Diversity and Multiplexing Tradeoff of Spatial Multiplexing MIMO Systems With CSI

Luis Garcia-Ordóñez, *Student Member, IEEE*, Alba Pagès-Zamora, *Member, IEEE*, and Javier Rodríguez Fonollosa, *Senior Member, IEEE*

***Abstract*—Following the seminal work of Zheng and Tse, this paper investigates the fundamental diversity and multiplexing tradeoff of multiple-input–multiple-output (MIMO) systems in which knowledge of the channel state at both sides of the link is employed to transmit independent data streams through the channel eigenmodes. First, the fundamental diversity and multiplexing tradeoff of each of the individual substreams is obtained and this result is then used to derive a tradeoff optimal scheme for rate allocation along channel eigenmodes. The tradeoff of spatial multiplexing is finally compared to the fundamental tradeoff of the MIMO channel and to the one of both space only codes and V-BLAST which do not require channel state information (CSI) at the transmit side.**

***Index Terms*—Capacity-achieving waterfilling, diversity and multiplexing tradeoff, linear multiple-input multiple-output (MIMO) transceivers, spatial multiplexing systems, Wishart matrices.**

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 multi-antenna channels [1]–[4], to wireline digital subscriber line (DSL) systems [5], and to single-antenna frequency-selective channels [6]. In particular, wireless MIMO systems 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) systems.

The gains obtained by the deployment of multiple antennas at both sides of the link can be characterized by the array gain, the diversity gain, and the multiplexing gain [7], [8]. The array gain is the improvement in signal-to-noise ratio (SNR) obtained by coherently combining the information signal on multiple transmit or multiple receive dimensions (specially important in the low-SNR regime) while the diversity gain is the improvement in link reliability obtained by receiving different replicas of the information signal through independently fading dimensions (specially important in the high-SNR regime). 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 spatial multiplexing gain, which refers to the increase of rate at no additional power or bandwidth consumption, is a unique characteristic of MIMO channels.

The focus of this paper is on the high-SNR regime, where the array gain becomes less significant and the performance of a given MIMO system can be effectively characterized in terms of the diversity and the spatial multiplexing gain. Recently, Zheng and Tse showed in [9] that both gains can be simultaneously obtained, but there is a fundamental tradeoff between how much of each type of gain any coding scheme can extract. The main result of [9] is a simple characterization of the optimal tradeoff curve between the diversity gain and the spatial multiplexing gain for any coding scheme when perfect channel state information (CSI) is available only at the receiver. This tradeoff framework does not only enlighten the fundamental limits of MIMO channels but also provides a very interesting procedure to compare the performance of existing diversity-based and multiplexing-based practical MIMO schemes by jointly analyzing their reliability and rate accommodation properties.

In this paper, we deal with the diversity and multiplexing tradeoff of MIMO systems with perfect CSI at the transmitter (CSI-T) and at the receiver (CSI-R)[1] in contrast to [9], where only perfect CSI-R is assumed. More specifically, we concentrate on spatial multiplexing MIMO systems, which divide the incoming data stream into multiple independent substreams without any temporal coding of the data symbols. When the transmitter is not aware of the channel realization, each substream is transmitted on a different antenna, such as the well-known V-BLAST scheme [10]. However, when perfect CSI-T is available, performance can be further improved by transmitting the established substreams through the strongest channel eigenmodes. The resulting spatial multiplexing MIMO system with CSI-T is optimal in the sense of achieving the ergodic channel capacity [3], and also arises in the joint linear transmitter-receiver design of practical MIMO systems, e.g., [6], [11]–[17].

