High-SNR Analytical Performance of Spatial Multiplexing MIMO Systems With CSI

Luis G. Ordóñez, *Student Member, IEEE*, Daniel P. Palomar, *Member, IEEE*, Alba Pagès-Zamora, *Member, IEEE*, and Javier Rodríguez Fonollosa, *Senior Member, IEEE*

***Abstract*—In this paper, we investigate the average and outage performance of spatial multiplexing multiple-input multiple-output (MIMO) systems with channel state information at both sides of the link. Such systems result, for example, from exploiting the channel eigenmodes in multiantenna systems. Due to the complexity of obtaining the exact expression for the average bit error rate (BER) and the outage probability, we derive approximations in the high signal-to-noise ratio (SNR) regime assuming an uncorrelated Rayleigh flat-fading channel. More exactly, capitalizing on previous work by Wang and Giannakis, the average BER and outage probability versus SNR curves of spatial multiplexing MIMO systems are characterized in terms of two key parameters: the array gain and the diversity gain. Finally, these results are applied to analyze the performance of a variety of linear MIMO transceiver designs available in the literature.**

***Index Terms*—Channel eigenmodes, diversity gain, linear MIMO transceivers, ordered eigenvalues, spatial multiplexing, Wishart.**

I. Introduction

MULTIPLE-INPUT MULTIPLE-OUTPUT (MIMO) channels are an abstract and general way to model many different communication systems of diverse physical nature; ranging from wireless multiantenna channels [1]–[4], to wireline digital subscriber line (DSL) systems [5], and to single-antenna frequency-selective channels [6]. In particular, wireless MIMO channels have been recently attracting a great interest since they provide significant improvements in terms of spectral efficiency and reliability with respect to single-input single-output (SISO) channels.

The gains obtained by the deployment of multiple antennas at both sides of the link are the array gain, the diversity gain, and the multiplexing gain [7]. The array gain is the improvement in signal-to-noise ratio (SNR) obtained by coherently combining the signals on multiple-transmit or multiple-receive dimensions while the diversity gain is the improvement in link reliability obtained by receiving replicas of the information signal through independently fading dimensions. These gains are not exclusive of MIMO channels and also exist in single-input multiple-output (SIMO) and multiple-input single-output (MISO) channels. In contrast, the multiplexing gain, which refers to the increase of rate at no additional power consumption, is a unique characteristic of MIMO channels. The basic idea is to exploit the multiple dimensions to open up several parallel subchannels within the MIMO channel, also termed *channel eigenmodes*. This allows the transmission of several symbols simultaneously or, in other words, the establishment of several substreams for communication.

In order to simplify the study of MIMO systems, it is customary to divide them into an uncoded part, which transmits symbols drawn from some constellations, and a coded part that builds upon the uncoded system. Although the ultimate system performance depends on the combination of both parts, it is convenient to consider the uncoded and coded parts independently to simplify the analysis and design.

In this paper, we focus on the uncoded part of the system and, specifically, on spatial multiplexing MIMO systems with perfect channel state information (CSI) at both sides of the link. Spatial multiplexing is a simple MIMO transmit technique that allows a high spectral efficiency by dividing the incoming data into multiple independent substreams and transmitting each substream on a different antenna [8], [1] (no CSI at the transmitter is required for this approach). When perfect CSI is available at the transmitter, channel-dependent linear precoding of the data substreams can further improve performance by adapting the transmitted signal to the instantaneous channel eigenstructure. There are different degrees of adaptation as have been considered in the literature, namely the following.

i) Adapt only the linear precoder/power allocation among the different substreams, keeping the number of substreams and the constellations fixed. This is by far the most widely considered scenario, e.g., [6], [9]–[16].
ii) Adapt the precoder/power allocation among the different substreams, the number of substreams, and choice of constellations, keeping the data rate fixed. This has been only partially considered in a few papers. For instance, the adaptation of the precoder and number of substreams (with equal constellations) is considered in [17] and [18], and the precoder, constellations, and number of substreams are designed in [19] to minimize the transmitted power under a BER constraint.

Manuscript received April 12, 2006; revised February 8, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ananthram Swami. This work has been partially funded by the European Commission (27187 SURFACE), the Spanish Ministry of Education and Science and FEDER funds (TEC2006-06481 and TEC2004-04526), the Catalan Government (2005SGR-00639), and by the Research Grant DAG06/07.EG02. Part of this work was presented at the Sixth IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), New York City, NY, June 2005.

L. G. Ordóñez, A. Pagès-Zamora, and J. R. Fonollosa are with the Department of Signal Theory and Communications, Technical University of Catalonia (UPC), 08034 Barcelona, Spain (e-mail: luisg@gps.tsc.upc.es; alba@gps.tsc.upc.es; fono@gps.tsc.upc.es).

D. P. Palomar is ith the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: palomar@ust.hk).

Digital Object Identifier 10.1109/TSP.2007.896109

This paper concentrates on case i), which embraces the schemes that have received more attention in the literature and for which no analytical performance analysis was available. First, we analyze the high-SNR performance of the substreams transmitted through the channel eigenmodes, when fixed constellations are imposed beforehand and the channel is a Rayleigh flat-fading MIMO channel. Since simple closed-form expressions for the marginal pdf's of the ordered eigenvalues of the Wishart matrix are not known, we first derive a first order polynomial expansion of these marginal probability density functions (pdf's). Then, based on this result, we study the average and outage performance of each channel eigenmode at high SNR following the methodology introduced by Wang and Giannakis in [20]. To be more precise, we characterize the curves corresponding to average BER versus SNR and outage probability versus SNR in terms of the *diversity gain*, which determines the slope of the curve at high SNR in a log–log scale, and the *array gain*, which determines the horizontal shift of the BER curve. We also extend this characterization to the global performance that takes into account all the established substreams for a fixed number of substreams according to scenario i).

These general results are then applied to analyze the performance of linear MIMO transceivers which are low-complexity schemes composed of a linear precoder at the transmitter and a linear equalizer at the receiver. The design of linear transceivers in the context of i) has been studied for many years according to a variety of criteria [6], [9]–[16]. Recently, a general and unifying framework for the joint linear transmit-receive design was developed in [15], embracing a wide range of different design criteria.

The rest of the paper is organized as follows. Section II is devoted to introducing the performance metrics of a general digital communication system in a fading environment and to presenting how these performance measures can be approximated in the high-SNR regime. Section III describes the signal model corresponding to a spatial multiplexing scheme with CSI. In Section IV, we derive a closed-form characterization for the individual and global average BER and outage performance at high SNR of spatial multiplexing MIMO systems. Then, in Section V, we apply these results to analyze the performance of linear MIMO transceivers by dividing the designs proposed in [15] into diagonal schemes with fixed power allocation, diagonal schemes with nonfixed power allocation, and nondiagonal schemes. Finally, in Section VI, we summarize the main contribution of the paper.

II. Preliminaries: Performance Metrics of Digital Communication Systems

A. Average and Outage Performance

The ultimate performance metric of a digital communication system is given in terms of the error probability, defined as the fraction of symbols in error, or in terms of the BER, defined as the fraction of bits in error. For fading channels, different realizations of the random channel have to be taken into account, leading to the concept of average and outage error probabilities. The average BER measures the bit error probability averaged over different channel realizations, whereas the outage probability is the probability that the system performance is below a given threshold.

In the presence of additive white Gaussian noise, the instantaneous BER of a digital communication system (with a linear modulation) as a function of the instantaneous SNR, denoted by γ, can be analytically approximated, when using a Gray encoding procedure, as [21] (see the exact expression in [22] and [23])

$$\mathrm{BER}(\gamma) \approx \frac{\alpha}{\log_2 M} \mathcal{Q}(\sqrt{\beta\gamma}) \tag{1}$$

where $\mathcal{Q}(\,\cdot\,)$ is the Gaussian Q function defined as

$$\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt \tag{2}$$

and the parameters α and β depend on the constellation. For instance, when using an M-QAM modulation, α and β are given by

$$\alpha \triangleq \alpha(M) = 4\left(1 - \frac{1}{\sqrt{M}}\right) \tag{3}$$

$$\beta \triangleq \beta(M) = \frac{3}{M-1}. \tag{4}$$

The effect of the channel on the transmitted signal is, under the flat-fading assumption, in the form of a random multiplicative distortion [24]. According to this, the instantaneous SNR γ is given by the product of a channel-dependent parameter μ and a deterministic positive quantity $\bar{\gamma}$, i.e.,

$$\gamma = \mu\bar{\gamma} \tag{5}$$

where $\bar{\gamma}$ is the average SNR at the receiver whenever $\mathrm{E}\{\mu\} = 1$. As we are interested in the average BER incurred by the system, we need to take the expectation over all possible channel states:

$$\overline{\mathrm{BER}}(\bar{\gamma}) \triangleq \mathrm{E}_\mu\{\mathrm{BER}(\gamma)\} = \int_0^\infty \mathrm{BER}(\mu\bar{\gamma}) f_\mu(\mu)\, d\mu \tag{6}$$

where $f_\mu(\mu)$ is the pdf of the channel-dependent parameter μ.

The average BER in (6) has been analytically evaluated only in certain simple cases, such as when the channel parameter is Rayleigh distributed [25]. A general procedure for finding the error of linearly modulated signals is given in [26], consisting of deriving the pdf of an equivalent noise which comprises the effects of the additive noise and the multiplicative distortion. However, this requires the numerical computation of integrals, and thus the resulting expressions do not help to gain insight about the system behavior. In [27], the authors provide a unified method for calculating these error rates based on an alternative representation of the Gaussian Q-function. Again, the resulting expressions contain integrals that must be numerically solved or simply that are too difficult to work with.

The average BER is a useful performance measure when the transmission interval is long enough to reveal the long-term ergodic properties of the fading channel. The ergodicity assumption, however, is not always necessarily satisfied in practical communication systems operating over fading channels, because no significant channel variability may occur during the whole transmission. In these circumstances, the most convenient measure to capture the performance of the system is the outage BER, i.e., the minimum BER value guaranteed with a small given probability. However, the outage BER is difficult to obtain analytically, and the performance of communication systems over nonergodic fading channels is instead commonly measured with the outage probability [28], defined as the probability that the instantaneous SNR γ falls below a certain threshold $\gamma_{\rm th}$:

$$P_{\rm out}(\bar{\gamma}) \triangleq \Pr(\gamma \le \gamma_{\rm th}) = \int_0^{\gamma_{\rm th}/\bar{\gamma}} f_\mu(\mu) d\mu. \tag{7}$$

Given the limited availability of closed-form expressions for the performance measures in (6) and (7), there is an imperative need to characterize the performance of a system in a simple and insightful way. A convenient method to find simple expressions is to allow a certain degree of approximation. The most common approach is to shift the focus from exact performance to large SNR performance as done in [20]. To be more precise, we are interested in deriving high SNR closed-form expressions for the performance measures in (6) and (7).

B. High-SNR Performance

In the high-SNR regime, the average BER function can be approximated in most cases as [20]

$$\overline{\rm BER}(\bar{\gamma}) \approx (G_c \cdot \bar{\gamma})^{-G_d} \tag{8}$$

where G_d and G_c are referred to as the diversity and coding gains,[1] respectively, and $\bar{\gamma}$ is the average SNR. The diversity gain determines the slope of the BER versus $\bar{\gamma}$ curve at high SNR in a log–log scale, and the coding gain determines the shift of the curve with respect to the benchmark BER curve $\bar{\gamma}^{-G_d}$. This leads to a simple parameterized average BER characterization for high SNR that can provide meaningful insights related to the system behavior.

As shown in [20], the average BER and the outage probability of a communication system at high SNR with the instantaneous SNR given in (5) depend only on the behavior of the pdf $f_\mu(\mu)$ near the origin ($\mu \to 0^+$) and can be parameterized as stated next.

Lemma 1 [20, Prop. 1]: The average BER of a communication system as defined in (6), in which the instantaneous SNR is given by $\gamma = \bar{\gamma}\mu$ and the pdf of the channel-dependent parameter μ can be written as[2] $f_\mu(\mu) = a\mu^d + o(\mu^d)$, is

$$\overline{\rm BER}(\bar{\gamma}) = (G_c \cdot \bar{\gamma})^{-G_d} + o(\bar{\gamma}^{-G_d}) \tag{9}$$

where the diversity gain G_d and the coding gain G_c are given by

$$G_d = d + 1 \tag{10}$$

$$G_c = \beta \left(\frac{\alpha}{\log_2 M} \frac{a 2^d \Gamma(d + 3/2)}{\sqrt{\pi}(d+1)} \right)^{-1/(d+1)} \tag{11}$$

and $\Gamma(\cdot)$ denotes the Gamma function defined as $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ [45, eq. (6.1.1)].

Lemma 2 [20, Prop. 5]: The outage probability of a communication system as defined in (7), in which the instantaneous SNR is given by $\gamma = \bar{\gamma}\mu$ and the pdf of the channel-dependent parameter μ can be written as $f_\mu(\mu) = a\mu^d + o(\mu^d)$, is

$$P_{\rm out}(\bar{\gamma}) = (O_c \cdot \bar{\gamma})^{-O_d} + o\left(\bar{\gamma}^{-O_d}\right) \tag{12}$$

where the outage diversity gain O_d and the outage coding gain O_c are given by

$$O_d = d + 1 \tag{13}$$

$$O_c = \frac{1}{\gamma_{\rm th}} \left(\frac{a}{d+1} \right)^{-1/(d+1)}. \tag{14}$$

Lemmas 1 and 2 offer a simple and unifying approach to evaluate the average and outage performance of communication systems over random fading channels and allow the interpretation of the effect of the system parameters in the performance. These results can be easily extended to the case where the instantaneous SNR is given by $\gamma = \bar{\gamma}\mu + \phi$ and ϕ is a fixed deterministic parameter. The following corollary shows the generalization of Lemma 1 (Lemma 2 can be generalized in the same way).

