Optimum Linear Joint Transmit-Receive Processing for MIMO Channels with QoS Constraints

Daniel Pérez Palomar, *Member, IEEE*, Miguel Angel Lagunas, *Fellow, IEEE*, and John M. Cioffi, *Fellow, IEEE*

***Abstract*—This paper considers vector communications through multiple-input multiple-output (MIMO) channels with a set of quality of service (QoS) requirements for the simultaneously established substreams. Linear transmit-receive processing (also termed linear precoder at the transmitter and linear equalizer at the receiver) is designed to satisfy the QoS constraints with minimum transmitted power (the exact conditions under which the problem becomes unfeasible are given). Although the original problem is a complicated nonconvex problem with matrix-valued variables, with the aid of majorization theory, we reformulate it as a simple convex optimization problem with scalar variables. We then propose a practical and efficient multilevel water-filling algorithm to optimally solve the problem for the general case of different QoS requirements. The optimal transmit-receive processing is shown to diagonalize the channel matrix only after a very specific prerotation of the data symbols. For situations in which the resulting transmit power is too large, we give the precise way to relax the QoS constraints in order to reduce the required power based on a perturbation analysis. We also propose a robust design under channel estimation errors that has an important interest for practical systems. Numerical results from simulations are given to support the mathematical development of the problem.**

***Index Terms*—Array signal processing, beamforming, joint transmit-receive equalization, linear precoding, MIMO channels, space-time filtering, water-filling.**

I. Introduction

COMMUNICATIONS over multiple-input multiple-output (MIMO) channels have recently gained considerable attention [1]–[4]. They arise in many different scenarios such as when a bundle of twisted pair copper wires in digital subscriber lines (DSL) is treated as a whole [1], when multiple antennas are used at both sides of a wireless link [3], or simply when a time-dispersive or frequency-selective channel is properly modeled for block transmission by using, for example, transmit and receive filterbanks [4]. In particular, MIMO channels arising from the use of multiple antennas at both the transmitter and the receiver have recently attracted a significant interest because they provide an important increase in capacity over single-input single-output (SISO) channels under some uncorrelation conditions [5], [6].

The transmitter and the receiver may or may not have channel state information (CSI), although for slowly varying channels, CSI is generally assumed at the receiver (CSIR). Many publications have dealt with the case of no CSI at the transmitter (CSIT) such as the popular space-time coding techniques [7]–[9]. Another great block of research has been devoted to the situation with CSIT [2]–[4], [10]. We focus on the latter and, in particular, when linear processing is utilized for the sake of complexity.

In many situations, the transmitter is constrained on its average transmit power to limit the interference level of the system, and the transmit-receive processing is designed to somehow maximize the quality of the communication. Several different criteria have been utilized as a means to measure the quality of the link and to design the system such as the minimization of the (weighted) trace of the mean-square error (MSE) matrix [11], [12], [2], [4], [10], the maximization of the signal to interference-plus-noise ratio (SINR) with a zero-forcing (ZF) constraint [4], the minimization of the determinant of the MSE matrix [13], and the minimization of the average bit error rate (BER) [14], [15]. A general unifying framework was developed in [15] to treat all such criteria by classifying the variety of objective functions into Schur-concave and Schur-convex functions. In [16], some criteria were considered under a peak power constraint.

In other situations, however, the approach of maximizing the quality subject to a transmit power constraint may not be the desired objective, as argued next. From a system level point of view, it may be interesting to consider the opposite formulation of the problem. Given that several substreams are to be established through the MIMO channel and that each substream requires a (possibly different) quality of service (QoS), the communication system wishes to satisfy these QoS constraints with minimum transmitted power. This type of design has been considered in the literature mainly for multiuser scenarios, in which multiple distributed users coexist, and joint multiuser processing cannot be assumed at one side of the link since the users are geographically distributed. In [17], a multiuser scenario with a multielement base station was considered, and optimal beamvectors and power allocation were obtained for the uplink and downlink under QoS constraints in terms of

Manuscript received May 10, 2002; revised June 5, 2003. Part of the work was presented at the 36th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, Nov. 3–6, 2002 [23]. This work was supported in part by the Spanish Government (CICYT) under Grants TIC2001-2356, TIC2002-04594, FIT-070000-2000-649 (MEDEA + A105 UniLAN); the Catalan Government (DURSI) under Grants 1999FI 00588 and 2001SGR 00268; and Samsung Advanced Institute of Technology. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Zhi Ding.

D. P. Palomar was with the Department of Signal Theory and Communications, Technical University of Catalonia (UPC), Campus Nord UPC, 08034 Barcelona, Spain. He is now with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08844 USA (e-mail: danielp@princeton.edu).

M. A. Lagunas is with the Telecommunications Technological Center of Catalonia (CTTC), 08034 Barcelona, Spain, and also with Department of Signal Theory and Communications, Technical University of Catalonia (UPC), Campus Nord UPC, 08034 Barcelona, Spain (e-mail: m.a.lagunas@cttc.es).

J. M. Cioffi is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305-9515 USA (e-mail: cioffi@stanford.edu).

Digital Object Identifier 10.1109/TSP.2004.826164

SINR. In [18] and [19], the same problem was solved under the powerful framework of convex optimization theory, which allows the introduction of additional constraints (for example to control the dynamic range of the transmitted signal or to increase the robustness against channel estimation errors). In [20], the problem was generalized to the case of having multiple antennas at both sides of the link, although a global optimal solution was not found due to the nonconvexity of the problem (a suboptimal iterative optimization approach was taken). A similar signal model arising from a single-antenna multiuser multicarrier CDMA system was treated in [21], in which a suboptimal (due to the nonconvexity of the problem) gradient-type algorithm was used. In [22], a single-antenna multiuser CDMA system was characterized in terms of user capacity by optimally designing the CDMA codes of the users with SINR requirements.

The approach in this paper is similar to the aforementioned examples in that we deal with the optimization of a system subject to QoS requirements and different in that it is assumed that joint processing is possible at both the transmitter and the receiver. Of course, the previously considered multiuser scenario in which the users are geographically distributed is no longer valid[1] (note that the considered model does not correspond to a multiple-access channel since both sides of the link are allowed to cooperate). Hence, the considered scenario is just a point-to-point communication system where more than one substream are simultaneously established with (possibly different) QoS requirements.[2] In fact, this situation happens naturally in spectrally efficient systems that are designed to approach the capacity of the MIMO channel, as we now describe.

The capacity-achieving solution dictates that the channel matrix has to be diagonalized and that the power at the transmitter has to be allocated following a water-filling distribution on the channel eigenmodes [24], [3], [25]. In theory, this solution has the implication that an ideal Gaussian code should be used on each channel eigenmode according to its allocated power [24]. In practice, however, Gaussian codes are substituted with simple (and suboptimal) signal constellations and practical (and suboptimal) coding schemes (if any). Therefore, the uncoded part of a practical system basically transmits a set of (possibly different) constellations simultaneously through the MIMO channel. In light of these observations, an interesting way to design the uncoded part of a communication system is based on the gap approximation (e.g. [26]), which basically gives the optimal bit distribution (following a water-filling allocation similar to that of the capacity-achieving solution) under the assumption that practical constellations such as quadrature amplitude modulation (QAM) of different sizes are used. Of course, in order to reduce the complexity of a system employing different constellations and codes, it can be constrained to use the same constellation and code in all channel eigenmodes (possibly optimizing the utilized bandwidth to transmit only over those eigenmodes with a sufficiently high gain), i.e., an equal-rate transmission. Examples of this pragmatic and simple solution are found in the European standard HIPERLAN/2 [27] and in the U.S. standard IEEE 802.11 [28] for wireless local area networks (WLANs). In any case, once the constellations to be used at each of the substreams are known, the system can be further optimized such that each established substream satisfies, for example, a given BER.

Hence, this paper considers the transmission of a vector of data symbols through a channel matrix subject to (possibly different) QoS constraints given in terms of MSE, SINR, or BER. The coding and modulation schemes used on the different substreams are assumed given and are not involved in the optimization process (therefore, different services can employ different signal constellations and different error control coding schemes yielding a general multirate communication system). Linear transmit-receive processing is designed to satisfy the QoS constraints with minimum transmitted power (the exact conditions under which the problem becomes unfeasible are given). The original formulation of the problem is a complicated nonconvex optimization problem with matrix-valued variables. With the aid of majorization theory, however, the problem can be reformulated as a simple convex problem with scalar variables. We then propose a practical and efficient multilevel water-filling algorithm that obtains an optimal solution for the general case of different QoS requirements among the established substreams. The optimal solution is shown to diagonalize the channel matrix only after a very specific prerotation of the data symbols. In some situations, when the transmit power required to satisfy the QoS constraints results too large, it may be desirable to relax some QoS requirements. By using a sensitivity analysis of the perturbed system, we obtain the precise way to relax the QoS constraints in order to reduce the power needed. We also propose a robust design under channel estimation errors, which has an important interest for practical systems.

The paper is structured as follows. In Section II, the signal model is introduced. Section III is devoted to a brief introduction to majorization theory on which the paper is strongly based. The optimal receive matrix is obtained in Section IV. The main result of the paper is given in Sections V and VI, where the optimal transmit matrix is obtained for equal and different MSE-based QoS requirements, respectively. A simple suboptimal solution of practical interest is also considered in Section VII. Practical considerations such as the relaxation of the problem and the robust design under channel estimation errors are considered in Section VIII. Numerical results obtained from simulations are given in Section IX. Finally, in Section X, the main conclusions of the paper are given.

The following notation is used. Boldface uppercase letters denote matrices, boldface lowercase letters denote column vectors, and italics denote scalars. $\mathbb{C}^{m\times n}$ represents the $m \times n$ complex field. The superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote transpose, complex conjugate, and Hermitian operations, respectively. $[\mathbf{X}]_{i,j}$ (also $[\mathbf{X}]_{ij}$) and $[\mathbf{X}]_{:,j}$ denote the (ith, jth) element and jth column of matrix $\mathbf{X}$, respectively. $\mathbf{A} \geq \mathbf{B}$ denotes that $\mathbf{A}-\mathbf{B}$ is positive semidefinite, and $\mathbf{a} \geq \mathbf{b}$ refers to the elementwise relation $a_i \geq b_i$. By $\mathbf{d}(\mathbf{X})$ and $\boldsymbol{\lambda}(\mathbf{X})$,

[1]Some very specific multiuser scenarios allow for cooperation at both sides of the link such as in DSL systems, where both ends of the MIMO system are each terminated in a single physical location, e.g., links between central offices and remote terminals (and also private networks) [1] (see also [23]).

[2]A very simple example of a single-user communication with several established substreams, each with a different QoS requirement, arises when the user wants to transmit simultaneously different services, such as audio and video (since video typically requires a higher SINR than audio).

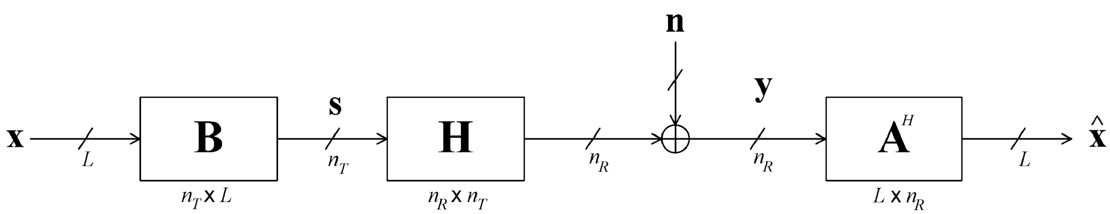

Fig. 1. Scheme of a general MIMO communication system.

we denote the vectors containing the diagonal elements and eigenvalues of matrix $\mathbf{X}$, respectively. With diag$(\{\mathbf{X}_k\})$, we denote a block-diagonal matrix with diagonal blocks given by the set $\{\mathbf{X}_k\}$. The trace and determinant of a matrix are denoted by Tr$(\cdot)$ and $|\cdot|$, respectively. $\mathbb{E}[\cdot]$ denotes mathematical expectation. We define $(x)^+ \triangleq \max(0,x)$ (for matrices it is defined elementwise).

II. Signal Model

Mathematically, a MIMO channel is conveniently and compactly represented by a channel matrix. This allows the utilization of an elegant signal vector notation. We consider the simultaneous transmission of L symbols through a general MIMO communication channel with n_T transmit and n_R receive dimensions (see Fig. 1). The formulation in this paper allows values of L greater than the rank of the channel matrix.[3] The signal model is

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n} \tag{1}$$

where $\mathbf{s} \in \mathbb{C}^{n_T \times 1}$ is the transmitted vector, $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ is the channel matrix, $\mathbf{y} \in \mathbb{C}^{n_R \times 1}$ is the received vector, and $\mathbf{n} \in \mathbb{C}^{n_R \times 1}$ is a zero-mean circularly symmetric complex Gaussian interference-plus-noise vector with arbitrary covariance matrix $\mathbf{R}_n$, i.e., $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_n)$.

Considering the utilization of a transmit linear processing matrix (commonly termed *linear precoder*), the transmitted vector can be written as

$$\mathbf{s} = \mathbf{B}\mathbf{x} \tag{2}$$

where $\mathbf{B} \in \mathbb{C}^{n_T \times L}$ is the transmit matrix, and $\mathbf{x} \in \mathbb{C}^{L \times 1}$ is the data vector that contains the L symbols to be transmitted (see Fig. 1). Note that the ith column of matrix $\mathbf{B}$ can be regarded as the beamvector associated to the ith data symbol x_i. The total average transmitted power (in units of energy per transmission) is

$$P_T = \mathbb{E}\left[\|\mathbf{s}\|^2\right] = \text{Tr}(\mathbf{B}\mathbf{B}^H) \tag{3}$$

where we have assumed zero-mean unit-energy uncorrelated symbols,[4] i.e., $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_L$.

Similarly, assuming a receive linear processing matrix (usually referred to as *linear equalizer*), the estimated data vector is

$$\hat{\mathbf{x}} = A^H \mathbf{y} \tag{4}$$

where $\mathbf{A}^H \in \mathbb{C}^{L \times n_R}$ is the receive matrix (see Fig. 1).

[3]When $L > \text{rank}(\mathbf{H})$, depending on the particular problem at hand, the problem may be unfeasible, i.e., without solution (c.f. Theorems 1 and 2).

[4]In the case of having correlated symbols, a prewhitening operation can be performed prior to precoding at the transmitter and the corresponding inverse operation at the receiver after the equalizer.

A. Multicarrier Signal Model

The previous general and abstract MIMO model can be readily particularized to the typical system in which there are n_T transmit and n_R receive dimensions (representing either antennas in a wireless system of copper wires in a DSL system), and the channel is frequency-selective. To deal easily with the frequency-selectivity of the channel, we take a multicarrier approach and assume the channel fixed during the transmission block. The signal model at each carrier k is

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{s}_k + \mathbf{n}_k \quad 1 \le k \le N \tag{5}$$

where N is the number of carriers, and $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{n_k})$ (noise from different carriers is assumed uncorrelated). To take into account the possibility of transmitting a different number of symbols through each carrier, we consider that L_k symbols are transmitted at the kth carrier, and therefore, a total of $L_T = \sum_{k=1}^{N} L_k$ symbols are transmitted. Note that the total number of transmit and receive dimensions is now $n_T N$ and $n_R N$, respectively. Regarding the linear processing at the transmitter and at the receiver, two different cases are considered, as we now describe (c.f. [15]):

- *Carrier-noncooperative scheme*: This is when an independent signal processing per carrier is used [see Fig. 2(a)]. The transmit and receive signal model is $\mathbf{s}_k = \mathbf{B}_k\mathbf{x}_k$ (with a total average transmitted power of $P_T = \sum_k \text{Tr}(\mathbf{B}_k\mathbf{B}_k^H)$) and $\hat{\mathbf{x}}_k = \mathbf{A}_k^H\mathbf{y}_k$.
- *Carrier-cooperative scheme*: This is a more general linear processing scheme that allows for cooperation among carriers (see Fig. 2(b)). The signal model is obtained [similarly to (1)–(4)] by stacking the vectors corresponding to all carriers (e.g., $\mathbf{x}^T = [\mathbf{x}_1^T, \cdots, \mathbf{x}_N^T]$), by considering global transmit and receive matrices $\mathbf{B} \in \mathbb{C}^{(n_T N) \times L_T}$ and $\mathbf{A}^H \in \mathbb{C}^{L_T \times (n_R N)}$, and by defining the global channel as $\mathbf{H} = \text{diag}(\{\mathbf{H}_k\}) \in \mathbb{C}^{(n_R N) \times (n_T N)}$. This general block processing scheme was used in [3] to obtain a capacity-achieving system.