The approach we adopt in this paper is to analyze the individual diversity and multiplexing tradeoff curves of the channel eigenmodes. Then, we obtain the fundamental diversity and multiplexing tradeoff of spatial multiplexing MIMO systems

Manuscript received November 2, 2004; revised December 17, 2007. This work was supported in part by the European Commission (27187 SURFACE), by the Spanish Ministry of Education and Science and FEDER funds (TEC2006-06481 and TEC2004-04526), by the Catalan Government (2005SGR-00639), and by the Research Grant DAG06/07.EG02. The material in this paper was presented in part at the 2005 IEEE International Symposium on Information Theory, Adelaide, Australia, September 2005.

The authors 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).

Communicated by H. Boche, Associate Editor for Communication.
Digital Object Identifier 10.1109/TIT.2008.924670

[1]The combination of CSI-T and CSI-R is henceforth referred to as just CSI-T.

with CSI-T by deriving the optimum rate allocation policy among these channel eigenmodes. The rest of the paper is organized as follows. Section II is devoted to introducing the system model and briefly reviewing the diversity and multiplexing tradeoff framework. Section III describes the signal model corresponding to spatial multiplexing MIMO systems with CSI-T. In Section IV we analyze the individual tradeoff curves of the different channel eigenmodes. Section V provides the fundamental diversity and multiplexing tradeoff of spatial multiplexing systems with CSI-T, which is compared to the fundamental limits offered by the channel in Section VI. Finally, the last section summarizes the main results of the paper.

II. Preliminaries: System Model and Problem Statement

A. System Model

We consider a wireless communication system with M transmit and N receive antennas. The corresponding MIMO channel is represented by an $N \times M$ matrix $\mathbf{H}$, containing the complex path gains $[\mathbf{H}]_{ij}$ between every transmit and receive antenna pair. We adopt an uncorrelated Rayleigh flat fading channel model and, hence, these coefficients are independent identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, $[\mathbf{H}]_{ij} \sim \mathcal{CN}(0,1)$. We also assume that the channel matrix $\mathbf{H}$ remains constant within a block of T symbols, i.e., the block length is significantly smaller than the channel coherence time. In this situation, the received signal within one block can be gathered in an $N \times T$ matrix $\mathbf{Y}$ related to the $M \times T$ transmitted matrix $\mathbf{X}$ as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W} \tag{1}$$

where $\mathbf{W}$ is the additive white Gaussian noise and has independent and identically distributed (i.i.d.) entries with zero mean and unit variance, $[\mathbf{W}]_{ij} \sim \mathcal{CN}(0,1)$. The transmitted signal $\mathbf{X}$ is normalized forcing the transmit power per channel use to satisfy

$$\frac{1}{T}\mathrm{E}\{\|\mathbf{X}\|_\mathrm{F}^2\} \leq \mathsf{SNR} \tag{2}$$

where SNR is the average SNR at each receive antenna.

In the work by Zheng and Tse [9], the channel is assumed to be perfectly known at the receiver only. In contrast, we focus on the situation where the instantaneous channel gains are perfectly known at both transmitter and receiver, so that the transmitter can adapt its transmission strategy relative to the instantaneous channel state under the short-term power constraint in (2).

B. Spatial Diversity and Spatial Multiplexing

The traditional role of multiple antennas was to provide spatial diversity to overcome channel fading by supplying the receiver with several independently faded replicas of the transmitted signal and, thus, increasing the link reliability. A MIMO channel with M transmit and N receive antennas has a maximal diversity order of MN, since there are a maximum of MN random fading coefficients to be averaged over in the reception process of one symbol. Mathematically, we define the diversity gain as the slope of the average error probability versus SNR curve at high SNR, i.e., the SNR exponent of the average error probability.

