

Robust Power Allocation Algorithms for MIMO OFDM Systems With Imperfect CSI

Francesc Rey, *Student Member, IEEE*, Meritxell Lamarca, *Member, IEEE*, and Gregori Vazquez, *Senior Member, IEEE*

Abstract—This paper presents a Bayesian approach to the design of transmit prefiltering matrices in closed-loop schemes robust to channel estimation errors. The algorithms are derived for a multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) system. Two different optimization criteria are analyzed: the minimization of the mean square error and the minimization of the bit error rate. In both cases, the transmitter design is based on the singular value decomposition (SVD) of the conditional mean of the channel response, given the channel estimate. The performance of the proposed algorithms is analyzed, and their relationship with existing algorithms is indicated. As with other previously proposed solutions, the minimum bit error rate algorithm converges to the open-loop transmission scheme for very poor CSI estimates.

Index Terms—Channel state information, diversity, fading channels, feedback, joint transceiver optimization, linear precoding, MIMO channels, OFDM, power allocations, wireless communications.

I. INTRODUCTION

IT is well known that the knowledge of the channel response at the transmitter can be used to improve the performance of digital communications systems, both in the case of single antenna [1] or multiple antenna transmission [2]. This has led to the design of closed-loop schemes where the receiver predicts and estimates the channel response and feeds back this information to the transmitter (e.g., [3]) so that one can adapt the modulation and the channel code of the transmitted signal according to the channel response. In this context, the application of orthogonal frequency division multiplexing (OFDM) modulation and the design of the transmitter based on the singular value decomposition (SVD) of the channel is particularly convenient since it can allocate power or bits on a subcarrier basis.

The problem of power allocation for single and multiple antenna transmission in fading channels has been widely studied in the literature under a variety of criteria, including zero-forcing (ZF) and minimum mean square error (MMSE) schemes [4], [5], maximum information rate [6], and minimum bit error rate (BER) [7], [8], providing, in all cases, a design based on the SVD of the channel estimate. However, all these algorithms, assume that the channel is perfectly known at the transmitter. This

hypothesis does not hold in real systems since neither the transmitter nor the receiver have access to ideal channel state information (CSI). The errors in the channel estimates can be originated from several sources, including the estimation variance due to noise, the time lag between channel estimation/prediction and its use for transmitter design, and the quantization error in the feedback channel. While the first one is common to the transmitter and the receiver, the second and third ones only appear in the transmitter. Thus, depending on the pace of channel variation and depending on the dimensioning of the feedback channel, the channel uncertainty at the transmitter can be larger than that one at the receiver, but in any case, none of them will have perfect CSI.

If the errors in the CSI are not taken into account in the transmitter design, the performance of the closed-loop algorithm will degrade and, eventually, may get worse than that one of open-loop transmission. Thus, the potential of linear precoding can only be fulfilled when the reliability of the channel estimates is considered in the cost function [9]. The design of algorithms that take into account partial knowledge of the CSI is still an open question. The previous contributions to this topic can be grouped into two categories: a first group that considers imperfect estimates of the channel impulse response to design the transmitter [7], [10] and a second group that designs optimal transmitter schemes based only on the knowledge of the channel statistics [7], [11]. The major contribution in [7] was the design of wideband signaling schemes under a minimum BER criterion. It was shown that when perfect CSI is available at the transmitter, a beamformer focused to the most dominant channel mode achieves the optimal solution, whereas a set of linearly independent signature codes that transmit on all channel modes are required when CSI is poor. See [10] and [11] for the design of linear precoders for space-time coded systems.

This paper proposes a Bayesian approach to design the optimal linear transformation when the channel estimates at the transmitter are noisy. Unbalanced CSI quality between the transmitter and the receiver is assumed. We do this by introducing the idea of the *a priori* design as the design carried out at the transmitter and based on the predicted channel at this side and the *a posteriori* design as the design carried out at the receiver and based on the updated CSI from the observation of the channel output. Two algorithms are proposed to allocate the available power across all subcarriers and channel modes in a multiple-input multiple-output (MIMO) OFDM channel: The first one minimizes the mean square error (MSE), whereas the second one follows a minimum BER criterion. As will be shown in Sections II–V, the Bayesian formulation leads to a design based on

Manuscript received December 23, 2002; revised November 20, 2003. This work was supported in part by the Spanish Government (CICYT) under Grants TIC2003-05482 and TIC2004-04526 and the Catalan Government (DURSI) under Grant 2001SGR-00268. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sergios Theodoridis.

The authors are with the Department of Signal Theory and Communications, Technical University of Catalonia, UPC Campus Barcelona, Spain (e-mail: frey@gps.tsc.upc.edu; xell@gps.tsc.upc.edu; gregori@gps.tsc.upc.edu).

Digital Object Identifier 10.1109/TSP.2004.842199

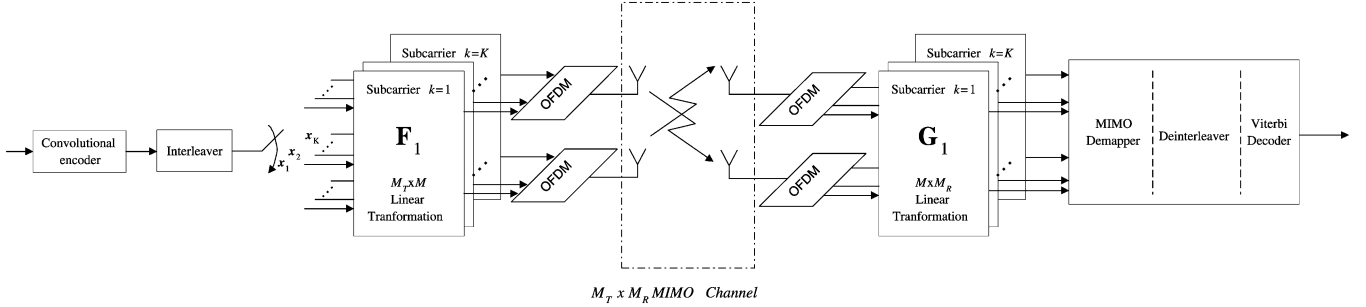


Fig. 1. Block diagram of the proposed MIMO OFDM system.

the SVD decomposition of the “equivalent channel,” that is, the result of averaging the channel estimate over the channel uncertainty. Both criteria lead to closed-form solutions whose complexity is similar to their nonrobust counterparts, although in the case of the BER optimization criterion, some constraints must be introduced. Besides, the algorithm for minimum BER can be regarded as a transmitter design that adapts automatically to the channel knowledge, moving from the open-loop scheme (the same power is allocated across all subcarriers and antennas) to the closed-loop scheme, according to channel uncertainty. This feature is in common with the results in [7], [10], and [11].

The proposed algorithms can accept different transmission rates; therefore, they can use the space diversity to improve channel reliability or to increase the transmission rate. The formulation is quite general, encompassing the single antenna transmission, the beamforming schemes, and the frequency flat-fading channels as particular cases.

The paper compares the proposed algorithms in terms of uncoded BER and coded BER for different degrees of CSI reliability. In order to focus the comparison on the different designs proposed for the transmitter, a maximum likelihood (ML) receiver has been always used. Note that the design of the pre-filtering matrices does not take into account the channel coding stage, even though, for the sake of completeness, the performance of the proposed algorithms has been compared in terms of coded BER for the particular channel code described in the HIPERLAN/2 standard.

The paper is organized as follows. First, the system architecture under analysis and the channel model are described in Section II. Afterwards, the two proposed algorithms are presented in Sections III and IV following the same structure: derivation of the cost function, closed-form solution, and asymptotic analysis of the solution for very high and very low uncertainty. Their performances are evaluated in Section V, and some conclusions are drawn in Section VI. Finally, some mathematical developments are detailed in Appendixes A–C.

The following notation is used in the paper. Superscripts $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, and $(\cdot)^\#$ stand for matrix transpose, matrix conjugate transpose, matrix conjugate, and matrix pseudoinverse, respectively. Uppercase and lowercase boldface denote matrices and vectors, respectively. The scalar $\mathbf{v}(i)$ denotes the i th element of vector \mathbf{v} . The symbols $\text{vec}\{\mathbf{A}\}$, \otimes , $|\mathbf{v}|$, $\text{Tr}\{\mathbf{A}\}$, and $\det(\mathbf{A})$ stand for the stacking vectorization operator, the Kronecker product, the Euclidean norm of a vector, the trace operator, and the determinant of a matrix, respectively.

II. PROBLEM STATEMENT

A. System Model

This section describes the signal model for an MIMO OFDM communications system over a frequency-selective fading channel, as shown in Fig. 1. The MIMO configuration consists of M_T transmit antennas and M_R receive antennas. M denotes the number of symbols to be transmitted per subcarrier that is bounded by $M \leq \min\{M_T, M_R\}$, and K is the total number of subcarriers.

Let \mathbf{x} be the $KM \times 1$ vector containing the KM symbols that are simultaneously transmitted in the M_T OFDM symbols (i.e., one OFDM symbol per antenna). The data are assumed to be i.i.d. symbols in a constellation $\{s_1, \dots, s_{\mathcal{N}}\}$ of cardinality \mathcal{N} , with zero mean and variance $E\{\mathbf{x}\mathbf{x}^H\} = \sigma_x^2 \mathbf{I}$. If the channel keeps invariant within one OFDM symbol and the cyclic prefix (whose length is appropriately chosen) is removed at the receiver, the system model can be written in the frequency domain as

$$\mathbf{r} = \mathbf{G}\mathbf{H}\mathbf{F}\mathbf{x} + \mathbf{G}\mathbf{n} \quad (1)$$

where \mathbf{H} is a $KM_R \times KM_T$ block diagonal channel matrix containing the frequency responses of the MIMO channels, \mathbf{F} is a $KM_T \times KM$ matrix that denotes the linear precoder matrix and allocates the power across the K subcarriers and M_T antennas, \mathbf{G} is a $KM \times KM_R$ matrix that combines the signal received at the M_R antennas, and \mathbf{n} is the $KM \times 1$ noise vector after the FFT, which has the same Gaussian statistic as its time-domain counterpart (i.e., zero mean and covariance $E\{\mathbf{n}\mathbf{n}^H\} = \sigma_n^2 \mathbf{I}$).

The frequency-selective MIMO channel can be decoupled into K MIMO frequency-flat channels by imposing a certain structure in the transmitter and receiver matrices. As shown in Fig. 1, the linear transformations \mathbf{F} and \mathbf{G} are split into K submatrices that process each subcarrier independently. Hence, \mathbf{G} , \mathbf{H} , \mathbf{F} matrices and \mathbf{x} , \mathbf{r} vectors involved in (1) can be structured as

$$\begin{aligned} \mathbf{H} &= \text{diag}\{[\mathbf{H}_1 \mathbf{H}_2 \dots \mathbf{H}_K]\} \\ \mathbf{F} &= \text{diag}\{[\mathbf{F}_1 \mathbf{F}_2 \dots \mathbf{F}_K]\} \\ \mathbf{G} &= \text{diag}\{[\mathbf{G}_1 \mathbf{G}_2 \dots \mathbf{G}_K]\} \\ \mathbf{x}^T &= [\mathbf{x}_1^T \dots \mathbf{x}_K^T] \\ \mathbf{r}^T &= [\mathbf{r}_1^T \dots \mathbf{r}_K^T] \end{aligned} \quad (2)$$

where \mathbf{F}_k , \mathbf{H}_k , and \mathbf{G}_k are, respectively, $M_T \times M$, $M_R \times M_T$, and $M \times M_R$ matrices processing the k th subcarrier and \mathbf{x}_k ,

\mathbf{r}_k are $M \times 1$ vectors containing the transmitted and received symbols by the k th subcarrier through the different antennas. Following the structure in (2) the input–output relation in (1) is rewritten as a set of equations that contains only one subcarrier each:

$$\mathbf{r}_k = \mathbf{G}_k \mathbf{H}_k \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{n}_k \quad k = 1 \dots K. \quad (3)$$

The architecture proposed in (3) is the basis for the study in this paper. Note that the block diagonal structure proposed for the linear transformations \mathbf{F} and \mathbf{G} in (2) has the advantage of providing a scheme where the spatial prefiltering matrix is applied individually to each subcarrier, reducing to the classic OFDM scheme when only one antenna is used at the transmitter. It is worth mentioning that the receiver filtering matrix \mathbf{G} has been included in the model for the sake of completeness and because it is required in the design of the minimum MSE algorithm, although the simulation results presented in the paper do not implement it because a ML detector was used to make a fair comparison of the proposed algorithms. This issue is reviewed in Section III when the cost function for minimum MSE is addressed.

The block diagram in Fig. 1 incorporates a convolutional encoder and an interleaver at the transmitter, and a Viterbi decoder at the receiver, that correspond to the channel coding stage. Although this stage has not been taken into account in the design of \mathbf{F}_k and \mathbf{G}_k matrices, it has been incorporated in the model to evaluate the performance of the proposed algorithms in terms of coded BER.