It is important to realize that the carrier-noncooperative signal model can be obtained from the more general carrier-cooperative signal model by setting $\mathbf{B} = \text{diag}(\{\mathbf{B}_k\})$ and $\mathbf{A} = \text{diag}(\{\mathbf{A}_k\})$, i.e., by imposing a block-diagonal structure on $\mathbf{B}$ and $\mathbf{A}$. In fact, it is this block-diagonal structure that makes the carrier-noncooperative scheme less general and, therefore, with a worse performance than the carrier-cooperative one, as will be seen in the numerical simulations of Section IX. Intuitively speaking, the carrier-cooperative scheme can reallocate the symbols among the carriers in an intelligent way (e.g., if one carrier is in a deep fading, it will try to use

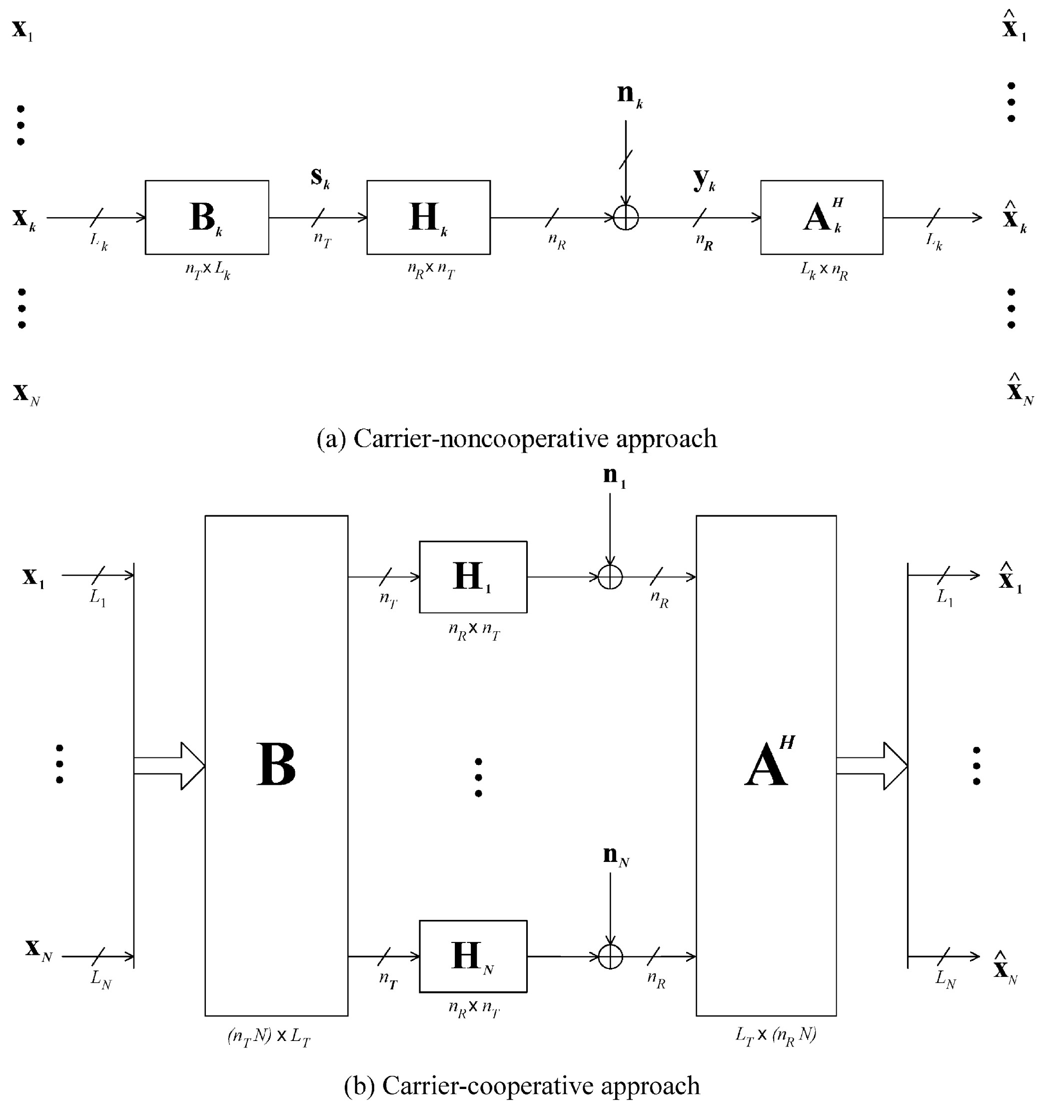

Fig. 2. Multicarrier MIMO channel (carrier-cooperative versus carrier-noncooperative approaches). (a) Carrier-noncooperative approach. (b) Carrier-cooperative approach.

other carriers instead), whereas the noncooperative scheme will always transmit L_k symbols through the kth carrier, regardless of the fading state of the carriers.

III. Majorization Theory

The transmit-receive processing design addressed in this paper results in complicated nonconvex constrained optimization problems that involve matrix-valued variables. Majorization theory is a key tool that will allow us to convert these problems into simple convex problems with scalar variables that can be optimally solved. In this section, we introduce the basic notion of majorization and state some basic results that will be needed in the sequel (see [29] for a complete reference of the subject). Those not interested in the details of the proofs of the results in this paper can safely skip the rest of this section.

Majorization makes precise the vague notion that the components of a vector $\mathbf{x}$ are "less spread out" or "more nearly equal" than are the components of a vector $\mathbf{y}$.

Definition 1: For any $\mathbf{x} \in \mathbb{R}^n$, let

$$x_{(1)} \leq \cdots \leq x_{(n)}$$

denote the components of vector $\mathbf{x}$ in increasing order.

Definition 2 [29, 1.A.1]: For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x}$ is majorized by $\mathbf{y}$ (or $\mathbf{y}$ majorizes $\mathbf{x}$) if[5]

$$\sum_{i=1}^{k} x_{(i)} \geq \sum_{i=1}^{k} y_{(i)}, \quad 1 \leq k \leq n-1$$

$$\sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}$$

and it is denoted by $\mathbf{x} \prec \mathbf{y}$ or, equivalently, by $\mathbf{y} \succ \mathbf{x}$.

Definition 3 [29, 1.A.2]: For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{x}$ is weakly majorized by $\mathbf{y}$ (or $\mathbf{y}$ weakly majorizes $\mathbf{x}$) if

$$\sum_{i=1}^{k} x_{(i)} \geq \sum_{i=1}^{k} y_{(i)}, \quad 1 \leq k \leq n$$

and it is denoted by $\mathbf{x} \prec^w \mathbf{y}$ or, equivalently, by $\mathbf{y} \succ^w \mathbf{x}$.

[5]Note that if the components of vectors $\mathbf{x}$ and $\mathbf{y}$ are assumed in decreasing order, the inequality sign of the majorization definition is "$\leq$" rather than "$\geq$" [29, 1.A.1].

Note that $\mathbf{x} \prec \mathbf{y}$ implies $\mathbf{x} \prec^w \mathbf{y}$; in other words, majorization is a more restrictive definition than weakly majorization.

Lemma 1 [29, p. 7]: For any $\mathbf{x} \in \mathbb{R}^n$, let $\mathbf{1} \in \mathbb{R}^n$ denote the constant vector with the ith element given by $1_i \triangleq \sum_{j=1}^n x_j/n$; then

$$\mathbf{x} \succ \mathbf{1}.$$

Lemma 1 is simply stating the obvious fact that a vector of equal components has the "least spread out" or the "most equal" components.

Lemma 2 [29, 9.B.1]: Let $\mathbf{R}$ be an $n \times n$ Hermitian matrix with diagonal elements denoted by the vector $\mathbf{d}$ and eigenvalues denoted by the vector $\boldsymbol{\lambda}$; then

$$\boldsymbol{\lambda} \succ \mathbf{d}.$$

Lemma 3 [29, 9.B.2]: For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ satisfying $\mathbf{x} \prec \mathbf{y}$, there exists a real symmetric (and therefore Hermitian) matrix with diagonal elements given by $\mathbf{x}$ and eigenvalues given by $\mathbf{y}$.

Lemma 3 is the converse of Lemma 2 (in fact, it is stronger than the converse since it guarantees the existence of a real symmetric matrix instead of just a Hermitian matrix). A recursive algorithm to obtain a matrix with a given vector of eigenvalues and vector of diagonal elements is indicated in [29, 9.B.2]. In this paper, however, we consider the practical and simple method given in [30, Sect. IV-A].

Corollary 1: For any $\boldsymbol{\lambda} \in \mathbb{R}^n$, there exists a real symmetric (and therefore Hermitian) matrix with equal diagonal elements and eigenvalues given by $\boldsymbol{\lambda}$.

Proof: The proof is straightforward from Lemmas 1 and 3. ■

As before, such a matrix can be obtained using the method given in [29, 9.B.2] or [30, Sect. IV-A]. However, for this particular case and allowing the desired matrix to be complex, it is easy to see that any unitary matrix $\mathbf{Q}$ satisfying the condition $|[\mathbf{Q}]_{ik}| = |[\mathbf{Q}]_{il}| \ \forall i,k,l$ provides a valid solution given by $\mathbf{Q}^H \text{diag}(\boldsymbol{\lambda})\mathbf{Q}$. As an example, the unitary discrete Fourier transform (DFT) matrix and the Hadamard matrix (when the dimensions are appropriate such as a power of two [31, p. 66]) satisfy this condition. Nevertheless, the method in [30, Sect. IV-A] has the nice property that the obtained matrix $\mathbf{Q}$ is real-valued and can be naturally decomposed (by construction) as the product of Givens rotations (where each term performs a single rotation). This simple structure plays a key role for practical implementation. Interestingly, an iterative approach to construct a matrix with equal diagonal elements and with a given set of eigenvalues was already obtained in [32], based on a sequence of rotations as well.

Lemma 4 [29, 5.A.9.a]: For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ satisfying $\mathbf{y} \succ^w \mathbf{x}$, there exists a vector $\mathbf{u}$ such that

$$\mathbf{u} \leq \mathbf{x} \quad \text{and} \quad \mathbf{y} \succ \mathbf{u}.$$

IV. Optimum Receive Processing

In this section, we obtain the optimum receive matrix $\mathbf{A}$ for a given transmit matrix $\mathbf{B}$. To start with, consider QoS constraints given in terms of MSE for each of the L established links or substreams through the MIMO channel:

$$\text{MSE}_i \triangleq \mathbb{E}\left[|\hat{x}_i - x_i|^2\right] \leq \rho_i, \quad 1 \leq i \leq L. \tag{6}$$

Note that $\rho_i < 1$ (otherwise, we could simply ignore the transmission of the ith symbol and set $\hat{x}_i = 0$ at the receiver) and that $\rho_i > 0$ (since a zero MSE can never be achieved unless the channel is noiseless). The MSE matrix is defined as the covariance matrix of the error vector (given by $\mathbf{e} \triangleq \hat{\mathbf{x}} - \mathbf{x}$):

$$\begin{aligned}\mathbf{E} &\triangleq \mathbb{E}\left[(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^H\right] \\ &= (\mathbf{A}^H\mathbf{H}\mathbf{B} - \mathbf{I})(\mathbf{B}^H\mathbf{H}^H\mathbf{A} - \mathbf{I}) + \mathbf{A}^H\mathbf{R}_n\mathbf{A}\end{aligned} \tag{7}$$

from which we can write $\text{MSE}_i = [\mathbf{E}]_{ii}$. The MSE of the ith substream depends only on the receive vector $\mathbf{a}_i$ (ith column of $\mathbf{A}$) and is independent of the other receive vectors (assuming $\mathbf{B}$ given and fixed). Therefore, it can be optimized independently of the others. Since the MSE_i is a quadratic function of $\mathbf{a}_i$, its minimum value can be found by setting its gradient to zero. The solution is given by

$$\mathbf{A} = (\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{R}_n)^{-1}\mathbf{H}\mathbf{B}. \tag{8}$$

Expression (8) is the linear minimum MSE (LMMSE) filter or *Wiener filter* [33] and minimizes simultaneously all the diagonal elements of the MSE matrix $\mathbf{E}$, as can be seen by "completing the squares" as follows (from (7)):

$$\begin{aligned}\mathbf{E} = &\left(\mathbf{A} - (\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{R}_n)^{-1}\mathbf{H}\mathbf{B}\right)^H (\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{R}_n) \\ &\times \left(\mathbf{A} - (\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{R}_n)^{-1}\mathbf{H}\mathbf{B}\right) \\ &+ \mathbf{I} - \mathbf{B}^H\mathbf{H}^H(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{R}_n)^{-1}\mathbf{H}\mathbf{B} \\ \geq &\, \mathbf{I} - \mathbf{B}^H\mathbf{H}^H(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{R}_n)^{-1}\mathbf{H}\mathbf{B}\end{aligned} \tag{9}$$

where we have used the fact that $\mathbf{X} + \mathbf{Y} \geq \mathbf{X}$ when $\mathbf{Y}$ is positive semidefinite. The lower bound is clearly achieved by (8). The concentrated MSE matrix is obtained by plugging (8) into (7) [or directly from (9)] as

$$\begin{aligned}\mathbf{E} &= \mathbf{I} - \mathbf{B}^H\mathbf{H}^H(\mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{R}_n)^{-1}\mathbf{H}\mathbf{B} \\ &= (\mathbf{I} + \mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}\end{aligned} \tag{10}$$

where $\mathbf{R}_H \triangleq \mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H}$, and we have used the matrix inversion lemma.[6]

Thus, for any given feasible transmit matrix $\mathbf{B}$ (a $\mathbf{B}$ such that there exists some $\mathbf{A}$ with which the QoS constraints can be satisfied), expression (8) will always give a feasible solution [clearly, if the $\mathbf{A}$ of (8) does not satisfy some of the QoS constraints, no other $\mathbf{A}$ will]. Note that in the case where (8) produces an MSE matrix $\mathbf{E}$ in which some QoS constraints are satisfied with strict inequality, there must be some other feasible solution; however, we stick to (8) since it guarantees that for any feasible $\mathbf{B}$, it will always yield a feasible solution.

[6]Matrix Inversion Lemma: $(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{D}\mathbf{A}^{-1}$.

The QoS constraints can also be given in terms of SINR:

$$\text{SINR}_i \triangleq \frac{\left|\mathbf{a}_i^H \mathbf{H}\mathbf{b}_i\right|^2}{\mathbf{a}_i^H \mathbf{R}_{n_i} \mathbf{a}_i} \geq \gamma_i \quad 1 \leq i \leq L, \tag{11}$$

where $\mathbf{R}_{n_i} \triangleq \mathbf{H}\mathbf{B}\mathbf{B}^H\mathbf{H}^H + \mathbf{R}_n - \mathbf{H}\mathbf{b}_i\mathbf{b}_i^H\mathbf{H}^H$ is the interference-plus-noise covariance matrix seen by the ith substream. It turns out that the $\mathbf{A}$ given in (8) also maximizes each of the SINRs [15], [34] (although an arbitrary scaling factor on each $\mathbf{a}_i$ could be included), and the resulting SINR is

$$\text{SINR}_i = \mathbf{b}_i^H \mathbf{H}^H \mathbf{R}_{n_i}^{-1} \mathbf{H}\mathbf{b}_i. \tag{12}$$

Using the matrix inversion lemma, the ith diagonal element of (10) can be related to (12) as $[(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}]_{ii} = 1/(1+\mathbf{b}_i^H\mathbf{H}^H\mathbf{R}_{n_i}^{-1}\mathbf{H}\mathbf{b}_i)$, i.e.,

$$\text{MSE}_i = \frac{1}{1+\text{SINR}_i}. \tag{13}$$

Therefore, for any set of QoS constraints given in terms of SINR, an equivalent set of QoS constraints given in terms of MSE can be straightforwardly obtained with (13).

Finally, consider QoS constraints given in terms of BER:

$$\text{BER}_i \leq p_i \quad 1 \leq i \leq L. \tag{14}$$

Under the Gaussian assumption,[7] the symbol error probability P_e can be analytically expressed as a function of the SINR [35]:

$$P_e(\text{SINR}) = \alpha \mathcal{Q}(\sqrt{\beta\, \text{SINR}}) \tag{15}$$

where α and β are constants that depend on the signal constellation, and $\mathcal{Q}$ is the $\mathcal{Q}$-function defined as $\mathcal{Q}(x) \triangleq (1/\sqrt{2\pi})\int_x^\infty e^{-\lambda^2/2}d\lambda$ [35]. The BER can be approximately obtained from the symbol error probability (assuming that a Gray encoding is used to map the bits into the constellation points) as

$$\text{BER} \approx \frac{P_e}{k} \tag{16}$$

where $k = \log_2 M$ is the number of bits per symbol, and M is the constellation size. Therefore, if the QoS constraints are given in terms of BER, they can also be equivalently expressed in terms of SINR or MSE. Note that we only consider the uncoded part of the communication system, and therefore, the BER always refers to the uncoded BER (in practice, an outer code should always be used on top of the uncoded part). It is worth mentioning that constraining the BER averaged over a set of substreams using the same constellation is equivalent to constraining each substream independently to the same BER [15].