Additionally, the simultaneous use of multiple antennas at the transmitter and receiver enables also the exploitation of multiple parallel channels which can operate independently. This property is a consequence of the MIMO channel ergodic capacity which can be approximated in the high SNR regime as [4]

$$C(\mathsf{SNR}) \approx \min\{M,N\}\log\left(\frac{\mathsf{SNR}}{M}\right). \tag{3}$$

The channel capacity increases with the SNR as $\log(\mathsf{SNR})$ with the prelog factor $\min\{M,N\}$, in contrast to the prelog factor 1 corresponding to SISO channels. In order to achieve a certain nontrivial fraction of the capacity in the high SNR regime, we consider schemes that support a data rate which also increases with the SNR. Hence, we define a scheme as a family of codes $\{\mathcal{C}(\mathsf{SNR})\}$ of block length T, which employs a different code $\mathcal{C}(\mathsf{SNR})$ for each SNR level. Let $R(\mathsf{SNR})$ be the rate of a code $\mathcal{C}(\mathsf{SNR})$, then a scheme is said to achieve a spatial multiplexing gain r if the supported data rate can be approximated in the high-SNR regime as

$$R(\mathsf{SNR}) \approx r\log\mathsf{SNR}. \tag{4}$$

These intuitive definitions regarding spatial diversity and multiplexing gain can be formalized as follows:

Definition: A MIMO coding scheme $\{\mathcal{C}(\mathsf{SNR})\}$ is said to achieve a spatial multiplexing gain r and a diversity gain d if the data rate satisfies

$$\lim_{\mathsf{SNR}\to\infty}\frac{R(\mathsf{SNR})}{\log\mathsf{SNR}} = r \tag{5}$$

and the error probability follows[2]

$$\lim_{\mathsf{SNR}\to\infty}\frac{\log \mathrm{P}_e(R(\mathsf{SNR}))}{\log\mathsf{SNR}} = -d. \tag{6}$$

For each r, we define $d^*(r)$ to be the supremum of the diversity gain achieved over all schemes.

Notation: Let us define the symbol "$\doteq$" to denote exponential equality, i.e., $P_i(\mathsf{SNR})) \doteq P_j(\mathsf{SNR})$ denotes

$$\lim_{\mathsf{SNR}\to\infty}\frac{\log P_i(\mathsf{SNR})}{\log\mathsf{SNR}} = \lim_{\mathsf{SNR}\to\infty}\frac{\log P_j(\mathsf{SNR})}{\log\mathsf{SNR}}. \tag{7}$$

The symbols "$\dot{\geq}$," "$\dot{\leq}$" are similarly defined. Hence, (6), for instance, can be expressed as

$$\mathrm{P}_e(R(\mathsf{SNR})) \doteq \mathsf{SNR}^{-d}. \tag{8}$$

[2]This definition (introduced in [9, Def. 1]) differs from the standard definition of diversity gain widely used in the space-time coding literature (see e.g., [18]) in the fact that we consider the error probability of a family of coding schemes $\{\mathcal{C}(\mathsf{SNR})\}$, which employs a different code for each SNR level, instead of considering the error probability of a fixed coding scheme.

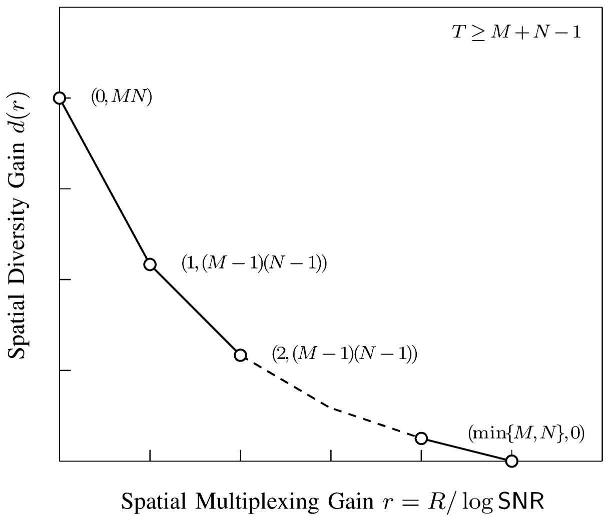

Fig. 1. Fundamental diversity and multiplexing tradeoff of MIMO channels with perfect CSI-R.