Corollary 1: The average BER of a communication system as defined in (6), in which the instantaneous SNR is given by $\gamma = \bar{\gamma}\mu + \phi$ and the pdf of the channel-dependent parameter μ can be written as $f_\mu(\mu) = a\mu^d + o(\mu^d)$, is

$$\overline{\rm BER}(\bar{\gamma}) = (G_c \cdot \bar{\gamma})^{-G_d} + o(\bar{\gamma}^{-G_d}) \tag{15}$$

where the diversity gain G_d and the coding gain G_c are given by

$$G_d = d + 1 \tag{16}$$

$$G_c = \beta \left(\frac{\alpha}{\log_2 M} \frac{a I(d, \beta\phi)}{\sqrt{2\pi}(d+1)} \right)^{-1/(d+1)} \tag{17}$$

and $I(d, \beta\phi)$ is defined as[3]

$$I(d, \beta\phi) = \int_{\sqrt{\beta\phi}}^\infty e^{-\frac{x^2}{2}} (x^2 - \beta\phi)^{(d+1)} dx. \tag{18}$$

Proof: See Appendix I.

III. Signal Model: Spatial Multiplexing MIMO Systems With CSI

The signal model corresponding to a transmission through a general MIMO channel with $n_{\rm T}$ transmit and $n_{\rm R}$ receive dimensions is

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{w} \tag{19}$$

where $\mathbf{s} \in \mathbb{C}^{n_{\rm T} \times 1}$ is the transmitted vector, $\mathbf{H} \in \mathbb{C}^{n_{\rm R} \times n_{\rm T}}$ is the channel matrix, $\mathbf{y} \in \mathbb{C}^{n_{\rm R} \times 1}$ is the received vector, and $\mathbf{w} \in \mathbb{C}^{n_{\rm R} \times 1}$ is a spatially white zero-mean circularly

[1]The coding gain is also known as the array gain in the context of multiantenna systems [29].

[2]We say that $f(x) = o(g(x))$ if $f(x)/g(x) \to 0$ as $x \to 0$ [30, eq. (1.3.1)].

[3]A closed-form expression for this integral does not exist for a general value of the parameter d; however, it can be easily evaluated for the most common values of d (integers).

symmetric complex Gaussian noise vector normalized so that $\mathrm{E}\{\mathbf{w}\mathbf{w}^\dagger\} = \mathbf{I}_{n_\mathrm{R}}$. The channel matrix $\mathbf{H}$ contains the complex path gains $[\mathbf{H}]_{ij}$ between every transmit and receive antenna pair. We adopt an uncorrelated Rayleigh flat-fading channel model and, consequently, these coefficients are independent identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, i.e., $[\mathbf{H}]_{ij} \sim \mathcal{CN}(0,1)$.

Following the singular value decomposition (SVD) analysis, the channel matrix $\mathbf{H}$ can be written as

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^\dagger \tag{20}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and $\mathbf{\Lambda}$ is a diagonal matrix containing the singular values of $\mathbf{H}$ sorted in descending order. This way, the channel matrix is effectively decomposed into $\mathrm{rank}\{\mathbf{H}\} = \min\{n_\mathrm{T}, n_\mathrm{R}\}$ independent orthogonal modes of excitation, which are referred to as channel eigenmodes [2], [7], [31].

Assuming that the channel is perfectly known at the transmitter and that $K \leq \min\{n_\mathrm{T}, n_\mathrm{R}\}$ data symbols per channel use have to be communicated, the transmitted vector can be written as

$$\mathbf{s} = \mathbf{V}_K\sqrt{\mathbf{P}}\mathbf{x} \tag{21}$$

where $\mathbf{x}$ gathers the K data symbols (zero mean, unit energy and uncorrelated, i.e., $\mathrm{E}\{\mathbf{x}\mathbf{x}^\dagger\} = \mathbf{I}_K$), $\mathbf{V}_K$ is formed with the K columns of $\mathbf{V}$ associated with the K strongest channel eigenmodes and $\mathbf{P} = \mathrm{diag}\left(\{p_k\}_{k=1,\ldots,K}\right)$ is a diagonal matrix containing the power allocated to each established substream. The transmitted power is constrained such that

$$\mathrm{E}\{\|\mathbf{s}\|^2\} = \sum_{k=1}^{K} p_k \leq \mathsf{snr} \tag{22}$$

where snr is the SNR per receive antenna. Assuming perfect channel knowledge also at the receiver, the symbols transmitted through the channel eigenmodes are recovered from the received signal $\mathbf{y}$ with matrix $\mathbf{U}_K$, similarly defined to $\mathbf{V}_K$, as

$$\hat{\mathbf{x}} = \mathbf{U}_K^\dagger(\mathbf{H}\mathbf{V}_K\sqrt{\mathbf{P}}\mathbf{x} + \mathbf{w}) = \mathbf{\Lambda}_K\sqrt{\mathbf{P}}\mathbf{x} + \mathbf{n} \tag{23}$$

where $\mathbf{\Lambda}_K$ is a diagonal matrix that contains the K largest singular values in descending order, and the noise vector $\mathbf{n} = \mathbf{U}_K^\dagger\mathbf{w}$ has the same statistical properties as $\mathbf{w}$, possibly with a reduced dimension. Each substream experiences then an instantaneous SNR given by

$$\rho_k = \lambda_k p_k \quad k = 1, \ldots, K \tag{24}$$

where λ_k denotes the kth largest channel eigenvalue (squared modulus of the kth channel singular value) and p_k defines the power allocation policy. Observe that most linear MIMO transceiver schemes proposed in the literature can be written as in (24), e.g., [6], [12], and [15] (other schemes include an additional pre- and postprocessing of the data symbols $\mathbf{x}$ and will also be treated in Section V-D).

IV. High-SNR Analytical Performance of Spatial Multiplexing MIMO Systems With CSI

In this section, we apply the results presented in Lemmas 1 and 2 to analyze the performance offered by the spatial multiplexing MIMO systems with CSI described in Section III in terms of array gain and diversity gain. We first obtain a parameterized analytic expression for the average BER and the outage probability of each *individual* substream, distinguishing between fixed and nonfixed power allocations. Then, we derive the *global* average BER and outage probability of the spatial multiplexing MIMO system, taking into account all established substreams.

A. Individual Performance With Fixed Power Allocation

We consider the diagonalized MIMO system obtained in (23). As the additive noise is assumed to be Gaussian distributed, the instantaneous uncoded BER of the substream transmitted through the kth channel eigenmode can be analytically approximated as

$$\mathrm{BER}_k(\rho_k) \approx \frac{\alpha_k}{\log_2 M_k}\mathcal{Q}(\sqrt{\beta_k\rho_k}) \tag{25}$$

where α_k, β_k, and M_k are the constellation parameters, and ρ_k is the instantaneous SNR in (24). Assuming that the power allocation is fixed, i.e.,

$$p_k = \phi_k\mathsf{snr} \tag{26}$$

where ϕ_k is a positive constant independent of the channel with $\sum_{i=1}^{K}\phi_k = 1$, the instantaneous SNR of (24) can be rewritten as

$$\rho_k = \lambda_k\phi_k\mathsf{snr}. \tag{27}$$

The average BER is then

$$\begin{aligned}\overline{\mathrm{BER}}_k(\mathsf{snr}) &\triangleq \mathrm{E}_{\lambda_k}\{\mathrm{BER}_k(\rho_k)\} \\ &= \int_0^\infty \mathrm{BER}_k\left(\phi_k\lambda_k\mathsf{snr}\right) f_{\lambda_k}(\lambda_k)d\lambda_k\end{aligned} \tag{28}$$

where $f_{\lambda_k}(\lambda_k)$ is the marginal pdf of the kth largest eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$. Due to the uncorrelated Rayleigh flat-fading assumption, $\mathbf{H}^\dagger\mathbf{H}$ is complex central Wishart distributed [32, Sec. 3.7]. As the marginal pdf of the kth eigenvalue of a Wishart matrix is not known in closed form and only some results for the largest and smallest eigenvalues are available in the literature (see, e.g., [33] and [34]), the average BER in (28) cannot be analytically computed. However, the marginal pdf $f_{\lambda_k}(\lambda_k)$ can be approximated for $\lambda_k \to 0^+$ as shown in the following theorem.[4]

Theorem 1: Let the entries of the $n_\mathrm{R} \times n_\mathrm{T}$ matrix $\mathbf{H}$ be i.i.d. complex Gaussian with zero mean and unit variance and let $n = \min\{n_\mathrm{T}, n_\mathrm{R}\}$ and $m = \max\{n_\mathrm{T}, n_\mathrm{R}\}$. The first-order expansion of the marginal pdf of the kth largest eigenvalue of the complex central Wishart matrix $\mathbf{H}\mathbf{H}^\dagger$ is given by

$$f_{\lambda_k}(\lambda_k) = a_k\lambda_k^{d_k} + o(\lambda_k^{d_k}) \quad k = 1, \ldots, n \tag{29}$$

[4]The Wishart distribution appears in other fields in signal processing (e.g., [35]–[38]), physics, and applied mathematics (see [34, Ch. 2] and references therein); thus, our result could have applications to other problems.

where

$$d_k = (m-k+1)(n-k+1)-1 \tag{30}$$
$$a_k = K_{m,n}^{-1}|\mathbf{A}(k)||\mathbf{B}(k)| \tag{31}$$

with $K_{m,n} = \prod_{i=1}^{n}(n-i)!(m-i)!$. Matrix $\mathbf{A}(k)$ is defined for $k \neq 1$ as

$$[\mathbf{A}(k)]_{ij} = (b(i+j)+2(n-k))! \quad i,j = 1,\ldots,(k-1) \tag{32}$$

and $\mathbf{A}(1) = 1, \mathbf{B}(k)$ is defined for $k \neq n$ as

$$[\mathbf{B}(k)]_{ij} = \frac{2}{(b(i+j)^2-1)b(i+j)} \quad i,j = 1,\ldots,(n-k) \tag{33}$$

and $\mathbf{B}(n) = 1$, and $b(i) = m-n+i$.

Proof: See Appendix II.

Now, using Theorem 1 and following the approach of [20] presented in Lemma 1, the average BER performance of the substream transmitted through the kth channel eigenmode can be parameterized in terms of array gain and diversity gain as follows.

Theorem 2: The average BER of a substream transmitted through the kth strongest eigenmode of a Rayleigh flat-fading $n_{\mathrm{R}} \times n_{\mathrm{T}}$ MIMO channel, when the power allocation is fixed as in (26), is

$$\overline{\mathrm{BER}}_k(\mathsf{snr}) = (G_a(k,\phi_k)\cdot \mathsf{snr})^{-G_d(k)} + o(\mathsf{snr}^{-G_d(k)}) \tag{34}$$

where the diversity gain and the array gain are given by

$$G_d(k) = (n_{\mathrm{T}}-k+1)(n_{\mathrm{R}}-k+1) \tag{35}$$
$$G_a(k,\phi_k) = \beta_k\phi_k\left(\frac{\alpha_k}{\log_2 M_k}\frac{a_k 2^{d_k}\Gamma(d_k+3/2)}{\sqrt{\pi}(d_k+1)}\right)^{-1/(d_k+1)} \tag{36}$$

and the parameters a_k and d_k model the fading distribution as given in Theorem 1.

Proof: The proof follows from Lemma 1 and Theorem 1.

In Fig. 1, we provide the average BER curves (obtained through numerical simulation) attained by the established substreams through the eigenmodes of a MIMO channel with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$. We assume all substreams active ($K = 4$), a uniform power allocation ($\phi_k = 1/K = 1/4$), and all data symbols drawn from a quadrature phase-shift keying (QPSK) constellation. We can see that the result given in Theorem 2 predicts correctly the diversity and array gain and, thus, approximates the average BER performance at medium to high SNR. Observe that, when the diversity gain is high, as for the first substreams ($k = 1,2$) in Fig. 1, the BER decreases with the SNR so rapidly that the given approximation is only accurate for very small BER values.

Theorem 2 can be easily particularized for the case of having multiple antennas at only one side of the link ($n_{\mathrm{T}} = 1$ or $n_{\mathrm{R}} = 1$), i.e., MISO systems and SIMO systems.