Thus, regardless of whether the QoS requirements of the system specifications are given in terms of MSE, SINR, or BER, the Wiener filter is always optimum, and then, the problem can be formulated in terms of MSE constraints, as we consider in the rest of the paper without loss of generality.

[7]Note that even when the transmitted symbols are not Gaussian distributed, the interference-plus-noise contribution seen by each substream tends to have a Gaussian distribution as the number of transmit dimensions grows.

V. Optimum Transmit Processing with Equal MSE QoS Constraints

In this section, equal QoS constraints expressed in terms of MSE as in (6) are considered:[8]

$$\left[(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}\right]_{ii} \leq \rho, \quad 1 \leq i \leq L. \tag{17}$$

Each of the previous constraints is nonconvex in $\mathbf{B}$ (even in the simple scalar and real case, $1/(1+b^2r) \leq \rho$ is a nonconvex region in b). Therefore, the minimization of the transmit power subject to (17) seems at first to be a formidable problem to solve. With the aid of majorization theory [29] (see Section III), we can reformulate the original nonconvex problem with matrix-valued variables as a simple convex problem with scalar variables that can be optimally solved. We state this result formally in the following theorem (along with the feasibility condition).

Theorem 1: The following nonconvex optimization problem subject to equal MSE QoS constraints:

$$\begin{aligned} \min_{\mathbf{B}} \quad & \text{Tr}(\mathbf{B}\mathbf{B}^H) \\ \text{s.t.} \quad & \left[(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}\right]_{ii} \leq \rho, \quad 1 \leq i \leq L \end{aligned} \tag{18}$$

can be optimally solved by first solving the simple convex optimization problem:

$$\begin{aligned} \min_{\{z_i\}} \quad & \sum_{i=1}^{\check{L}} z_i \\ \text{s.t.} \quad & \frac{1}{L}\left(L_0 + \sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_{H,i}}\right) \leq \rho \\ & z_i \geq 0, \quad 1 \leq i \leq \check{L} \end{aligned} \tag{19}$$

where L is the number of established links, $\check{L} \triangleq \min(L, \text{rank}(\mathbf{R}_H))$ is the number of effective channel eigenvalues used, $L_0 \triangleq L - \check{L}$ is the number of links associated to zero eigenvalues, and the set $\{\lambda_{H,i}\}_{i=1}^{\check{L}}$ contains the $\check{L}$ largest eigenvalues of $\mathbf{R}_H$ in increasing order.

The optimal solution to (18) satisfies all QoS constraints with equality and is given by $\mathbf{B} = \mathbf{U}_{H,1}\mathbf{\Sigma}_{B,1}\mathbf{Q}$, where $\mathbf{U}_{H,1} \in \mathbb{C}^{n_T \times \check{L}}$ has as columns the eigenvectors of $\mathbf{R}_H$ corresponding to the $\check{L}$ largest eigenvalues in increasing order, $\mathbf{\Sigma}_{B,1} = [\mathbf{0}\ \text{diag}(\{\sigma_{B,i}\})] \in \mathbb{C}^{\check{L}\times L}$ has zero elements except along the rightmost main diagonal (assumed real w.l.o.g.), which are given by $\sigma_{B,i}^2 = z_i$, $1 \leq i \leq \check{L}$ [the z_i's are the solution to the convex problem (19)], and $\mathbf{Q}$ is a unitary matrix such that the diagonal elements of $(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}$ are equal (see [30, Sect. IV-A] for a practical algorithm to obtain $\mathbf{Q}$ or use the unitary DFT matrix or the Hadamard matrix, as discussed in Section III). The problem is feasible if and only if $\rho > L_0/L$.

Proof: See Appendix A. ■

Now that the original nonconvex problem has been reformulated as a simple convex problem, we know that a global optimal solution can be obtained in practice by using, for example,

[8]Note that if equal QoS constraints are considered in terms of BER when using different constellations, the resulting QoS constraints in terms of MSE are different, as considered in Section VI.

interior-point methods [36], [37]. Nevertheless, the particular convex problem (19) obtained in Theorem 1 can be optimally solved by a simple water-filling algorithm (Algorithm 1), as stated in the following proposition (by setting $\tilde{\rho} = \rho L - L_0$).

Proposition 1: The optimal solution to the following convex optimization problem:

$$\begin{aligned} \min_{\{z_i\}} \quad & \sum_{i=1}^{\check{L}} z_i \\ \text{s.t.} \quad & \sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_i} \le \tilde{\rho} \\ & z_i \ge 0, \quad 1 \le i \le \check{L} \end{aligned}$$

is given (if feasible) by the water-filling solution $z_i = \left(\mu^{1/2}\lambda_i^{-1/2} - \lambda_i^{-1}\right)^+$ (it is tacitly assumed that all the λ_i's are strictly positive), where $\mu^{1/2}$ is the water-level chosen such that the MSE constraint is satisfied with equality: $\sum_{i=1}^{\check{L}}(1/(1+z_i\lambda_i)) = \tilde{\rho}$.

Furthermore, the optimal water-filling solution can be efficiently obtained in practice with Algorithm 1 in no more than $\check{L}$ iterations (worst-case complexity). (It is assumed that $\tilde{\rho} < \check{L}$; otherwise, the optimal solution is trivially given by $z_i = 0\ \forall i$, i.e., by not transmitting anything.)

Proof: See Appendix B. ■

Algorithm 1: The following is the practical water-filling algorithm that solves the convex problem corresponding to the design with equal MSE QoS requirements of Theorem 1.

Input: Number of available positive eigenvalues $\check{L}$, set of eigenvalues $\{\lambda_i\}_{i=1}^{\check{L}}$, and MSE constraint $\tilde{\rho}$.
Output: Set of allocated powers $\{z_i\}_{i=1}^{\check{L}}$ and water-level $\mu^{1/2}$.

0) Reorder the λ_i's in decreasing order (define $\lambda_{\check{L}+1} \triangleq 0$). Set $\tilde{L} = \check{L}$.
1) Set $\mu = \lambda_{\tilde{L}}^{-1}$ (if $\lambda_{\tilde{L}} = \lambda_{\tilde{L}+1}$, then set $\tilde{L} = \tilde{L} - 1$ and go to step 1).
2) If $\mu^{1/2} \ge \sum_{i=1}^{\tilde{L}} \lambda_i^{-1/2}/(\tilde{\rho} - (\check{L} - \tilde{L}))$, then set $\tilde{L} = \tilde{L} - 1$ and go to step 1.

Otherwise, obtain the definitive water-level $\mu^{1/2}$ and allocated powers as

$$\mu^{1/2} = \frac{\sum_{i=1}^{\tilde{L}} \lambda_i^{-1/2}}{\tilde{\rho} - (\check{L} - \tilde{L})} \quad \text{and}$$
$$z_i = \left(\mu^{1/2}\lambda_i^{-1/2} - \lambda_i^{-1}\right)^+$$

undo the reordering done at step 0, and finish.

Thus, the problem of finding a transmit matrix $\mathbf{B}$ that minimizes the required transmit power subject to equal MSE QoS requirements has been completely solved in a practical and optimal way. It is remarkable that, in general, the optimal solution does not consist of transmitting each symbol through a channel eigenmode in a parallel fashion (diagonal transmission); instead, the symbols are transmitted in a distributed way over all the channel eigenmodes. The next section is devoted to the generalization of this result to the case of different MSE QoS requirements.

VI. Optimum Transmit Processing with Different MSE QoS Constraints

This section generalizes the results obtained in Section V by allowing different QoS constraints in terms of MSE:

$$\left[(\mathbf{I} + \mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}\right]_{ii} \le \rho_i, \quad 1 \le i \le L. \tag{20}$$

This problem is more general and far more complicated than the one with equal MSE constraints of the previous section. Nevertheless, using majorization theory [29] (see Section III), we can still reformulate the original complicated nonconvex problem as a simple convex optimization problem that can be optimally solved as is formally stated in the following theorem (along with the feasibility condition).

Theorem 2: The following nonconvex optimization problem subject to different MSE QoS constraints (assumed bounded by $0, \rho_i < 1$ and in decreasing order $\rho_i \ge \rho_{i+1}$ w.l.o.g.):

$$\begin{aligned} \min_{\mathbf{B}} \quad & \text{Tr}(\mathbf{B}\mathbf{B}^H) \\ \text{s.t.} \quad & \left[(\mathbf{I} + \mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}\right]_{ii} \le \rho_i, \quad 1 \le i \le L \end{aligned} \tag{21}$$

can be optimally solved by first solving the simple convex optimization problem:

$$\begin{aligned} \min_{\{z_i\}} \quad & \sum_{i=1}^{\check{L}} z_i \\ \text{s.t.} \quad & \sum_{i=k}^{\check{L}} \frac{1}{1+z_i\lambda_{H,i}} \le \sum_{i=k+L_0}^{L} \rho_i \quad 1 \le k \le \check{L} \\ & \sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_{H,i}} \le \sum_{i=1}^{L} \rho_i - L_0 \\ & z_k \ge 0, \quad 1 \le k \le \check{L} \end{aligned} \tag{22}$$

where L is the number of established links, $\check{L} \triangleq \min(L, \text{rank}(\mathbf{R}_H))$ is the number of effective channel eigenvalues used, $L_0 \triangleq L - \check{L}$ is the number of links associated to zero eigenvalues, and the set $\{\lambda_{H,i}\}_{i=1}^{\check{L}}$ contains the $\check{L}$ largest eigenvalues of $\mathbf{R}_H$ in increasing order.

The optimal solution to (21) satisfies all QoS constraints with equality and is given by $\mathbf{B} = \mathbf{U}_{H,1}\mathbf{\Sigma}_{B,1}\mathbf{Q}$, where $\mathbf{U}_{H,1} \in \mathbb{C}^{n_T \times \check{L}}$ has as columns the eigenvectors of $\mathbf{R}_H$ corresponding to the $\check{L}$ largest eigenvalues in increasing order, $\mathbf{\Sigma}_{B,1} = [\mathbf{0}\ \text{diag}(\{\sigma_{B,i}\})] \in \mathbb{C}^{\check{L} \times L}$ has zero elements except along the rightmost main diagonal (assumed real w.l.o.g.), which are given by $\sigma_{B,i}^2 = z_i$, $1 \le i \le \check{L}$ (the z_i's are the solution to the convex problem (22)), and $\mathbf{Q}$ is a unitary matrix such that $[(\mathbf{I} + \mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}]_{ii} = \rho_i$, $1 \le k \le L$ (see [30, Sect. IV-A] for a practical algorithm to obtain $\mathbf{Q}$). The problem is feasible if and only if $\sum_{i=1}^{L} \rho_i > L_0$ (a simple sufficient condition for the feasibility of the problem is the rule of thumb: $L \le \text{rank}(\mathbf{H})$[9]).

Proof: See Appendix C. ■

[9]In a practical situation, a threshold should be considered in the rank determination to avoid extremelly small eigenvalues to artificially increase the rank.

Noting that $\sum_{i=1}^{L}\rho_i - L_0 < \sum_{i=1+L_0}^{L}\rho_i$ (recall that $\rho_i < 1$), problem (22) can be rewritten more compactly as

$$\begin{aligned} \min_{\{z_i\}} \quad & \sum_{i=1}^{\check{L}} z_i \\ \text{s.t.} \quad & \sum_{i=k}^{\check{L}} \frac{1}{1+z_i\lambda_{H,i}} \leq \sum_{i=k}^{\check{L}} \tilde{\rho}_i, \quad 1 \leq k \leq \check{L} \\ & z_k \geq 0 \end{aligned} \tag{23}$$

where $\tilde{\rho}_i \triangleq \begin{cases} \sum_{k=1}^{L_0+1}\rho_k - L_0, & \text{for } i=1 \\ \rho_{i+L_0}, & \text{for } 1 < i \leq \check{L} \end{cases}$ (note that the resulting $\tilde{\rho}_i$'s need not be positive and in decreasing ordering as are the ρ_i's).

By Theorem 2, the original formidable problem has been reformulated as a convex problem that can always be solved in practice using, for example, interior-point methods [36], [37]. However, as happened with the case of equal MSE QoS requirements in Section V, this problem can be solved with a multilevel water-filling algorithm (Algorithms 2 and 3) as shown next.

Proposition 2: The optimal solution to the following convex optimization problem:

$$\begin{aligned} \min_{\{z_i\}} \quad & \sum_{i=1}^{\check{L}} z_i \\ \text{s.t.} \quad & \sum_{i=k}^{\check{L}} \frac{1}{1+z_i\lambda_i} \leq \sum_{i=k}^{\check{L}} \tilde{\rho}_i, \quad 1 \leq k \leq \check{L} \\ & z_k \geq 0 \end{aligned}$$

is given (if feasible) by the multilevel water-filling solution $z_i = \left(\tilde{\mu}_i^{1/2}\lambda_i^{-1/2} - \lambda_i^{-1}\right)^+$ (it is tacitly assumed that all the λ_i's are strictly positive), where the multiple water-levels $\tilde{\mu}_i^{1/2}$'s are chosen to satisfy

$$\begin{aligned} & \sum_{i=k}^{\check{L}} \frac{1}{1+z_i\lambda_i} \leq \sum_{i=k}^{\check{L}} \tilde{\rho}_i, \qquad 1 < k \leq \check{L} \\ & \sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_i} = \sum_{i=1}^{\check{L}} \tilde{\rho}_i \\ & \tilde{\mu}_k \geq \tilde{\mu}_{k-1} (\tilde{\mu}_0 \triangleq 0) \\ & (\tilde{\mu}_k - \tilde{\mu}_{k-1}) \left(\sum_{i=k}^{\check{L}} \frac{1}{1+z_i\lambda_i} - \sum_{i=k}^{\check{L}} \tilde{\rho}_i \right) = 0. \end{aligned}$$

Furthermore, the optimal multilevel water-filling solution can be efficiently obtained in practice with Algorithm 2 (or with the equivalent Algorithm 3) in no more (worst-case complexity) than $\check{L}(\check{L}+1)/2$ inner iterations (simple water-fillings) or, more exactly, $\check{L}^2(\check{L}+1)/6$ basic iterations (iterations within each simple waterfilling). (It is assumed that $\tilde{\rho}_i < 1$; otherwise, the optimal solution is trivially given by $z_i = 0 \ \forall i$, i.e., by not transmitting anything.)

Proof: See Appendix D. ■

Algorithm 2: The following is the practical multilevel water-filling algorithm that solves the convex problem corresponding to the design with different MSE QoS requirements of Theorem 2. Version 1. (See Fig. 3 for an illustrative example of the application of the algorithm.)

Input: Number of available positive eigenvalues $\check{L}$, set of eigenvalues $\{\lambda_i\}_{i=1}^{\check{L}}$, and set of MSE constraints $\{\tilde{\rho}_i\}_{i=1}^{\check{L}}$ (note that the appropriate ordering of the λ_i's and of the $\tilde{\rho}_i$'s is independent of this algorithm).
Output: Set of allocated powers $\{z_i\}_{i=1}^{\check{L}}$ and set of water-levels $\{\tilde{\mu}_i^{1/2}\}_{i=1}^{\check{L}}$.

0) Set $\tilde{L} = \check{L}$.
1) Perform an outer iteration.
2) If $k_0 = 1$, then finish. Otherwise set $\tilde{L} = k_0 - 1$ and go to step 1.

Outer iteration:

0) Set $k_0 = 1$.
1) Perform an inner iteration: Solve the equal MSE QoS constrained problem in $[k_0, \tilde{L}]$ using Algorithm 1 with the set of $\tilde{L} - k_0 + 1$ eigenvalues $\{\lambda_i\}_{i=k_0}^{\tilde{L}}$ and with the MSE constraint given by $\tilde{\rho} = \sum_{i=k_0}^{\tilde{L}} \tilde{\rho}_i$.
2) If all intermediate constraints are also satisfied (i.e., if $\sum_{i=k}^{\tilde{L}}(1/(1+z_i\lambda_i)) \leq \sum_{i=k}^{\tilde{L}} \tilde{\rho}_i$, $k_0 < k \leq \tilde{L}$), then finish.
   Otherwise, set k_0 equal to the smallest index whose constraint is not satisfied, and go to step 1.