During subsequent developments, it will be useful to store the complete channel response in a vector. The MIMO channel response for the k th subcarrier and for the multicarrier system will be denoted \mathbf{h}_k and \mathbf{h} , respectively

$$\begin{aligned} \mathbf{h}_k &= \text{vec} \{ \mathbf{H}_k \} \\ \mathbf{h} &= [\mathbf{h}_1^T \quad \dots \quad \mathbf{h}_K^T]^T. \end{aligned} \quad (4)$$

According to this new notation and making use of the identity [12]

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}) \quad (5)$$

the received vector for the k th subcarrier \mathbf{r}_k can be rewritten as

$$\mathbf{r}_k = \mathbf{A}_k \mathbf{h}_k + \mathbf{G}_k \mathbf{n}_k \quad k = 1 \dots K \quad (6)$$

where the matrix \mathbf{A}_k is defined as

$$\mathbf{A}_k = (\mathbf{F}_k \mathbf{x}_k)^T \otimes \mathbf{G}_k. \quad (7)$$

B. Channel Model

This section introduces the model for the channel response and channel estimates used to design and simulate the power allocation algorithms. The propagation channel is modeled as a fading channel with uncorrelated coefficients for all taps in the impulse response and identical power delay profile for all subchannels. This model encompasses the Rayleigh fading and

Ricean fading channels as particular cases. For the sake of completeness, spatial correlation is considered in the channel model and is also introduced into the formulation of the algorithms, although the uncorrelation assumption will be required to derive closed-form solutions for the proposed algorithms. The channel vector \mathbf{h} is modeled as a multivariate Gaussian process with mean and covariance

$$\begin{aligned} E \{ \mathbf{h} \} &= \mathbf{m}_h \\ E \{ (\mathbf{h} - \mathbf{m}_h) (\mathbf{h} - \mathbf{m}_h)^H \} &= \mathbf{C}_{hh} = \mathbf{P} \otimes \mathbf{R} \end{aligned} \quad (8)$$

where \mathbf{R} is the antenna correlation matrix, and \mathbf{P} is the circulant matrix built as an Hermitian Toeplitz matrix whose first row is the DFT of the variance in the channel impulse response coefficients, which is assumed to be known.

The design of the linear precoder in the presence of channel estimation errors requires the definition of a model for the estimated channel $\hat{\mathbf{h}}$. It will be assumed that the channel estimation error $\boldsymbol{\varepsilon}$ can be modeled as an additive term

$$\hat{\mathbf{h}} = \sqrt{\frac{\sigma_h^2}{\sigma_h^2 + \sigma_\varepsilon^2}} (\mathbf{h} + \boldsymbol{\varepsilon}) \quad (9)$$

where $\sigma_h^2 = E \{ \mathbf{h}^H \mathbf{h} \}$, and $\sigma_\varepsilon^2 = E \{ \boldsymbol{\varepsilon}^H \boldsymbol{\varepsilon} \}$.

This model includes CSI errors both at the transmitter and at the receiver. The transmitter may update CSI based on a feedback link and the prediction of the future channel state from previous CSI when channel is time-varying [3]; thus, the feedback delay and the prediction error are the most important causes of imperfect channel estimates at this side. At the receiver side, the degradation of channel estimate quality is mainly caused by the additive noise. These sources of CSI degradation can be collected in the channel estimation error $\boldsymbol{\varepsilon}$, which is modeled as a zero mean Gaussian process independent of the true channel \mathbf{h} and with covariance matrix

$$E \{ \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^H \} = \mathbf{E} \otimes \mathbf{I} \quad (10)$$

where matrix \mathbf{E} , with the same structure as matrix \mathbf{P} in (8), contains the DFT of the variance of the channel estimation error for each tap, and it is also assumed to be known.

The Bayesian approach adopted in this paper relies on the hypothesis that the channel \mathbf{h} and its estimate $\hat{\mathbf{h}}$ are jointly Gaussian. Under this assumption, the statistics of \mathbf{h} given $\hat{\mathbf{h}}$ are also Gaussian with mean and covariance given by [13, p. 324]:

$$\begin{aligned} \mathbf{m}_{\mathbf{h}|\hat{\mathbf{h}}} &= \mathbf{m}_h + \mathbf{C}_{h\hat{h}} \mathbf{C}_{\hat{h}\hat{h}}^{-1} (\hat{\mathbf{h}} - \mathbf{m}_{\hat{h}}) \\ \mathbf{C}_{\mathbf{h}|\hat{\mathbf{h}}} &= \mathbf{C}_{hh} - \mathbf{C}_{h\hat{h}} \mathbf{C}_{\hat{h}\hat{h}}^{-1} \mathbf{C}_{\hat{h}h}. \end{aligned} \quad (11)$$

In our case, following (8)–(10), the conditional mean becomes

$$\begin{aligned} \mathbf{m}_{\mathbf{h}|\hat{\mathbf{h}}} &= \sqrt{1 + \frac{\sigma_\varepsilon^2}{\sigma_h^2}} \left((\mathbf{P} \otimes \mathbf{R}) (\mathbf{P} \otimes \mathbf{R} + \mathbf{E} \otimes \mathbf{I})^{-1} \right) \hat{\mathbf{h}} \\ &\quad + (\mathbf{E} \otimes \mathbf{I}) (\mathbf{P} \otimes \mathbf{R} + \mathbf{E} \otimes \mathbf{I})^{-1} \mathbf{m}_h \end{aligned} \quad (12)$$

where $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$ [12] has been used. In the same way, using twice the same identity, the covariance

matrix $\mathbf{C}_{\mathbf{h}|\hat{\mathbf{h}}}$ that gives a measure of the degree of the channel uncertainty can be expanded as

$$\mathbf{C}_{\mathbf{h}|\hat{\mathbf{h}}} = (\mathbf{P} \otimes \mathbf{R})(\mathbf{P} \otimes \mathbf{R} + \mathbf{E} \otimes \mathbf{I})^{-1}(\mathbf{E} \otimes \mathbf{I}). \quad (13)$$

In the particular case of uncorrelated antennas (i.e., $\mathbf{R} = \mathbf{I}$), the conditional mean and conditional covariance expressions reduce to

$$\begin{aligned} \mathbf{m}_{\mathbf{h}|\hat{\mathbf{h}}} &= \sqrt{1 + \frac{\sigma_\epsilon^2}{\sigma_h^2}} \left(\mathbf{P}(\mathbf{P} + \mathbf{E})^{-1} \otimes \mathbf{I} \right) \hat{\mathbf{h}} \\ &\quad + \left(\mathbf{E}(\mathbf{P} + \mathbf{E})^{-1} \otimes \mathbf{I} \right) \mathbf{m}_{\mathbf{h}} \end{aligned} \quad (14)$$

$$\mathbf{C}_{\mathbf{h}|\hat{\mathbf{h}}} = \left(\mathbf{P}(\mathbf{P} + \mathbf{E})^{-1} \mathbf{E} \right) \otimes \mathbf{I}. \quad (15)$$

Sections II-C and D will focus on the conditional covariance for a specific subcarrier \mathbf{h}_k under the hypothesis of uncorrelated antennas. In that case, matrix $\mathbf{C}_{\mathbf{h}_k|\hat{\mathbf{h}}}$ is given by

$$\mathbf{C}_{\mathbf{h}_k|\hat{\mathbf{h}}} = \left[\mathbf{P}(\mathbf{P} + \mathbf{E})^{-1} \mathbf{E} \right]_{[k,k]} \mathbf{I} \quad (16)$$

where subindex $[\cdot]_{[k,k]}$ refers to the (k, k) th element. Recalling the circulant structure of the \mathbf{P} and \mathbf{E} matrices, all the elements in the diagonal of $\mathbf{P}(\mathbf{P} + \mathbf{E})^{-1} \mathbf{E}$ are equal, and thus, the conditional covariance is the same for all subcarriers:

$$\mathbf{C}_{\mathbf{h}_k|\hat{\mathbf{h}}} = \omega \mathbf{I} \quad (17)$$

where ω is a scalar whose value depends on \mathbf{P} and \mathbf{E} . The asymptotic values when channel is perfectly known and when it is completely unknown are

$$\omega = \begin{cases} 0, & \text{Perfect CSI at the Transmitter} \\ \sigma_h^2, & \text{No CSI at the Transmitter.} \end{cases} \quad (18)$$

C. Equivalent Channel

A detailed analysis of (14) shows that the conditional mean can be understood as an equivalent channel that exploits the correlation between subcarriers and the channel uncertainty structure to mitigate the degradation due to CSI errors. We will denote this equivalent channel as $\hat{\mathbf{h}}_k^{\text{eq}} = \mathbf{m}_{\mathbf{h}_k|\hat{\mathbf{h}}}$. Assuming uncorrelated antennas, the equivalent channel for a particular subcarrier can be expressed as a linear combination of the mean and channel estimates for all subcarriers:

$$\hat{\mathbf{h}}_k^{\text{eq}} = \sum_{j=1}^K \hat{\beta}_k(j) \hat{\mathbf{h}}_j + \bar{\beta}_k(j) \mathbf{m}_{\mathbf{h}_j} \quad (19)$$

where $\hat{\beta}_k(j)$ and $\bar{\beta}_k(j)$ are the j th elements of vectors $\hat{\boldsymbol{\beta}}_k$ and $\bar{\boldsymbol{\beta}}_k$, respectively

$$\begin{aligned} \hat{\beta}_k^T &= \sqrt{1 + \frac{\sigma_\epsilon^2}{\sigma_h^2}} \mathbf{p}_k (\mathbf{P} + \mathbf{E})^{-1} \\ \bar{\beta}_k^T &= \mathbf{e}_k (\mathbf{P} + \mathbf{E})^{-1} \end{aligned} \quad (20)$$

and vectors \mathbf{p}_k and \mathbf{e}_k are the k th rows of \mathbf{P} and \mathbf{E} matrices, respectively. The weighting vectors $\hat{\boldsymbol{\beta}}_k$ and $\bar{\boldsymbol{\beta}}_k$ exploit the subcarrier correlation structure defined by the power delay profile to mitigate the mismatch between the real and the estimated channels. Note that when perfect channel knowledge is available (i.e., matrix $\mathbf{E} = \mathbf{0}$), the coefficients $\bar{\beta}_k(j)$ are all zero, and $\hat{\beta}_k(j) = \delta(k - j)$, leading to the equivalent channel: $\hat{\mathbf{h}}_k^{\text{eq}} = \hat{\mathbf{h}}_k$. In all other cases, the coefficients $\hat{\beta}_k(j)$ and $\bar{\beta}_k(j)$ weight vectors $\hat{\mathbf{h}}_j$ and $\mathbf{m}_{\mathbf{h}_j}$ for all subcarriers to reduce the uncertainty in the specific k th subcarrier. The asymptotic values when channel is perfectly known and when it is completely unknown are

$$\hat{\mathbf{h}}_k^{\text{eq}} = \begin{cases} \hat{\mathbf{h}}_k = \mathbf{h}_k, & \text{Perfect CSI at the Transmitter} \\ 0, & \text{No CSI at the Transmitter.} \end{cases} \quad (21)$$

Hereafter, we will refer to the equivalent channel in any of its two forms: $\hat{\mathbf{h}}_k^{\text{eq}} = \text{vec} \left\{ \hat{\mathbf{H}}_k^{\text{eq}} \right\}$, and its SVD will be denoted as

$$\hat{\mathbf{H}}_k^{\text{eq}} = \mathbf{U}_k \boldsymbol{\Lambda}_k \mathbf{V}_k^H. \quad (22)$$

D. Unbalanced CSI Between the Transmitter and the Receiver

Some algorithms, as will be the case of the minimum mean square error criterion, allow the design, in a single step, of the linear transformations \mathbf{F}_k and \mathbf{G}_k when the same imperfect CSI is assumed at both sides of the link. However, when the CSI quality is unbalanced between the transmitter and the receiver, there is no way to introduce, into the transmitter cost function, different levels of CSI. Alternatively, we divide the design in two steps. A first step is carried out at the transmitter, which we will call the *a priori* design, and whose aim is to design the transmitter matrix \mathbf{F}_k based on the prediction of the channel impulse response, aided from the receiver by means of the feedback channel link [3], before the signal is actually transmitted. A second step carried out at the receiver and called the *a posteriori* design, whose aim is to design the optimal receiver based on the channel knowledge at the receiver and the specific linear transformation \mathbf{F}_k once the transmitted signal has been received and the CSI has been updated.

III. MMSE DESIGN

This section aims at designing the linear transmitter \mathbf{F}_k that minimizes the mean square error (MSE) at the output of the linear receiver \mathbf{G}_k subject to an average power constraint across all subcarriers and antennas when channel estimates are noisy. The MMSE cost function requires a linear receiver \mathbf{G}_k that implicitly assumes that the same CSI quality is available at the transmitter and at the receiver. However, in most cases, the CSI quality is unbalanced, and thus, the linear transformations \mathbf{F}_k and \mathbf{G}_k cannot be designed in a single step. Hence, this section, focusing on the *a priori* design, introduces into the optimization problem a linear receiver \mathbf{G}_k that is never used at the receiver. The *a posteriori* design, which designs the optimal receiver based on the ML receiver, is described in Section V.

The optimization criterion is

$$\min_{\mathbf{F}_k, \mathbf{G}_k} \sum_{k=1}^K E \left\{ |\mathbf{r}_k - \mathbf{x}_k|^2 \right\} \text{ subject to } \sum_{k=1}^K \text{Tr} \{ \mathbf{F}_k \mathbf{F}_k^H \} = P_0 \quad (23)$$

where the expectation in (23) is computed over the additive Gaussian noise \mathbf{n}_k , the discrete-time sequence \mathbf{x}_k , and the true channel impulse response \mathbf{h}_k given the channel estimate at the transmitter $\hat{\mathbf{h}}$. For the sake of generality, it is first assumed that the transmitted symbols are known, which is a realistic assumption when training sequences are used, and the purpose is to improve the channel estimation [i.e., in data-aided (DA) schemes], and the result is later averaged over the vector \mathbf{x}_k to extend the solution to schemes where the transmitted sequences are not known at the receiver.

