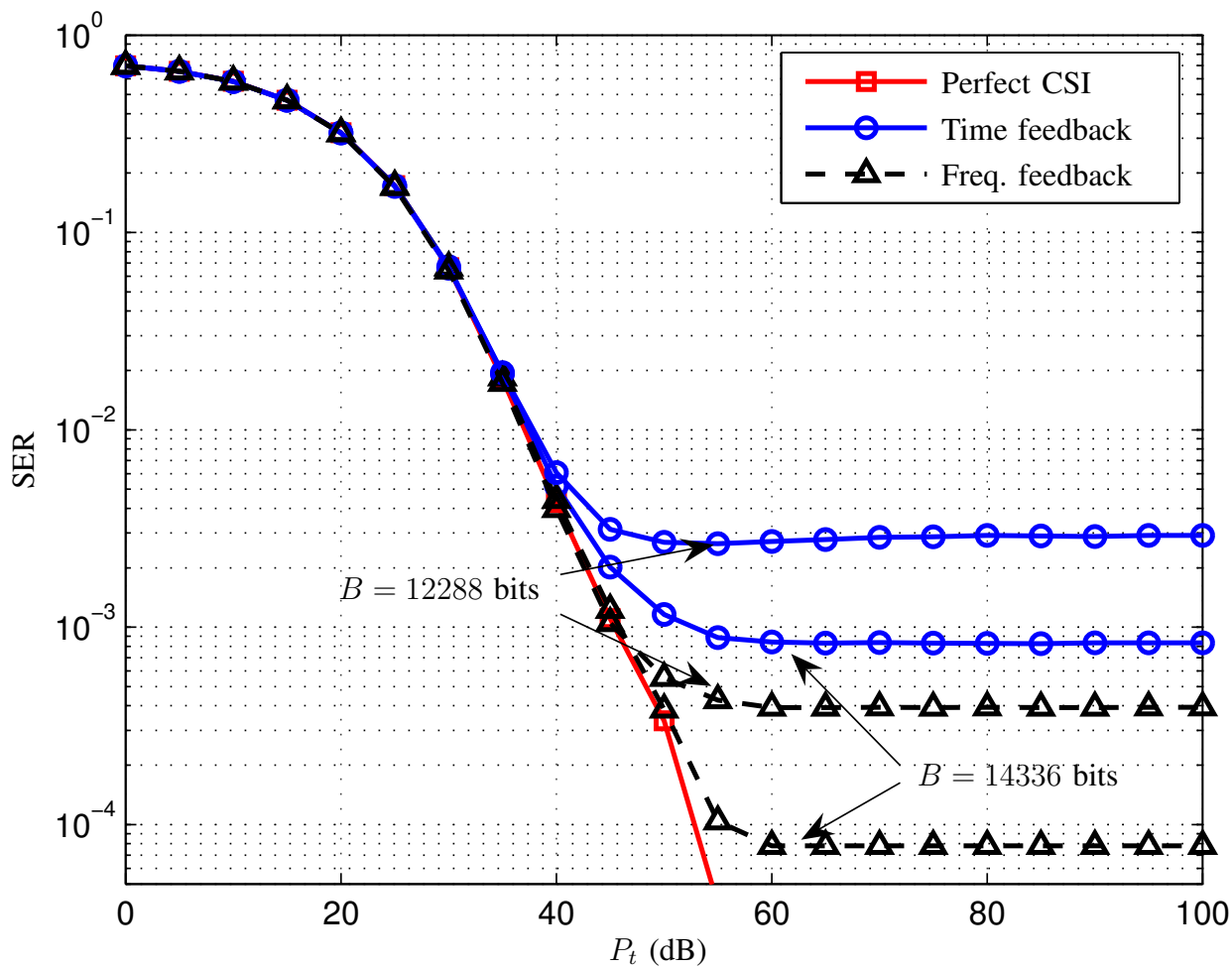


Fig. 10. Feedback of the frequency domain CSI versus feedback of the time domain CSI. Scenario with $L = 8$ taps, $F = 64$ carriers, and different values of the feedback overhead in number of bits. The transmit power P_t is spread over all 64 carriers and all 4 antennas, and 256 QPSK symbols transmitted simultaneously.