Note that each outer iteration implicitly computes the water-level for the set $[k_0, \tilde{L}]$ denoted by $\mu^{1/2}([k_0, \tilde{L}])$; in other words, $\tilde{\mu}_k = \mu([k_0, \tilde{L}])$, $k_0 \leq k \leq \tilde{L}$.

Algorithm 2 was conveniently written to prove its optimality. For a practical implementation, however, it can be rewritten in a much simpler way as in Algorithm 3.

Algorithm 3: The following is a practical multilevel water-filling algorithm that solves the convex problem corresponding to the design with different MSE QoS requirements of Theorem 2. Version 2.

Input: Number of available positive eigenvalues $\check{L}$, set of eigenvalues $\{\lambda_i\}_{i=1}^{\check{L}}$, and set of MSE constraints $\{\tilde{\rho}_i\}_{i=1}^{\check{L}}$ (note that the appropriate ordering of the λ_i's and of the $\tilde{\rho}_i$'s is independent of this algorithm).
Output: Set of allocated powers $\{z_i\}_{i=1}^{\check{L}}$ and set of water-levels $\{\tilde{\mu}_i^{1/2}\}_{i=1}^{\check{L}}$.

0) Set $k_0 = 1$ and $\tilde{L} = \check{L}$.
1) Solve the equal MSE QoS constrained problem in $[k_0, \tilde{L}]$ using Algorithm 1 with the set of $\tilde{L} - k_0 + 1$ eigenvalues $\{\lambda_i\}_{i=k_0}^{\tilde{L}}$ and with the MSE constraint given by $\tilde{\rho} = \sum_{i=k_0}^{\tilde{L}} \tilde{\rho}_i$.
2) If any intermediate constraint $(\sum_{i=k}^{\tilde{L}}(1/(1+z_i\lambda_i)) \leq \sum_{i=k}^{\tilde{L}} \tilde{\rho}_i$, $k_0 < k \leq \tilde{L})$ is not satisfied, then set k_0 equal to the smallest index whose constraint is not satisfied and go to step 1.
   Otherwise, if $k_0 = 1$ finish and if $k_0 > 1$ set $\tilde{L} = k_0 - 1$, $k_0 = 1$, and go to step 1.

This section has obtained the main result of the paper: an efficient and optimal way to solve in practice the problem of finding a transmit matrix $\mathbf{B}$ that minimizes the required transmit power

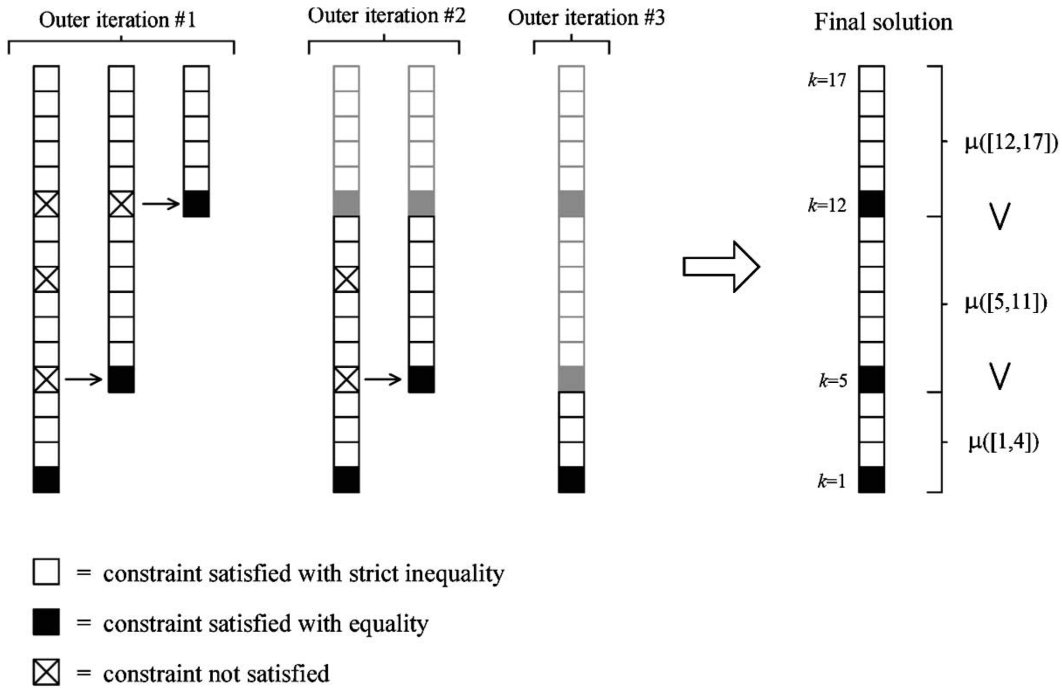

Fig. 3. Example of the execution of Algorithm 2 to solve the problem with different QoS requirements.

while satisfying different MSE QoS requirements. It suffices to use the multilevel water-filling algorithm (Algorithm 3) and then to find the proper rotation matrix $\mathbf{Q}$, as described in Theorem 2.

As happened in the case of equal MSE QoS constraints, the optimal solution does not generally consist of transmitting each symbol through a channel eigenmode in a parallel fashion (diagonal transmission); instead, the symbols are transmitted in a distributed way over all the channel eigenmodes.

VII. Suboptimum Transmit Processing: A Simple Approach

At this point, it is interesting to consider a suboptimum but very simple solution to the considered problem. The simplicity of the solution comes from imposing a diagonality constraint in the MSE matrix, i.e., from forcing $\mathbf{E} = (\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}$ to have a diagonal structure. Imposing such a structure implies that the transmission is performed in a parallel fashion through the channel eigenmodes. In Lemma 5, we obtain such a simple solution and give the feasibility condition. Then, in Lemma 6, we state the exact conditions under which such a constrained solution happens to be the optimum of the original problem considered in Theorem 2.

Lemma 5: The following nonconvex optimization problem subject to different MSE QoS constraints (assumed in decreasing order $\rho_i \geq \rho_{i+1}$ w.l.o.g. and bounded by $0 < \rho_i < 1$):

$$\begin{aligned} \min_{\mathbf{B}} \quad & \text{Tr}(\mathbf{B}\mathbf{B}^H) \\ \text{s.t.} \quad & \left[(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}\right]_{ii} \leq \rho_i, \quad 1 \leq i \leq L \\ & \mathbf{B}^H\mathbf{R}_H\mathbf{B} \quad \text{diagonal} \end{aligned}$$

is feasible if and only if the number of established links L satisfy $L \leq \text{rank}(\mathbf{R}_H)$, and the optimal solution is then given by $\mathbf{B} = \mathbf{U}_{H,1}\boldsymbol{\Sigma}_{B,1}$, where $\mathbf{U}_{H,1} \in \mathbb{C}^{n_T \times L}$ has as columns the eigenvectors of $\mathbf{R}_H$ corresponding to the L largest eigenvalues in increasing order, denoted by $\{\lambda_{H,i}\}$, and $\boldsymbol{\Sigma}_{B,1} \in \mathbb{C}^{L\times L}$ is a diagonal matrix with squared-diagonal elements given by

$$z_i = \lambda_{H,i}^{-1}\left(\rho_i^{-1} - 1\right), \quad 1 \leq i \leq L.$$

Proof: See Appendix E. ■

Lemma 6: The optimal solution obtained in Lemma 5 under the diagonality constraint of the MSE matrix $\mathbf{E} = (\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}$ is the optimal solution to the problem considered in Theorem 2 without the diagonality constraint on $\mathbf{E}$ if and only if

$$\lambda_{H,i}\rho_i^2 \geq \lambda_{H,i+1}\rho_{i+1}^2, \quad 1 \leq i < L \tag{24}$$

where the ρ_i's are in decreasing order, and the $\lambda_{H,i}$'s are the L largest eigenvalues of $\mathbf{R}_H$ in increasing order. Note that this condition implies the feasibility condition $L \leq \text{rank}(\mathbf{R}_H)$.

The conditions in (24) can be stated in words by saying that the singular values of the whitened channel $\lambda_{H,i}^{1/2}$ have to increase at a slower rate than the decrease of the MSE constraints ρ_i.

Proof: See Appendix F. ■

As an example, the conditions of Lemma 6 are always satisfied for channels with equal singular values (this corresponds to a diagonal squared channel matrix $\mathbf{R}_H$), and consequently, a parallel transmission is always the optimum structure for such channels.

Another interesting example arises for systems with equal MSE constraints, as treated in Theorem 1, for which the conditions in (24) are never satisfied (unless the squared-channel has equal eigenvalues).

VIII. Practical Implementation Issues

In this section, we consider interesting issues for practical implementations such as the relaxation of the QoS constraints and the robust design under channel estimation errors (imperfect CSI).

A. Relaxation of the QoS Requirements

In situations where the problem is feasible but the required transmit power exceeds some prespecified maximum level, the system may be forced to relax some QoS requirements so that the required power is reduced. For that purpose, we identify which QoS constraints produce the largest reduction in transmit power when relaxed by means of a perturbation analysis. The questions of whether and when these relaxations are necessary is a high-level decision[10] that may depend on factors as disparate as the energy left on the batteries at the transmitter or the number of services/users requesting a link. Such high-level decisions are out of the scope of this paper.

It is well known from convex optimization theory [36], [37] that the optimal dual variables (Lagrange multipliers) of a convex optimization problem give useful information about the sensitivity of the optimal objective value with respect to perturbations of the constraints. Consider the following relaxation of the original QoS constraints of (6):

$$\mathrm{MSE}_i \leq \rho_i + u_i(\delta_i) \tag{25}$$

where $u_i(\delta_i)$ is a positive differentiable function parameterized with δ_i such that $u_i(\delta_i) \underset{\delta_i \to 0^+}{\longrightarrow} 0$. By Theorem 2, the relaxed problem in convex form is [for simplicity of exposition, we consider $L \leq \mathrm{rank}(\mathbf{H})$ in (22)]

$$\begin{array}{lll} \min\limits_{\{z_i\}} & \sum\limits_{i=1}^{L} z_i & \\ \text{s.t.} & \sum\limits_{i=k}^{L} \dfrac{1}{1+z_i\lambda_{H,i}} \leq \sum\limits_{i=k}^{L} \rho_i + \check{u}_k(\boldsymbol{\delta}), & 1 \leq k \leq L \\ & z_k \geq 0 & \end{array}$$

where $\check{u}_k(\boldsymbol{\delta}) \triangleq \sum_{i=k}^{L} u_i(\delta_i)$, and $\boldsymbol{\delta} = [\delta_1, \cdots, \delta_L]^T$. Defining $p^\star(\check{\mathbf{u}})$ as the optimal objective value of the relaxed problem as a function of $\check{\mathbf{u}} = [\check{u}_1, \cdots, \check{u}_L]^T$, the following local sensitivity result holds [36], [37] (note that the problem satisfies the Slater's condition, and therefore, strong duality holds):

$$\mu_k^\star = -\left.\frac{\partial p^\star}{\partial \check{u}_k}\right|_{\boldsymbol{\delta}=\mathbf{0}} \tag{26}$$

where $\mu_k^\star$ is the Lagrange multiplier of the Lagrangian of the relaxed problem [similar to (36)] at an optimal point. Using the chain rule for differentiation $\partial p^\star/\partial \delta_i = \sum_{k=1}^{L} (\partial p^\star/\partial \check{u}_k)(\partial \check{u}_k/\partial \delta_i)$, noting that

$$\partial \check{u}_k(\boldsymbol{\delta})/\partial \delta_i = \begin{cases} u_i(\delta_i), & i \geq k \\ 0, & \text{otherwise} \end{cases}$$

[10]Advanced communication systems are envisioned to exploit cross-layer signaling to further control the performance of the whole system.

and using (26), it follows that

$$\left.\frac{\partial p^\star}{\partial \delta_i}\right|_{\boldsymbol{\delta}=\mathbf{0}} = -\tilde{\mu}_i^\star \left.\frac{\partial u_i(\delta_i)}{\partial \delta_i}\right|_{\boldsymbol{\delta}=\mathbf{0}} \tag{27}$$

where we have defined $\tilde{\mu}_i^\star \triangleq \sum_{k=1}^{i} \mu_k^\star$.

The largest value of $-(\partial p^\star/\partial \delta_i)|_{\boldsymbol{\delta}=\mathbf{0}}$ for $1 \leq i \leq L$ indicates the QoS constraint that should be relaxed in order to get the largest reduction of the required transmitted power. Note that the optimal $\tilde{\mu}_i$'s used in (27) are readily given by the water-levels implicitly obtained in Algorithms 2 and 3. For each subblock $[k_1, k_2]$ of the partition on $[1, L]$, choose $\tilde{\mu}_k = \mu([k_1, k_2])$ $k_1 \leq k \leq k_2$. The term $(\partial u_i(\delta_i)/\partial \delta_i)|_{\boldsymbol{\delta}=\mathbf{0}}$ in (27) depends on the particular cost function that relates the QoS in terms of MSE as used in the problem formulation and the QoS as seen by the service/user, which can be in terms of MSE, SINR, or BER. We consider now a few examples.

Example 1—$\mathit{MSE}_i \leq \rho_i + \delta_i$*:* In this case, it follows that $\partial u_i(\delta_i)/\partial \delta_i = 1$, and therefore, $(\partial p^\star/\partial \delta_i)|_{\boldsymbol{\delta}=\mathbf{0}} = -\tilde{\mu}_i^\star$. The largest $\tilde{\mu}_i^\star$ is given by $\tilde{\mu}_L^\star$ (or any other $\tilde{\mu}_i^\star$ belonging to the same water-filling subblock as obtained in Appendix D). This means that the best way to relax the constraints in this case is by relaxing the tightest constraint.

Example 2—$\mathit{MSE}_i \leq \mathit{BER}^{-1}(p_i(1+\delta_i))$*:* In this example, the QoS are given in terms of BER. The relaxation can be expressed as $\mathrm{MSE}_i \leq \rho_i + u_i(\delta_i)$ by defining $\rho_i \triangleq \mathrm{BER}^{-1}(p_i)$ and $u_i(\delta_i) \triangleq \mathrm{BER}^{-1}(p_i(1+\delta_i)) - \mathrm{BER}^{-1}(p_i)$. It follows that $\partial u_i(\delta_i)/\partial \delta_i = \partial \mathrm{BER}^{-1}(p_i(1+\delta_i))/\partial \delta_i$.

As a final remark, note that the previous sensitivity analysis is local, and therefore, the relaxation of the QoS constraints has to be performed in sufficiently small steps (recall that the decreasing order of the ρ_i's must be kept at each step).

B. Robust Design Under Channel Estimation Errors

Hitherto, perfect CSI has been assumed at both sides of the link. A common problem in practical communication systems arises from channel estimation errors (see [38] and references therein). Whereas a sufficiently accurate CSIR) can be assumed in many cases, CSIT will be far from perfect in any realistic situation (see [19] for a description of channel estimation strategies in real systems). Hence, as a first approximation, we assume perfect CSIR and imperfect CSIT (further work has do be done to consider imperfect CSIR as well, c.f. [34]).