Under the previous premises, the design of matrices \mathbf{F}_k and \mathbf{G}_k given a transmitted sequence \mathbf{x}_k and a channel estimate $\hat{\mathbf{h}}_k$ is done by minimizing

$$\xi = \sum_{k=1}^K E_{\mathbf{h}_k | \hat{\mathbf{h}}} \left\{ E_{\mathbf{n}_k} \left\{ |\mathbf{r}_k - \mathbf{x}_k|^2 \right\} \right\} \quad (24)$$

subject to the average power constraint. Note that this minimization problem differs from those proposed in [4] and [14] on the conditional expectation $E_{\mathbf{h}_k | \hat{\mathbf{h}}}$, which mitigates the impact of the channel uncertainty on the algorithm performance, providing a robust design that adapts to the channel estimation quality.

A. Cost Function

This section expands the cost function in (24), making use of the conditional mean and conditional covariance given in (12) and (13). Assuming that the transmitted symbols are known, it can be shown that the robust cost function for k th subcarrier is

$$\zeta_k^{\text{Robust-DA}} = \left| \mathbf{A}_k \hat{\mathbf{h}}_k^{\text{eq}} - \mathbf{x}_k \right|^2 + \sigma_n^2 \text{Tr} \{ \mathbf{G}_k^H \mathbf{G}_k \} + \text{Tr} \left\{ \mathbf{A}_k \mathbf{C}_{\mathbf{h}_k | \hat{\mathbf{h}}} \mathbf{A}_k^H \right\}. \quad (25)$$

According to (4) and (7) and applying the identity (5), the contribution of the k th subcarrier to the cost function can be rewritten as

$$\zeta_k^{\text{Robust-DA}} = \left| \left(\mathbf{G}_k \hat{\mathbf{H}}_k^{\text{eq}} \mathbf{F}_k - \mathbf{I} \right) \mathbf{x}_k \right|^2 + \sigma_n^2 \text{Tr} \{ \mathbf{G}_k^H \mathbf{G}_k \} + \text{Tr} \left\{ \mathbf{A}_k \mathbf{C}_{\mathbf{h}_k | \hat{\mathbf{h}}} \mathbf{A}_k^H \right\} \quad (26)$$

and the optimization criterion is finally expressed as

$$\min_{\mathbf{F}_k, \mathbf{G}_k} \sum_{k=1}^K \zeta_k^{\text{Robust-DA}} \text{ subject to } \sum_{k=1}^K \text{Tr} \{ \mathbf{F}_k \mathbf{F}_k^H \} = P_0. \quad (27)$$

This cost function is only useful when sequence \mathbf{x}_k is known at the receiver; however, in most cases, the transmitted symbols are not known. Hence, when only the statistics of the transmitted symbols are known at the receiver, the expectation over the transmitted sequence allows the design of the optimal linear transformations \mathbf{F}_k and \mathbf{G}_k independently of the specific information symbols. Computing the expected value of (27) over the

transmitted data, the contribution to the robust cost function of the k th subcarrier becomes

$$\begin{aligned} \zeta_k^{\text{Robust}} &= \sigma_x^2 \text{Tr} \left\{ \left(\mathbf{G}_k \hat{\mathbf{H}}_k^{\text{eq}} \mathbf{F}_k - \mathbf{I} \right)^H \left(\mathbf{G}_k \hat{\mathbf{H}}_k^{\text{eq}} \mathbf{F}_k - \mathbf{I} \right) \right\} \\ &\quad + \sigma_n^2 \text{Tr} \{ \mathbf{G}_k^H \mathbf{G}_k \} \\ &\quad + \sigma_x^2 \text{Tr} \left\{ \mathbf{C}_{\mathbf{h}_k | \hat{\mathbf{h}}} \left(\mathbf{F}_k^* \mathbf{F}_k^T \otimes \mathbf{G}_k^H \mathbf{G}_k \right) \right\} \end{aligned} \quad (28)$$

and the whole cost function is finally expressed as

$$\min_{\mathbf{F}_k, \mathbf{G}_k} \sum_{k=1}^K \zeta_k^{\text{Robust}} \text{ subject to } \sum_{k=1}^K \text{Tr} \{ \mathbf{F}_k \mathbf{F}_k^H \} = P_0. \quad (29)$$

It is interesting to analyze the similarities between the derived expression and the cost function appearing in the nonrobust solution in [4] and [14], where the MSE for the k th subcarrier is written as

$$\begin{aligned} \zeta_k^{\text{Nonrobust}} &= \sigma_x^2 \text{Tr} \left\{ \left(\mathbf{G}_k \hat{\mathbf{H}}_k \mathbf{F}_k - \mathbf{I} \right)^H \left(\mathbf{G}_k \hat{\mathbf{H}}_k \mathbf{F}_k - \mathbf{I} \right) \right\} \\ &\quad + \sigma_n^2 \text{Tr} \{ \mathbf{G}_k^H \mathbf{G}_k \}. \end{aligned} \quad (30)$$

The nonrobust cost function, as shown in (30), contains two terms: The first one is a measure of the ISI, whereas the second one introduces into the cost function the contribution of the noise. Note that the residual ISI and the noise also appear in the robust cost function derived in (28), but there, the channel estimate $\hat{\mathbf{H}}_k$ has been replaced by the equivalent channel $\hat{\mathbf{H}}_k^{\text{eq}}$. Thus, the robust cost function exploits the correlation between subcarriers to mitigate the channel uncertainty, improving its performance. A second difference between the robust and the nonrobust solutions is the third term in (28) that does not appear in (30). As the covariance matrix $\mathbf{C}_{\mathbf{h}_k | \hat{\mathbf{h}}}$ is a measure of the channel estimation error, the third term introduces into the score function the mismatch between the real and the estimated channel due to the estimation errors.

B. Closed-Form Solution

A solution for the optimization of (28) can only be obtained by means of numerical techniques. Unfortunately, a closed-form solution is feasible only for uncorrelated antennas. Hence, in this section, we will assume that the channel is spatially uncorrelated. Substituting (17) into (28), the cost function simplifies to

$$\begin{aligned} \zeta_k^{\text{Robust}} &= \sigma_x^2 \text{Tr} \left\{ \left(\mathbf{G}_k \hat{\mathbf{H}}_k^{\text{eq}} \mathbf{F}_k - \mathbf{I} \right)^H \left(\mathbf{G}_k \hat{\mathbf{H}}_k^{\text{eq}} \mathbf{F}_k - \mathbf{I} \right) \right\} \\ &\quad + \sigma_n^2 \text{Tr} \{ \mathbf{G}_k^H \mathbf{G}_k \} + \omega \sigma_x^2 \text{Tr} \{ \mathbf{F}_k \mathbf{F}_k^H \} \text{Tr} \{ \mathbf{G}_k^H \mathbf{G}_k \}. \end{aligned} \quad (31)$$

In Appendix A, it is shown that the cost function in (31) is minimized when the \mathbf{F}_k and \mathbf{G}_k matrices have the following structure according to the SVD of the equivalent channel matrix $\hat{\mathbf{H}}_k^{\text{eq}}$ (22) for each subcarrier:

$$\mathbf{F}_k = \mathbf{V}_k \mathbf{\Phi}_k \quad \mathbf{G}_k = \mathbf{\Gamma}_k \mathbf{\Lambda}_k^{\#} \mathbf{U}_k^H \quad (32)$$

where $\mathbf{\Phi}_k$ is a diagonal matrix that sets the power distribution policy (the square of its elements $\phi_k^2(l)$ define the power allocated to each channel mode), and $\mathbf{\Gamma}_k$ is also a diagonal matrix applied to the symbols at the channel output. According to

(32), the function to minimize, including the average power constraint, becomes

$$\begin{aligned} \xi = & \sum_{k=1}^K \sigma_x^2 \text{Tr} \left\{ (\mathbf{\Gamma}_k \mathbf{\Phi}_k - \mathbf{I})^H (\mathbf{\Gamma}_k \mathbf{\Phi}_k - \mathbf{I}) \right\} \\ & + \sigma_n^2 \text{Tr} \left\{ \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H \mathbf{\Lambda}_k^{\#2} \right\} \\ & + \omega \sigma_x^2 \text{Tr} \left\{ \mathbf{\Phi}_k^H \mathbf{\Phi}_k \right\} \text{Tr} \left\{ \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H \mathbf{\Lambda}_k^{\#2} \right\} \\ & - \mu \left[\left(\sum_{k=1}^K \text{Tr} \left\{ \mathbf{\Phi}_k^H \mathbf{\Phi}_k \right\} \right) - P_0 \right] \end{aligned} \quad (33)$$

where μ is the Lagrange multiplier.

The optimum elements in the diagonal matrices $\mathbf{\Phi}_k$ and $\mathbf{\Gamma}_k$ can be found differentiating the cost function with respect to the matrices $\mathbf{\Phi}_k^H$ and $\mathbf{\Gamma}_k^H$ and equating it to zero. Thus, the following set of equations is obtained:

$$\begin{aligned} & \left(\omega \sigma_x^2 \text{Tr} \left\{ \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H \mathbf{\Lambda}_k^{\#2} \right\} + \mu \right) \mathbf{\Phi}_k + \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H \mathbf{\Phi}_k = \mathbf{\Gamma}_k^H \\ & \left(\omega \sigma_x^2 \text{Tr} \left\{ \mathbf{\Phi}_k \mathbf{\Phi}_k^H \right\} \mathbf{\Lambda}_k^{\#2} + \sigma_n^2 \mathbf{\Lambda}_k^{\#2} \right) \mathbf{\Gamma}_k + \mathbf{\Phi}_k \mathbf{\Phi}_k^H \mathbf{\Gamma}_k = \mathbf{\Phi}_k^H. \end{aligned} \quad (34)$$

In order to solve this equation with respect to the matrices $\mathbf{\Phi}_k$ and $\mathbf{\Gamma}_k$ and the Lagrange multiplier μ , the first step is to multiply first equation by the matrix $\mathbf{\Phi}_k^H$ and second equation by the the matrix $\mathbf{\Gamma}_k^H$ to get the following identity:

$$\begin{aligned} & \left(\omega \sigma_x^2 \text{Tr} \left\{ \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H \mathbf{\Lambda}_k^{\#2} \right\} + \mu + \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H \right) \mathbf{\Phi}_k \mathbf{\Phi}_k^H \\ & = \left(\omega \sigma_x^2 \text{Tr} \left\{ \mathbf{\Phi}_k \mathbf{\Phi}_k^H \right\} \mathbf{\Lambda}_k^{\#2} + \sigma_n^2 \mathbf{\Lambda}_k^{\#2} + \mathbf{\Phi}_k \mathbf{\Phi}_k^H \right) \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H. \end{aligned} \quad (35)$$

From this point, μ can be solved as

$$\mu = \frac{\sigma_n^2}{P_0} \sum_{j=1}^K \text{Tr} \left\{ \mathbf{\Gamma}_j^H \mathbf{\Gamma}_j \mathbf{\Lambda}_j^{\#2} \right\}. \quad (36)$$

Replacing μ into (35), the next relation between the diagonal matrices $\mathbf{\Gamma}_k$ and $\mathbf{\Phi}_k$ is derived:

$$\mathbf{\Phi}_k = C_k \mathbf{\Lambda}_k^{\#} \mathbf{\Gamma}_k \quad (37)$$

where C_k is a scalar that must be found. Substituting (37) in (35), it can be seen that the identity is only accomplished when the scalar C_k is constant for all subcarriers, i.e., $C_k = C$.

Introducing (36) and (37) into (34), and after some manipulations, it is straightforward to verify that both equations in (34) collapse to one:

$$\left(\omega \text{Tr} \left\{ \mathbf{\Phi}_k \mathbf{\Phi}_k^H \right\} + \frac{\sigma_n^2}{\sigma_x^2} \right) \mathbf{I} + \mathbf{\Lambda}_k^2 \mathbf{\Phi}_k \mathbf{\Phi}_k^H = C \mathbf{\Lambda}_k. \quad (38)$$

For each subcarrier, this equality provides a set of M equations that are linear on the $M + 1$ unknowns given by the M elements of the diagonal matrix $\mathbf{\Phi}_k \mathbf{\Phi}_k^H$ and the scalar C . Hence, all the unknowns can be found when simultaneously solving the K sets of equations given by (38) combined with the power constraint. Alternatively, their resolution can be simplified by solving the problem for each subcarrier in two stages, as shown next.

First, (38) is solved for each subcarrier, where we have (39), shown at the bottom of the page, where $\lambda_k^{\text{eq}}(i)$ is the i th singular value of $\hat{\mathbf{H}}_k^{\text{eq}}$ matrix. If the diagonal terms of matrix $\mathbf{\Phi}_k \mathbf{\Phi}_k^H$ are stored in vector ϕ_k^2 , and \mathbf{B} matrix is defined, the set of equations can be written as

$$\mathbf{B} \begin{bmatrix} \phi_k^2 \\ C \end{bmatrix} = -\frac{\sigma_n^2}{\sigma_x^2} \mathbf{1}_{M \times 1} \quad (40)$$

where vector $\mathbf{1}$ stands for the all-ones vector. The solution to the set of equations can be found as the summation of the particular solution given by (40) and the kernel of matrix \mathbf{B} :

$$\begin{bmatrix} \phi_k^2 \\ C \end{bmatrix} = -\frac{\sigma_n^2}{\sigma_x^2} \mathbf{B}^{\#} \mathbf{1} + D_k \mathbf{v} \quad (41)$$

where \mathbf{v} is a vector spanning the null-space of matrix \mathbf{B} , and the scalar D_k is an arbitrary constant. Afterwards, the values for the constant D_k are found by solving a second set of equations, including the power constraint

$$\sum_{k=0}^{N-1} \mathbf{1}^T \phi_k^2 = P_0. \quad (42)$$

Notice that the power allocated to each subcarrier in vector ϕ_k^2 should satisfy $\phi_k^2(l) \geq 0$. If any of the components obtained when solving (41) were negative, the power allocated to that subcarrier should be set to zero $\phi_k^2(l) = 0$, and the described procedure should be repeated for the rest of subcarriers until all the elements in vector ϕ_k^2 satisfy $\phi_k^2(l) \geq 0$.