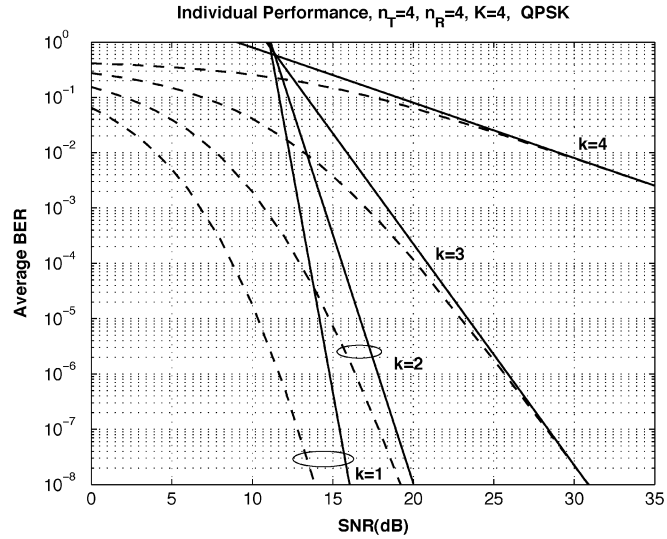

Fig. 1. Individual simulated average BER (dashed line) and parameterized average BER (solid line) of the substreams transmitted through the channel eigenmodes ($k = 1,2,3,4$) with $n_{\mathrm{T}} = 4$ and $n_{\mathrm{R}} = 4$.

Corollary 2: The average BER of the single substream that can be established in a MISO system with m transmit antennas and one receive antenna or in a SIMO system with one transmit antenna and m receive antennas is

$$\overline{\mathrm{BER}}(\mathsf{snr}) = (G_a\cdot\mathsf{snr})^{-G_d} + o(\mathsf{snr}^{-G_d}) \tag{37}$$

where the diversity gain and the array gain are given by

$$G_d = m \tag{38}$$
$$G_a = \beta\left(\frac{\alpha}{\log_2 M}\frac{2^{m-1}\Gamma(m+1/2)}{m!\sqrt{\pi}}\right)^{-1/m}. \tag{39}$$

The performance attained by an MISO and SIMO system coincide when perfect CSI is available at the transmitter and the receiver,[5] since in both cases the m signal components can be coherently combined (either at the transmitter or at the receiver) leading to the m-fold diversity in (38). This equivalence has been already pointed out in the literature, when analyzing the capacity of MIMO channels [3]. It is insightful to further particularize Corollary 2 to a SISO channel ($m = 1$), obtaining $G_d = 1$ and $G_a = \beta(\alpha/\log_2 M)^{-1}$, i.e., no diversity gain and no additional array gain.

Consider now the outage probability, given for the kth established substream by

$$P_{\mathrm{out},k}(\mathsf{snr}) \triangleq \Pr(\rho_k \le \rho_{\mathrm{th}}) = \int_0^{\mathsf{snr}\phi_k/\rho_{\mathrm{th}}} f_{\lambda_k}(\lambda_k)d\lambda_k \tag{40}$$

which depends on the marginal pdf of the kth largest eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$ as in the case of average BER.

Theorem 3: The outage probability of a substream using the kth eigenmode of a Rayleigh flat-fading $n_{\mathrm{R}} \times n_{\mathrm{T}}$ MIMO channel, when the power allocation is fixed as in (26), is

$$P_{\mathrm{out},k}(\mathsf{snr}) = (O_a(k,\phi_k)\cdot\mathsf{snr})^{-O_d(k)} + o(\mathsf{snr}^{-O_d(k)}) \tag{41}$$

[5]Observe that without CSI at the transmitter this equivalence does not hold anymore and the SIMO systems outperforms the MISO system.

where the outage diversity and the outage array gain are given by

$$O_d(k) = (n_{\rm T} - k + 1)(n_{\rm R} - k + 1) \tag{42}$$

$$O_a(k, \phi_k) = \frac{\phi_k}{\rho_{\rm th}} \left(\frac{a_k}{d_k + 1} \right)^{-1/(d_k+1)} \tag{43}$$

and the parameters a_k and d_k model the fading distribution as given in Theorem 1.

Proof: The proof follows from Lemma 2 and Theorem 1.

B. Individual Performance With Nonfixed Power Allocation

We have so far analyzed fixed power allocation strategies as expressed in (26). We now consider power allocation strategies that depend on the eigenvalues associated with the K active channel eigenmodes.

Theorem 4: Consider that the power allocated to a substream transmitted through the kth strongest eigenmode of a Rayleigh flat-fading $n_{\rm R} \times n_{\rm T}$ MIMO channel is a function of the K strongest channel eigenvalues and the SNR, i.e., $p_k = g_k(\lambda_1, \ldots, \lambda_K, \mathsf{snr})$ under the short-term power constraint in (22), and that the power allocation satisfies

$$\Pr(p_k \leq \phi_k \mathsf{snr}) = a(\phi_k)\mathsf{snr}^{-d(\phi_k)} + o(\mathsf{snr}^{-d(\phi_k)}) \tag{44}$$

where $\phi_k, a(\phi_k)$, and $d(\phi_k)$ are positive deterministic parameters. Then, the average BER of the kth substream is

$$\overline{\rm BER}_k(\mathsf{snr}) = (\tilde{G}_a(k) \cdot \mathsf{snr})^{-\tilde{G}_d(k)} + o(\mathsf{snr}^{-\tilde{G}_d(k)}) \tag{45}$$

where the diversity gain is given by

$$\tilde{G}_d(k) = d_k + 1 = (n_{\rm T} - k + 1)(n_{\rm R} - k + 1) \tag{46}$$

whenever[6] $d(\phi_k) \geq d_k + 1$ and the array gain $\tilde{G}_a(k)$ can be bounded by distinguishing between the following two cases.

i) If there exists $\phi_k \in (0,1)$ such that $d(\phi_k) > d_k + 1$, the array gain is bounded as

$$G_a(k, \phi_k) \leq \tilde{G}_a(k) < G_a(k, 1). \tag{47}$$

ii) If $d(0) = d_k + 1$ and there exists $\phi_k \in (0,1)$ such that $d(\phi_k) = d_k + 1$, the array gain is bounded as

$$\left(G_a(k, \phi_k)^{-(d_k+1)} + \frac{\alpha_k}{2 \log_2 M_k} a(\phi_k) \right)^{-1/(d_k+1)} \leq \tilde{G}_a(k) < G_a(k, 1) \tag{48}$$

where $G_a(k, \phi_k)$ is the array gain obtained when using a fixed power allocation and is defined in (36).

Proof: See Appendix III.

[6]The condition $d(\phi_k) \geq d_k + 1$ is satisfied by any reasonable power allocation. Namely, if $d(\phi_k) < d_k + 1$, the average BER performance is inherently limited by the power allocation and not by the statistical properties of the channel.

C. Global Performance

We now consider the global performance of the spatial multiplexing MIMO system described in Section III, i.e., when transmitting over the K strongest channel eigenmodes. For this purpose, we define first the global instantaneous BER as the arithmetic mean of the instantaneous BER of the K established substreams:

$${\rm BER}(\{\rho_k\}_{k=1,\ldots,K}) = \frac{1}{K} \sum_{i=1}^{K} {\rm BER}_k(\rho_k). \tag{49}$$

Note that the global instantaneous BER in (49) takes into account the instantaneous BER performance experienced by each one of the K data symbols to be transmitted (i.e., the number of substreams K is fixed regardless of the power assigned to each substream, which can even be zero for some power allocation strategies under poor propagation conditions). Finally, we obtain the global average BER performance of the spatial multiplexing MIMO system by averaging the global instantaneous BER in (49) over all possible channel states:

$$\begin{aligned} \overline{\rm BER}(\mathsf{snr}) &= {\rm E}\left\{{\rm BER}(\{\rho_k\}_{k=1,\ldots,K})\right\} \\ &= \frac{1}{K} \sum_{i=1}^{K} \overline{\rm BER}_k(\mathsf{snr}). \end{aligned} \tag{50}$$

Since $G_d(1) > G_d(2) > \cdots > G_d(K)$ (see Theorem 2 for fixed and Theorem 4 for non-fixed power allocations), the global average BER is dominated by the average BER associated with the Kth substream:

$$\overline{\rm BER}(\mathsf{snr}) = \frac{1}{K} \left(G_a(K) \cdot \mathsf{snr} \right)^{-G_d(K)} + o(\mathsf{snr}^{-G_d(K)}). \tag{51}$$

This result is summarized in the following theorem.

Theorem 5: The global average BER attained when transmitting through the K strongest eigenmodes of a Rayleigh flat-fading $n_{\rm R} \times n_{\rm T}$ MIMO channel with either a fixed power allocation (see Theorem 2) or a nonfixed power allocation (see Theorem 4) is

$$\overline{\rm BER}(\mathsf{snr}) = (G_a \cdot \mathsf{snr})^{-G_d} + o(\mathsf{snr}^{-G_d}) \tag{52}$$

where the diversity gain G_d and the array gain G_a are given by

$$G_d = (n_{\rm T} - K + 1)(n_{\rm R} - K + 1) \tag{53}$$

$$G_a = \left(\frac{1}{K} \right)^{-1/G_d} G_a(K) \tag{54}$$

and $G_a(K)$ is the array gain of the Kth substream given either by (36) for fixed power allocations or by (47) or (48) for nonfixed power allocations.

It is interesting to note that when particularizing Theorem 5 for $K = 1$, i.e., a beamforming strategy, we obtain the full diversity of the channel $n_{\rm T} n_{\rm R}$, as has been widely observed in the literature, e.g., [39] and [40]. The diversity order of the case $K = \min\{n_{\rm T}, n_{\rm R}\}$ has been also previously documented in [41].

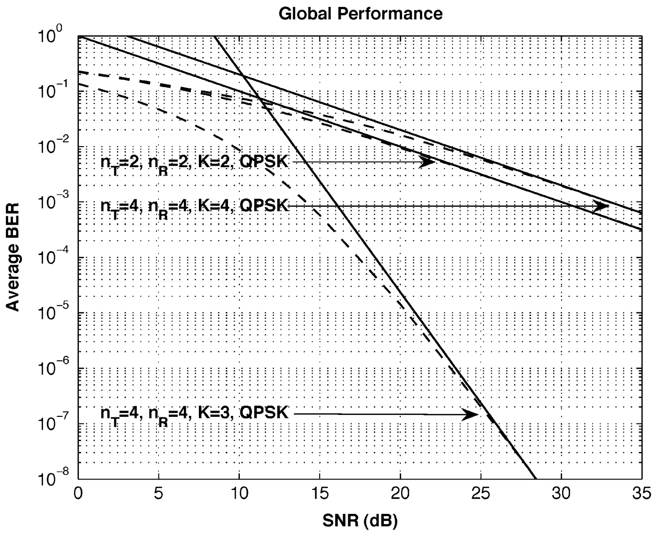

Fig. 2. Global simulated average BER (dashed line) and parameterized average BER (solid line) of the spatial multiplexing MIMO system with CSI.

In Fig. 2, we provide the global average BER curve (obtained through numerical simulation) of the spatial multiplexing MIMO system with $n_{\mathrm{T}} = n_{\mathrm{R}} = 2$ and $K = 2$, and with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$ and $K = 3$ and $K = 4$. In all three cases, the theoretical global average BER performance is correctly approximated by the parameterized characterization proposed in Theorem 5.

The global outage probability is defined similarly to the global average BER, i.e., we say that the global performance of the spatial multiplexing MIMO system is below some certain threshold, when the performance averaged over the established substreams is below this threshold. Hence, the outage performance, analogously to the average BER performance, is dictated by the behavior of the Kth channel eigenmode and the generalization of the individual outage probability characterization given in Theorem 3 to the global outage probability can be easily obtained.

V. High-SNR Analytical Performance Analysis of Linear MIMO Transceivers

The performance of linear MIMO transceivers has been always analyzed numerically, due to the difficulty of finding a closed-form expression for the average bit error probability. For instance, in [12], we can find the simulated average BER curves for linear precoding schemes designed under the weighted minimum mean-square error (MMSE) criterion in a Rayleigh flat-fading channel. Other numerical results can also be found in [6], [10], and [15]. The advantage of obtaining numerical results via computer simulation is that they provide the performance in realistic environments. However, they do not give insight on the behavior of the system as analytical expressions do. In this section, we fill the gap by applying the results obtained in Section IV to analytically characterize the high-SNR performance of the linear MIMO transceivers given in the unifying framework of [15].

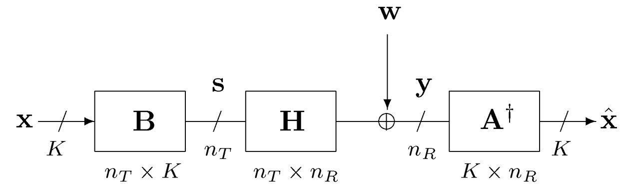

Fig. 3. Linear MIMO transceivers system model.

A. System Model: Linear MIMO Transceivers

Suppose that the general MIMO communication system of (19) is equipped with a linear transceiver (linear precoder and linear equalizer) as shown in Fig. 3. The transmitted vector is given by

$$\mathbf{s} = \mathbf{B}\mathbf{x} \tag{55}$$

where $\mathbf{B} \in \mathbb{C}^{n_{\mathrm{T}} \times K}$ is the transmit matrix (precoder), and $\mathbf{x} \in \mathbb{C}^{K \times 1}$ gathers the $K \leq \min\{n_{\mathrm{T}}, n_{\mathrm{R}}\}$ data symbols to be transmitted (zero mean, unit energy and uncorrelated, i.e., $\mathrm{E}\{\mathbf{x}\mathbf{x}^{\dagger}\} = \mathbf{I}_K$) drawn from a set of constellations. The average transmit power is constrained to satisfy

$$\mathrm{E}\{\|\mathbf{s}\|^2\} = \mathrm{Tr}\{\mathbf{B}\mathbf{B}^{\dagger}\} \leq \mathsf{snr} \tag{56}$$

where snr is the average SNR at each receive antenna. Similarly, the estimated data vector at the receiver is

$$\hat{\mathbf{x}} = \mathbf{A}^{\dagger}\mathbf{y} \tag{57}$$

where $\mathbf{A}^{\dagger} \in \mathbb{C}^{K \times n_{\mathrm{R}}}$ is the receive matrix (equalizer).