Under the assumption of perfect CSIR, the optimal receive matrix is still given by (8) for any transmit matrix $\mathbf{B}$. The receiver, however, instead of taking the channel estimate $\hat{\mathbf{H}}$ as perfect (naive approach), will take a robust approach by considering that the real channel $\mathbf{H}$ can be written as

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{H}_\Delta \tag{28}$$

where $\mathbf{H}_\Delta$ represents the channel estimation error bounded as $\|\mathbf{H}_\Delta\| \leq \epsilon_H$ [18], [19] or as $\|\mathbf{H}_\Delta\|/\|\hat{\mathbf{H}}\| \leq \epsilon$ for some small ϵ_H or ϵ (both constraints are equivalent if $\epsilon_H \triangleq \epsilon\|\hat{\mathbf{H}}\|$). Assuming that the channel has already been whitened (estimation errors on the noise covariance matrix can be explicitly considered, as in [34]), we can write $\mathbf{R}_H = \hat{\mathbf{R}}_H + \mathbf{R}_\Delta$, where $\hat{\mathbf{R}}_H \triangleq \hat{\mathbf{H}}^H\hat{\mathbf{H}}$, and $\mathbf{R}_\Delta \triangleq \hat{\mathbf{H}}^H\mathbf{H}_\Delta + \mathbf{H}_\Delta^H\hat{\mathbf{H}} + \mathbf{H}_\Delta^H\mathbf{H}_\Delta$. Taking the maximum singular-value norm $\|\mathbf{X}\| \triangleq \sigma_{\max}(\mathbf{X})$, it follows that $|\lambda_i(\mathbf{R}_\Delta) - \epsilon_H^2| \le 2\epsilon_H\sigma_{\max}(\hat{\mathbf{H}})$ [34]. Since $\epsilon_H \ll \sigma_{\max}(\hat{\mathbf{H}})$, we can ignore the quadratic term and finally write

$$\mathbf{R}_H = \hat{\mathbf{R}}_H + \mathbf{R}_\Delta, \quad \begin{array}{r} |\lambda_i(\mathbf{R}_\Delta)| \le \epsilon_R \\ \mathbf{R}_H = \mathbf{R}_H^H \ge \mathbf{0} \end{array} \tag{29}$$

where $\epsilon_R = 2\epsilon_H\sigma_{\max}(\hat{\mathbf{H}})$, and we have made explicit the inherent constraints on $\mathbf{R}_\Delta$ to guarantee the positive semidefiniteness of $\mathbf{R}_H$. (If the Frobenius norm is used instead of the maximum singular-value norm, similar bounds on $\lambda_i(\mathbf{R}_\Delta)$ are obtained.)

Therefore, the robust design can be formulated [similarly to (21)] as

$$\begin{array}{ll} \min\limits_{\mathbf{B}} & \text{Tr}(\mathbf{B}\mathbf{B}^H) \\ \text{s.t.} & \max\limits_{\mathbf{R}_H} \left[(\mathbf{I} + \mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}\right]_{ii} \le \rho_i, \quad 1 \le i \le L \end{array}$$

$$\mathbf{R}_H = \hat{\mathbf{R}}_H + \mathbf{R}_\Delta, \quad \mathbf{R}_{H,\Delta} : \begin{array}{r} |\lambda_i(\mathbf{R}_\Delta)| \le \epsilon_R \\ \mathbf{R}_H = \mathbf{R}_H^H \ge \mathbf{0}. \end{array} \tag{30}$$

From (29) and using the eigendecomposition $\hat{\mathbf{R}}_H = \hat{\mathbf{U}}_H\hat{\mathbf{D}}_H\hat{\mathbf{U}}_H^H$, the eigenvalues of $\mathbf{R}_H$ can be shown to be lower bounded as $\lambda_i(\mathbf{R}_H) \ge (\lambda_i(\hat{\mathbf{R}}_H) - \epsilon_R)^+$ or, equivalently

$$\mathbf{R}_H \ge \check{\mathbf{R}}_H \triangleq \hat{\mathbf{U}}_H(\hat{\mathbf{D}}_H - \epsilon_R\mathbf{I})^+\hat{\mathbf{U}}_H^H.$$

Using the fact that $\mathbf{R}_H \ge \check{\mathbf{R}}_H \Rightarrow (\mathbf{I} + \mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1} \le (\mathbf{I} + \mathbf{B}^H\check{\mathbf{R}}_H\mathbf{B})^{-1}$, the robust problem formulation of (30) can be equivalently and compactly written as

$$\begin{array}{ll} \min\limits_{\mathbf{B}} & \text{Tr}(\mathbf{B}\mathbf{B}^H) \\ \text{s.t.} & \left[(\mathbf{I} + \mathbf{B}^H\check{\mathbf{R}}_H\mathbf{B})^{-1}\right]_{ii} \le \rho_i, \quad 1 \le i \le L. \end{array}$$

Thus, a robust design can be easily implemented in practice simply by using the lower bound $\check{\mathbf{R}}_H$ instead of the estimated channel matrix $\hat{\mathbf{R}}_H$ in the solution obtained in Sections V and VI.[11] As a final remark, it is worth pointing out that a different robust approach based on the Bayesian philosophy can also be taken as opposed to the considered worst-case robust approach (c.f. [34]).

IX. Simulations

For the numerical simulations, we consider a wireless communication system with multiple antennas at both sides of the link (in particular four transmit and four receive antennas).[12] Unless otherwise stated, perfect CSI is assumed at both sides of the communication system. The channel model includes the frequency-selectivity and the spatial correlation as measured in real scenarios. The MIMO channels were generated using the parameters of the European standard for Wireless Local Area Networks (WLAN) HIPERLAN/2 [27], which is based on the multicarrier modulation orthogonal frequency division multiplexing (OFDM) (64 carriers were used in the simulations). The frequency selectivity of the channel was modeled using the power delay profile type C for HIPERLAN/2, as specified in [39], which corresponds to a typical large open space indoor environment for nonline-of-sight conditions with a 150-ns average r.m.s. delay spread, and a 1050-ns maximum delay (the sampling period is 50 ns [27]). The spatial correlation of the MIMO channel was modeled according to the *Nokia* model defined in [40] (for the uplink) as specified by the correlation matrices of the envelope of the channel fading at the transmit and receive side. It provides a large open indoor environment with two floors, which could easily illustrate a conference hall or a shopping galleria scenario (see [40] for details of the model). The matrix channel generated was normalized so that $\sum_n \mathbb{E}[|[\mathbf{H}(n)]_{ij}|^2] = 1$.

The results are given in terms of required transmit power at some outage probability P_{out}, i.e., the transmit power that will not suffice to satisfy the QoS constraints only with a small probability P_{out} (it will suffice for $(1 - P_{\text{out}})$ of the time). In particular, we consider an outage probability of 5%. The reason for using the outage power instead of the average power is that for typical systems with delay constraints, the former is more realistic than the latter, which only makes sense when the transmission coding block is long enough to reveal the long-term ergodic properties of the fading process (no delay constraints). Instead of plotting absolute values of the required transmit power, we plot relative values of the transmitted power normalized with the noise spectral density N_0. We call this normalized transmitted power SNR, which is defined as SNR $= \text{Tr}(\mathbf{B}\mathbf{B}^H)/(N \times N_0)$,[13] where N is the number of carriers. Note that the plots are valid only for the channel normalization used.

In the following, numerical results for the proposed methods obtained in Sections V and VI are given (Algorithm 1 for equal MSE QoS requirements and Algorithm 3 for different MSE QoS requirements). As a means of comparison, the simple benchmark obtained by imposing the diagonality of $\mathbf{E}$ as obtained in Section VII is also simulated. We consider two multicarrier approaches as described in Section II: the carrier-cooperative scheme and the carrier-noncooperative scheme. Note that the total number of established substreams is $N \times L$, where L denotes the number of spatial dimensions used per carrier.

A. Equal QoS Constraints

We now consider equal QoS requirements in terms of BER, which correspond to equal QoS requirements in terms of MSE because the same constellation [quaternary phase shift keying (QPSK)] is used on all the substreams.

In Fig. 4, the SNR per spatial dimension SNR$/L$ is plotted as a function of L subject to equal QoS requirements given by BER $= 10^{-3}$. For $L = 3$, the gain over the benchmark is about 2 dB. For $L = 4$, the gain is of 7 dB for the carrier-noncooperative

[11] In a practical implementation, care has to be taken so that the problem does not become unfeasible due to the reduction of the eigenvalues. To avoid this, the value of ϵ_R should be monitored and, if necessary, decreased so that the feasibility condition obtained in Theorems 1 and 2 is always satisfied.

[12] In [23], numerical simulations were performed for a DSL system using realistic channel models. First, a bit distribution was performed using the gap approximation, and then, the proposed methods (Algorithms 1–3) were utilized to obtain the optimum transmit-receive processing to achieve a symbol error probability of 10^{-7} with minimum transmitted power.

[13] Note that such an SNR definition is a measure of the total transmitted power per symbol (normalized with N_0) and does not correspond to the SNR at each receive antenna.

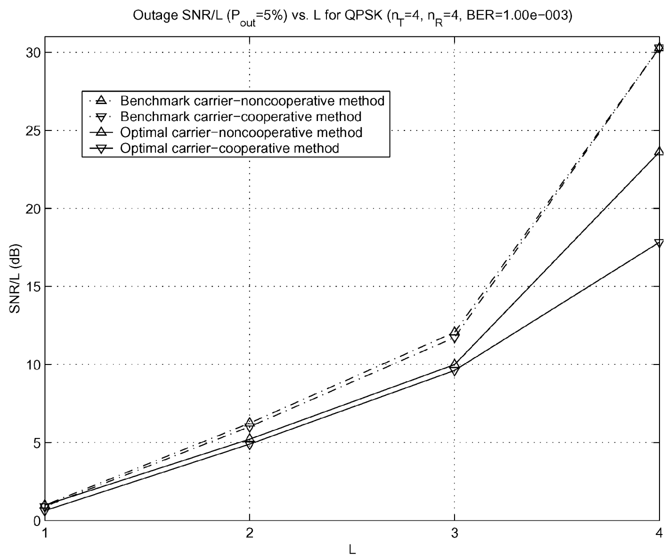

Fig. 4. Outage SNR per spatial dimension versus the number of spatial dimensions utilized L when using QPSK in a multicarrier 4×4 MIMO channel with an equal QoS for all substreams given by BER $\leq 10^{-3}$.

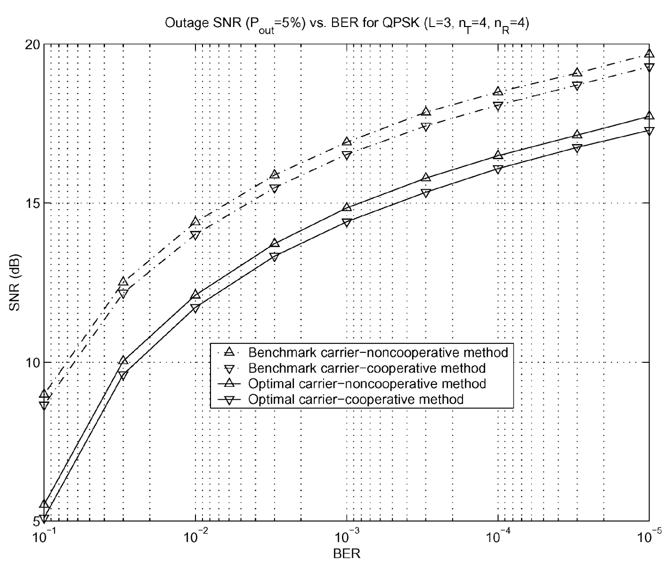

Fig. 5. Outage SNR versus the equal QoS constraints given in terms of BER when using QPSK in a multicarrier 4×4 MIMO channel with $L = 3$.

scheme and of 12 dB for the carrier-cooperative scheme. The required power for $L = 4$ increases significantly with respect to $L = 3$, and therefore, we choose the latter for the rest of the simulations.

In Fig. 5, the SNR is given as a function of the (equal) QoS constraints in terms of BER for $L = 3$. It can be observed that the gain over the benchmark is about 2–3 dB and constant for the entire range of the BER. Cooperation among carriers improves the performance by no more than 0.5 dB. Given that the carrier-noncooperative scheme has an attractive parallel implementation (since each carrier is independently processed), it may be an interesting solution for a practical and efficient implementation.

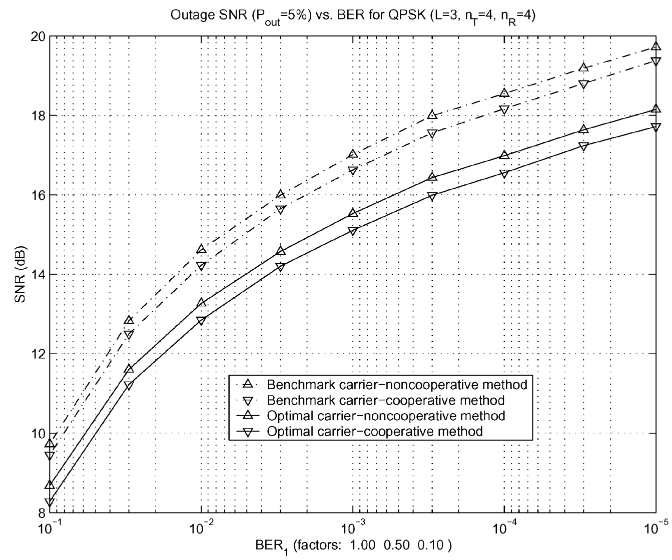

Fig. 6. Outage SNR versus the different QoS constraints given in terms of BER when using QPSK in a multicarrier 4×4 MIMO channel with $L = 3$. (The BER of the first substream is along the x axis, and the BER of the second and third substreams are given by scaling with the factors 0.5 and 0.1.)

B. Different QoS Constraints

We now generalize the setup by considering different QoS constraints both in terms of BER and MSE (since the same constellations are used on all the substreams).

In Fig. 6, the SNR is plotted as a function of the nominal QoS corresponding to the first substream (the QoS constraints for the others substreams are obtained by scaling this nominal QoS constraint with the factors 0.5 and 0.1). Similar observations to those corresponding to Fig. 5 hold in this case.

C. Relaxation of the QoS Requirements

In Fig. 7, an example of a relaxed system is given (as explained in Section VIII). The relaxed-I case consists of ten relaxations per carrier with $\delta = 0.1$ of the form $\text{MSE}_i \leq \text{BER}^{-1}(p_i 10^{\delta_i})$ (note that an extreme case of this series of relaxations amounts to increasing the BER of a single substream one order of magnitude), and the relaxed-II case consists of ten additional relaxations of the same type. As observed, a significant reduction of the required transmit power can be achieved at the expense of the relaxation of some QoS requirements (recall that these relaxations are done in an optimal way).

D. Robust Design Under Channel Estimation Errors

The estimated channel matrix contains a random estimation error drawn according to a Gaussian pdf with zero mean and uncorrelated elements with variance $\sigma_H^2 = 0.01$. The noise covariance matrix is assumed white and perfectly known. In a real system, the upper bound on the norm of the channel estimation error for the robust design [see (28)] should be chosen such that it is satisfied with high probability; otherwise, an outage event is declared. For an outage probability of 5%, the upper bound on the norm should be chosen as $\epsilon_H \simeq \sqrt{14.6645 \times \sigma_H^2}$

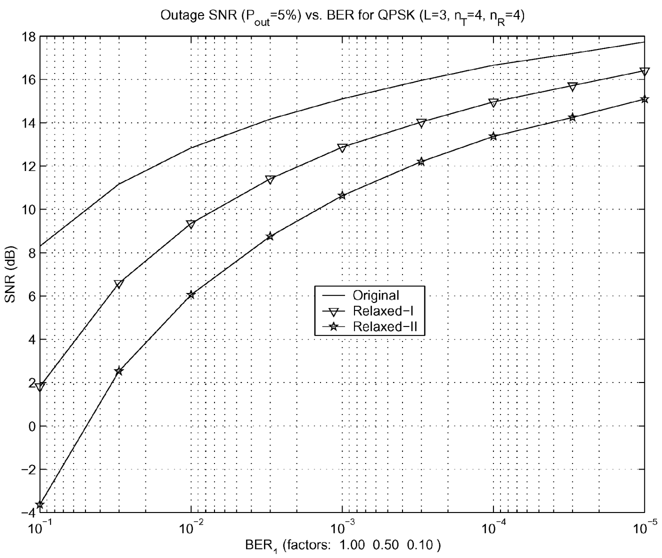

Fig. 7. Outage SNR versus the different QoS constraints given in terms of BER for the original carrier-cooperative method and for two succesive series of relaxations when using QPSK in a multicarrier 4 × 4 MIMO channel with $L = 3$. (The BER of the first substream is along the x axis and the BER of the second and third substreams are given by scaling with the factors 0.5 and 0.1.)