C. Asymptotic Performance

In this section, the robust MMSE algorithm is analyzed for the extreme cases where the uncertainty is very high or very low. When channel knowledge at the transmitter is perfect, i.e., when $\hat{\mathbf{H}}_k^{\text{eq}} = \mathbf{H}$ and $\omega = 0$ [see (18) and (21)], the robust cost

$$\begin{bmatrix} \omega + |\lambda_k^{\text{eq}}(1)|^2 & \omega & \dots & \omega & -|\lambda_k^{\text{eq}}(1)| \\ \omega & \omega + |\lambda_k^{\text{eq}}(2)|^2 & \ddots & \omega & -|\lambda_k^{\text{eq}}(2)| \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \omega & \dots & \dots & \omega + |\lambda_k^{\text{eq}}(M)|^2 & -|\lambda_k^{\text{eq}}(M)| \end{bmatrix} \begin{bmatrix} \phi_k^2(1) \\ \vdots \\ \phi_k^2(M) \\ C \end{bmatrix} = \begin{bmatrix} -\frac{\sigma_n^2}{\sigma_x^2} \\ \vdots \\ -\frac{\sigma_n^2}{\sigma_x^2} \end{bmatrix} \quad (39)$$

function in (28) converges to the nonrobust solution in (30), and therefore, both coincide. When the channel uncertainty is very high, the MMSE design is meaningless since the design is based on the assumption that neither the transmitter nor the receiver know the channel.

IV. MINIMUM BER DESIGN

This section proposes the *a priori* design for the set of linear precoder matrices \mathbf{F}_k that minimize the BER in an ML receiver subject to an average power constraint. This algorithm considers the reliability of channel estimation in the cost function, adapting to the channel uncertainty and providing a solution that goes from the configuration for open-loop to the closed-loop with perfect CSI as estimation errors diminish (in a similar way as [7] and [10]). The filtering matrix at the receiver is set to the identity matrix $\mathbf{G}_k = \mathbf{I}$ since the BER cannot be improved by means of a linear filtering stage at this side. The design is formulated for the general MIMO OFDM case, including single antenna transmission and beamforming [7], [10] as particular cases.

The cost function to be optimized is described for a general constellation $\{s_1 \dots s_{\mathcal{N}}\}$ of size \mathcal{N} , even though final equations are shown for the particular case of quadrature phase shift keying (QPSK) modulation. In order to be able to derive a closed-form solution, two main assumptions have been made.

- A1) The receiver is operating at high SNR.
- A2) The function $\mathcal{Q}(\sqrt{x})$ can be approximated with small error as (see Appendix B).

$$\mathcal{Q}(\sqrt{x}) \simeq \delta e^{-\alpha x}. \quad (43)$$

For a given channel realization \mathbf{H}_k , the BER for the ML receiver can be written in terms of the pairwise error probability (PEP) of detecting symbol \mathbf{x}_{kj} when the symbol \mathbf{x}_{ki} was transmitted:

$$\text{BER}(\mathbf{H}_1 \dots \mathbf{H}_K) = \frac{1}{\mathcal{N}^{2M} K} \sum_{k=1}^K \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \times P(\mathbf{x}_{ki} \rightarrow \mathbf{x}_{kj} | \mathbf{H}_1 \dots \mathbf{H}_K) \quad (44)$$

where $B(\mathbf{x}_{ki}, \mathbf{x}_{kj})$ denotes the number of bits that are different in vectors \mathbf{x}_{ki} and \mathbf{x}_{kj} .

The robustness of the algorithm is obtained averaging the BER over the channel uncertainty using a Bayesian formulation. Hence, the optimization criterion subject to an average power constraint across all antennas and subcarriers becomes

$$\begin{aligned} \min \quad & \xi = \frac{1}{\mathcal{N}^{2M} K} \sum_{k=1}^K \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \\ & \times E_{\mathbf{h}/\hat{\mathbf{h}}} \{P(\mathbf{x}_{ki} \rightarrow \mathbf{x}_{kj} | \mathbf{H}_1 \dots \mathbf{H}_K)\} \\ \text{subject to} \quad & \sum_{k=1}^K \text{Tr} \{\mathbf{F}_k^H \mathbf{F}_k\} = P_0. \end{aligned} \quad (45)$$

A. Cost Function

The pairwise error probability in (45) can be written in terms of the Euclidean distance between the transmitted codewords \mathbf{x}_{ki} and \mathbf{x}_{kj} as they appear at the receiver and is upper bounded by

$$P(\mathbf{x}_{ki} \rightarrow \mathbf{x}_{kj}) \leq \mathcal{Q} \left(\sqrt{\frac{1}{2\sigma_n^2} |\mathbf{H}_k \mathbf{F}_k \mathbf{d}_{ijk}|^2} \right) \quad (46)$$

where vector $\mathbf{d}_{ijk} = (\mathbf{x}_{ki} - \mathbf{x}_{kj})$ contains the distance between the codewords at the transmitter before the prefiltering. Under assumption A2), the pairwise probability can be approximated by the exponential function

$$P(\mathbf{x}_{ki} \rightarrow \mathbf{x}_{kj}) \simeq \delta_k e^{-\alpha_k (1/2\sigma_n^2) |\mathbf{H}_k \mathbf{F}_k \mathbf{d}_{ijk}|^2}. \quad (47)$$

This equation can be rewritten in a more compact form in order to simplify the notation in the subsequent equations

$$P(\mathbf{x}_{ki} \rightarrow \mathbf{x}_{kj}) \simeq \delta_k e^{-\mathbf{h}^H \mathbf{M}_{ijk} \mathbf{h}} \quad (48)$$

where \mathbf{M}_{ijk} is the $KM_T M_R \times KM_T M_R$ matrix defined as

$$\mathbf{M}_{ijk} = \gamma_k \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{I}_{M_T M_R} \\ \vdots \\ \mathbf{0} \end{bmatrix} \left[\left((\mathbf{F}_k \mathbf{d}_{ijk}) (\mathbf{F}_k \mathbf{d}_{ijk})^H \right) \otimes \mathbf{I}_{M_R} \right] \times [\mathbf{0} \dots \mathbf{I}_{M_T M_R} \dots \mathbf{0}] \quad (49)$$

and $\gamma_k = (\alpha_k/2\sigma_n^2)$, $\mathbf{0}$ is a square all zero matrix, $\mathbf{I}_{M_T M_R}$ is the identity matrix, both of size $M_T M_R$, and \mathbf{I}_{M_R} is the identity matrix of size M_R .

The approximation in (47) simplifies the evaluation of the expectation over channel response, as shown next. Using the results in (14) and (15) to expand the Gaussian probability density function (pdf) $f_{\mathbf{h}/\hat{\mathbf{h}}}$, the averaged PEP becomes

$$\begin{aligned} E_{\mathbf{h}/\hat{\mathbf{h}}} \{P(\mathbf{x}_{ki} \rightarrow \mathbf{x}_{kj})\} &= \frac{\delta_k}{\pi^K \det(\mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}})} \\ &\times \int_{\mathbf{h} \in \mathbf{C}} e^{-\mathbf{h}^H \mathbf{M}_{ijk} \mathbf{h}} e^{-(\mathbf{h} - \hat{\mathbf{h}}^{\text{eq}})^H \mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}}^{-1} (\mathbf{h} - \hat{\mathbf{h}}^{\text{eq}})} d\mathbf{h} \end{aligned} \quad (50)$$

where $\hat{\mathbf{h}}^{\text{eq}}$, as defined in (19), denotes the equivalent channel over all subcarriers. This integral can be easily solved rewriting its integrand as

$$\int_{\mathbf{h} \in \mathbf{C}} e^{-(\mathbf{h} - \mu)^H \beta (\mathbf{h} - \mu) - \eta} d\mathbf{h} \quad (51)$$

where

$$\begin{aligned} \beta &= \mathbf{M}_{ijk} + \mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}}^{-1} \\ \mu &= \left(\mathbf{M}_{ijk} \mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}} + \mathbf{I} \right)^{-1} \hat{\mathbf{h}}^{\text{eq}} \\ \eta &= \hat{\mathbf{h}}^{\text{eq}H} \left(\mathbf{M}_{ijk} \mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}} + \mathbf{I} \right)^{-1} \mathbf{M}_{ijk} \hat{\mathbf{h}}^{\text{eq}}. \end{aligned} \quad (52)$$

The solution to the integral in (51) can be found by comparing its integrand with a complex Gaussian pdf, whose integral equals to one. Thus, the averaged PEP in (50) becomes

$$E_{\mathbf{h}/\hat{\mathbf{h}}}\{P(\mathbf{x}_{ki} \rightarrow \mathbf{x}_{kj})\} = \frac{\delta_k}{\det(\mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}}\mathbf{M}_{ijk} + \mathbf{I})} \times e^{-\hat{\mathbf{h}}^{eqH}(\mathbf{M}_{ijk}\mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}} + \mathbf{I})^{-1}\mathbf{M}_{ijk}\hat{\mathbf{h}}^{eq}}. \quad (53)$$

The optimization of this cost function for any spatial correlation between antennas can only be obtained by means of numerical techniques. The derivation of a closed-form solution for (53) is feasible only when antennas are uncorrelated. Hence, the conditional covariance $\mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}}$ will be hereafter specified for the particular case of uncorrelated antennas. Replacing \mathbf{M}_{ijk} and $\mathbf{C}_{\mathbf{h}/\hat{\mathbf{h}}}$ by its values in (49) and (17), the function to minimize, subject to a power constraint, becomes

$$\xi = \frac{1}{\mathcal{N}^{2M}K} \sum_{k=1}^K \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} \frac{B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \delta_k}{(1 + \gamma_k \omega |\mathbf{F}_k \mathbf{d}_{ijk}|^2)^{M_R}} \times e^{-(\gamma_k/(1+\gamma_k\omega|\mathbf{F}_k\mathbf{d}_{ijk}|^2))|\hat{\mathbf{H}}_k^{eq}\mathbf{F}_k\mathbf{d}_{ijk}|^2} \quad (54)$$

where the identity (5) has been applied.

B. Closed-Form Solution

This section derives a suboptimal solution for the minimization of (54). As the direct optimization of (54) leads to very intricate equations, the minimization of the cost function has been obtained under the assumption that the linear transformation at the transmitter has the following structure:

$$\mathbf{F}_k = \mathbf{V}_k \mathbf{\Phi}_k \mathbf{T} \quad (55)$$

where matrix \mathbf{V}_k contains the right singular vectors of the equivalent channel matrix, $\mathbf{\Phi}_k$ is a diagonal matrix, and \mathbf{T} is a $M \times M$ unitary matrix whose properties are described next. This configuration has been proved to be optimal for minimum BER in a ZF receiver with perfect CSI [8] and can also be shown to be optimal for minimum BER in a ML receiver when multiple transmit antennas are used for beamforming (i.e., only a single symbol is spreaded on all antennas for each subcarrier) [15], but it has not been proved to be optimal when multiple symbols are transmitted at the same time. However, there are several reasons to support its choice. It keeps the same structure as the MMSE solution and other papers published previously. Besides, it leads to a simple closed-form design, as will be shown next. Finally, from the point of view of BER minimization, it can also be argued that the introduction of matrix \mathbf{T} alleviates the main drawback of the use of matrices \mathbf{V}_k and $\mathbf{\Phi}_k$: the loss of space diversity caused by the decomposition of the MIMO channel into a set of parallel multiplicative subchannels. If $\mathbf{T} = \mathbf{I}$ and matrix \mathbf{G}_k is selected as the left singular vectors of the equivalent channel matrix, then the MIMO channel reduces to a set of KM parallel independent flat-fading subchannels:

$$\mathbf{r}_k(i) = \lambda_k^{eq}(i)\boldsymbol{\phi}_k(i)\mathbf{x}_k(i) + \mathbf{w}_k(i) \quad (56)$$

so that the symbols corresponding to small values of $\lambda_k^{eq}(i)\boldsymbol{\phi}_k(i)$ are systematically lost. In the context of frequency flat Rayleigh fading channels, it was shown in [16]–[18] that the receiver could benefit from the diversity of the fading channel, provided the transmitter used a linear transformation that spread the symbols in time, obtaining significant performance gains both in terms of uncoded and coded BER. Similar conclusions would be obtained here if the Rayleigh fading channel statistics were replaced by those of the eigenvalues of the $\hat{\mathbf{H}}_k^{eqH}\hat{\mathbf{H}}_k^{eq}$ matrix, but to the best of the authors knowledge, the statistics of the eigenvalues of noncentral Wishart matrices are not known. In this paper, the unitary matrix \mathbf{T} has been set to the DFT matrix, given that it was shown in [17, eq. 17] that the DFT or the Walsh–Hadamard matrices could provide the desired fading diversity with minimum complexity, and besides, the same condition was required for the design of the optimal matrix \mathbf{T} in the ZF receiver [8, Lemma 1] and for the minimum MSE design in a linear receiver [16, eq. 39].