The general problem of designing the optimal linear MIMO transceiver under perfect CSI knowledge is formulated in [15] as the minimization of some cost function of mean-square errors (MSEs), since the other common system quality measures such as the SNR, or the BER can be easily related to the MSE. Assuming that K data symbols have to be communicated at each channel use, [15] shows that i) the optimum receive matrix $\mathbf{A}$, for a given transmit matrix $\mathbf{B}$, is given by the Wiener filter solution

$$\mathbf{A} = (\mathbf{H}\mathbf{B}\mathbf{B}^{\dagger}\mathbf{H}^{\dagger} + \mathbf{I}_{n_{\mathrm{R}}})^{-1}\mathbf{H}\mathbf{B} \tag{58}$$

and ii) the optimum transmit matrix $\mathbf{B}$, for a wide family of criteria (Schur-concave and Schur-convex cost functions), has the following form:

$$\mathbf{B} = \mathbf{U}\sqrt{\mathbf{P}}\mathbf{Q} \tag{59}$$

where $\mathbf{U} \in \mathbb{C}^{n_{\mathrm{T}} \times K}$ has as columns the eigenvectors of $\mathbf{H}^{\dagger}\mathbf{H}$ corresponding to the K largest nonzero eigenvalues, $\mathbf{Q} \in \mathbb{C}^{K \times K}$ is a unitary matrix, and $\mathbf{P} \in \mathbb{C}^{K \times K}$ is a diagonal matrix with diagonal entries equal to $\{p_k \in \mathbb{R}^{+}\}_{k=1,\ldots,K}$, that represent the power allocated to each established substream and satisfy (due to the power constraint in (56))

$$\sum_{k=1}^{K} p_k \leq \mathsf{snr}. \tag{60}$$

For Schur-concave objective functions, $\mathbf{Q} = \mathbf{I}_K$ and the global communication process including pre- and postprocessing is fully diagonalized. For Schur-convex objective functions, however, $\mathbf{Q}$ is a unitary matrix such that $(\mathbf{I}_K + \mathbf{B}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{B})^{-1}$ has identical diagonal elements (see [15] for details). In this case, the communication process is diagonalized up to a very specific rotation of the data symbols.

Given the transmit matrix in (59) and the receive matrix in (58), the components of the estimated signal $\hat{\mathbf{x}}$ are equal to (possibly with an additional pre- and postprocessing of the data symbols x_k in the case of Schur-convex cost functions)

$$\hat{x}_k = \frac{p_k\lambda_k}{1+p_k\lambda_k}x_k + \frac{\sqrt{p_k\lambda_k}}{1+p_k\lambda_k}n_k \quad k=1,\ldots,K \tag{61}$$

with instantaneous SNR given by

$$\rho_k = \lambda_k p_k \quad k=1,\ldots,K \tag{62}$$

where $\{\lambda_k \in \mathbb{R}^+\}_{k=1,\ldots,K}$ are the K largest nonzero eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$ in decreasing order and the complex K-dimensional vector $\mathbf{n} = [n_1, n_2, \ldots, n_K]^{\mathrm{T}}$ is a normalized equivalent noise vector with i.i.d. zero-mean, unit variance, Gaussian entries. In summary, linear MIMO transceivers transforms the MIMO channel into K SISO channels, in which each signal component (possibly after a rotation) corresponds to a different substream transmitted in parallel through a different channel eigenmode.

B. Performance of Diagonal Schemes With Fixed Power Allocation

Several design criteria (with a Schur-concave cost function) found in the literature fall within the class of diagonal schemes with fixed power allocations. Examples are the maximization of the (weighted) sum of SNRs and the maximization of the (exponentially weighted) product of the SNRs [15]. Furthermore, if the short-term power constraint is substituted by a peak-power constraint (see, e.g., [13]), the optimum spatial multiplexing system transmits always at full power through each active channel eigenmode independently of the channel state. For all these schemes, the high-SNR individual performance is given in Theorems 2 and 3, and the high-SNR global performance (for a general number of active substreams) in Theorem 5.

C. Performance of Diagonal Schemes With Nonfixed Power Allocation

Some other design criteria (with Schur-concave cost functions) found in the literature, which still have a diagonal structure, use a nonfixed power allocation. For instance, we consider first design criteria that lead to waterfilling power allocations of the type[7]

$$p_{\mathrm{wf},k} = \left(\mu - \lambda_k^{-1}\right)^+ \quad k=1,\ldots,K \tag{63}$$

where μ is chosen to satisfy the power constraint in (60) and $a^+ = \max(0,a)$. These criteria include the minimization of the determinant of the MSE matrix, the minimization of the (exponentially weighted) product of the MSEs, and the maximization of the mutual information [15]. The global average BER performance achieved with the waterfilling in (63) can be analyzed combining Theorem 5 with the results for nonfixed power allocations given in Theorem 4-ii). In addition, tighter bounds can be obtained as presented in the following result.

[7]Similarly, one can consider a more general waterfilling power allocation of the form $p_{\mathrm{wf},k} = (\phi_k\mu - \lambda_k^{-d})^+$.

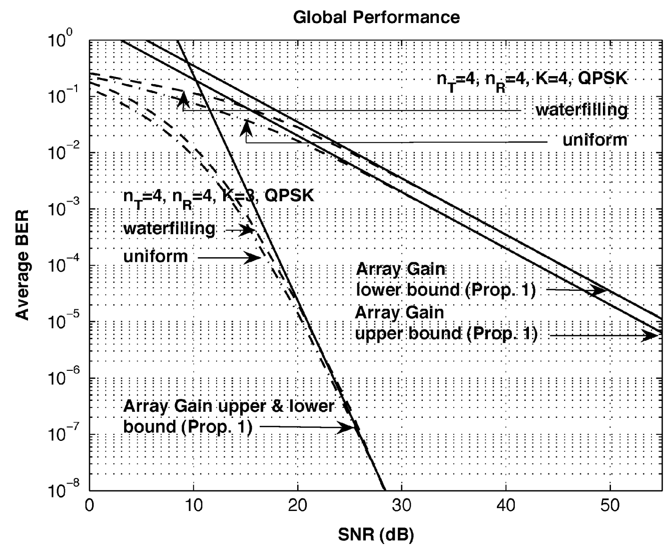

Fig. 4. Global simulated average BER of different MIMO linear transceivers (dashed line) and parameterized average BER (solid line).

Proposition 1: The global average BER attained by a diagonal MIMO linear transceiver when K data symbols have to be communicated and the power is allocated in a waterfilling fashion as in (63) is

$$\overline{\mathrm{BER}}(\mathsf{snr}) = (G_{\mathrm{wf},a}\cdot\mathsf{snr})^{-G_{\mathrm{wf},d}} + o(\mathsf{snr}^{-G_{\mathrm{wf},d}}) \tag{64}$$

where the diversity gain is given by

$$G_{\mathrm{wf},d} = (n_{\mathrm{T}} - K + 1)(n_{\mathrm{R}} - K + 1) \tag{65}$$

and the array gain is bounded as

$$\left(G_a^{-G_{\mathrm{wf},d}} + \left(\frac{\alpha_K}{2K\log_2 M_K}\right) \times \left(\frac{a_K}{d_K+1}\right)(K-1)^{G_{\mathrm{wf},d}}\right)^{-1/G_{\mathrm{wf},d}} < G_{\mathrm{wf},a} < G_a \tag{66}$$

where G_a is the global array gain when using a uniform power allocation (see (54) in Theorem 5) and the parameters a_K and d_K model the fading distribution as given in Theorem 1.

Proof: See Appendix IV.

In Fig. 4, we show the average BER attained by a linear MIMO transceiver with a uniform power allocation over the K active substreams and with the waterfilling power allocation in (63). Note that the average BER performance is always measured as the BER averaged over the K data symbols to be transmitted even when the waterfilling power allocation assigns zero power (or a very small amount of power) to the worst substreams. In particular, we provide the results for a MIMO system with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$ antennas, $K = 4$ and $K = 3$ active substreams with equal QPSK constellations. The simulation results in Fig. 4 demonstrate how the average BER curve for both power allocation policies is correctly approximated by the results presented in Theorem 5 and in Proposition 1, respectively.

In addition, there are other diagonal linear MIMO transceivers with waterfilling-type power allocations that, in contrast to the waterfilling power allocation in (63), do not asymptotically allocate equal power to all the active substreams. For instance, we consider the design that minimizes the (weighted) sum of the MSEs [6], [9], [10], [12], [15] with power allocation given by

$$p_{\mathrm{mse},k} = (\mu\lambda_k^{-1/2} - \lambda_k^{-1})^{+} \quad k = 1, \ldots, K \tag{67}$$

where μ is chosen to satisfy the power constraint in (60). The global average BER performance is analyzed in the following proposition.

Proposition 2: The global average BER attained by a diagonal MIMO linear transceiver when K data symbols have to be communicated and the power is allocated in a waterfilling fashion as in (67) is

$$\overline{\mathrm{BER}}(\mathsf{snr}) = (G_{a,\mathrm{mse}} \cdot \mathsf{snr})^{-G_{d,\mathrm{mse}}} + o(\mathsf{snr}^{-G_{d,\mathrm{mse}}}) \tag{68}$$

where the diversity gain is given by

$$G_{d,\mathrm{mse}} = (n_{\mathrm{T}} - K + 1)(n_{\mathrm{R}} - K + 1) \tag{69}$$

and the array gain is bounded as

$$G_a < G_{a,\mathrm{mse}} < KG_a \tag{70}$$

where G_a is the global array gain when using a uniform power allocation (see (54) in Theorem 5).

Proof: It follows from Theorem 5 and Theorem 4-i) with $\phi_k = 1/K$, since the exponent of $\Pr(p_{\mathrm{mse},K} \leq \mathsf{snr}/K)$ is greater than $G_{d,\mathrm{mse}}$ (see Appendix V).

D. Performance of Nondiagonal Schemes With Nonfixed Power Allocation

In this section, we complete the performance analysis of linear MIMO transceivers by focusing on the nondiagonal scheme with the nonfixed power allocation obtained from Schur-convex cost functions. We consider, for instance, the minimum BER design with equal constellations (independently derived in [16] and [15]). Other examples of design criteria with a Schur-convex cost function are the minimization of the maximum MSE, the maximization of the minimum SNR, or the minimization of the maximum BER (cf. [15]).

When the cost function is Schur-convex, the global communication system including pre- and postprocessing is diagonalized only up to a rotation of the data symbols, which ensures that all substreams have the same MSE, and the optimal power allocation is independent of the particular cost function and coincides with the power allocation in (67) (cf. [15]). Due to the rotation of the data symbols, Theorem 4 can not be directly applied and the global average BER performance is analyzed in the following proposition.

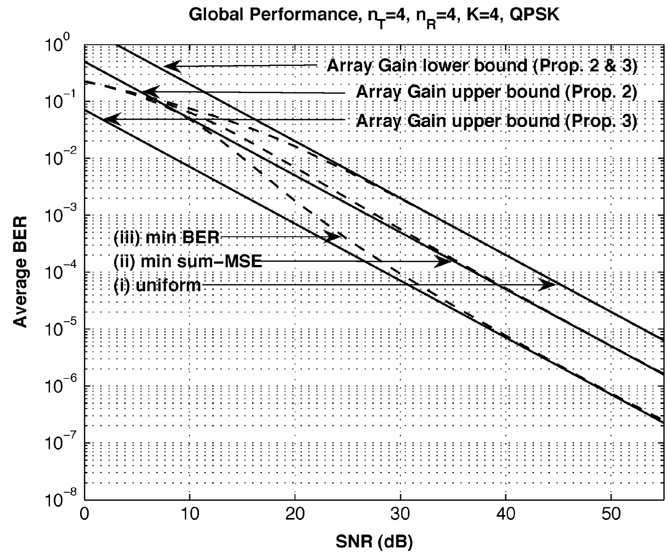

Fig. 5. Global simulated average BER of different MIMO linear transceivers (dashed line) and parameterized average BER bounds (solid line).

Proposition 3: The global average BER attained by the nondiagonal MIMO linear transceiver derived from Schur-convex cost functions when K data symbols have to be communicated is

$$\overline{\mathrm{BER}}(\mathsf{snr}) = (G_{a,\mathrm{ber}} \cdot \mathsf{snr})^{-G_{d,\mathrm{ber}}} + o(\mathsf{snr}^{-G_{d,\mathrm{ber}}}) \tag{71}$$

where the diversity gain is given by

$$G_{d,\mathrm{ber}} = (n_{\mathrm{T}} - K + 1)(n_{\mathrm{R}} - K + 1) \tag{72}$$

and the array gain can be bounded as

$$G_a < G_{a,\mathrm{ber}} < G_{a,\mathrm{ub}} \tag{73}$$

where G_a is the global array gain when using a uniform power allocation (see (54) in Theorem 5) and $G_{a,\mathrm{ub}}$ is defined as

$$G_{a,\mathrm{ub}} = \beta_K \left(\frac{\alpha_K}{\log_2 M_K} \frac{a_K I(d_K, \beta_K(K-1))}{\sqrt{2\pi}(d_K + 1)} \right)^{-1/(d_K+1)} \tag{74}$$

where $I(\cdot,\cdot)$ is given in (18) and the parameters a_K and d_K model the fading distribution as given in Theorem 1.