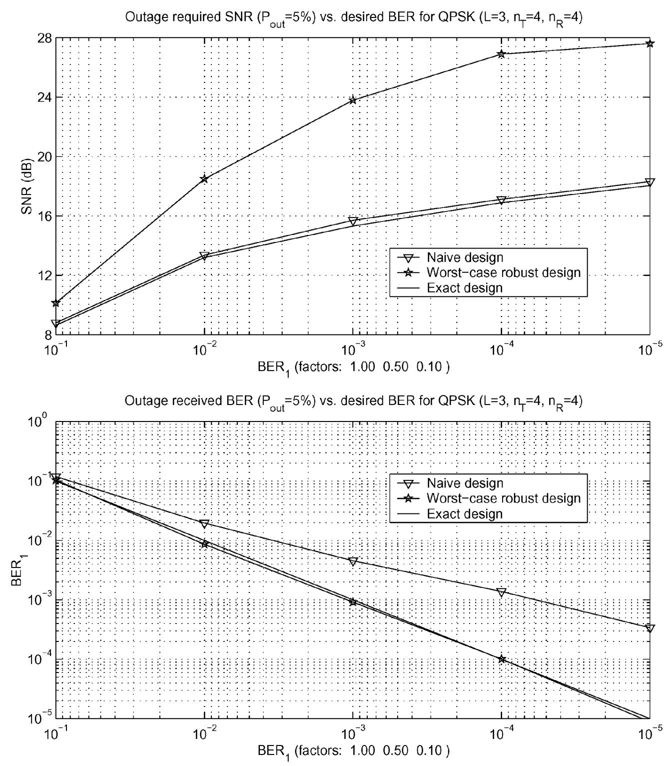

Fig. 8. Outage required SNR and obtained BER versus the required BER for the (carrier-noncooperative) exact, naive, and worst-case robust methods when using QPSK in a multicarrier 4 × 4 MIMO channel with $L = 3$. (The BER of the first substream is along the x axis, and the BER of the second and third substreams is given by scaling with the factors 0.5 and 0.1.)

for the case of four transmit and four receive antennas. In practice, however, this value is unnecessarily high due to the pessimism inherent in all bounds used in Section VIII-B to derive the worst-case design. Therefore, to overcome such excessive pessimism, it is necessary to include in ϵ_H an additional factor α_H that has to be found by numerical simulations (in this case, $\alpha_H = 10^{-3}$).

In Fig. 8, the required SNR and the obtained BER are plotted as a function of the desired BER given as a QoS constraint for three different cases: exact design with perfect CSIT, naive design with imperfect CSIT, and worst-case robust design against imperfect CSIT (in all cases, perfect CSIR is assumed). The exact design achieves exactly the desired BER, as expected. The naive design simply uses the estimated channel and does not satisfy the QoS constraints (in more than one order of magnitude for a BER of 10^{-4} and 10^{-5}). The robust design, however, manages to satisfy the QoS requirements due to its robustness at the expense of an increased transmit power (which is the price to be paid for robustness).

X. Conclusions

This paper has dealt with the optimization of the transmission of a vector signal through a matrix channel using linear processing both at the transmitter and at the receiver. In particular, the minimization of the transmitted power has been considered subject to (possibly different) QoS requirements for each of the established substreams in terms of MSE, SINR, and BER. Although the original problem formulation is a complicated nonconvex problem with matrix-valued variables, by using majorization theory, we have been able to reformulate it as a simple convex optimization problem with scalar variables. To optimally solve the convex problem in practice, we have proposed a practical and efficient multilevel water-filling algorithm. For situations in which the required power results too large, a perturbation analysis has been conducted to obtain the optimal way in which the QoS requirements should be relaxed in order to reduce power needed. Furthermore, we have proposed a simple robust design under channel estimation errors that has an important interest in practical systems.

Appendix A
Proof of Theorem 1

We first present a lemma and then proceed to the proof of Theorem 1.

Lemma 7: [15] Given a matrix $\mathbf{B} \in \mathbb{C}^{n_T \times L}$ and a positive semidefinite Hermitian matrix $\mathbf{R}_H \in \mathbb{C}^{n_T \times n_T}$ such that $\mathbf{B}^H \mathbf{R}_H \mathbf{B}$ is a diagonal matrix with diagonal elements in increasing order (possibly with some zero diagonal elements), it is always possible to find another matrix of the form $\tilde{\mathbf{B}} = \mathbf{U}_{H,1} \mathbf{\Sigma}_{B,1}$ of at most rank $\check{L} \triangleq \min(L, \text{rank}(\mathbf{R}_H))$ that satisfies $\tilde{\mathbf{B}}^H \mathbf{R}_H \tilde{\mathbf{B}} = \mathbf{B}^H \mathbf{R}_H \mathbf{B}$ with $\text{Tr}(\tilde{\mathbf{B}}\tilde{\mathbf{B}}^H) \leq \text{Tr}(\mathbf{B}\mathbf{B}^H)$, where $\mathbf{U}_{H,1} \in \mathbb{C}^{n_T \times \check{L}}$ has as columns the eigenvectors of $\mathbf{R}_H$ corresponding to the $\check{L}$ largest eigenvalues in increasing order, and $\mathbf{\Sigma}_{B,1} = [\mathbf{0}_{\check{L} \times (L-\check{L})} \text{diag}(\{\sigma_{B,i}\})_{\check{L} \times \check{L}}] \in \mathbb{C}^{\check{L} \times L}$ has zero elements except along the rightmost main diagonal (which can be assumed real w.l.o.g.).

Proof of Theorem 1: First, rewrite the original problem as

$$\begin{aligned} \min_{\mathbf{B}} \quad & \text{Tr}(\mathbf{B}\mathbf{B}^H) \\ \text{s.t.} \quad & \max_i \left[(\mathbf{I} + \mathbf{B}^H \mathbf{R}_H \mathbf{B})^{-1} \right]_{ii} \leq \rho. \end{aligned}$$

Note that this problem is exactly the opposite formulation of one of the design criteria considered in [15] (the MAX-MSE criterion), where $\max_i[(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}]_{ii}$ was minimized subject to a power constraint $\text{Tr}(\mathbf{B}\mathbf{B}^H) \leq P_T$.[14] We claim that matrix $\mathbf{E} = (\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}$ must have equal diagonal elements at an optimal point and that the QoS constraints must be all satisfied with equality. Otherwise, by Corollary 1, we could find a unitary matrix $\mathbf{Q}$ with $\mathbf{Q}^H\mathbf{E}\mathbf{Q}$ having identical diagonal elements equal to $(1/L)\text{Tr}(\mathbf{E})$ (this amounts to using $\mathbf{BQ}$ as transmit matrix instead of $\mathbf{B}$). Since $(1/L)\text{Tr}(\mathbf{E}) \leq \max_i[\mathbf{E}]_{ii}$ with equality if and only if all diagonal elements of $\mathbf{E}$ are equal, using $\mathbf{BQ}$ would satisfy all QoS constraints with strict inequality, and the objective value could be further minimized by scaling down the whole transmit matrix. Therefore, the problem can be rewritten as

$$\begin{aligned} \min_{\mathbf{B}} \quad & \text{Tr}(\mathbf{B}\mathbf{B}^H) \\ \text{s.t.} \quad & (\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1} \quad \text{with equal diagonal elements} \\ & \frac{1}{L}\text{Tr}(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1} \leq \rho. \end{aligned}$$

By Corollary 1, it follows that for any given $\mathbf{B}$, we can always find a unitary matrix $\mathbf{Q}$ such that $\mathbf{Q}^H(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}\mathbf{Q}$ has equal diagonal elements. Therefore, we can decompose $\mathbf{B}$ as $\mathbf{B} = \tilde{\mathbf{B}}\mathbf{Q}$, where $\tilde{\mathbf{B}}$ does not have to be chosen such that $(\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1}$ has equal diagonal elements since the unitary matrix $\mathbf{Q}$ is responsible for that. Furthermore, to simplify the problem, we can impose the condition that $\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}}$ is diagonal w.l.o.g. It is important to remark that imposing a diagonal structure on $\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}}$ does not affect the problem since the objective value remains the same as $\text{Tr}(\mathbf{B}\mathbf{B}^H) = \text{Tr}(\tilde{\mathbf{B}}\tilde{\mathbf{B}}^H)$, and the QoS constraint also remains unchanged as $\text{Tr}(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1} = \text{Tr}(\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1}$. In addition, since $\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}}$ is diagonal, the unitary matrix $\mathbf{Q}$ such that $\mathbf{Q}^H(\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1}\mathbf{Q}$ has equal diagonal elements can be found with the practical algorithm given in [30, Sect. IV-A] or simply by using the unitary DFT matrix or the Hadamard matrix, as discussed in Section III.

The problem can be finally written as

$$\begin{aligned} \min_{\tilde{\mathbf{B}}} \quad & \text{Tr}(\tilde{\mathbf{B}}\tilde{\mathbf{B}}^H) \\ \text{s.t.} \quad & \tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}} \quad \text{diagonal} \\ & \frac{1}{L}\text{Tr}(\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1} \leq \rho. \end{aligned}$$

Since $\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}}$ is diagonal, $\tilde{\mathbf{B}}$ can be assumed without loss of optimality (by Lemma 7) of the form $\tilde{\mathbf{B}} = \mathbf{U}_{H,1}\mathbf{\Sigma}_{B,1}$, where $\mathbf{U}_{H,1} \in \mathbb{C}^{n_T \times \check{L}}$ has as columns the eigenvectors of $\mathbf{R}_H$ corresponding to the $\check{L} \triangleq \min(L, \text{rank}(\mathbf{R}_H))$ largest eigenvalues in increasing order, and $\mathbf{\Sigma}_{B,1} = [\mathbf{0} \ \text{diag}(\{\sigma_{B,i}\})] \in \mathbb{C}^{\check{L}\times L}$ has zero elements, except along the right-most main diagonal (assumed real w.l.o.g.). The problem formulation of (19) follows by defining $z_i \triangleq \sigma_{B,i}^2$ and denoting with the set $\{\lambda_{H,i}\}_{i=1}^{\check{L}}$ the $\check{L}$ largest eigenvalues of $\mathbf{R}_H$ in increasing order. Note that the term $L_0 \triangleq L - \check{L}$ in (19) arises from the zero diagonal elements of $\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}}$.

Rewriting the MSE constraint in (19) as $\sum_{i=1}^{\check{L}}(1/(1+z_i\lambda_{H,i})) \leq \rho L - L_0$, it becomes clear that it can be satisfied for sufficiently large values of the z_i's (equivalently, the problem is feasible) if and only if $\rho > L_0/L$. ∎

[14]Both problems are equivalent in the sense that they both describe the same curve of required power P_T for a given MSE constraint ρ (in the considered case, the curve is parameterized with respect to ρ, $P_T(\rho)$, and in [15], the curve is parameterized with respect to P_T, $\rho(P_T)$).

APPENDIX B
PROOF OF PROPOSITION 1

We first obtain the closed-form solution to the problem using convex optimization theory [36], [37] and then proceed to prove the optimality of Algorithm 1.

Optimal Solution. The Lagrangian corresponding to the constrained convex problem is

$$\mathcal{L} = \sum_{i=1}^{\check{L}} z_i + \mu\left(\sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_i} - \tilde{\rho}\right) - \sum_{i=1}^{\check{L}} \gamma_i z_i,]$$

where μ and the γ_i's are the dual variables or Lagrange multipliers [36], [37]. The water-filling solution is easily found from the sufficient and necessary Karush–Kuhn–Tucker (KKT) optimality conditions (the problem satisfies the Slater's condition and therefore strong duality holds) [36], [37]:

$$\sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_i} \leq \tilde{\rho}, \quad z_i \geq 0 \tag{31}$$

$$\mu \geq 0, \quad \gamma_i \geq 0 \tag{32}$$

$$\mu\frac{\lambda_i}{(1+z_i\lambda_i)^2} + \gamma_i = 1 \tag{33}$$

$$\mu\left(\sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_i} - \tilde{\rho}\right) = 0, \quad \gamma_i z_i = 0. \tag{34}$$

Note that if $\mu = 0$, then $\gamma_i = 1 \ \forall i$ and $z_i = 0 \ \forall i$, which cannot be since the MSE constraint would not be satisfied because it was assumed that $\tilde{\rho} < \check{L}$. If $z_i > 0$, then $\gamma_i = 0$ (by the complementary slackness condition $\gamma_i z_i = 0$), $\mu(\lambda_i/(1+z_i\lambda_i)^2) = 1$ (note that $\mu\lambda_i > 1$), and $z_i = \mu^{1/2}\lambda_i^{-1/2} - \lambda_i^{-1}$. If $z_i = 0$, then $\mu\lambda_i + \gamma_i = 1$ (note that $\mu\lambda_i \leq 1$). Equivalently

$$z_i = \begin{cases} \mu^{1/2}\lambda_i^{-1/2} - \lambda_i^{-1}, & \text{if } \mu\lambda_i > 1 \\ 0, & \text{if } \mu\lambda_i \leq 1 \end{cases}$$

or, more compactly

$$z_i = \left(\mu^{1/2}\lambda_i^{-1/2} - \lambda_i^{-1}\right)^+$$

where $\mu^{1/2}$ is the water-level chosen such that $\sum_{i=1}^{\check{L}}(1/(1+z_i\lambda_i)) = \tilde{\rho}$. This solution satisfies all KKT conditions and is therefore optimal.

Optimal Algorithm. Algorithm 1 is based on hypothesis testing. It first makes the assumption that all $\check{L}$ eigenmodes are active ($z_i > 0$ for $1 \leq i \leq \check{L}$) and then checks whether the MSE constraint could be satisfied with less power, in which case, the current hypothesis is rejected, a new hypothesis with one less active eigenmode is made, and so forth.

In more detail, Algorithm 1 first reorders the eigenvalues in decreasing order. With this ordering, since $\lambda_i z_i = \left(\mu^{1/2}\lambda_i^{1/2} - 1\right)^+$, a hypothesis is completely described by the set of active eigenmodes $\tilde{L}$ (such that $z_i > 0$ for $1 \leq i \leq \tilde{L}$ and zero otherwise). This allows a reduction of the total number of hypotheses from $2^{\check{L}}$ to $\check{L}$. The initial hypothesis chooses the highest number of active eigenmodes $\tilde{L} = \check{L}$.

For each hypothesis, the water-level $\mu^{1/2}$ must be such that the considered $\tilde{L}$ eigenmodes are indeed active, whereas the rest remain inactive:

$$\begin{cases} \mu^{1/2}\lambda_i^{-1/2} - \lambda_i^{-1} > 0, & 1 \leq i \leq \tilde{L} \\ \mu^{1/2}\lambda_i^{-1/2} - \lambda_i^{-1} \leq 0, & \tilde{L} < i \leq \check{L} \end{cases}$$

or, more compactly

$$\lambda_{\tilde{L}}^{-1/2} < \mu^{1/2} \leq \lambda_{\tilde{L}+1}^{-1/2}$$

where we define $\lambda_{\check{L}+1} \triangleq 0$ for simplicity of notation. Assuming that $\lambda_{\tilde{L}} \neq \lambda_{\tilde{L}+1}$ (otherwise, the hypothesis is clearly rejected since the set of possible water-levels is empty), the algorithm checks whether the QoS constraint can be satisfied with the highest water-level of the subsequent hypothesis $\mu^{1/2} = \lambda_{\tilde{L}}^{-1/2}$, in which case, the current hypothesis is rejected since the QoS constraint can be satisfied with a lower water-level and, equivalently, with lower z_i's (a reduced power). To be more precise, the algorithm checks whether $\sum_{i=1}^{\check{L}}(1/(1+z_i\lambda_i)) = (\check{L}-\tilde{L}) + \lambda_{\tilde{L}}^{1/2}\sum_{i=1}^{\tilde{L}}\lambda_i^{-1/2} \leq \tilde{\rho}$ or, equivalently, whether $\lambda_{\tilde{L}}^{-1/2} \geq \sum_{i=1}^{\tilde{L}}\lambda_i^{-1/2}/(\tilde{\rho}-(\check{L}-\tilde{L}))$.

If the current hypothesis is rejected, the algorithm forms a new hypothesis by decreasing $\tilde{L}$ to account for the decrease of the water-level. Otherwise, the current hypothesis is accepted since it contains the optimum water-level that satisfies the QoS constraint with equality (removing more active eigenmodes would keep the QoS constraint away of being satisfied, and the addition of more active eigenmodes has already been tested and rejected for requiring a higher power to satisfy the QoS constraint). This reasoning can be applied as many times as needed for each remaining set of active eigenmodes. Once the optimal set of active eigenmodes is known, the active z_i's are obtained such that the QoS constraint is satisfied with equality, and the definitive water-level is then

$$\mu^{1/2} = \frac{\sum_{i=1}^{\tilde{L}}\lambda_i^{-1/2}}{\tilde{\rho}-(\check{L}-\tilde{L})}.$$

By the nature of the algorithm, the maximum number of iterations (worst-case complexity) is $\check{L}$. ■

APPENDIX C
PROOF OF THEOREM 2

We prove the theorem in two steps. First, we show the equivalence of the original complicated problem and a simpler problem, and then, we solve the simple problem.