Once the structure of \mathbf{F}_k has been set to (55), the design of \mathbf{F}_k reduces to the design of matrix $\mathbf{\Phi}_k$. This paper derives a closed-form solution for the cost function resulting from (54) and (55) for high SNR replacing the true BER by an upper bound, as shown next. The proposed solution is optimum when $M = 1$. When $M > 1$, this closed-form solution is suboptimum; hence, a gradient algorithm [19] has been computed as a benchmark.

As a result of the application of Jensen's inequality to the convex function $\mathcal{Q}(\sqrt{x})$, Appendix C shows that the following inequality holds for high SNR:

$$\frac{1}{\mathcal{N}^{2M}} \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \mathcal{Q}\left(\sqrt{\frac{|\boldsymbol{\Lambda}_k \mathbf{\Phi}_k \mathbf{T} \mathbf{d}_{ijk}|^2}{2\sigma_n^2}}\right) \leq \frac{1}{\mathcal{N}^{2M}} \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \mathcal{Q}\left(\sqrt{\frac{|\boldsymbol{\Lambda}_k \mathbf{\Phi}_k \mathbf{d}_{ijk}|^2}{2\sigma_n^2}}\right) \quad (57)$$

where equality applies if and only if the product matrix $\boldsymbol{\Lambda}_k \mathbf{\Phi}_k$ is proportional to the identity matrix (ZF solution) or $\mathbf{T} = \mathbf{I}$. Therefore, if the true BER depending on $|\boldsymbol{\Lambda}_k \mathbf{\Phi}_k \mathbf{T} \mathbf{d}_{ijk}|$ is replaced by an upper bound depending on $|\boldsymbol{\Lambda}_k \mathbf{\Phi}_k \mathbf{d}_{ijk}|$, an expression is obtained that does not depend on matrix \mathbf{T} . This upper bound will allow to get a simple closed-form expression for the diagonal matrix $\mathbf{\Phi}_k$, as shown next. Note that in the particular case of $M = 1$, the closed-form solution will be the optimum solution because the matrix \mathbf{T} reduces to one, and thus, the equality in (57) applies.

Expanding the cost function in (54) according to the structure of the matrix \mathbf{F}_k , the new cost function, including the average power constraint, becomes

$$\xi = \frac{1}{\mathcal{N}^{2M}K} \sum_{k=1}^K \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} \frac{B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \delta_k}{(1 + \gamma_k \omega |\boldsymbol{\Phi}_k \mathbf{d}_{ijk}|^2)^{M_R}} \times e^{-((\gamma_k |\boldsymbol{\Lambda}_k^{eq} \boldsymbol{\Phi}_k \mathbf{d}_{ijk}|^2)/(1+\gamma_k\omega|\boldsymbol{\Phi}_k\mathbf{d}_{ijk}|^2))} - \mu \left[\sum_{k=1}^K \sum_{l=1}^M |\boldsymbol{\phi}_k(l)|^2 - P_0 \right] \quad (58)$$

where μ is the Lagrange multiplier. As the cost function depends on $|\phi_k(l)|^2$, we will choose $\phi_k(l)$ as a real positive variable. The minimization problem follows, differentiating (58) with respect to $\phi_k(l)$:

$$\begin{aligned} \frac{\partial \xi}{\partial \phi_k(l)} = & -\frac{1}{\mathcal{N}^{2M} K} \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} \frac{B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \delta_k \gamma_k |\mathbf{d}_{ijk}(l)|^2 \phi_k(l)}{\left(1 + \gamma_k \omega |\Phi_k \mathbf{d}_{ijk}|^2\right)^{M_R+1}} \\ & \times e^{-((\gamma_k |\Lambda_k^{\text{eq}} \Phi_k \mathbf{d}_{ijk}|^2)/(1 + \gamma_k \omega |\Phi_k \mathbf{d}_{ijk}|^2))} \\ & \times \left[M_R \omega + |\lambda_k^{\text{eq}}(l)|^2 - \frac{\gamma_k \omega |\Lambda_k^{\text{eq}} \Phi_k \mathbf{d}_{ijk}|^2}{1 + \gamma_k \omega |\Phi_k \mathbf{d}_{ijk}|^2} \right] \\ & - \mu \phi_k(l) \end{aligned} \quad (59)$$

where $\mathbf{d}_{ijk}(l)$ denotes the l th element in vector \mathbf{d}_{ijk} , and $\lambda_k^{\text{eq}}(l)$ is the l th singular value of channel matrix $\hat{\mathbf{H}}_k^{\text{eq}}$. Under the high SNR assumption A1), the error probability is dominated by the minimum distance between any pairs of symbols (\mathbf{x}_{ki} , \mathbf{x}_{kj}). Therefore, the summation in (59) can be approximated, considering the terms where $\mathbf{x}_{ki} - \mathbf{x}_{kj} = d\mathbf{i}_m$, where \mathbf{i}_m is one of the columns of the identity matrix, and d is a scalar that refers to the minimum distance between any two constellation symbols. Using this approximation in (59) and setting the derivative to zero, the following equation is obtained for QPSK constellation ($d = \sqrt{2}\sigma_x$):

$$\begin{aligned} -\mathcal{K} \frac{2\sigma_x^2 \delta_k \gamma_k \phi_k(l)}{(1 + 2\sigma_x^2 \gamma_k \omega \phi_k(l)^2)^{M_R+1}} \\ \times e^{-((2\sigma_x^2 \gamma_k |\lambda_k^{\text{eq}}(l)|^2 \phi_k(l)^2)/(1 + 2\sigma_x^2 \gamma_k \omega \phi_k(l)^2))} \\ \times \left[M_R \omega + \frac{|\lambda_k^{\text{eq}}(l)|^2}{1 + 2\sigma_x^2 \gamma_k \omega \phi_k(l)^2} \right] - \mu \phi_k(l) = 0 \end{aligned} \quad (60)$$

where \mathcal{K} is a nonrelevant constant. After some manipulations, this equality can be rewritten as

$$\begin{aligned} -\frac{2\sigma_x^2 \gamma_k \phi_k(l)^2 |\lambda_k^{\text{eq}}(l)|^2}{1 + 2\sigma_x^2 \gamma_k \omega \phi_k(l)^2} - (M_R + 1) \ln(1 + 2\sigma_x^2 \gamma_k \omega \phi_k(l)^2) \\ + \ln\left(M_R \omega + \frac{|\lambda_k^{\text{eq}}(l)|^2}{1 + 2\sigma_x^2 \gamma_k \omega \phi_k(l)^2}\right) \\ = \mu - \ln(2\sigma_x^2 \delta_k \gamma_k). \end{aligned} \quad (61)$$

A closed-form solution for this identity cannot be derived. However, under the assumption that the channel uncertainty is low, the following approximations can be used: $\ln(1 + x) \simeq x$, and $1/(1 + x) \simeq 1$. Thus, (61) is simplified as

$$\begin{aligned} -2\sigma_x^2 \gamma_k |\lambda_k^{\text{eq}}(l)|^2 \phi_k(l)^2 - (M_R + 1) \sigma_x^2 \gamma_k \omega \phi_k(l)^2 \\ + \ln(M_R \omega + |\lambda_k^{\text{eq}}(l)|^2) = \mu - \ln(2\sigma_x^2 \delta_k \gamma_k) \end{aligned} \quad (62)$$

obtaining a closed-form solution for $\phi_k^2(l)$:

$$\phi_k^2(l) = \left[-\frac{\mu - \ln(2\sigma_x^2 \delta_k \gamma_k) - \ln(M_R \omega + |\lambda_k^{\text{eq}}(l)|^2)}{2\sigma_x^2 \gamma_k ((M_R + 1) \omega + |\lambda_k^{\text{eq}}(l)|^2)} \right]^+ \quad (63)$$

where μ is determined forcing the power constraint (45), and the function $[x]^+$ is defined as $\max\{x, 0\}$.

C. Asymptotic Performance

In this section, the robust BER algorithm is analyzed for the extreme cases where the uncertainty is very high or very low.

When channel knowledge at the transmitter is perfect, i.e., when $\hat{\mathbf{H}}_k^{\text{eq}} = \mathbf{H}$ and $\omega = 0$ [see (18) and (21)], and only one symbol is transmitted per subcarrier (i.e., $M = 1$), the solution for the robust cost function in (63) coincides with that one of [20]. Note that the solution does not coincide with the solution proposed in [8] since one assumes a ML receiver and the other one a ZF receiver.

When the channel uncertainty is very high, in the extreme case of $\hat{\mathbf{H}}_k^{\text{eq}} = 0$ and $\omega = 1$, the average PER function converges to

$$E_{\mathbf{h}/\hat{\mathbf{h}}} \{P(\mathbf{x}_{ki} \rightarrow \mathbf{x}_{kj})\} = \frac{\delta_k}{\left(1 + \gamma_k \omega |\mathbf{F}_k \mathbf{d}_{ijk}|^2\right)^{M_R}} \quad (64)$$

which is the same one found in [11] for uncorrelated antennas. Note that in this case, the optimization criterion is simpler, and therefore, there is no need to resort to the high SNR approximations in order to get a closed-form solution. Indeed, it can be seen that the power allocation tends to the open-loop solution (i.e., the same power is allocated to all antennas and subcarriers), as it will be shown in Section V.

V. SIMULATION RESULTS

In order to illustrate the performance of the proposed closed-loop algorithms in the presence of imperfect channel estimates, several simulations are presented for different scenarios and channel estimation errors, assuming that CSI quality is unbalanced between the transmitter and the receiver. To validate the closed-form solutions in (41) and (42) and (63), uncorrelated antennas are always assumed. Four algorithms are compared: the proposed algorithm for MMSE (“*MMSE Robust*”) and for minimum uncoded BER (“*BER Robust*”), and its nonrobust counterparts that assume $\omega = 0$ (“*MMSE Non-Robust*” and “*BER Non-Robust*”). The simulations compare the performance of these algorithms in terms of uncoded BER and coded BER.

The simulation parameters are selected according to the HIPERLAN/2 standard [21]. As depicted in Fig. 1, the bit stream to be transmitted is applied to the classical convolutional code of rate 1/2 and generator polynomials 133_{OCT} and 171_{OCT}.¹ The encoder is initialized to the zero state and returned to it after encoding 3456 bits (i.e., eight DLC-PDUs of 54 bytes—according to HL/2) by appending six tail bits. The coded bits are mapped into a QPSK constellation and interleaved with a symbol-block interleaver of depth 16M. Afterwards, the symbols are multiplexed in the K subcarriers (grouping M symbols per subcarrier), prefiltered by the matrices \mathbf{F}_k , and finally modulated in OFDM symbols (including pilot tones and empty subcarriers according to HIPERLAN/2).

In order to have a fair comparison, all algorithms are simulated using an ML detector that takes into account the channel uncertainty at the receiver. As the channel is not perfectly known

¹This channel code is not specific for HIPERLAN/2, but it is widely used in several standards

at the transmitter and/or the receiver, the relationship between the transmitted and the received data (3) is not equivalent to a set of independent flat-fading subchannels (i.e., $\mathbf{G}_k \mathbf{H}_k \mathbf{F}_k$ is not diagonal). Therefore, the optimum receiver requires the joint estimation of all bits transmitted in the same symbol \mathbf{x}_k . This procedure is used to evaluate the uncoded BER, detecting the symbol $\hat{\mathbf{x}}_k$ that maximizes (65), shown at the bottom of the page, where \mathcal{K} refers to a nonrelevant constant, and the channel uncertainty ω and the equivalent channel $\hat{\mathbf{H}}_k^{\text{eq}}$ are computed, making use of the channel uncertainty at the receiver.

For the evaluation of the coded BER, it is too complex to apply (65) due to the presence of the block interleaver. Hence, as the goal of this paper is not on the receiver architecture but on the design of the transmitter, a simplified decoder has been implemented that provides a bound on the performance achieved by practical decoding schemes. The coded BER results presented in this section have been obtained using a genie decoder that computes the log-likelihood of a symbol $\mathbf{x}_k(l)$, assuming that all other symbols that were transmitted simultaneously in in vector \mathbf{x}_k are known to the receiver. The BER achieved by this scheme bounds an iterative decoder like the one proposed in [22], performing an infinite number of iterations and assuming that no decoding errors were made.