Proof: See Appendix VI.

In summary, Propositions 1, 2, and 3 show that linear MIMO transceivers with nonfixed power allocation policies (with or without additional pre- and postprocessing of the data symbols) do not provide any diversity advantage with respect to diagonal schemes with fixed power allocation policies but a possibly higher array gain, which results in non-negligible average performance differences. This statement is confirmed by Fig. 5, where we show the global performance of linear MIMO transceivers with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$ antennas, all substreams active, and all symbols drawn from a QPSK modulation for the following cases: i) diagonal scheme with uniform power allocation, ii) diagonal scheme with the power allocation that minimizes the sum of the MSEs, and iii) the nondiagonal scheme obtained for Schur-convex cost functions. Similarly to Fig. 4, the average BER performance is always measured as the BER averaged over the K transmitted data symbols even when the corresponding power allocation assigns zero power (or a very small amount of power) to the worst substreams. We also provide the parameterized upper and lower bounds derived from Propositions 2 and 3. It turns out that the proposed array gain upper bounds are in fact very tight and approximate perfectly the high-SNR performance of the diagonal and nondiagonal designs with the corresponding nonfixed power allocation.

VI. Conclusion

This paper has analyzed the performance of spatial multiplexing MIMO systems in terms of average BER and outage probability in a Rayleigh flat-fading channel. First, the individual performance of each one of the spatial substreams transmitted through the channel eigenmodes has been characterized in terms of the array and the diversity gains. Then, the global performance of spatial multiplexing MIMO systems with fixed number of substreams and fixed constellations has been also considered, which happens to be dominated by the performance of the worst active substream. The proposed parameterized characterization fully identifies the high SNR behavior of the BER versus SNR curve and the outage probability versus SNR curve and, thus, provides a correct approximation for practical performance values.

Based on this parameterized characterization, we have analyzed the global performance of most linear MIMO transceivers existing in the literature with adaptive linear precoder but fixed number of data symbols and fixed constellations. Independently of the design criterion, it turns out that all linear transceivers (even with a minimum BER approach) have a diversity order limited by that of the worst eigenmode used $(n_{\mathrm{T}}-K+1)(n_{\mathrm{R}}-K+1)$, which can be far from the full diversity of $n_{\mathrm{T}}n_{\mathrm{R}}$ provided by the channel. This shows that fixing *a priori* the number of independent data streams to be transmitted, a very common assumption in the linear transceiver design literature, inherently limits the average BER performance of the system. As a consequence, it seems reasonable to optimize the number of substreams (and thus the constellation size of each of them assuming constant data rate) jointly with the linear precoder for each channel realization.

Appendix I
Proof of Corollary 1

This proof is strongly based on the proof given in [20] to Lemma 1, thus some repetitive parts are omitted. Let ϵ be a small positive number so that $p_\mu(\mu)$ can be approximated by its first order expansion for $0 \leq \mu \leq \epsilon$, then the average BER can be written as

$$\begin{aligned}\overline{\mathrm{BER}}(\bar{\gamma}) =& \frac{\alpha}{\log_2 M}\int_0^{\epsilon} \mathcal{Q}(\sqrt{\beta(\bar{\gamma}\mu+\phi)})p_\mu(\mu)d\mu \\ &+ \frac{\alpha}{\log_2 M}\int_{\epsilon}^{\infty} \mathcal{Q}(\sqrt{\beta(\bar{\gamma}\mu+\phi)})p_\mu(\mu)d\mu \\ =& \frac{\alpha}{\log_2 M}\int_{\epsilon}^{\infty} \mathcal{Q}(\sqrt{\beta(\bar{\gamma}\mu+\phi)})p_\mu(\mu)d\mu \qquad (75)\\ &- \frac{\alpha}{\log_2 M}\frac{1}{\sqrt{2\pi}}\int_{\epsilon}^{\infty}\int_{\sqrt{\beta(\bar{\gamma}\mu+\phi)}}^{\infty} e^{-\frac{x^2}{2}} \\ &\times (a\mu^d + o(\mu^d))dxd\mu \qquad (76)\\ &+ \frac{\alpha}{\log_2 M}\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\int_{\sqrt{\beta(\bar{\gamma}\mu+\phi)}}^{\infty} e^{-\frac{x^2}{2}} \\ &\times (a\mu^d + o(\mu^d))dxd\mu. \qquad (77)\end{aligned}$$

The terms (75) and (76) are shown in [20] to be a $o(\bar{\gamma}^{-(d+1)})$ and, computing the integral in (77) by interchanging the integration order, it follows

$$\begin{aligned}\overline{\mathrm{BER}}(\bar{\gamma}) =& \frac{\alpha}{\log_2 M}\frac{a}{\sqrt{2\pi}} \\ &\times \int_{\sqrt{\beta\phi}}^{\infty}\int_0^{\frac{x^2-\beta\phi}{\beta\bar{\gamma}}} e^{-\frac{x^2}{2}}\mu^d d\mu dx + o(\bar{\gamma}^{-(d+1)}) \\ =& \frac{\alpha}{\log_2 M}\frac{a}{\sqrt{2\pi}(d+1)(\beta\bar{\gamma})^{d+1}} \\ &\int_{\sqrt{\beta\phi}}^{\infty} e^{-\frac{x^2}{2}}(x^2-\beta\phi)^{(d+1)}dx + o(\bar{\gamma}^{-(d+1)}) \\ =& \left(\frac{\alpha}{\log_2 M}\frac{aI(d,\beta\phi)}{\sqrt{2\pi}(d+1)\beta^{d+1}}\right)\bar{\gamma}^{-(d+1)} + o(\bar{\gamma}^{-(d+1)})\end{aligned} \qquad (78)$$

where $I(d,\beta\phi)$ is defined as

$$I(d,\beta\phi) = \int_{\sqrt{\beta\phi}}^{\infty} e^{-\frac{x^2}{2}}(x^2-\beta\phi)^{(d+1)}dx \qquad (79)$$

and this completes the proof. □

Appendix II
Proof of Theorem 1

Let the entries of the $n_{\mathrm{R}} \times n_{\mathrm{T}}$ matrix $\mathbf{H}$ be i.i.d. complex Gaussian with zero mean and unit variance and let $n = \min\{n_{\mathrm{R}}, n_{\mathrm{T}}\}$ and $m = \max\{n_{\mathrm{R}}, n_{\mathrm{T}}\}$. The joint pdf of the ordered strictly positive eigenvalues of the complex central Wishart matrix $\mathbf{H}^\dagger\mathbf{H}, \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$, equals [42, eq. (95)]

$$\begin{aligned}f(\lambda_1,\lambda_2,\ldots,\lambda_n) &= f(\boldsymbol{\lambda}) \\ &= K_{m,n}^{-1}\left[\prod_{i=1}^{n} e^{-\lambda_i}\lambda_i^{m-n}\right]\left[\prod_{i<j}^{n}(\lambda_i-\lambda_j)^2\right] \qquad (80)\end{aligned}$$

where the normalizing constant $K_{m,n}$ is given by $K_{m,n} = \prod_{i=1}^{n}(n-i)!(m-i)!$.

The marginal pdf of the kth eigenvalue, denoted by $f_{\lambda_k}(\lambda_k)$, is obtained from the joint distribution of the ordered eigenvalues $f(\boldsymbol{\lambda})$ as

$$\begin{aligned}f_{\lambda_k}(\lambda_k) =& \int_{\lambda_k}^{\infty}\int_{\lambda_k}^{\lambda_1}\cdots\int_{\lambda_k}^{\lambda_{k-2}}\int_0^{\lambda_k}\cdots \\ &\int_0^{\lambda_{n-1}} f(\boldsymbol{\lambda})d\lambda_n\cdots d\lambda_{k+1}d\lambda_{k-1}\cdots d\lambda_2 d\lambda_1. \qquad (81)\end{aligned}$$

Thus, the derivation of the marginal pdf λ_k involves the integration with respect to the eigenvalues larger than λ_k and smaller than λ_k. The difficulty lies in the fact that both groups of integrals must be calculated over an ordered domain. Based on the continuous counterpart of the Cauchy–Binet theorem [43, Sec. 2.1.10], the authors of [37] showed how to transform a multiple integral over an ordered domain with a particular structure into a determinant of a matrix, whose elements are given by simple integrals. Due to the importance of this result in the proof of Theorem 1, we present it in the following lemma.

Lemma 3 [37, Cor. 2]: Given two $n \times n$ arbitrary matrices $\mathbf{F}(\mathbf{x})$ and $\mathbf{G}(\mathbf{x})$ with elements be $[\mathbf{F}(\mathbf{x})]_{ij} = f_i(x_j)$ and $[\mathbf{G}(\mathbf{x})]_{ij} = g_i(x_j)$, and an arbitrary function $h(\cdot)$, the following identity holds:

$$\int_{\mathcal{D}_{\rm ord}} |\mathbf{F}(\mathbf{x})||\mathbf{G}(\mathbf{x})| \prod_{i=1}^{n} h(x_i) d\mathbf{x} = |\mathbf{A}| \tag{82}$$

where $\mathcal{D}_{\rm ord} = \{b \geq x_1 \geq x_2 \geq \cdots \geq x_n \geq a\}$ and $\mathbf{A}$ is an $n \times n$ matrix with elements

$$[\mathbf{A}]_{ij} = \int_a^b f_i(x) g_j(x) h(x) dx. \tag{83}$$

The proof of Theorem 1 basically consists of rewriting the joint pdf of the eigenvalues in order to be able to apply Lemma 3 twice to obtain a more tractable group of integrals. Then, we derive the first-order expansion of these new integrals and finally of the marginal pdf.

The joint distribution of the ordered eigenvalues $f(\boldsymbol{\lambda})$ of (80) can be rewritten as[8]

$$\begin{aligned} f(\boldsymbol{\lambda}) &= K_{m,n}^{-1} \left[\prod_{i=1}^{n} e^{-\lambda_i} \lambda_i^{m-n}\right] \left[\prod_{i=1}^{n-1} \prod_{j=i+1}^{n} (\lambda_i - \lambda_j)^2\right] \\ &= K_{m,n}^{-1} e^{-\lambda_k} \lambda_k^{m-n} \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2\right] \\ &\quad \times \left[\prod_{i=1}^{k-2} \prod_{j=i+1}^{k-1} (\lambda_i - \lambda_j)^2\right] \\ &\quad \times \left[\prod_{i=k+1}^{n-1} \prod_{j=i+1}^{n} (\lambda_i - \lambda_j)^2\right] \\ &\quad \times \left[\prod_{i=k+1}^{n} e^{-\lambda_i} \lambda_i^{m-n} \prod_{j=1}^{k} (\lambda_i - \lambda_j)^2\right]. \end{aligned} \tag{84}$$

We define the Vandermonde Matrix of order $v, \mathbf{V}(\boldsymbol{\phi}) = \mathbf{V}(\phi_1, \phi_2, \ldots, \phi_v)$, as the $v \times v$ matrix given by

$$\mathbf{V}(\boldsymbol{\phi}) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \phi_1 & \phi_2 & \cdots & \phi_v \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1^{v-1} & \phi_2^{v-1} & \cdots & \phi_v^{v-1} \end{bmatrix} \tag{85}$$

and we denote by $\boldsymbol{\lambda}_k^+$ the set of eigenvalues larger than λ_k, i.e., $\{\lambda_i\}_{i=1,\ldots,k-1}$, and by $\boldsymbol{\lambda}_k^-$ the set of eigenvalues smaller than λ_k, i.e., $\{\lambda_i\}_{i=k+1,\ldots,n}$. The determinant of a Vandermonde matrix of order v can be calculated as [44, eq. (6.1.33)]

$$|\mathbf{V}(\boldsymbol{\phi})| = \prod_{i=1}^{v-1} \prod_{j=i+1}^{v} (\phi_j - \phi_i). \tag{86}$$

Then, the pdf in (84) can be expressed alternatively in terms of the Vandermonde matrix of order $k-1, \mathbf{V}(\boldsymbol{\lambda}_k^+)$, and the Vandermonde matrix of order $n-k, \mathbf{V}(\boldsymbol{\lambda}_k^-)$, as

$$\begin{aligned} f(\boldsymbol{\lambda}) &= K_{m,n}^{-1} e^{-\lambda_k} \lambda_k^{m-n} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \\ &\quad \times \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2\right] \\ &\quad \times |\mathbf{V}(\boldsymbol{\lambda}_k^-)|^2 \left[\prod_{i=k+1}^{n} e^{-\lambda_i} \lambda_i^{m-n} \prod_{j=1}^{k} (\lambda_i - \lambda_j)^2\right] \\ &= K(\boldsymbol{\lambda}_k^+, \lambda_k) |\mathbf{V}(\boldsymbol{\lambda}_k^-)|^2 \prod_{i=k+1}^{n} \xi_k^-(\lambda_i) \end{aligned} \tag{87}$$

where

$$\begin{aligned} K(\boldsymbol{\lambda}_k^+, \lambda_k) &= K_{m,n}^{-1} e^{-\lambda_k} \lambda_k^{m-n} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \\ &\quad \times \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2\right] \end{aligned} \tag{88}$$

$$\begin{aligned} \xi_k^-(\lambda_i) &\triangleq \xi_k^-(\lambda_i, \boldsymbol{\lambda}_k^+, \lambda_k) \\ &= e^{-\lambda_i} \lambda_i^{m-n} \prod_{j=1}^{k} (\lambda_i - \lambda_j)^2. \end{aligned} \tag{89}$$