The original problem in (21) (problem P1) is equivalent to the following problem (problem P2):

$$\begin{aligned} \min_{\tilde{\mathbf{B}}} \quad & \text{Tr}(\tilde{\mathbf{B}}\tilde{\mathbf{B}}^H) \\ \text{s.t.} \quad & \tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}} \quad \text{diagonal} \\ & \mathbf{d}\left((\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1}\right) \succ^w \boldsymbol{\rho}. \end{aligned}$$

Intuitively, the second constraint will guarantee the existence of a unitary matrix $\mathbf{Q}$ such that $\mathbf{d}(\mathbf{Q}^H(\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1}\mathbf{Q}) \leq \boldsymbol{\rho}$ (note the difference with respect to the case with equal MSE constraints for which there always exists $\mathbf{Q}$ such that $\mathbf{E}$ has equal diagonal elements by Corollary 1). To prove the equivalence of both problems, it suffices to show that for any feasible point $\mathbf{B}$ of problem P1 (i.e., a point that satisfies the constraints of the problem), there is a corresponding feasible point $\tilde{\mathbf{B}}$ in problem P2 with the same objective value, i.e., $\text{Tr}(\mathbf{B}\mathbf{B}^H) = \text{Tr}(\tilde{\mathbf{B}}\tilde{\mathbf{B}}^H)$, and vice-versa. Therefore, solving one problem is tantamount to solving the other problem.

We prove first one direction. Let $\mathbf{B}$ be a feasible point of problem P1 with objective value $\text{Tr}(\mathbf{B}\mathbf{B}^H)$. Define $\boldsymbol{\lambda}_B \triangleq \boldsymbol{\lambda}((\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1})$ and $\mathbf{d}_B \triangleq \mathbf{d}((\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1})$. Since $\mathbf{B}$ is feasible, $\mathbf{d}_B \leq \boldsymbol{\rho}$, and therefore, $\mathbf{d}_B \succ^w \boldsymbol{\rho}$ (see Definition 3). It then follows by Lemma 2 that $\boldsymbol{\lambda}_B \succ^w \boldsymbol{\rho}$ ($\boldsymbol{\lambda}_B \succ \mathbf{d}_B \Rightarrow \boldsymbol{\lambda}_B \succ^w \mathbf{d}_B$). Find a unitary matrix $\mathbf{Q}$ that diagonalizes $\mathbf{B}^H\mathbf{R}_H\mathbf{B}$, and define $\tilde{\mathbf{B}} = \mathbf{B}\mathbf{Q}$. It is straightforward to check that $\tilde{\mathbf{B}}$ is a feasible point of problem P2 (clearly, $\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}}$ is diagonal by selection of $\mathbf{Q}$, and therefore, $\mathbf{d}_{\tilde{B}} = \boldsymbol{\lambda}_{\tilde{B}} = \boldsymbol{\lambda}_B \succ^w \boldsymbol{\rho}$) with the same objective value ($\text{Tr}(\tilde{\mathbf{B}}\tilde{\mathbf{B}}^H) = \text{Tr}(\mathbf{B}\mathbf{B}^H)$).

We prove now the other direction. Let $\tilde{\mathbf{B}}$ be a feasible point of problem P2. Since $\boldsymbol{\lambda}_{\tilde{B}} = \mathbf{d}_{\tilde{B}} \succ^w \boldsymbol{\rho}$, by Lemma 4, there exists a vector $\check{\boldsymbol{\rho}}$ such that $\check{\boldsymbol{\rho}} \leq \boldsymbol{\rho}$ and $\boldsymbol{\lambda}_{\tilde{B}} \succ \check{\boldsymbol{\rho}}$. We can now invoke Lemma 3 to show that there exists a unitary matrix $\mathbf{Q}$ such that $\mathbf{d}(\mathbf{Q}^H(\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1}\mathbf{Q}) = \check{\boldsymbol{\rho}} \leq \boldsymbol{\rho}$. Defining $\mathbf{B} = \tilde{\mathbf{B}}\mathbf{Q}$ (choosing $\mathbf{Q}$ such that the diagonal elements of $(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}$ are in decreasing order), we have that $\mathbf{B}$ is a feasible point of problem P1 ($\mathbf{d}((\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}) \leq \boldsymbol{\rho}$ or, equivalently, $[(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}]_{ii} \leq \rho_i$). Note that if $\tilde{\mathbf{B}}$ is such that $\mathbf{d}_{\tilde{B}} \succ \boldsymbol{\rho}$, then $\mathbf{B}$ will satisfy the constraints of problem P1 with equality and vice-versa.

Now that problems P1 and P2 have been shown to be equivalent, we focus on solving problem P2, which is much simpler than problem P1. Since, in problem P2, matrix $\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}}$ is diagonal with diagonal elements in increasing order (recall that the diagonal elements of $(\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1}$ are considered in decreasing order because the ρ_i's are in decreasing order by definition), Lemma 7 can be invoked to show that $\tilde{\mathbf{B}}$ can be assumed without loss of optimality of the form $\tilde{\mathbf{B}} = \mathbf{U}_{H,1}\boldsymbol{\Sigma}_{B,1}$, where $\mathbf{U}_{H,1} \in \mathbb{C}^{n_T\times\check{L}}$ has as columns the eigenvectors of $\mathbf{R}_H$ corresponding to the $\check{L} \triangleq \min(L, \text{rank}(\mathbf{R}_H))$ largest eigenvalues in increasing order, and $\boldsymbol{\Sigma}_{B,1} = [\mathbf{0}\ \text{diag}(\{\sigma_{B,i}\})] \in \mathbb{C}^{\check{L}\times L}$ has zero elements, except along the right-most main diagonal (assumed real w.l.o.g.). Writing the weak majorization constraint of problem P2 explicitly according to Definition 3 (note that $\sum_{i=1}^{k}\rho_{(i)} = \sum_{i=L-k+1}^{L}\rho_i$ because the ρ_i's and the $\rho_{(i)}$'s are in decreasing and increasing ordering, respectively, and the same applies to the diagonal elements of $(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}$), defining $z_i \triangleq \sigma_{B,i}^2$, and denoting with the set $\{\lambda_{H,i}\}_{i=1}^{\check{L}}$ the $\check{L}$ largest eigenvalues of $\mathbf{R}_H$ in increasing order, the problem reduces to

$$\begin{aligned}\min_{\{z_i\}} \quad & \sum_{i=1}^{\check{L}} z_i \\ \text{s.t.} \quad & \sum_{i=k-L_0}^{\check{L}} \frac{1}{1+z_i\lambda_{H,i}} \leq \sum_{i=k}^{L} \rho_i, \qquad L_0 < k \leq L \\ & (L_0-k+1)+\sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_{H,i}} \leq \sum_{i=k}^{L} \rho_i, \quad 1\leq k\leq L_0 \\ & z_k \geq 0.\end{aligned}$$

Note that the term $L_0 \triangleq L-\check{L}$ for the range $1\leq k\leq L_0$ arises from the zero diagonal elements of $\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}}$.

To be exact, the explicit weak majorization constraint should also include the ordering constraints

$$\frac{1}{1+z_k\lambda_{H,k}} \geq \frac{1}{1+z_{k+1}\lambda_{H,k+1}} \quad 1\leq k<\check{L}. \tag{35}$$

Note that the remaining ordering constraints are trivially verified since $1\geq (1/(1+z_k\lambda_{H,k}))$. However, it is not necessary to include such ordering constraints since an optimal solution always satisfies them. Otherwise, we could reorder the terms $(z_i\lambda_{H,i})$'s to satisfy (35), and the solution obtained this way would still satisfy the other constraints of the convex problem with the same objective value. At this point, however, the $\lambda_{H,i}$'s would not be in increasing order and, by Lemma 7, this is not an optimal solution since the terms $(z_i\lambda_{H,i})$'s could be put back in increasing order with a lower objective value.

The problem formulation of (22) follows by noting that since $\rho_i<1$, the constraints for $1\leq k\leq L_0$ imply and are implied by the constraint for $k=1: \sum_{i=1}^{\check{L}}(1/(1+z_i\lambda_{H,i}))\leq\sum_{i=1}^{L}\rho_i - L_0$. This constraint can be satisfied for sufficiently large values of the z_i's (equivalently, the problem is feasible) if and only if $\sum_{i=1}^{L}\rho_i > L_0$ (the constraints for $L_0<k\leq L$ can always be satisfied).

It is straightforward to see that $\sum_{i=1}^{\check{L}}(1/(1+z_i\lambda_{H,i})) \leq \sum_{i=1}^{L}\rho_i - L_0$ must be satisfied with equality at an optimal point. Otherwise, z_1 could be decreased until it is satisfied with equality or z_1 becomes zero (in which case, the same reasoning applies to z_2 and so forth). This means that an optimal solution to problem P2 must satisfy $\mathbf{d}((\mathbf{I}+\tilde{\mathbf{B}}^H\mathbf{R}_H\tilde{\mathbf{B}})^{-1}) \succ \boldsymbol{\rho}$, which in turn implies that the QoS constraints in problem P1 must be satisfied with equality: $[(\mathbf{I}+\mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}]_{ii}=\rho_i$. ∎

Appendix D
Proof of Proposition 2

We first present a lemma and then proceed to prove Proposition 2.

Lemma 8: In Algorithm 2 (for different QoS constraints), the water-level of each outer iteration (if more than one) is strictly lower than that of the previous iteration.

Proof: Let $\mu([k_1,k_2])$ denote the squared water-level when applying the single water-filling of Algorithm 1 on $[k_1,k_2]$.

For any outer iteration that has more than one inner iteration (by inner iteration, we refer to one execution of steps 1 and 2 of the outer iteration), after the first inner iteration, we obtain k_0, which is the smallest index whose constraint $\sum_{i=k_0}^{\tilde{L}}(1/(1+z_i\lambda_i)) \leq \sum_{i=k_0}^{\tilde{L}}\tilde{\rho}_i$ is not satisfied. It follows that $\mu([1,\tilde{L}]) < \mu([k_0,\tilde{L}])$ because the squared water-level obtained in the first inner iteration $\mu([1,\tilde{L}])$ was not high enough to satisfy the constraint at k_0 and has to be strictly increased. It also follows that $\mu([k,\tilde{L}]) \leq \mu([1,\tilde{L}])$ for $1<k\leq k_0$ since the water-filling over $[1,\tilde{L}]$ also satisfies the constraints for $1<k<k_0$. Therefore, we have that $\mu([k,\tilde{L}])<\mu([k_0,\tilde{L}])$ for $1\leq k<k_0$. The same reasoning applies to all subsequent inner iterations. Thus, after each outer iteration on [1, $\tilde{L}$], we have that $\mu([k,\tilde{L}])<\mu([k_0,\tilde{L}])$ for $1\leq k<k_0$.

The following outer iteration (if any) is on [1, k_0-1]. Any water-filling performed in this outer iteration verifies $\mu([k,k_0-1])<\mu([k,\tilde{L}])$ for $1\leq k\leq k_0-1$, as we now show. Using $\mu([k,\tilde{L}])$ on $[k,\tilde{L}]$ implies that the constraint is satisfied with equality $\sum_{i=k}^{\tilde{L}}(1/(1+z_i\lambda_i)) = \sum_{i=k}^{\tilde{L}}\tilde{\rho}_i$. We have shown before that $\mu([k_0,\tilde{L}]) > \mu([k,\tilde{L}])$ for $1\leq k<k_0$; therefore, using $\mu([k,\tilde{L}])$ only on $[k,k_0-1]$ and $\mu([k_0,\tilde{L}])$ on $[k_0,\tilde{L}]$, the constraint is satisfied with strict inequality $\sum_{i=k}^{\tilde{L}}(1/(1+z_i\lambda_i)) < \sum_{i=k}^{\tilde{L}}\tilde{\rho}_i$, which can be rewritten as $\sum_{i=k}^{k_0-1}(1/(1+z_i\lambda_i)) < \sum_{i=k}^{k_0-1}\tilde{\rho}_i$ (recalling that $\sum_{i=k_0}^{\tilde{L}}(1/(1+z_i\lambda_i)) = \sum_{i=k_0}^{\tilde{L}}\tilde{\rho}_i$ since we are using $\mu([k_0,\tilde{L}])$ in that range). Since using $\mu([k,\tilde{L}])$ only on $[k,k_0-1]$ implies $\sum_{i=k}^{k_0-1}(1/(1+z_i\lambda_i)) < \sum_{i=k}^{k_0-1}\tilde{\rho}_i$ and using $\mu([k,k_0-1])$ implies $\sum_{i=k}^{k_0-1}(1/(1+z_i\lambda_i)) = \sum_{i=k}^{k_0-1}\tilde{\rho}_i$ (by definition), it follows that $\mu([k,\tilde{L}])>\mu([k,k_0-1])$.

Finally, since $\mu([k,k_0-1])<\mu([k,\tilde{L}])$ and $\mu([k,\tilde{L}])<\mu([k_0,\tilde{L}])$, it follows that $\mu([k,k_0-1])<\mu([k_0,\tilde{L}])$ for $1\leq k<k_0$. In other words, the water-level of each outer iteration is strictly lower than that of the previous one. ∎

Proof of Proposition 2: We first obtain the closed-form solution to the problem using convex optimization theory [36], [37] and then proceed to prove the optimality of Algorithm 2.

Optimal Solution. The Lagrangian corresponding to the constrained convex problem is

$$\mathcal{L}=\sum_{k=1}^{\check{L}} z_k+\sum_{k=1}^{\check{L}}\mu_k\left(\sum_{i=k}^{\check{L}}\frac{1}{1+z_i\lambda_i}-\sum_{i=k}^{\check{L}}\tilde{\rho}_i\right)-\sum_{k=1}^{\check{L}}\gamma_k z_k \tag{36}$$

where the μ_k's and the γ_k's are the dual variables or Lagrange multipliers [36], [37]. The water-filling solution is easily found from the sufficient and necessary KKT optimality conditions (the problem satisfies the Slater's condition and therefore strong duality holds) [36], [37]:

$$\sum_{i=k}^{\check{L}}\frac{1}{1+z_i\lambda_i}\leq\sum_{i=k}^{\check{L}}\tilde{\rho}_i, \quad z_k\geq 0 \tag{37}$$

$$\mu_k\geq 0, \qquad \gamma_k\geq 0 \tag{38}$$

$$\left(\sum_{i=1}^{k}\mu_i\right)\frac{\lambda_k}{(1+z_k\lambda_k)^2}+\gamma_k=1 \tag{39}$$

$$\mu_k\left(\sum_{i=k}^{\check{L}}\frac{1}{1+z_i\lambda_i}-\sum_{i=k}^{\check{L}}\tilde{\rho}_i\right)=0, \qquad \gamma_k z_k=0. \tag{40}$$

It is important to point out here that μ_1 cannot be zero at an optimal solution, as we now show. If $\mu_1 = 0$, then $\gamma_1 = 1$, and $z_1 = 0$. It follows then, from the inequality $\sum_{i=1}^{\check{L}}(1/(1+z_i\lambda_i)) \le \sum_{i=1}^{\check{L}} \tilde{\rho}_i$ and using the trivial assumption that $\tilde{\rho}_1 < 1$, that $\sum_{i=2}^{\check{L}}(1/(1+z_i\lambda_i)) < \sum_{i=2}^{\check{L}} \tilde{\rho}_i$ (note the strict inequality), which in turn implies that $\mu_2 = 0$. This reasoning can be repeatedly applied for $k = 2, \cdots, \check{L}$ to show that if $\mu_1 = 0$, then $\mu_k = 0\ \forall k$, but this cannot be since it would imply that $z_k = 0\ \forall k$, and then, the constraints $\sum_{i=k}^{\check{L}}(1/(1+z_i\lambda_i)) \le \sum_{i=k}^{\check{L}} \tilde{\rho}_i\ 1 \le k \le \check{L}$ would not be satisfied. Thus, it must be that $\mu_1 > 0$, which implies that $\sum_{i=1}^{\check{L}}(1/(1+z_i\lambda_i)) = \sum_{i=1}^{\check{L}} \tilde{\rho}_i$.