The simulated Rayleigh MIMO channel obeys an exponential power delay profile with 50 ns of delay spread (RMS delay spread of the discrete channel 45.6 ns [23]), modeling a typical office indoor scenario. The simulations model the channel uncertainties due to the errors in the channel estimation process and the errors in the channel prediction when channels are time-varying. In both cases, the variance in the channel estimation error is assumed to be constant for all taps of the channel impulse response. For the sake of clarity, we will denote in this section $\sigma_{\epsilon_{Tx}}^2$ and $\sigma_{\epsilon_{Rx}}^2$ the variances in the channel estimation error at the transmitter and at the receiver, respectively. It is assumed that the main contribution to channel uncertainty at the transmitter is the channel tracking error in fast linear time-varying channels; hence, it is considered to be independent of the SNR, as would be the case of a channel tracker based on a linear predictor [3]. Simulations are carried out with $\sigma_{\epsilon_{Tx}}^2 = 0.12$. This value models a realistic scenario where a user is moving at 4 Km/h, and the channel is updated at the transmitter each 2 ms (as would be the case of HIPERLAN/2) [3]. The degree of channel uncertainty at the transmitter is measured by the coefficient ρ , which is defined as

$$\rho = \frac{\sigma_{\epsilon_{Tx}}^2}{\sigma_{\epsilon_{Tx}}^2 + \sigma_h^2} \quad (66)$$

where $\rho = 0$ denotes perfect CSI, whereas $\rho = 1$ means no channel knowledge. At the receiver, the main contribution to channel uncertainty is the estimation variance due to the presence of additive noise. Hence, $\sigma_{\epsilon_{Rx}}^2$ is proportional to the noise variance σ_n^2 , as would be the case of a linear channel

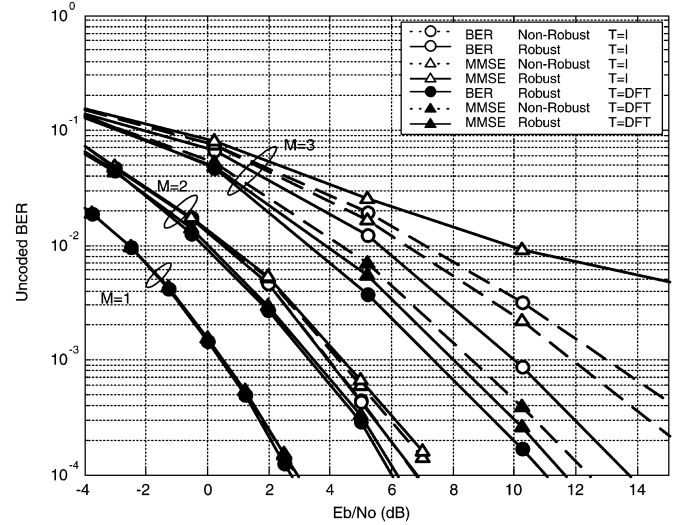


Fig. 2. Uncoded BER comparison between different power allocation strategies. $M_T = 3$, $M_R = 3$. Transmitter uncertainty: $\sigma_{\epsilon_{Tx}}^2 = 0.12$ ($\rho = 0.1$) Receiver uncertainty: Perfect CSI.

estimator [24], where the factor of proportionality depends on the training sequence length. Simulations are carried out with $\sigma_{\epsilon_{Rx}}^2 = 0.375\sigma_n^2$, which approximately corresponds to the estimation of the simulated channel from a linear channel estimator when the training sequence consists of one OFDM symbol according to HIPERLAN/2 [15].

A. Uncoded BER Performance versus Eb/No

Fig. 2 illustrates the performance of the two proposed algorithms in terms of uncoded BER for the 3×3 antenna configuration. In order to focus on the transmitter design when CSI is noisy at this side, it is assumed that the receiver has perfect CSI knowledge, whereas the channel estimate at the transmitter has a constant variance equal to $\sigma_{\epsilon_{Tx}}^2 = 0.12$. The robust (solid lines) and nonrobust (dotted lines) algorithms are plotted for $M = \{1, 2, 3\}$, which correspond to 12, 24, and 36 Mbps, respectively, comparing its performance when the unitary matrix \mathbf{T} is omitted (i.e., $\mathbf{T} = \mathbf{I}$) and when matrix \mathbf{T} is set to the DFT matrix. In the particular case of $\mathbf{T} = \mathbf{I}$, the elements of the diagonal matrix Φ_k are obtained according to (41) and (42) for the MMSE solution and according to (63) for the BER solution. When matrix \mathbf{T} is set to the DFT matrix, the same solution provided by (41) and (42) is used because, as it has been proved in Appendix A, the MMSE cost function is insensitive to any unitary matrix at the right side of the linear transformation \mathbf{F}_k . On the contrary, the minimum BER algorithm (54) has been minimized applying an iterative solution since (63) is suboptimum when the matrix \mathbf{T} is different from the identity matrix.

The performance comparison in terms of M exhibits the tradeoff between BER and transmission rate. Observe that the difference between the algorithms is greater as M increases. When $M = 1$ (i.e., one symbol per subcarrier is transmitted

$$f_{\mathbf{r}_k | \hat{\mathbf{h}}}(\mathbf{r}_k | \hat{\mathbf{x}}_k, \hat{\mathbf{h}}, \mathbf{F}_k) = \mathcal{K} e^{-((\mathbf{r}_k - \hat{\mathbf{H}}_k^{\text{eq}} \mathbf{F}_k \hat{\mathbf{x}}_k)^H (\mathbf{r}_k - \hat{\mathbf{H}}_k^{\text{eq}} \mathbf{F}_k \hat{\mathbf{x}}_k)) / (\omega \hat{\mathbf{x}}_k^H \mathbf{F}_k^H \mathbf{F}_k \hat{\mathbf{x}}_k + \sigma_n^2)} \quad (65)$$

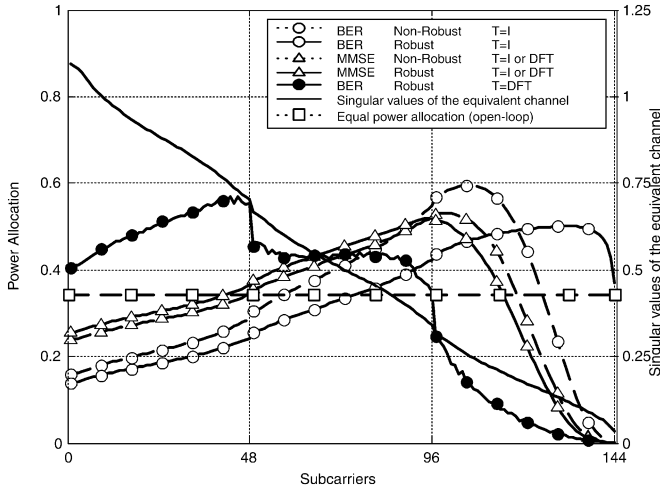


Fig. 3. Example of singular values (sorted in decreasing order) of the equivalent channel $\hat{H}_k^{\text{eq}}(k = 1 \dots K)$ for a specific channel realization and its associated power allocation elements (i.e., $\text{diag}\{\Phi_k \Phi_k^H\}$) for the different algorithms. $M_T = 3$, $M_R = 3$, and $M = 3$. Transmitter uncertainty: $\sigma_{\epsilon_{Tx}}^2 = 0.12$ ($\rho = 0.1$). SNR = 0 dB.

focused on the best channel mode), all algorithms have similar performances. On the contrary, when $M = 3$, the algorithms exhibit different performances, showing that the way to distribute the total power is crucial.

Some conclusions can also be drawn on the importance of the unitary matrix \mathbf{T} when analyzing the performance losses if this unitary matrix is omitted. As it can be shown, the performance losses are greater as M or, equivalently, the dispersion of the channel singular values increases. This result confirms that the unitary matrix alleviates loss of space diversity caused by the channel diagonalization.

Finally Fig. 2 shows that the “BER Robust” algorithm has best performance in terms of uncoded BER, as it was expected. However, when matrix \mathbf{T} is set to the DFT, the “MMSE Robust” algorithm exhibits good performance in terms of BER. The relative performance of the BER and MMSE algorithms is better understood when comparing in Fig. 3 the power allocation per subcarrier (i.e., the elements of the diagonal matrix $\Phi_k \Phi_k^H$) for each one of the four algorithms. As shown in the figure, when matrix \mathbf{T} is omitted, the algorithms that minimize the uncoded BER assign less power to the stronger channel modes and more power to the deepest faded ones, trying to compensate the fading, whereas the MMSE algorithms assign less power to the faded subchannels. This behavior leads the MMSE algorithms to have very high uncoded BER in some of the subcarriers, explaining the poor performance of the “MMSE Robust” algorithm when the unitary matrix \mathbf{T} is omitted. On the contrary, when matrix \mathbf{T} is set to the DFT matrix, as the unitary matrix combines all symbols into all channel modes, the optimal way to distribute the power is to allocate more power to the stronger channel modes, just as “BER Robust” when \mathbf{T} is the DFT matrix and “MMSE Robust” do.

B. Coded BER Performance versus E_b/N_0

Despite the fact the neither the MMSE algorithm nor the minimum BER algorithm have been optimized taking into account a channel code, it is interesting to evaluate the proposed

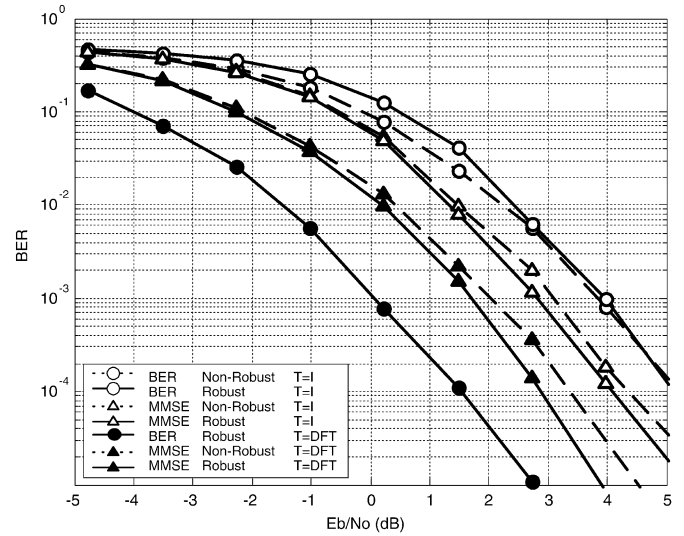


Fig. 4. Coded BER comparison between different power allocation strategies. $M_T = 3$, $M_R = 3$. Transmitter uncertainty: $\sigma_{\epsilon_{Tx}}^2 = 0.12$ ($\rho = 0.1$). Receiver uncertainty: $\sigma_{\epsilon_{Rx}}^2 = 0.375 \sigma_n^2$.

algorithms in terms of coded BER since most communication schemes include this stage. The performance of the algorithms has been compared for the specific channel code described in HIPERLAN/2, and they have also been compared in terms of the capacity and the cut-off rate as measures of the coded BER performance, which are independent of the specific channel code.

Fig. 4 depicts the performance of the proposed algorithms in terms of BER at the decoder output for the 3×3 antenna configuration when the receiver has an estimation variance proportional to σ_n^2 . It can be seen that when matrix \mathbf{T} is omitted, “MMSE Robust” algorithm achieves best performance, followed by “MMSE Non-Robust,” “BER Robust,” and “BER Non-Robust” solutions, whereas when matrix \mathbf{T} is set to the DFT matrix, the “BER Robust” algorithm outperforms the others. This behavior can be intuitively understood comparing, in Fig. 3, the power allocation strategy followed by the different algorithms. Note that the best algorithms in terms of coded BER are those that allocate more power to the good subchannels and penalize the weakest ones. This strategy, as opposed to that one followed by the algorithms that minimize uncoded BER, induces some systematic errors in the deepest faded channels. However, if the amount of errors is within the correction capability of the channel code, these errors can be corrected by the Viterbi decoder, explaining the overall increased performance at the decoder output in terms of coded BER.

To get more insight in the performance of the proposed power allocation algorithms in terms of coded BER, Fig. 5 evaluates the capacity and cut-off rate for each algorithm. Both parameters can be used as an indicator of the coded BER performance that is independent of the specific channel code. While the capacity is the theoretical upper limit on data rates where arbitrarily small BER can be achieved with coding, from the practical point of view, it is difficult to attain this upper limit. However, the cut-off rate provides a lower bound on the capacity and, until the proposal of turbo codes, it was considered to be the limit for the coding techniques that had practical interest.

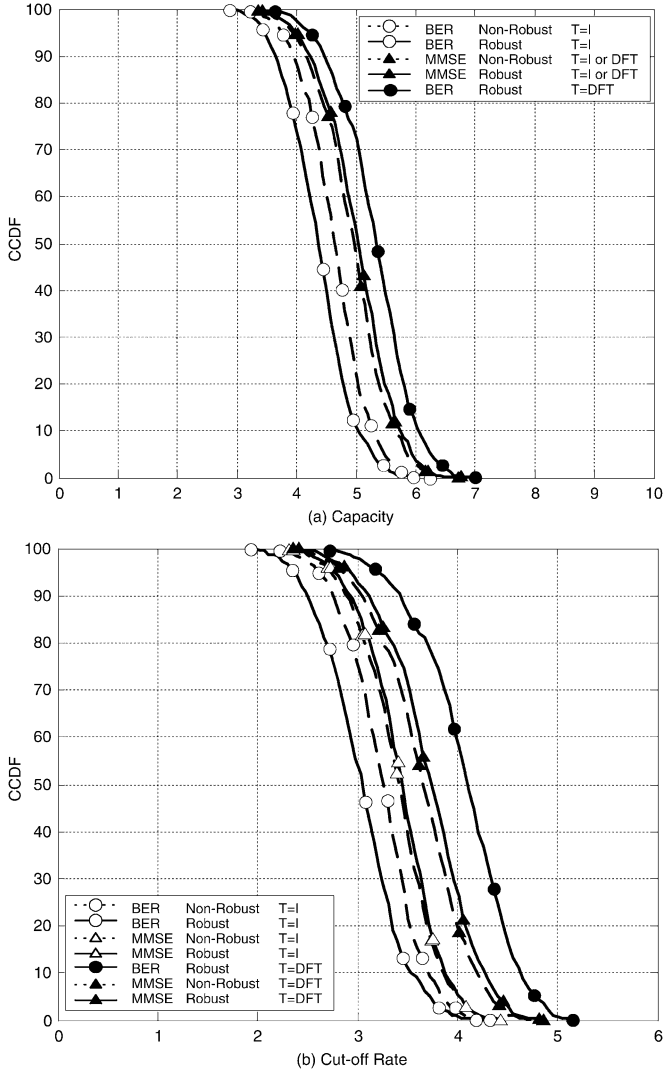


Fig. 5. CCDF curves of (a) the capacity and (b) cut-off rate for different power allocation strategies. $M_T = 3$, $M_R = 3$. Transmitter uncertainty: $\sigma_{\epsilon_{T_x}}^2 = 0.12$ ($\rho = 0.1$). SNR = 0 dB.