First, we integrate the joint pdf of the ordered eigenvalues over the domain of $\boldsymbol{\lambda}_k^-$, defined as $\mathcal{D}_k^- \triangleq \mathcal{D}_k^-(\boldsymbol{\lambda}_k^-, \lambda_k) = \{0 < \lambda_n \leq \cdots \leq \lambda_{k+1} \leq \lambda_k\}$

$$f(\boldsymbol{\lambda}_k^+, \lambda_k) = K(\boldsymbol{\lambda}_k^+, \lambda_k) \int_{\mathcal{D}_k^-} |V(\boldsymbol{\lambda}_k^-)|^2 \prod_{i=k+1}^{n} \xi_k^-(\lambda_i) d\boldsymbol{\lambda}_k^-. \tag{90}$$

Applying Lemma 3, the previous integral can be computed as

$$f(\boldsymbol{\lambda}_k^+, \lambda_k) = K(\boldsymbol{\lambda}_k^+, \lambda_k) |\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)| \tag{91}$$

where

$$[\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)]_{ij} = \int_0^{\lambda_k} e^{-x} x^{m-n+j+i-2} \prod_{s=1}^{k} (x - \lambda_s)^2 dx \qquad i,j = 1, \ldots, n-k. \tag{92}$$

Then, the marginal pdf of λ_k is obtained by integrating $f(\boldsymbol{\lambda}_k^+, \lambda_k)$ over the ordered domain of $\boldsymbol{\lambda}_k^+$, defined as $\mathcal{D}_k^+ \triangleq \mathcal{D}_k^+(\boldsymbol{\lambda}_k^+, \lambda_k) = \{\lambda_k \leq \lambda_{k-1} \leq \cdots \leq \lambda_1 < \infty\}$

$$\begin{aligned} f_{\lambda_k}(\lambda_k) &= \int_{\mathcal{D}_k^+} K(\boldsymbol{\lambda}_k^+, \lambda_k) |\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)| d\boldsymbol{\lambda}_k^+ \\ &= K(\lambda_k) \int_{\mathcal{D}_k^+} |\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)| |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \\ &\quad \times \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2\right] d\boldsymbol{\lambda}_k^+ \end{aligned} \tag{93}$$

[8]For the sake of notation, we follow the common assumption that an empty product is the unity, and that the determinant of an empty matrix equals one.

where

$$K(\lambda_k) = K_{m,n}^{-1} e^{-\lambda_k} \lambda_k^{m-n}. \tag{94}$$

The determinant of a generic $n \times n$ matrix $\mathbf{M}$ is given by

$$|\mathbf{M}| = \sum_{\boldsymbol{\sigma}} \text{sgn}(\boldsymbol{\sigma}) \prod_{i=1}^{n} [\mathbf{M}]_{\sigma_i, i} \tag{95}$$

where the summation over $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \dots, \sigma_n]$ is for all permutations of the integers $1, \dots, n$ and $\text{sgn}(\boldsymbol{\sigma})$ denotes the sign of the permutation. Expanding the determinant of $\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)$, we obtain that

$$\begin{aligned} f_{\lambda_k}(\lambda_k) = K(\lambda_k) \sum_{\boldsymbol{\sigma}} \text{sgn}(\boldsymbol{\sigma}) \int_{\mathcal{D}_k^+} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \\ \times \left[\prod_{i=1}^{n-k} \int_0^{\lambda_k} e^{-x} x^{m-n+i+\sigma_i-2} \prod_{j=1}^{k} (x-\lambda_j)^2 dx \right] \\ \times \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2 \right] d\boldsymbol{\lambda}_k^+. \end{aligned} \tag{96}$$

Now, we rewrite the product of $n-k$ integrals over x as a multiple integral over $\mathbf{x} = \{x_s\}_{s=1,\dots,n-k}$ in the domain $\mathcal{D}_{\mathbf{x}} = \{0 < x_s \leq \lambda_k\}_{s=1,\dots,n-k}$

$$\begin{aligned} f_{\lambda_k}(\lambda_k) &= K(\lambda_k) \sum_{\boldsymbol{\sigma}} \text{sgn}(\boldsymbol{\sigma}) \\ &\quad \times \int_{\mathcal{D}_{\mathbf{x}}} \left[\prod_{s=1}^{n-k} e^{-x_s} x_s^{m-n+s+\sigma_s-2} (x_s - \lambda_k)^2 \right] \\ &\quad \times \int_{\mathcal{D}_k^+} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2 \right. \\ &\quad \left. \times \prod_{s=1}^{n-k} (x_s - \lambda_i)^2 \right] d\boldsymbol{\lambda}_k^+ d\mathbf{x} \\ &= K(\lambda_k) \sum_{\boldsymbol{\sigma}} \text{sgn}(\boldsymbol{\sigma}) \\ &\quad \times \int_{\mathcal{D}_{\mathbf{x}}} \left[\prod_{s=1}^{n-k} e^{-x_s} x_s^{m-n+s+\sigma_s-2} (x_s - \lambda_k)^2 \right] \\ &\quad \times \left[\int_{\mathcal{D}_k^+} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \prod_{i=1}^{k-1} \xi_k^+(\lambda_i) d\boldsymbol{\lambda}_k^+ \right] d\mathbf{x} \end{aligned} \tag{97}$$

where

$$\begin{aligned} \xi_k^+(\lambda_i) &\triangleq \xi_k^+(\lambda_i, \lambda_k, \mathbf{x}) \\ &= e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2 \prod_{s=1}^{n-k} (\lambda_i - x_s)^2. \end{aligned} \tag{98}$$

The integral over $\boldsymbol{\lambda}_k^+$ can be solved applying again Lemma 3:

$$\int_{\mathcal{D}_k^+} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \prod_{i=1}^{k-1} \xi_k^+(\lambda_i) d\boldsymbol{\lambda}_k^+ = |\mathbf{W}(\lambda_k, \mathbf{x})| \tag{99}$$

where

$$\begin{aligned} &[\mathbf{W}(\lambda_k, \mathbf{x})]_{ij} \\ &\quad = \int_{\lambda_k}^{\infty} e^{-y} y^{m-n+i+j-2} (y - \lambda_k)^2 \prod_{s=1}^{n-k} (y - x_s)^2 dy \\ &\qquad i, j = 1, \dots, k-1. \end{aligned} \tag{100}$$

Then, expanding the polynomial $\prod_{s=1}^{n-k} (y - x_s)^2$ as

$$\prod_{s=1}^{n-k} (y - x_s)^2 = \sum_{v=0}^{2(n-k)} C_v(\mathbf{x}) y^v \tag{101}$$

we have that

$$[\mathbf{W}(\lambda_k, \mathbf{x})]_{ij} = \sum_{v=0}^{2(n-k)} C_v(\mathbf{x}) \int_{\lambda_k}^{\infty} e^{-y} y^{b(i+j+v-2)} (y - \lambda_k)^2 dy \tag{102}$$

where $b(i) = m - n + i$. The integral in (102) has the same structure as the upper incomplete gamma function, defined as [45, eq. (6.5.3)]

$$\Gamma(a, \lambda) = \int_{\lambda}^{\infty} e^{-t} t^{a-1} dt. \tag{103}$$

As we are only interested in the first-order expansion of $f_{\lambda_k}(\lambda_k)$ in λ_k, noting that, for $a \in \mathbb{N}$,

$$\Gamma(a, \lambda) = (a-1)! e^{-\lambda} \sum_{i=0}^{a-1} \frac{\lambda^i}{i!} = (a-1)! + o(1) \tag{104}$$

where $o(1)$ tends to zero as $\lambda \to 0$, we can expand $(y - \lambda_k)^2$ in (102) and neglect the terms with λ_k and λ_k^2, since they are also $o(1)$ terms as $\lambda_k \to 0$. Hence, it follows that

$$[\mathbf{W}(\lambda_k, \mathbf{x})]_{ij} = \sum_{v=0}^{2(n-k)} C_v(\mathbf{x}) b(i+j+v)! + o(1). \tag{105}$$

Now, expanding the determinant of $\mathbf{W}(\lambda_k, \mathbf{x})$, the marginal pdf of λ_k is given by

$$\begin{aligned} f_{\lambda_k}(\lambda_k) =& K(\lambda_k) \sum_{\boldsymbol{\sigma}} \sum_{\boldsymbol{\mu}} \text{sgn}(\boldsymbol{\sigma}) \text{sgn}(\boldsymbol{\mu}) \\ &\times \int_{\mathcal{D}_{\mathbf{x}}} \left[\prod_{s=1}^{n-k} e^{-x_s} x_s^{b(s+\sigma_s-2)} (x_s - \lambda_k)^2 \right] \\ &\times \left[\prod_{u=1}^{k-1} \sum_{v=0}^{2(n-k)} (C_v(\mathbf{x}) b(u+\mu_u+v)! + o(1)) \right] d\mathbf{x}. \end{aligned} \tag{106}$$

All the integrals over $\mathbf{x} = \{x_s\}_{s=1,\ldots,n-k}$ have the same structure as the lower incomplete gamma function defined as [45, eq. (6.5.2)]

$$\gamma(a,\lambda) = \int_0^\lambda e^{-t} t^{a-1} dt. \tag{107}$$

Noting that for $a \in \mathbb{N}$,

$$\begin{aligned}\gamma(a,\lambda) &= (a-1)!\left(1 - e^{-\lambda}\sum_{i=0}^{a-1}\frac{\lambda^i}{i!}\right)\\ &= (a-1)!\left(1 - e^{-\lambda}\left(e^{\lambda} - \sum_{i=a}^{\infty}\frac{\lambda^i}{i!}\right)\right)\\ &= \frac{1}{a}\lambda^a + o(\lambda^a)\end{aligned} \tag{108}$$

the term with the lowest λ_k exponent of $f_{\lambda_k}(\lambda_k)$ in (106) is found for all x_s having the lowest possible exponent. This occurs for $C_v(\mathbf{x}) = 1$, i.e., for $v = 2(n-k)$ in the summation of (106). Thus, it follows that

$$\begin{aligned}f_{\lambda_k}(\lambda_k) = K(\lambda_k)&\sum_{\boldsymbol{\sigma}}\sum_{\boldsymbol{\mu}}\mathrm{sgn}(\boldsymbol{\sigma})\mathrm{sgn}(\boldsymbol{\mu})\\ &\times\left[\prod_{u=1}^{k-1} b(u+\mu_u+2(n-k))! + o(1)\right]\\ &\times\left[\int_{\mathcal{D}_{\mathbf{x}}}\prod_{s=1}^{n-k} e^{-x_s} x_s^{b(s+\sigma_s-2)}(x_s-\lambda_k)^2 d\mathbf{x}\right].\end{aligned} \tag{109}$$

Then, rewriting the multiple integral over $\mathbf{x}$ as a product of $n-k$ single integrals, we have that

$$\begin{aligned}&f_{\lambda_k}(\lambda_k)\\ &= K(\lambda_k)\left[\sum_{\boldsymbol{\mu}}\mathrm{sgn}(\boldsymbol{\mu})\prod_{u=1}^{k-1} b(u+\mu_u 2(n-k))! + o(1)\right]\\ &\quad\times\left[\sum_{\boldsymbol{\sigma}}\mathrm{sgn}(\boldsymbol{\sigma})\prod_{s=1}^{n-k}\int_0^{\lambda_k} e^{-x} x^{b(s+\sigma_s-2)}(x-\lambda_k)^2 dx\right]\\ &= K(\lambda_k)\left(|\mathbf{A}(k)| + o(1)\right)\\ &\quad\times\left[\sum_{\boldsymbol{\sigma}}\mathrm{sgn}(\boldsymbol{\sigma})\prod_{s=1}^{n-k}\int_0^{\lambda_k} e^{-x} x^{b(s+\sigma_s-2)}(x-\lambda_k)^2 dx\right]\end{aligned} \tag{110}$$

where the $(k-1)\times(k-1)$ matrix $\mathbf{A}(k)$ is given by

$$[\mathbf{A}(k)]_{ij} = b(i+j+2(n-k))!. \tag{111}$$

The integral in (110) can be evaluated using (107) and (108) as

$$\begin{aligned}&\int_0^{\lambda_k} e^{-x} x^{b(i+j-2)}(x-\lambda_k)^2 dx\\ &\quad= \lambda_k^2\gamma(b(i+j-1),\lambda_k) - 2\lambda_k\gamma(b(i+j),\lambda_k)\\ &\qquad+ \gamma(b(i+j+1),\lambda_k)\\ &\quad= B(i+j)\lambda_k^{b(i+j+1)} + o\left(\lambda_k^{b(i+j+1)}\right)\end{aligned} \tag{112}$$

with $B(i) = 2/((b(i)^2-1)b(i))$. Substituting back (112) into (110), we obtain that

$$\begin{aligned}&f_{\lambda_k}(\lambda_k)\\ &\quad= K(\lambda_k)|\mathbf{A}(k)|\left[\sum_{\boldsymbol{\sigma}}\mathrm{sgn}(\boldsymbol{\sigma})\prod_{s=1}^{n-k}\left(B(s+\sigma_s)\right.\right.\\ &\qquad\left.\left.\times\,\lambda_k^{b(s+\sigma_s+1)} + o\left(\lambda_k^{b(s+\sigma_s+1)}\right)\right)\right]\\ &\quad= K(\lambda_k)|\mathbf{A}(k)|\left[\sum_{\boldsymbol{\sigma}}\mathrm{sgn}(\boldsymbol{\sigma})\prod_{s=1}^{n-k}B(s+\sigma_s)\right]\\ &\qquad\left[\lambda_k^S + o\left(\lambda_k^S\right)\right]\\ &\quad= K(\lambda_k)|\mathbf{A}(k)||\mathbf{B}(k)|\left[\lambda_k^S + o\left(\lambda_k^S\right)\right]\end{aligned} \tag{113}$$

where the $(n-k)\times(n-k)$ matrix $\mathbf{B}(k)$ is defined as $[\mathbf{B}(k)]_{ij} = B(i+j)$ and

$$\begin{aligned}S &= \sum_{i=1}^{n-k}(b(i+\sigma_i)+1) = \sum_{i}^{n-k}(m-n+i+\sigma_i+1)\\ &= (m-k+1)(n-k) + (n-k).\end{aligned} \tag{114}$$