By defining $\tilde{\mu}_k \triangleq \sum_{i=1}^{k} \mu_i$, the KKT conditions involving the μ_k's can be more compactly rewritten as (define $\tilde{\mu}_0 \triangleq 0$)

$$\begin{aligned} &\tilde{\mu}_k \ge \tilde{\mu}_{k-1}, \qquad 1 \le k \le L \\ &\tilde{\mu}_k \frac{\lambda_k}{(1+z_k\lambda_k)^2} + \gamma_k = 1 \\ &(\tilde{\mu}_k - \tilde{\mu}_{k-1}) \left(\sum_{i=k}^{\check{L}} \frac{1}{1+z_i\lambda_i} - \sum_{i=k}^{\check{L}} \tilde{\rho}_i \right) = 0. \end{aligned} \tag{41}$$

If $z_k > 0$, then $\gamma_k = 0$ (by the complementary slackness condition $\gamma_k z_k = 0$), $\tilde{\mu}_k(\lambda_k/(1+z_k\lambda_k)^2) = 1$ (note that $\tilde{\mu}_k\lambda_k > 1$), and $z_k = \tilde{\mu}_k^{1/2}\lambda_k^{-1/2} - \lambda_k^{-1}$. If $z_k = 0$, then $\tilde{\mu}_k\lambda_k + \gamma_k = 1$ (note that $\tilde{\mu}_k\lambda_k \le 1$). Equivalently

$$z_k = \begin{cases} \tilde{\mu}_k^{1/2}\lambda_k^{-1/2} - \lambda_k^{-1}, & \text{if } \tilde{\mu}_k\lambda_k > 1 \\ 0, & \text{if } \tilde{\mu}_k\lambda_k \le 1 \end{cases}$$

or, more compactly

$$z_k = \left(\tilde{\mu}_k^{1/2}\lambda_k^{-1/2} - \lambda_k^{-1} \right)^+$$

where the water-levels $\tilde{\mu}_k^{1/2}$'s are chosen to satisfy the remaining KKT conditions

$$\begin{aligned} &\sum_{i=k}^{\check{L}} \frac{1}{1+z_i\lambda_i} \le \sum_{i=k}^{\check{L}} \tilde{\rho}_i, \qquad 1 < k \le \check{L} \\ &\sum_{i=1}^{\check{L}} \frac{1}{1+z_i\lambda_i} = \sum_{i=1}^{\check{L}} \tilde{\rho}_i \\ &\tilde{\mu}_k \ge \tilde{\mu}_{k-1} \quad (\tilde{\mu}_0 \triangleq 0) \\ &(\tilde{\mu}_k - \tilde{\mu}_{k-1}) \left(\sum_{i=k}^{\check{L}} \frac{1}{1+z_i\lambda_i} - \sum_{i=k}^{\check{L}} \tilde{\rho}_i \right) = 0. \end{aligned}$$

This solution satisfies all KKT conditions and is, therefore, optimal.

Optimal Algorithm. We now prove the optimality of the solution given by Algorithm 2 (and of the equivalent Algorithm 3) by showing that the solution it gives satisfies the KKT conditions (37)–(40). Note that after running Algorithm 2, the set $[1, \check{L}]$ is partitioned into subsets, where each one is solved by a single water-filling given by Algorithm 1. By the construction of the algorithm, the constraints $\sum_{i=k}^{\check{L}}(1/(1+z_i\lambda_i)) \le \sum_{i=k}^{\check{L}} \rho_i$ are clearly satisfied. Since the algorithm produces a partition on the set $[1, \check{L}]$ (each subset solved by a single water-filling), the following conditions are necessarily satisfied [from (31), (32), and (34)]:

$$z_k \ge 0, \quad \gamma_k \ge 0, \quad \text{and} \quad \gamma_k z_k = 0.$$

The remaining conditions are given by (41). If, for each subblock of the partition on $[1, \check{L}]$, $[k_1, k_2]$, we choose $\tilde{\mu}_k = \mu([k_1, k_2])$ for $k_1 \le k \le k_2$, it can be readily checked that they are satisfied. From Lemma 8, it follows that $\tilde{\mu}_k \ge \tilde{\mu}_{k-1}$ is verified. In addition, $\tilde{\mu}_k(\lambda_k/(1+z_k\lambda_k)^2) + \gamma_k = 1$ is satisfied by the nature of the single water-filling solution [see (33)]. Finally, since on each subblock $[k_1, k_2]$ the water-level is fixed, it follows that $(\tilde{\mu}_k - \tilde{\mu}_{k-1}) = 0$ for $k_1 < k \le k_2$ and that $\sum_{i=k_1}^{k_2}(1/(1+z_i\lambda_i)) = \sum_{i=k_1}^{k_2} \tilde{\rho}_i$; hence, condition $(\tilde{\mu}_k - \tilde{\mu}_{k-1})(\sum_{i=k}^{\check{L}}(1/(1+z_i\lambda_i)) - \sum_{i=k}^{\check{L}} \tilde{\rho}_i) = 0$ is satisfied for $k_1 \le k \le k_2$.

The worst-case number of outer iterations in Algorithm 2 is $\check{L}$, and the worst-case number of inner iterations (simple water-fillings) for an outer iteration on $[1, \tilde{L}]$ is $\tilde{L}$; consequently, the worst-case number of total inner iterations is $\check{L}(\check{L}+1)/2$. If, instead, we evaluate the complexity in terms of basic iterations (iterations within each simple water-filling), the worst-case number of these basic iterations is approximately equal to $\check{L}^2(\check{L}+1)/6$. ∎

APPENDIX E
PROOF OF LEMMA 5

Since the MSE matrix $\mathbf{E} = (\mathbf{I} + \mathbf{B}^H\mathbf{R}_H\mathbf{B})^{-1}$ is constrained to be diagonal (recall that by definition the ρ_i's are in decreasing order and the diagonal elements of $\mathbf{E}$ can be assumed in decreasing order w.l.o.g.), it follows from Lemma 7 that an optimal solution can be expressed as $\mathbf{B} = \mathbf{U}_{H,1}\mathbf{\Sigma}_{B,1}$, where $\mathbf{U}_{H,1} \in \mathbb{C}^{n_T \times \check{L}}$ has as columns the eigenvectors of $\mathbf{R}_H$ corresponding to the $\check{L} \triangleq \min(L, \text{rank}(\mathbf{R}_H))$ largest eigenvalues in increasing order ($L_0 \triangleq L - \check{L}$ is the number of zero eigenvalues used), and $\mathbf{\Sigma}_{B,1} = [\mathbf{0}\ \text{diag}(\{\sigma_{B,i}\})] \in \mathbb{C}^{\check{L} \times L}$ has zero elements, except along the right-most main diagonal (assumed real w.l.o.g.). Defining $z_i \triangleq \sigma_{B,i}^2$ and denoting with the set $\{\lambda_{H,i}\}_{i=1}^{\check{L}}$ the $\check{L}$ largest eigenvalues of $\mathbf{R}_H$ in increasing order, the original problem is then simplified to the convex problem

$$\begin{aligned} \min_{\{z_i\}} \quad & \sum_{i=1}^{\check{L}} z_i \\ \text{s.t.} \quad & \frac{1}{1+z_i\lambda_{H,i}} \le \rho_{i+L_0}, \qquad 1 \le i \le \check{L} \\ & 1 \le \rho_i, \qquad \check{L} < i \le L \\ & z_i \ge 0. \end{aligned}$$

The problem is clearly feasible if and only if $\rho_i \ge 1$ for $\check{L} < i \le L$, but this cannot be since, by definition, we know that $\rho_i < 1$. Therefore, the problem is feasible if and only if $\check{L} = L$ or, equivalently, $L \le \text{rank}(\mathbf{R}_H)$. In such a case, the optimal solution to the problem is trivially given by

$$z_i = \lambda_{H,i}^{-1}\left(\rho_i^{-1} - 1\right). \quad ∎$$

APPENDIX F
PROOF OF LEMMA 6

To study the optimality of the solution obtained in Lemma 5 (under the diagonality constraint of the MSE matrix) with respect the original problem in Theorem 2 (without the diagonality constraint), it suffices to check under which conditions the solution obtained in Lemma 5 satisfies the KKT conditions obtained in the proof of Proposition 2 (which solves the convex problem obtained in Theorem 2).

Since $z_i > 0$ for $1 \leq i \leq L$, it must be that $\gamma_i = 0$ for $1 \leq i \leq L$, and therefore, $\tilde{\mu}_i(\lambda_i/(1+z_i\lambda_i)^2) = 1 \Longrightarrow \tilde{\mu}_i = ((1+z_i\lambda_i)^2/\lambda_i) = (1/(\lambda_i\rho_i^2))$ (recall that $\lambda_i \triangleq \lambda_{H,i}$). At this point, all KKT conditions (37)–(40) are clearly satisfied, except $\tilde{\mu}_i \geq \tilde{\mu}_{i-1}$ for $1 < i \leq L$, which is satisfied if and only if

$$\lambda_i\rho_i^2 \geq \lambda_{i+1}\rho_{i+1}^2, \quad 1 \leq i < L$$

which clearly cannot be satisfied if $L > \text{rank}(\mathbf{R}_H)$ since the λ_i's are in increasing order. ∎

ACKNOWLEDGMENT

The authors wish to thank M. Bengtsson for reviewing the paper and for his helpful comments. They would also like to thank the anonymous reviewers for their comments on the paper.

REFERENCES

[1] M. L. Honig, K. Steiglitz, and B. Gopinath, "Multichannel signal processing for data communications in the presence of crosstalk," *IEEE Trans. Commun.*, vol. 38, pp. 551–558, Apr. 1990.

[2] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, pp. 3221–3231, Dec. 1994.

[3] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[4] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers Part I: unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2006, July 1999.

[5] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.

[6] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, 1998.

[7] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[8] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[9] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[10] 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.

[11] K. H. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. COM-24, no. 12, pp. 1283–1290, Dec. 1976.

[12] J. Salz, "Digital transmission over cross-coupled linear channels," *At&T Tech. J.*, vol. 64, no. 6, pp. 1147–1159, July–Aug. 1985.

[13] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1334–1347, Sept. 1994.

[14] Y. Ding, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Minimum BER block precoders for zero-forcing equalization," *IEEE Trans. Signal Processing*, vol. 51, pp. 2410–2423, Sept. 2003.

[15] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Signal Processing*, vol. 51, pp. 2381–2401, Sept. 2003.

[16] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051–1064, May 2002.

[17] F. Rashid-Farrokhi, K. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1437–1450, Oct. 1998.

[18] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. 37th Annu. Allerton Conf. Commun., Contr., Comput.*. Monticello, IL: Allerton House, Sept. 22–24, 1999, pp. 987–996.

[19] ——, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. Boca Raton, FL: CRC, 2001.

[20] J.-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter receiver diversity for efficient space division multiaccess," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 16–27, Jan. 2002.

[21] T. M. Lok and T. F. Wong, "Transmitter and receiver optimization in multicarrier CDMA systems," *IEEE Trans. Commun.*, vol. 48, pp. 1197–1207, July 2000.

[22] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1968–1983, Sept. 1999.

[23] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum joint transmit-receive linear processing for vectored DSL transmission with QoS requirements," in *Proc. 36th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 3–6, 2002.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[25] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1019–1032, Apr. 1999.

[26] G. Ginis and J. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE J. Select. Areas Commun. (Special Issue on Twisted Pair Transmission)*, vol. 20, pp. 1085–1104, June 2002.

[27] ETSI, Broadband radio access networks (BRAN); HIPERLAN type 2; physical (PHY) layer, in ETSI TS 101 475 V1.2.2, pp. 1–41, Feb. 2001.

[28] IEEE, Part 11: wireless LAN medium access control (MAC) and physical layer (PHY), in IEEE Std 802.11a, Dec. 1999.

[29] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic, 1979.

[30] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1984–1991, Sept. 1999.

[31] S. Verdú, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.

[32] C. T. Mullis and R. A. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-23, pp. 551–562, Sept. 1976.

[33] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[34] D. P. Palomar, "A Unified framework for communications through MIMO channels," Ph.D. dissertation, Tech. Univ. Catalonia (UPC), Barcelona, Spain, May 2003.

[35] J. G. Proakis, *Digital Communications*, 3rd ed: McGraw-Hill, 1995.

[36] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[37] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003, Available: http://www.stanford.edu/\~boyd/cvxbook.html.

[38] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1365–1376, Oct. 1987.

[39] ETSI, Channel models for HIPERLAN/2 in different indoor scenarios, in ETSI EP BRAN 3ERI085B, pp. 1–8, Mar. 1998.

[40] L. Schumacher, J. P. Kermoal, F. Frederiksen, K. I. Pedersen, A. Algans, and P. E. Mogensen, "MIMO channel characterization,", Deliverable D2 V1.1 of IST-1999-11 729 METRA Projecthttp://www.ist-metra.org, Feb. 2001.

Daniel Pérez Palomar (S'99–M'03) was born in Barcelona, Spain, in 1975. He received the Electrical Engineering and Ph.D. degrees from the Technical University of Catalonia (UPC), Barcelona, in 1998 and 2003, respectively.

During 1998, he was with the Department of Electronic Engineering, King's College London (KCL), London, U.K. From January 1999 to December 2003, he was a Research Assistant with the Department of Signal Theory and Communications, UPC. From April to November 2001, he held a visiting research appointment at the Department of Electrical Engineering, Stanford University, Stanford, CA. From January to December 2002, he was a visiting researcher with the Telecommunications Technological Center of Catalonia (CTTC), Barcelona. From August to November 2003, he was a Guest Researcher with the Department of Signals, Sensors, and Systems, Royal Institute of Technology (KTH), Stockholm, Sweden. From November 2003 to February 2004, he was a Visiting Researcher with the INFOCOM Department, University of Rome "La Sapienza," Rome, Italy. He is currently a Fulbright Research Fellow at Princeton University, Princeton, NJ. He has participated in several European projects such as ACTS-SUNBEAM (1999), IST-METRA (2000–2001), IST I-METRA (2001–2003), and IST ROMANTIK (2002–2004). His primary research interests include information-theoretic and communication aspects of wireless MIMO channels and array signal processing, with special emphasis on convex optimization theory applied to communications systems.

Dr. Palomar received the 2002–2003 Rosina Ribalta first prize for the Best Doctoral Thesis within the areas of information technologies and communications from the Epson Foundation. He has also received a Fulbright Research Fellowship.

Miguel Angel Lagunas (S'73–M'78–SM'89–F'97) was born in Madrid, Spain, in 1951. He received the Telecommunications Engineer degree from the Polytechnic University of Madrid (UPM) in 1973 and the Ph.D. degree in telecommunications from the Technical University of Catalonia (UPC), Barcelona, Spain.

From 1971 to 1973, he was a Research Assistant, from 1973 to 1979, a Teacher Assistant in Network Synthesis and Semiconductor Electronics, from 1979 to 1982, an Associate Professor of digital signal processing, and since 1983, Full Professor, teaching courses in signal processing, array processing, and digital commincations, all at the Semiconductor Lab., ETSIT, Madrid. He was Project Leader of high-speed SCMA from 1987 to 1989 and ATM cable networks from 1994 to 1995. He was co-director of the first projects for the European Spatial Agency and the European Union, providing engineering demonstration models on smart antennas for satellite communications using DS and FH systems (in 1986) and mobile communications GSM (Tsunami, in 1994). His research is devoted to spectral estimation, adaptive systems, and array processing. His technical activities are in advanced front-ends for digital communications combining spatial with frequency-time and coding diversity. He was Vice President for Research at UPC from 1986 to 1989 and Vice Secretary General for Research at CICYT from 1995 to 1996. Currently, he is Director of the Telecommunications Technological Center of Catalonia (CTTC), Barcelona.

Dr. Lagunas is a member-at-large of Eurasip. He is an elected member of the Academy of Engineers of Spain and of the Academy of Science and Art of Barcelona. He was a Fullbright scholar at the University of Colorado, Boulder.

John M. Cioffi (S'77–M'78–SM'90–F'96) received the B.S.E.E. degree in 1978 from the University of Illinois, Urbana-Champaign, and the Ph.D.E.E. degree in 1984 from Stanford University, Stanford, CA.

He was with Bell Laboratories, Holmdel, NJ, from 1978 to 1984; IBM Research, San Jose, CA, from 1984 to 1986; and has been a Professor of electrical engineering at Stanford since 1986. He founded Amati Communications Corp. (which was purchased by Texas Instrumnets in 1997), in 1991 and was officer/director from 1991 to 1997. He currently is on the Board of Directors of Marvell, IteX, Teknovus, Clariphy, and Teranetics. He is on the advisory boards of Actelis, Ikanos, Charter Ventures, and Portview Ventures and a is member of the US National Research Council's CSTB. His specific interests are in the area of high-performance digital transmission. He has published over 200 papers and holds over 40 patents.

Dr. Cioffi has held the Hitachi America Professorship in Electrical Engineering since 2002 and has been a Member of the National Academy of Engineering since 2001. He received the IEEE Kobayashi Medal in 2001, the IEEE Millennium Medal in 2000, the IEE J. J. Tomson Medal in 2000, was the 1999 University of Illinois Outstanding Alumnus, received the 1991 IEEE COMMUNICATIONS MAGAZINE best paper award, received the ANSI T1 Outstanding Achievement Award in 1995, and was an NSF Presidential Investigator from 1987 to 1992.