Fig. 5 shows the complementary cumulative distribution function (CCDF) curves of the capacity (in *bps/Hz*), which are computed as

$$C = \frac{1}{K} \sum_{k=1}^K \log_2 \left(\det \left(\mathbf{I} + \frac{\mathbf{F}_k^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{F}_k}{\sigma_x^2} \right) \right) \quad (67)$$

and the cut-off rate (in *bits/channel use*), which are computed as

$$R_0 = -\frac{1}{K} \sum_{k=1}^K \log_2 \left(\frac{1}{\mathcal{N}^{2M}} \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} e^{-(1/4\sigma_n^2) |\mathbf{H}_k \mathbf{F}_k \mathbf{d}_{i,j,k}|^2} \right). \quad (68)$$

When the unitary matrix \mathbf{T} is omitted, the MMSE algorithms outperform the algorithms that minimize BER in terms of capacity and in terms of cut-off rate, which explains why MMSE solutions result in better performance than minimum BER solutions. To explain the behavior of the proposed solutions when the unitary matrix \mathbf{T} is set to the DFT, note that the capacity expression in (67) is insensitive to any unitary matrix at the right

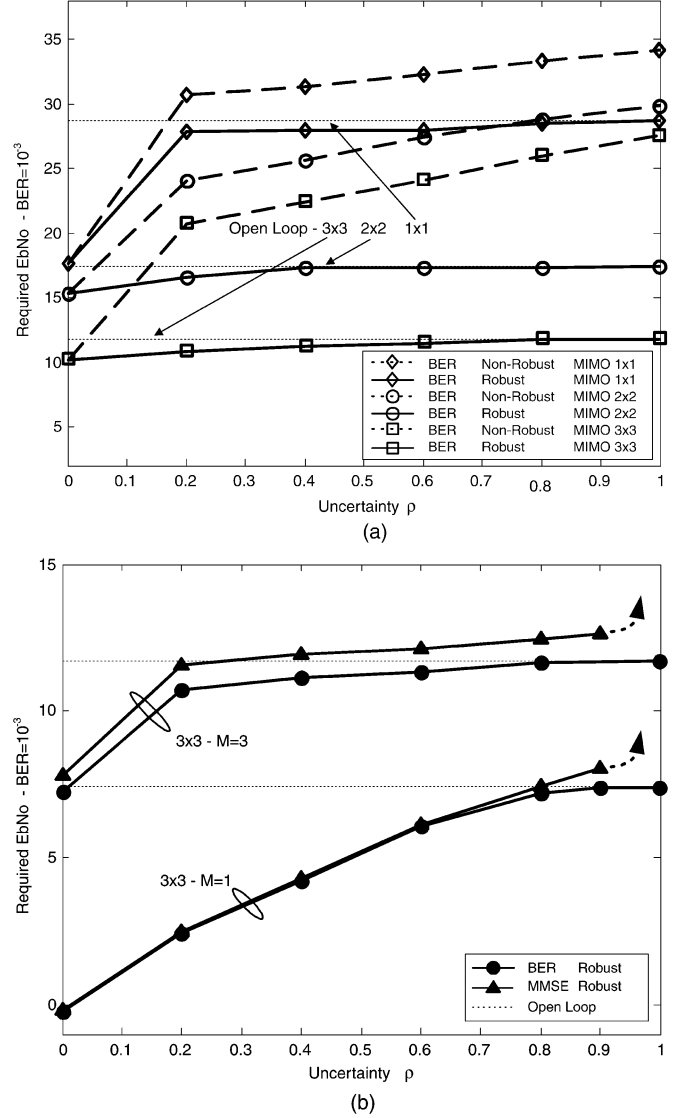


Fig. 6. Minimum Eb/No that achieves an uncoded BER $\leq 10^{-3}$ versus the transmitter uncertainty ρ . Receiver uncertainty: $\sigma_{\epsilon_{R_x}}^2 = 0.375\sigma_n^2$. (a) BER algorithm comparison for different MIMO configurations: 1×1 $M = 1$, 2×2 $M = 2$, 3×3 $M = 3$. (b) BER and MMSE algorithms comparison for 3×3 $M = \{1, 3\}$.

side of the linear transformation \mathbf{F}_k , and hence, this parameter cannot be used in our case as a measure of the coded BER performance. On the contrary, the cut-off rate is sensitive to the matrix \mathbf{T} and justifies the relevance of this matrix. As shown in Fig. 5, when the unitary matrix is set to the DFT, the cut-off rate is increased, which justifies that the BER at the decoder output is decreased. Moreover, the performance of the “BER Robust” algorithm \mathbf{T} is the DFT matrix as the best solution in terms of coded BER is explained since this algorithm exhibits the best cut-off rate.

C. Performance versus CSI Quality at the Transmitter

To test the performance of the proposed algorithms when channel uncertainty increases, Fig. 6 shows the minimum Eb/No required to achieve an uncoded BER $\leq 10^{-3}$. The required Eb/No is plotted as a function of the channel uncertainty degree at the transmitter ρ , while keeping the receiver uncertainty

proportional to the noise variance. As a reference, the point $\rho = 0$ was simulated assuming perfect CSI at both transmitter and receiver. Two simulations are presented. A first plot analyzes the minimum BER design for different MIMO configurations: 1×1 , 2×2 , 3×3 , assuming the maximum transmission rate (i.e., $M = \{1, 2, 3\}$). A second simulation evaluates the MMSE and the minimum BER algorithms for the MIMO configuration 3×3 with $M = \{1, 3\}$.

As shown in Fig. 6(a), the robustness of the proposed algorithm is evidenced since the “*BER Robust*” solution always outperforms the “*BER Non-Robust*” one. Note that when CSI quality at the transmitter degrades (i.e., $\rho \rightarrow 1$), the “*BER Robust*” design tends to the open-loop solution (i.e., equally power allocation for all subcarriers and antennas), whereas the performance of the “*BER Non-Robust*” algorithm is degraded.

Comparing in Fig. 6(b) the MMSE and the minimum BER algorithms, it is shown that the “*MMSE Robust*” algorithm only exhibits a good performance in terms of uncoded BER for low channel uncertainties. Note that for $M = 3$, the performance of the “*MMSE Robust*” algorithm is very close to the “*BER Robust*” algorithms when the CSI quality is high (i.e., $\rho < 0.2$), but it performs worse than open-loop when the CSI quality degrades. When the transmission rate is decreased (i.e., $M = 1$), the difference in the performances is reduced, but still, when $\rho \rightarrow 1$, the “*MMSE Robust*” algorithm degrades rapidly. According to the previous results, and recalling that the “*BER Robust*” algorithm has been computed applying an iterative solution since the closed-form solution in (63) is suboptimum when $M > 1$, it can be concluded that when CSI quality at the transmitter is good (i.e., $\rho < 0.2$), the “*MMSE Robust*” algorithm, which is computed by means of the closed-form solution given by (41) and (42), is an alternative to the “*BER Robust*” algorithm in terms of uncoded BER. On the contrary, when the CSI quality is decreased, the performance of the “*MMSE Robust*” algorithm degrades rapidly, especially for $M = 3$, and thus, the best solution is the “*BER Robust*” algorithm that adapts to the channel uncertainty and tends to the open-loop solution.

VI. CONCLUSIONS

This paper has presented a Bayesian approach for the design of linear precoding schemes that are robust to channel estimation errors. The proposed linear transformations have been designed according to a minimum MSE and a minimum uncoded BER criterion, both of them subject to an average power constraint across all antennas and subcarriers.

Closed-form solutions that are based on the SVD of the so-called equivalent channel have been derived for both algorithms. The prefiltering matrix derived for the minimum MSE criterion has been shown to be optimum, whereas the matrix structure given for the minimum BER cost function is optimum when one symbol per subcarrier is transmitted (i.e., $M = 1$) and SNR is high. The algorithms converge to the results published previously in the literature in the extreme cases of very high or very low uncertainty in the CSI. Besides, it has been shown that the BER design can be regarded as a reconfigurable

algorithm that adapts the transmitted signal to the available channel knowledge, providing a solution that converges to the open-loop design (i.e., the same power is allocated across all antennas and subcarriers) for the case of no channel knowledge and to the closed-loop design with perfect CSI for the case of no uncertainty. The minimum BER design has been formulated based on a generalized exponential bound of the function $Q(\sqrt{x})$ that includes the Chernoff bound as a particular case.

The two algorithms have been compared in terms of uncoded BER and coded BER for the specific channel code described in the HIPERLAN/2 standard using a ML detector at the receiver. Moreover, the capacity and the cut-off rate have also been used as measures of the coded BER performance, which are independent of the specific channel code. The results have shown that the robust algorithms exhibit a lower sensitivity to channel estimation errors when compared with the nonrobust techniques. In terms of uncoded BER, it has been proved that although the algorithm that minimizes the BER has the best performance, the MMSE algorithm exhibits a good performance for low uncertainty in the CSI. In terms of coded BER, it has been shown that the robust algorithm that minimizes the BER outperforms the rest of algorithms, which is a conclusion that has been compared with the results given by the capacity and cut-off rate parameters.

APPENDIX A PROOF OF (32)

This Appendix proves that the structures given in (32) for the matrices \mathbf{F}_k and \mathbf{G}_k , which are based on the SVD of the equivalent channel, minimize the robust cost function (31). Denote the SVD of \mathbf{F}_k , \mathbf{G}_k and $\hat{\mathbf{H}}_k^{\text{eq}}$ matrices as

$$\mathbf{F}_k = \mathbf{U}_F \mathbf{\Lambda}_F \mathbf{V}_F^H; \quad \mathbf{G}_k = \mathbf{U}_G \mathbf{\Lambda}_G \mathbf{V}_G^H; \quad \hat{\mathbf{H}}_k^{\text{eq}} = \mathbf{U}_H \mathbf{\Lambda}_H \mathbf{V}_H^H \quad (69)$$

where the subcarrier index “ k ” and the label “eq” are omitted for the sake of clarity in the notation.

Differentiating (31) with respect to matrix \mathbf{G}_k^H and equating the result to zero, the function to minimize is solved for \mathbf{G}_k to obtain

$$\mathbf{G}_k = \mathbf{F}_k^H \mathbf{H}_k^H (\mathbf{H}_k \mathbf{F}_k \mathbf{F}_k^H \mathbf{H}_k^H + \sigma_n^2 \mathbf{I} + \omega \sigma_n^x \text{Tr}\{\mathbf{F}_k \mathbf{F}_k^H\} \mathbf{I})^{-1}. \quad (70)$$

Substituting the isolated matrix \mathbf{G}_k in (31), the new cost function, that only depends on the linear transformation \mathbf{F}_k , is

$$\begin{aligned} \xi_k = & \left[\frac{\sigma_n^2}{\sigma_x^2} + \omega \text{Tr}\{\mathbf{F}_k \mathbf{F}_k^H\} \right] \\ & \times \text{Tr} \left\{ \left(\hat{\mathbf{H}}_k^{\text{eq}} \mathbf{F}_k \mathbf{F}_k^H \hat{\mathbf{H}}_k^{\text{eq}H} + \mathcal{K}_1 \mathbf{I} \right)^{-1} \right\} \\ & - \mu \left[\left(\sum_{k=1}^K \text{Tr}\{\mathbf{F}_k^H \mathbf{F}_k\} \right) - P_0 \right] \end{aligned} \quad (71)$$

where the scalar $\mathcal{K}_1 = (\sigma_n^2/\sigma_x^2) + \omega \text{Tr}\{\mathbf{F}_k \mathbf{F}_k^H\}$. Next, the optimum linear transformation \mathbf{F}_k will be obtained differentiating

(71) with respect to the matrix \mathbf{F}_k^H and equating the result to zero. Isolating the matrix \mathbf{F}_k , we arrive at

$$\mathbf{F}_k = \mathcal{K}_2 \underbrace{\hat{\mathbf{H}}_k^{eqH} \left(\hat{\mathbf{H}}_k^{eq} \mathbf{F}_k \mathbf{F}_k^H \hat{\mathbf{H}}_k^{eqH} + \mathcal{K}_1 \mathbf{I} \right)^{-2}}_{\mathbf{P}_F} \hat{\mathbf{H}}_k^{eq} \mathbf{F}_k \quad (72)$$

where \mathcal{K}_2 is a nonrelevant constant, and the underbraced matrix \mathbf{P}_F is the projection matrix of \mathbf{F}_k . Substituting the SVDs defined in (69), the projection matrix \mathbf{P}_F must satisfy $\mathbf{P}_F = \mathbf{U}_F \mathbf{U}_F^H$, and consequently, the next equality must be accomplished:

$$\begin{aligned} \frac{1}{\mathcal{K}_2} \mathbf{U}_F \mathbf{U}_F^H &= \mathbf{V}_H \mathbf{\Lambda}_H^H \mathbf{U}_H^H \\ &\times \left(\mathbf{U}_H \mathbf{\Lambda}_H \mathbf{V}_H^H \mathbf{U}_F \mathbf{\Lambda}_F \mathbf{\Lambda}_F^H \mathbf{U}_F^H \mathbf{V}_H \mathbf{\Lambda}_H^H \mathbf{U}_H^H \right. \\ &\quad \left. + \mathcal{K}_1 \mathbf{I} \right)^{-2} \\ &\times \mathbf{U}_H \mathbf{\Lambda}_H \mathbf{V}_H^H \end{aligned} \quad (73)$$

which only holds if $\mathbf{V}_H^H \mathbf{U}_F = \mathbf{P}$ where \mathbf{P} is any permutation matrix. If the permutation matrix is set to the identity matrix, then $\mathbf{U}_F = \mathbf{V}_H$.