Finally, substituting (94) into (113), the marginal pdf of λ_k is given by

$$\begin{aligned}&f_{\lambda_k}(\lambda_k)\\ &\quad= K_{m,n}^{-1}|\mathbf{A}(k)||\mathbf{B}(k)|\lambda_k^{(m-k+1)(n-k+1)-1}\\ &\qquad+ o\left(\lambda_k^{(m-k+1)(n-k+1)-1}\right)\end{aligned} \tag{115}$$

and the proof is completed. □

APPENDIX III
PROOF OF THEOREM 4

The average BER of the kth substream with $p_k = g_k(\lambda_1,\ldots,\lambda_K,\mathsf{snr})$ can be expressed as

$$\begin{aligned}\overline{\mathrm{BER}}_k(\mathsf{snr}) &= \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k > \phi_k\mathsf{snr})(1-\Pr(p_k \le \phi_k\mathsf{snr}))\\ &\quad+\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k \le \phi_k\mathsf{snr})\Pr(p_k \le \phi_k\mathsf{snr})\end{aligned} \tag{116}$$

where $\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k > \phi_k\mathsf{snr})$ denotes the BER averaged over the channel states that imply $p_k > \phi_k\mathsf{snr}$ and $\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k \le \phi_k\mathsf{snr})$ is analogously defined. Using the expression for $\Pr(p_k \le \phi_k\mathsf{snr})$ given in (44), the average BER can be rewritten as

$$\begin{aligned}\overline{\mathrm{BER}}_k(\mathsf{snr}) =&\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k > \phi_k\mathsf{snr})\\ &\times\left(1 - a(\phi_k)\mathsf{snr}^{-d(\phi_k)} + o\left(\mathsf{snr}^{-d(\phi_k)}\right)\right)\end{aligned} \tag{117}$$

$$\begin{aligned}&+\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k \le \phi_k\mathsf{snr})\\ &\times\left(a(\phi_k)\mathsf{snr}^{-d(\phi_k)} + o\left(\mathsf{snr}^{-d(\phi_k)}\right)\right).\end{aligned} \tag{118}$$

In the following, we distinguish between two cases: i) when $d(\phi_k) > d_k + 1$ for a given $\phi_k \in (0,1)$ and ii) when $d(\phi_k) = d_k + 1$ for a given $\phi_k \in (0,1)$ and $d(0) = d_k + 1$.

i) If there exists $\phi_k \in (0,1)$ such that $d(\phi_k) > d_k + 1$, the term in (118) is $o(\mathsf{snr}^{-(d_k+1)})$, since it holds that

$$\begin{aligned}\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k \leq \phi_k \mathsf{snr}) a(\phi_k)\mathsf{snr}^{-d(\phi_k)} \\ < \frac{\alpha_k}{2\log_2 M_k} a(\phi_k)\mathsf{snr}^{-d(\phi_k)} \\ = o\left(\mathsf{snr}^{-(d_k+1)}\right). \end{aligned} \tag{119}$$

Using the short-term power constraint in (22), the power allocated to the kth substream is upper-bounded as $p_k \leq \mathsf{snr}$ and, hence, $\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k > \phi_k\mathsf{snr})$ satisfies

$$\begin{aligned}\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = \mathsf{snr}) \\ < \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k > \phi_k\mathsf{snr}) \\ \leq \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = \phi_k\mathsf{snr}). \end{aligned} \tag{120}$$

The upper and lower bound in (120) can be analyzed applying Theorem 2. Both result in the same diversity gain, $\tilde{G}_d(k) = d_k + 1$ and

$$\begin{aligned}\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k > \phi_k\mathsf{snr})\left(1 - a(\phi_k)\mathsf{snr}^{-d(\phi_k)}\right) \\ = \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k > \phi_k\mathsf{snr}) + o\left(\mathsf{snr}^{-(d_k+1)}\right). \end{aligned} \tag{121}$$

Finally, the average BER is

$$\overline{\mathrm{BER}}_k(\mathsf{snr}) = \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k > \phi_k\mathsf{snr}) + o\left(\mathsf{snr}^{-(d_k+1)}\right) \tag{122}$$

and the array gain bounds given in the theorem are obtained by deriving the array gain associated with the bounds in (120) recalling Theorem 2.

ii) If $d(0) = d_k + 1$ and there exists $\phi_k \in (0,1)$ such that $d(\phi_k) = d_k + 1$, the term in (118) can be bounded as

$$\begin{aligned}\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k \leq \phi_k\mathsf{snr}) a(\phi_k)\mathsf{snr}^{-(d_k+1)} \\ < \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = 0) a(\phi_k)\mathsf{snr}^{-(d_k+1)} \\ = \frac{\alpha_k}{2\log_2 M_k} a(\phi_k)\mathsf{snr}^{-(d_k+1)}. \end{aligned} \tag{123}$$

Then, proceeding as in case i), it follows that

$$\begin{aligned}\overline{\mathrm{BER}}_k(\mathsf{snr}) <& \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = \phi_k\mathsf{snr}) \\ &+ \frac{\alpha_k}{2\log_2 M_k} a(\phi_k)\mathsf{snr}^{-(d_k+1)} + o\left(\mathsf{snr}^{-(d_k+1)}\right) \end{aligned} \tag{124}$$

and the array gain lower bound given in the theorem is obtained by combining the array gain corresponding to $\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = \phi_k\mathsf{snr})$ with the term in (123). Similarly, we can lower-bound the average BER as

$$\begin{aligned}&\overline{\mathrm{BER}}_k(\mathsf{snr}) \\ &> \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = \mathsf{snr})\left(1 - a(\phi_k)\mathsf{snr}^{-(d_k+1)}\right) \\ &\quad + \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = \phi_k\mathsf{snr})\left(a(\phi_k)\mathsf{snr}^{-(d_k+1)}\right) \\ &\quad + o\left(\mathsf{snr}^{-(d_k+1)}\right) \\ &= \overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = \mathsf{snr}) + o\left(\mathsf{snr}^{-(d_k+1)}\right) \end{aligned} \tag{125}$$

and the array gain corresponding to $\overline{\mathrm{BER}}_k(\mathsf{snr}\,|\,p_k = \mathsf{snr})$ is the array gain upper bound given in the theorem. □

Appendix IV
Proof of Proposition 1

Observe that the waterfilling in (63) tends to a uniform power allocation over the K active substreams as $\mathsf{snr} \to \infty$, i.e.,

$$\lim_{\mathsf{snr}\to\infty} \frac{p_{\mathrm{wf},k}}{\mathsf{snr}/K} = 1 \tag{126}$$

whenever $p_{\mathrm{wf},K} > 0$. Hence, the average BER at high SNR can be expressed as

$$\begin{aligned}\overline{\mathrm{BER}}(\mathsf{snr}) =& \frac{1}{K}\overline{\mathrm{BER}}_K(\mathsf{snr})(1 - \Pr(p_{\mathrm{wf},K} = 0)) \\ &+ \frac{\alpha_K}{2K\log_2 M_K}\Pr(p_{\mathrm{wf},K} = 0) \end{aligned} \tag{127}$$

where $\overline{\mathrm{BER}}_K(\mathsf{snr})$ is the average BER when using a uniform power allocation and $\Pr(p_{\mathrm{wf},K} = 0)$ denotes the probability of not transmitting power through the Kth channel eigenmode. This probability is upper-bounded as

$$\begin{aligned}&\Pr(p_{\mathrm{wf},K} = 0) \\ &= \Pr\left(\left((K-1)\lambda_K^{-1} - \sum_{k=1}^{K-1}\lambda_k^{-1}\right) \geq \mathsf{snr}\right) \\ &\leq \Pr\left(\lambda_K^{-1} \geq \frac{\mathsf{snr}}{K-1}\right) = \Pr\left(\lambda_K \leq \frac{K-1}{\mathsf{snr}}\right) \\ &= \frac{a_K}{d_K+1}\left(\frac{\mathsf{snr}}{K-1}\right)^{-(d_K+1)} + o\left(\mathsf{snr}^{-(d_K+1)}\right) \end{aligned} \tag{128}$$

where the last equality comes from Theorem 1. Then, substituting $\overline{\mathrm{BER}}_K(\mathsf{snr})$ by its parameterized characterization applying Theorem 2 and $\Pr(p_{\mathrm{wf},K} = 0)$ by its upper bound derived in (128) back in (127), it follows that

$$\begin{aligned}\overline{\mathrm{BER}}(\mathsf{snr}) <& \left(G_a^{-(d_K+1)} + \left(\frac{\alpha_K}{2K\log_2 M_K}\right)\right. \\ &\left.\times\left(\frac{a_K}{d_K+1}\right)(K-1)^{(d_K+1)}\right)\mathsf{snr}^{-(d_K+1)} \\ &+ o\left(\mathsf{snr}^{-(d_K+1)}\right) \end{aligned} \tag{129}$$

where G_a denotes the global array gain achieved with a uniform power allocation. Finally, the array gain lower bound can be directly obtained from (129) and the upper bound simply comes from setting $\Pr(p_{\mathrm{wf},K}=0)=0$ in (127). Note that this proof is in essence the same as the proof of Theorem 4-ii), but here we use the asymptotic equivalence in (126), instead of bounding the power allocation. □

APPENDIX V
PROOF OF PROPOSITION 2. EXPONENT OF $\Pr(p_{\mathrm{mse},K} \le \mathsf{snr}/K)$

We want to prove that

$$\begin{aligned} d &= \lim_{\mathsf{snr}\to\infty} -\frac{\log \Pr(p_{\mathrm{mse},K} \le \mathsf{snr}/K)}{\log \mathsf{snr}} \\ &> G_{d,\mathrm{mse}} = d_K + 1 \end{aligned} \tag{130}$$

or, equivalently

$$d = \lim_{\mathsf{snr}\to\infty} -\frac{\log \Pr(p_{\mathrm{mse},K} \le \phi\mathsf{snr})}{\log \mathsf{snr}} > d_K + 1 \tag{131}$$

with $\phi < 1/K$ arbitrarily close to $1/K$, since $\Pr(p_{\mathrm{mse},K} = \mathsf{snr}/K) = \Pr(\lambda_1 = \lambda_2 = \cdots = \lambda_K) = 0$. Noting that $\Pr(p_{\mathrm{mse},K} \le \phi\mathsf{snr})$ can be bounded as

$$\begin{aligned} &\Pr(p_{\mathrm{mse},K} \le \phi\mathsf{snr}) \\ &= \Pr\left(\frac{\lambda_K^{-1/2}}{\sum_{k=1}^K \lambda_k^{-1/2}}\left(\mathsf{snr} + \sum_{k=1}^K \lambda_k^{-1}\right) - \lambda_K^{-1} \le \phi\mathsf{snr}\right) \\ &= \Pr\left(\lambda_K^{-1/2}\sum_{k=1}^{K-1}\lambda_k^{-1/2} - \sum_{k=1}^{K-1}\lambda_k^{-1}\right. \\ &\qquad \left.\ge \mathsf{snr}\left(1 - \phi\lambda_K^{1/2}\sum_{k=1}^K \lambda_k^{-1/2}\right)\right) \\ &\le \Pr\left(\lambda_K^{-1/2}\sum_{k=1}^{K-1}\lambda_k^{-1/2} \ge \mathsf{snr}(1-K\phi)\right) \\ &\le \Pr\left(\lambda_K\lambda_{K-1} \le \left(\frac{K-1}{1-K\phi}\right)^2 \mathsf{snr}^{-2}\right) \end{aligned} \tag{132}$$

it follows that

$$d \ge \lim_{\mathsf{snr}\to\infty} \frac{\log \Pr\left(\lambda_K\lambda_{K-1} \le \mathsf{snr}^{-2}\right)}{\log \mathsf{snr}^{-1}} \tag{133}$$

where we have omitted the term $((K-1)/(1-K\phi))^2$, because it does not affect the lower bound of the exponent whenever $\phi < 1/K$. Then, by defining $\mu = \lambda_{K-1}\lambda_K$, we can rewrite (133) as

$$\begin{aligned} d &\ge \lim_{x\to 0} \frac{\log \Pr(\mu \le x^2)}{\log x} \\ &= \lim_{x\to 0} \frac{\log\left(\int_0^{x^2} f_\mu(\mu) d\mu\right)}{\log x} \end{aligned} \tag{134}$$

where $f_\mu(\mu)$ is the pdf of a product of two random variables and is given by [46, Sec. 4.4, Th. 7]

$$f_\mu(\mu) = \int_0^\infty \frac{1}{x} f_{\lambda_{K-1},\lambda_K}(x, \mu/x)\, dx. \tag{135}$$