C. Fundamental Diversity and Multiplexing Tradeoff With Perfect CSI-R

The design of MIMO systems has been traditionally tackled from two different perspectives: either the maximization of the diversity gain (to increase the transmission reliability) or the maximization of the spatial multiplexing gain (to approach the capacity limits). The diversity and multiplexing tradeoff was conceived as a unified framework to consider simultaneously both design criteria. In fact, given a MIMO channel, both gains can be simultaneously obtained, and the fundamental tradeoff curve shows how much of each one any coding scheme can potentially extract or, equivalently, it provides the fundamental relation between the error probability and the normalized data rate in a system.

The tradeoff curve in [9] is based on analyzing the behavior of the error probability $\mathrm{P}_e(R)$ as function of the data rate in the high-SNR regime by deriving tight (exponentially equal) upper and lower bounds. In particular, the lower bound is given by the outage probability, denoted by $\mathrm{P}_{\mathrm{out}}(R)$ and defined as the probability that the mutual information between the input and the output of the channel is smaller than a given data rate [19]. On the other hand, the upper bound is obtained by conditioning the error probability on the outage event as

$$\begin{aligned}\mathrm{P}_e(R) &= \mathrm{P}_{\mathrm{out}}(R)\mathrm{Pr}(\text{error} \,|\, \text{outage}) + \mathrm{Pr}(\text{error, no outage}) \\ &\leq \mathrm{P}_{\mathrm{out}}(R) + \mathrm{Pr}(\text{error, no outage}). \end{aligned} \tag{9}$$

The second term in (9) can be upper-bounded by the pairwise error probability averaged over the no-outage channel states and the resulting SNR exponent coincides with that of the lower bound whenever the coding length T satisfies

$$T \geq M + N - 1. \tag{10}$$

This illustrates that, under the condition in (10), the typical error occurrence is caused by the outage event. The resulting tradeoff curve is given in the following lemma and plotted in Fig. 1.

Lemma 1 ([9, Th. 2]): Consider the MIMO system in (1). The fundamental diversity and multiplexing tradeoff $d^*(r)$ in an uncorrelated Rayleigh flat fading $N \times M$ MIMO channel is given by the piecewise-linear function connecting the points

$$(k, d^*(k)) \quad k = 0, 1, \ldots, \min\{M, N\} \tag{11}$$

where

$$d^*(k) = (M-k)(N-k) \tag{12}$$

whenever the coding block length condition in (10) is satisfied.

If the block length does not fulfill (10), the given upper bound can not be guaranteed to be tight anymore, since the outage occurrence is not the dominant event in the probability of error. Hence, we have to take into account the three circumstances leading to a detection error: 1) the channel matrix is atypically ill-conditioned, 2) the additive noise is atypically large, and 3) some codewords are atypically close together. When the block length is finite and satisfies (10), the two last events are averaged out and the probability of error is dominated by the bad channel occurrence. Otherwise, all three events come into play and the optimal tradeoff curve can be only partially obtained.

In the case of space-only coding schemes $(T = 1)$ tight bounds on the average error probability can be also derived. The tradeoff curve was indeed completely characterized in [9] when no CSI-T is available. This case is of particular interest in this paper, since we analyze spatial multiplexing schemes with CSI-T which do not perform any temporal encoding strategy, and, thus, it is introduced in Section VI for comparison purposes.

D. Fundamental Diversity and Multiplexing Tradeoff With Perfect CSI-T

The diversity and multiplexing tradeoff framework is still a meaningful framework when perfect CSI-T is available to analyze the performance of delay limited systems in which the data rate cannot depend on the channel variations except in outage states, where the channel cannot support the desired data rate and the data to be transmitted is lost. In fact, during outages the wisest strategy is to stop transmission and do not waste power (see [20, Remark 3]) but, of course, generating unrecoverable errors at the receiver.