The structure for the matrix \mathbf{G}_k is derived substituting in (70) the SVDs defined in (69) for the particular case of $\mathbf{U}_F = \mathbf{V}_H$. After some manipulations, it is straightforward to verify that $\mathbf{V}_G = \mathbf{U}_H$ and $\mathbf{U}_G = \mathbf{V}_F$ satisfy the solution for \mathbf{G}_k . Note that the MSE is independent of the unitary matrix \mathbf{V}_F ; hence, for the sake of simplicity, $\mathbf{V}_F = \mathbf{I}$ has been selected in (32).

APPENDIX B

EXPONENTIAL EXPANSION OF THE $\mathcal{Q}(\sqrt{x})$ FUNCTION

As the use of the exact $\mathcal{Q}(\cdot)$ function derives in complex solutions [25], this paper proposes an expansion of the $\mathcal{Q}(\sqrt{x})$ function in the neighborhood of the point $x = a$. By a generalized exponential expansion, the $\mathcal{Q}(\sqrt{x})$ function can be expressed as

$$\mathcal{Q}(\sqrt{x}) \simeq \delta e^{-\alpha x} \quad (74)$$

where constants α and δ can be chosen to set certain constraints. The expressions derived in this paper can be applied for any value of δ and α , including the Chernoff bound as a particular case for $\alpha = 1/2$ and $\delta = 1/2$. However, it is possible to derive a lower bound for $\mathcal{Q}(\sqrt{x})$ that is very tight in a wide range of values around a . This bound can be obtained rewriting $\mathcal{Q}(\sqrt{x}) = e^{\ln(\mathcal{Q}(\sqrt{x}))}$ and doing a Taylor expansion of the exponent in the neighborhood of the point a :

$$\ln(\mathcal{Q}(\sqrt{x})) \simeq \ln(\mathcal{Q}(\sqrt{a})) - \frac{e^{-|a|/2}}{2\sqrt{2\pi a} \mathcal{Q}(\sqrt{a})} (x - a). \quad (75)$$

Then, α and δ are given by

$$\alpha = \frac{e^{-|a|/2}}{2\sqrt{2\pi a} \mathcal{Q}(\sqrt{a})}; \quad \delta = \mathcal{Q}(\sqrt{a}) e^{\alpha a}. \quad (76)$$

It is straightforward to show that (76) provides a lower bound for $\mathcal{Q}(\cdot)$. As the exponential function is monotonically increasing, the lower bound is ensured if the Taylor expansion in (75) is always lower than $\ln(\mathcal{Q}(\sqrt{x}))$: a condition that always holds since $\ln(\mathcal{Q}(\sqrt{x}))$ is a concave function.

APPENDIX C

PROOF OF (57)

In this Appendix, we show that for high SNR, the following inequality holds:

$$\begin{aligned} \frac{1}{\mathcal{N}^{2M}} \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \mathcal{Q} \left(\sqrt{\frac{|\mathbf{\Lambda}_k \mathbf{\Phi}_k \mathbf{T} d_{ijk}|^2}{2\sigma_n^2}} \right) \\ \leq \frac{1}{\mathcal{N}^{2M}} \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \mathcal{Q} \left(\sqrt{\frac{|\mathbf{\Lambda}_k \mathbf{\Phi}_k d_{ijk}|^2}{2\sigma_n^2}} \right). \end{aligned} \quad (77)$$

Proof: Let s_i and s_j be two elements of a given constellation, let $R(b, d)$ denote the number of symbol pairs (s_i, s_j) that differ in b bits, and let $d = |s_i - s_j|$ be any of the D possible values of the difference between two constellation symbols. Let \mathbf{i}_m denote the m th column of the identity matrix of size M . Under the high SNR condition (assumption A1), the average BER can be approximated taking into account only the pairwise error probability between those pairs of symbols that differ in one coefficient. Thus, for high SNR, the BER for the k th subcarrier can be written as follows:

$$\begin{aligned} \overline{\text{BER}}_k &= \frac{1}{\mathcal{N}^{2M}} \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \\ &\times \mathcal{Q} \left(\sqrt{\frac{|\mathbf{\Lambda}_k \mathbf{\Phi}_k \mathbf{T} (\mathbf{x}_{ki} - \mathbf{x}_{kj})|^2}{2\sigma_n^2}} \right) \\ &\simeq \frac{1}{\mathcal{N}^{2M}} \sum_{m=1}^M \sum_{b=1}^{\log_2 \mathcal{N}} \sum_{l=1}^D b R(b, d_l) \\ &\times \mathcal{Q} \left(\sqrt{\frac{|\mathbf{\Lambda}_k \mathbf{\Phi}_k \mathbf{T} \mathbf{i}_m d_l|^2}{2\sigma_n^2}} \right) \end{aligned} \quad (78)$$

and reduces to the following expression for $\mathbf{T} = \mathbf{I}$:

$$\begin{aligned} \overline{\text{BER}}_k &= \frac{1}{\mathcal{N}^{2M}} \sum_{i=1}^{\mathcal{N}^M} \sum_{j=1}^{\mathcal{N}^M} B(\mathbf{x}_{ki}, \mathbf{x}_{kj}) \\ &\times \mathcal{Q} \left(\sqrt{\frac{|\mathbf{\Lambda}_k \mathbf{\Phi}_k (\mathbf{x}_{ki} - \mathbf{x}_{kj})|^2}{2\sigma_n^2}} \right) \\ &\simeq \frac{1}{\mathcal{N}^{2M}} \sum_{m=1}^M \sum_{b=1}^{\log_2 \mathcal{N}} \sum_{l=1}^D b R(b, d_l) \\ &\times \mathcal{Q} \left(\sqrt{\frac{|\mathbf{\Lambda}_k \mathbf{\Phi}_k d_l \mathbf{i}_m|^2}{2\sigma_n^2}} \right). \end{aligned} \quad (79)$$

Comparing (78) and (79), it is clear that proving the inequality in (77) is equivalent to proving a similar expression for the $\mathcal{Q}(\cdot)$

function that appears in them. It can be shown [8], [17] that the DFT matrix satisfies that $|\Lambda_k \Phi_k \mathbf{T} \mathbf{i}_m|$ is constant for all m . Therefore

$$\begin{aligned}
 & \frac{1}{M} \sum_{m=1}^M \mathcal{Q} \left(\sqrt{\frac{d_l^2}{2\sigma_n^2} |\Lambda_k \Phi_k \mathbf{T} \mathbf{i}_m|^2} \right) \\
 &= \mathcal{Q} \left(\sqrt{\frac{d_l^2}{2\sigma_n^2} \frac{1}{M} \sum_{m=1}^M |\Lambda_k \Phi_k \mathbf{T} \mathbf{i}_m|^2} \right) \\
 &= \mathcal{Q} \left(\sqrt{\frac{d_l^2}{2\sigma_n^2} \frac{1}{M} \text{Tr} \left\{ \Lambda_k \Phi_k \mathbf{T} \mathbf{T}^H \Phi_k^H \Lambda_k^H \right\}} \right) \\
 &= \mathcal{Q} \left(\sqrt{\frac{d_l^2}{2\sigma_n^2} \frac{1}{M} \text{Tr} \left\{ \Lambda_k \Phi_k \Phi_k^H \Lambda_k^H \right\}} \right) \\
 &= \mathcal{Q} \left(\sqrt{\frac{d_l^2}{2\sigma_n^2} \frac{1}{M} \sum_{m=1}^M |\Lambda_k \Phi_k \mathbf{i}_m|^2} \right) \\
 &\leq \frac{1}{M} \sum_{m=1}^M \mathcal{Q} \left(\sqrt{\frac{d_l^2}{2\sigma_n^2} |\Lambda_k \Phi_k \mathbf{i}_m|^2} \right) \quad (80)
 \end{aligned}$$

where the equality in second line applies because \mathbf{T} is a unitary matrix, and last inequality is a result of the application of Jensen's inequality [26] to the convex function $\mathcal{Q}(\sqrt{x})$. The combination of (80) and (78) and (79) leads to (77). ■

REFERENCES

- [1] R. G. Gallager, *Information Theory and Reliable Communications*. New York: Wiley, 1968.
- [2] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, Mar. 1998.
- [3] F. Rey, M. Lamarca, and G. Vázquez, "Transmitter channel tracking for optimal power allocation," in *Proc. ICASSP*, Salt Lake City, UT, May 2001, pp. 2137–2140.
- [4] S. Barbarossa and A. Scaglione, "Time-varying channels," in *Signal Processing Advances in Wireless Communications*, G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, Eds. Englewood Cliffs, NJ: Prentice-Hall, 2000, vol. II, Trends in Single- and Multiuser Systems, ch. 1.
- [5] A. M. Teherani, A. Hassibi, J. Cioffi, and S. Boyd, "An implementation of discrete multi-tone over slowly time-varying MIMO channels," in *Proc. Globecom*, Nov. 1998, pp. 2806–2811.
- [6] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 3, pp. 1019–1032, Apr. 1999.
- [7] E. N. Onggosanusi, A. M. Sayeed, and D. Van Veen, "Optimal antenna diversity signaling for wide-band systems utilizing channel side information," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 341–353, Feb. 2002.
- [8] Y. Ding, T. N. Davidson, J. Zang, Z. Luo, and K. M. Wong, "Minimum BER block precoders for zero-forcing equalization," in *Proceedings of ICASSP*, Orlando, FL, May 2002, pp. 2261–2264.
- [9] 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 J. Sel. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.
- [10] G. Jongren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.
- [11] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Lett.*, vol. 6, pp. 239–241, Jun. 2002.
- [12] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*. New York: Wiley, 1999.
- [13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [14] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.
- [15] A. Kuzminskiy, M. Lamarca, J. A. Lopez, I. Modonesi, and F. Rey, (2002, Nov., Deliverable D3.2.1 of FITNESS IST-2000-30116 project) Performance Analysis of Re-Configurable MTMR Transceivers for WLAN. [Online]. Available: <http://www.telecom.ece.ntua.gr/fitness>
- [16] D. Rainish, "Diversity transform for fading channels," *IEEE Trans. Commun.*, vol. 44, no. 12, pp. 1653–1661, Dec. 1996.
- [17] M. Lamarca and G. Vázquez, "Transform modulations for mobile communications," in *Proc. PIMRC*, Helsinki, Finland, Sep. 1997, pp. 462–466.
- [18] J. Boutros and E. Viterbo, "Signal space diversity: A power- and bandwidth- efficient diversity technique for the rayleigh fading channel," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1453–1467, Jul. 1998.
- [19] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.
- [20] E. N. Onggosanusi, A. M. Sayeed, and B. D. Van Veen, "High throughput wideband space-time signaling using channel state information," in *Proc. ICASSP*, Salt Lake City, UT, May 2001, pp. 2421–2424.
- [21] "Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Physical (PHY) Layer," ETSI, ETSI TS 101 475 v1.2.2, 2001.
- [22] A. M. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *Proc. Veh. Technol. Conf.*, Boston, MA, Sep. 2000, pp. 473–478.
- [23] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [24] T. Tung, K. Yao, and R. E. Hudson, "Channel estimation and adaptive power allocation for performance and capacity improvement of multiple-antenna OFDM systems," in *Proc. SPAWC*, Taoyuan, Taiwan, R.O.C., Mar. 2001, pp. 82–85.
- [25] J. G. Proakis, *Digital Communications*, Third ed. New York: McGraw-Hill, 1995.
- [26] T. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.



Francesc Rey (S'98) was born in Barcelona, Spain, in 1973. He received the electrical engineering degree from the Technical University of Catalonia (UPC), Barcelona, in 1997. He is currently pursuing the Ph.D. degree in electrical engineering with the Signal Processing and Communications Group, UPC.

In 1998, he joined the Department of Signal Theory and Communications, UPC, where he has been an Assistant Professor since April 2001. From 1998 to 2001, he held a grant from the Generalitat de Catalunya in support of the Ph.D. thesis. His current research interests lie in areas of signal processing and communications, with particular emphasis on wireless MIMO channels, multicarrier modulation and blind channel estimation, and optimal equalization. He has been involved in many research projects in the framework of the research programs of the European Union, the European Space Agency, and industry.



Meritxell Lamarca (M'97) was born in Barcelona, Spain, in 1967. She received the M.Sc. and Ph.D. degrees in telecommunication engineering from the Technical University of Catalonia (UPC), Barcelona, in 1992 and 1997, respectively.

In 1992, she joined the Department of Signal Theory and Communications, UPC, and soon after, she became a Lecturer. Since 1997, she has been an Associate Professor at UPC. Her general interests are signal processing and digital communications. She conducts research activities in the design of advanced frontends for digital communications combining coding, spatial, and frequency-time diversity, with particular emphasis in turbo detection schemes. She has been involved in many research projects in wireless and satellite communications under the research programs of the European Union and the European Space Agency.



Gregori Vazquez (SM'98) was born in Barcelona, Spain, in 1961. He received the M.S. and Ph.D. degrees in telecommunications engineering from the Technical University of Catalonia (UPC), Barcelona, in 1984 and 1988, respectively.

He is a Professor with the Department of Signal Theory and Communications, UPC. His general interests are Statistical Signal Processing and Digital Communications.

Dr. Vazquez received the 2003 Best Paper Award from the IEEE Signal Processing Society. He has served as Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING since 1999. He is also a member of the Editorial Board of the IEEE SIGNAL PROCESSING MAGAZINE and a member of the Signal Processing for Communications Technical Committee. He is the Chairman of the Research Program on Communications of the Spanish Science and Technology Ministry.