We are interested in $f_\mu(\mu)$ as $\mu \to 0$ and hence we only need to derive the joint pdf $f_{\lambda_{K-1},\lambda_K}(\lambda_{K-1},\lambda_K)$ as $\lambda_K \to 0$. Using the same procedure as in the proof of Theorem 1 (see Appendix II), it can be shown that

$$f_{\lambda_{K-1},\lambda_K}(\lambda_{K-1},\lambda_K) = g(\lambda_{K-1})\lambda_K^{d_K} + o\left(\lambda_K^{d_K}\right) \tag{136}$$

where $g(\lambda_{K-1})$ is a function of λ_{K-1}. Then, substituting back this result in the expression of $f_\mu(\mu)$ in (135), it follows that

$$f_\mu(\mu) = C\mu^{d_K} + o(\mu^{d_K}) \tag{137}$$

where C is a fixed constant in terms of μ. Finally, the exponent of $\Pr(p_{\mathrm{mse},K} \le \mathsf{snr}/K)$ can be bounded as

$$\begin{aligned} d &\ge \lim_{x\to 0} \frac{\log\left(\int_0^{x^2} C\mu^{d_K} d\mu\right)}{\log x} \\ &= \lim_{\mu\to 0} \frac{\log\left(\mu^{2d_K+2}\right)}{\log \mu} = 2d_K + 2 \end{aligned} \tag{138}$$

and this proves that $d > d_K + 1$. □

APPENDIX VI
PROOF OF PROPOSITION 3

The minimum BER scheme distributes the available power over the K active substreams in a waterfilling fashion as in (67). Since the exponent of the probability of not allocating power to the Kth substream is greater than $G_{d,\mathrm{ber}} = G_{d,\mathrm{mse}}$ (see Appendix V for details), we can focus just in the case that $\{p_{k,\mathrm{ber}} > 0\}_{k=1,\ldots,K}$.

First we lower-bound the instantaneous SNR of the minimum BER design using the uniform power allocation, because it leads to a higher sum MSE than the power allocation in (67):

$$\rho_{\mathrm{ber},k} > \left(\frac{1}{K}\sum_{i=1}^K \frac{1}{(\lambda_k/K)\mathsf{snr} + 1}\right)^{-1} - 1 \tag{139}$$

$$> (\lambda_K/K)\mathsf{snr} \tag{140}$$

where in (139) we have forced all substreams to experience the same MSE (cf. [15]) and (140) follows from lower-bounding each λ_k by λ_K. The lower bound in (140) corresponds to the instantaneous SNR achieved by the Kth substream of a diagonal scheme with a uniform power allocation and, hence, we can lower-bound the array gain by G_a.

Let us now consider a non-diagonal linear MIMO transceiver that allocates infinite power to all the substreams except to the Kth one, to which it assigns $p_K = \mathsf{snr}$. Due to the power constraint in (60), the instantaneous SNR of the BER minimizing design can be upper-bounded by the instantaneous SNR of this scheme:

$$\rho_{\mathrm{ber},k} < \left(\frac{1}{K\lambda_k \mathsf{snr} + K}\right)^{-1} - 1$$
$$= K\lambda_K \mathsf{snr} + (K-1). \tag{141}$$

Then, using Corollary 1 with the upper bound in (141), it follows the array gain upper bound given in Proposition 3. □

Acknowledgment

The authors would like to thank the anonymous reviewers for their comments which have significantly helped to improve the organization and contents of the paper.

References

[1] 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.

[2] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–365, Mar. 1998.

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

[4] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 3, no. 5, pp. 311–335, 1998.

[5] M. L. Honig, K. Steiglitz, and B. Gopinath, "Multi channel signal processing for data communications in the presence of crosstalk," *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 551–558, Apr. 1990.

[6] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers—Part I: Unification and optimal designs," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1988–2006, Jul. 1999.

[7] H. Bölcskei and A. J. Paulraj, "Multiple-input multiple-output (MIMO) wireless systems," in *The Communications Handbook*, J. Gibson, Ed., 2nd ed. Boca Raton, FL: CRC Press, 2002, pp. 90.1–90.14.

[8] A. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception (DTDR)," U. S. Patent no. 5 345 599, Sep. 1994.

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

[10] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[11] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1334–1347, Sep. 1994.

[12] 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, no. 12, pp. 2198–2206, Dec. 2001.

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

[14] E. N. Onggosanusi, A. M. Sayeed, and B. D. V. Veen, "Efficient signaling schemes for wideband space-time wireless channels using channel state information," *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, pp. 1–13, Jan. 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 Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[16] Y. Ding, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Minimum BER block precoders for zero-forcing equalization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2410–2423, Sep. 2003.

[17] N. Khaled, S. Thoen, and L. Deneire, "Optimizing the joint transmit and receive MMSE design using mode selection," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 730–737, Apr. 2005.

[18] D. J. Love and R. W. Heath, Jr., "Multimode precoding for MIMO wireless systems," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3674–3687, Oct. 2005.

[19] D. P. Palomar and S. Barbarossa, "Designing MIMO communication systems: Constellation choice and linear transceiver design," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3804–3818, Oct. 2005.

[20] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[21] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*. New York: Kluwer Academic, 1999.

[22] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, Jul. 2002.

[23] P. K. Vitthaladevuni and M. S. Alouini, "A closed-form expression for the exact BER of generalized PAM and QAM constellations," *IEEE Trans. Commun.*, vol. 52, no. 5, pp. 698–700, May 2004.

[24] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[25] J. G. Proakis, *Digital Communications*, 4th ed. : McGraw-Hill, 2001.

[26] M. G. Shayesteh, "On the error probability of linearly modulated signals on frequency-flat Ricean, Rayleigh and AWGN channels," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1454–1466, Feb./Mar./Apr. 1995.

[27] M. S. Alouni and A. J. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1324–1334, Sep. 1999.

[28] G. L. Stüber, *Principles of Mobile Communications*. New York: Kluwer Academic, 1996.

[29] J. B. Andersen, "Antenna arrays in mobile communications: Gain, diversity, and channel capacity," *IEEE Antennas Propag. Mag.*, vol. 42, no. 2, pp. 12–16, Apr. 2000.

[30] N. G. Bruijn, *Asymptotic Methods in Analysis*, 3rd ed. New York: Dover, 1981.

[31] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2172–2178, Nov. 2000.

[32] M. S. Srivastava and C. G. Khatri, *An Introduction to Multivariate Statistics*. Amsterdam, The Netherlands: Elsevier-North Holland, 1979.

[33] C. G. Khatri, "Distribution of the largest or smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *Ann. Math. Stat.*, vol. 35, no. 4, pp. 1807–1810, Dec. 1964.

[34] A. Edelman, "Eigenvalues and condition number of random matrices," Ph.D. dissertation, Dept. Mathematics, MIT, Cambridge, MA, 1989.

[35] P. Comon and D. T. Pham, "An error bound for a noise canceller," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 10, pp. 1513–1517, Oct. 1989.

[36] C. D. Richmond, "Derived pdf of maximum likelihood signal estimator which employs an estimated noise covariance," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 305–311, Feb. 1996.

[37] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[38] H. Shin and J. H. Lee, "Capacity of multi-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[39] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1458–1461, Oct. 1999.

[40] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 694–703, Apr. 2003.

[41] G. Burel, "Statistical analysis of the smallest singular value in MIMO transmission systems," presented at the WSEAS Int. Conf. Signal, Speech, Image Processing (ICOSSIP), Skiathos Island, Greece, Sep. 2002.

[42] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Stat.*, vol. 35, no. 2, pp. 475–501, Jun. 1964.

[43] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Hanover, MA: Now Publishers, 2004.

[44] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press.

[45] *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*, M. Abramowitz and I. A. Stegun, Eds. New York: Dover, 1972.

[46] V. K. Rohatgi, *An Introduction to Probability Theory and Mathematical Statistics*. New York: Wiley, 1976.

Luis G. Ordóñez (S'04) received the electrical engineering degree from the Telecom Barcelona Cchool of the Technical University of Catalonia (UPC), Barcelona, Spain, in 2003. He is currently working towards the Ph.D. degree at the Department of Signal Theory and Communications of UPC.

Since 2004, he has been the recipient of a Spanish Ministry of Education and Science Research Assistantship at the Department of Signal Theory and Communications of UPC. From 2001 to 2004, he participated in the European IST projects I-METRA and NEXWAY. Currently, he is involved in the Sixth Framework Programme project SURFACE. His research is devoted to studying the performance limits of wireless MIMO systems from the information-theoretic and the communication point-of-view.

Daniel P. Palomar (S'99–M'03) received the Electrical Engineering and Ph.D. degrees (both with honors) from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1998 and 2003, respectively.

He is an Assistant Professor in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology, Hong Kong, and a member of the Center for Wireless Information Technology at HKUST. He has held several research appointments, namely, at King's College London (KCL), London, U.K., during 1998; the Technical University of Catalonia (UPC), Barcelona, from January 1999 to December 2003; Stanford University, Stanford, CA, from April to November 2001; the Telecommunications Technological Center of Catalonia (CTTC), Barcelona, Spain, from January to December 2002; the Royal Institute of Technology (KTH), Stockholm, Sweden, from August to November 2003; the University of Rome "La Sapienza," Rome, Italy, from November 2003 to February 2004; and Princeton University, Princeton, NJ, from March 2004 to July 2006. His primary research interests include information-theoretic and signal processing aspects of MIMO channels, with special emphasis on convex optimization theory and majorization theory applied to communication systems.

Dr. Palomar received a 2004/06 Fulbright Research Fellowship; the 2004 Young Author Best Paper Award by the IEEE Signal Processing Society; the 2006 Best Student Paper Award at ICASSP'06 (corecipient); the 2002/03 best Ph.D. prize in Information Technologies and Communications by the Technical University of Catalonia (UPC); the 2002/03 Rosina Ribalta first prize for the Best Doctoral Thesis in Information Technologies and Communications by the Epson Foundation; and the 2004 prize for the best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT. He is an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (IEEE JSAC) 2008 Special Issue on Game Theory in Communication Systems, and the lead Guest Editor of the IEEE JSAC 2007 Special Issue on Optimization of MIMO Transceivers for Realistic Communication Networks.

Alba Pagès-Zamora (M'06) received the M.S. and Ph.D. degrees, both in electrical engineering, from the Universitat Polithcnica de Catalunya (UPC), Spain, in 1992 and 1996, respectively.

In 1992, she joined the Department of Signal Theory and Communications at the same university and became an Associate Professor in 2001. She teaches graduate and undergraduate courses related to communications and signal processing. She has published one book chapter, four papers in international periodic journals and about 35 papers in international conferences. Her current research interests include MIMO channel wireless communications and communication in wireless sensor networks.

Dr. Pagès-Zamora has been also involved in six national scientific projects, two European IST projects, and one R+D project with the European Space Agency. In addition, she has led two R+D projects with the national industry and was the UPC's technical leader in the IST I-METRA and the IST NEXWAY projects. Currently, she is involved in the 6th Framework Programme IST project SURFACE and she is the UPC's technical leader of the 6th Framework Programme IST project WINSOC devoted to develop Wireless Sensor Networks with Self-Organization Capabilities for Critical and emergency Applications.

Javier Rodríguez Fonollosa (S'90–M'92–SM'98) received the Ph.D. degree in electrical and computer engineering at Northeastern University, Boston, MA, in 1992.

In 1993, he joined the Department of Signal Theory and Communications of the Universitat Polithcnica de Catalunya (UPC), Spain, where he became Associate Professor in 1996, Professor in 2003 and Department Head in 2006. He is the author of 17 journal and more than 80 conference papers in the area of signal processing and communications. In 1995 he lead UPCs participation in the European Commission funded ACTS Mobile projects TSUNAMI(II) and SUNBEAM that included the analysis of adaptive antennas in 2on and 3rd generation cellular mobile communication systems. In January 2000 he was appointed technical and project coordinator of the IST project METRA dedicated to the introduction of multiantenna terminals in UMTS. This project continued until 2003 under the name of I-METRA looking into more advanced systems and Systems beyond 3G. Since January 2006, he has coordinated the Sixth Framework Programme IST project SURFACE which evaluates the performance of a generalized air interface with self-configuration capabilities. Also, since October 2006 he has been project coordinator of the five-year Type C-Consolider project Fundamental bounds in Network Information Theory of the National Research Plan of Spain.

Dr. Rodríguez Fonollosa was Co-Chairman and organizer of the IEEE Signal Processing/ATHOS Workshop on Higher-Order Statistics held in Begur, Girona, Spain, in June 1995 and September 2001 and of the IST Mobile Communications Summit 2001 held in Sitges, Barcelona, Spain. He was elected member of the Signal Processing for Communications (SPCOM) Technical Committee of the IEEE Signal Processing Society in January 1999. His research interests include many different aspects of statistical signal processing for communications and information theory. Since May 2005, he has been a member of the Editorial Board of the *EURASIP Signal Processing Journal*.