When perfect CSI-T is available, the coding strategy can be adapted to the instantaneous channel state by properly tuning the input distribution. As in [9], the input distribution can be taken to be Gaussian with a covariance matrix $\mathbf{Q} \triangleq \mathbf{Q}(\mathbf{H})$, where the optimum $\mathbf{Q}$ maximizes the mutual information, i.e.

$$\mathbf{Q}^* = \arg \max_{\mathbf{Q} \geq 0, \mathrm{Tr}\{\mathbf{Q}\} \leq \mathsf{SNR}} \log\det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger). \tag{13}$$

Then, the outage probability defined as [3], [20]

$$\begin{aligned}\mathrm{P}_{\mathrm{out}}(R) &= \inf_{\mathbf{Q} \geq 0, \mathrm{Tr}\{\mathbf{Q}\} \leq \mathsf{SNR}} \mathrm{Pr}(\log\det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger) \leq R) \\ &= \mathrm{Pr}(\log\det(\mathbf{I} + \mathbf{H}\mathbf{Q}^*\mathbf{H}^\dagger) \leq R) \end{aligned} \tag{14}$$

can be upper-bounded by choosing $\mathbf{Q} = (\mathsf{SNR}/M)\mathbf{I}$ (without exploiting the available CSI-T) and lower-bounded by choosing $\mathbf{Q} = \mathsf{SNR}\mathbf{I}$, since $(\mathsf{SNR}\mathbf{I} - \mathbf{Q}^*) \geq 0$, due to the short-term power constraint, and this implies that $\log\det(\mathbf{I}+\mathsf{SNR}\mathbf{H}\mathbf{H}^\dagger) \geq \log\det(\mathbf{I}+\mathsf{SNR}\mathbf{H}\mathbf{Q}^*\mathbf{H}^\dagger)$. Both bounds were shown in [9] to be tight and the corresponding exponent was derived. Finally, using Fano's inequality, the outage probability was shown to provide a lower bound for the error probability in the high-SNR regime, independently of $\mathbf{Q}$.

In consequence, the fundamental tradeoff presented in [9] for $T \geq M+N-1$ holds for any $\mathbf{Q}$ satisfying (13) with or without CSI-T. This result is not surprising, since the diversity and multiplexing tradeoff framework focuses in the high-SNR regime, where the system is degree-of-freedom limited [9] and, thus, the additional power gain obtained by adapting $\mathbf{Q}$ to the instantaneous CSI-T does not modify the fundamental tradeoff of the channel whenever $T \geq M+N-1$.

III. Spatial Multiplexing MIMO Systems With CSI-T

A. Signal Model

Let us consider the block-fading MIMO system presented in (1), i.e.,

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}. \tag{15}$$

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

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^\dagger \tag{16}$$

where $\mathbf{U} \in \mathbb{C}^{N\times N}$ and $\mathbf{V} \in \mathbb{C}^{M\times M}$ are unitary matrices, and $\boldsymbol{\Lambda} \in \mathbb{R}^{N\times M}$ is a diagonal[3] matrix containing the nonnegative singular values of $\mathbf{H}$ sorted in descending order. Assuming that perfect CSI-T is available, we can rewrite the signal model in (15) without loss of generality as

$$\hat{\mathbf{S}} = \boldsymbol{\Lambda}\mathbf{S} + \mathbf{N} \tag{17}$$

where $\hat{\mathbf{S}} = \mathbf{U}^\dagger\mathbf{Y}$, $\mathbf{S} = \mathbf{V}^\dagger\mathbf{X}$, and $\mathbf{N} = \mathbf{U}^\dagger\mathbf{W}$ with i.i.d. Gaussian entries with zero mean and unit variance, i.e., $[\mathbf{N}]_{ij} \sim \mathcal{CN}(0,1)$, as the distribution of $\mathbf{W}$ is invariant under unitary transformations [21, Ex. 2.4]. Since $\mathrm{E}\{\|\mathbf{X}\|_\mathrm{F}^2\} = \mathrm{E}\{\|\mathbf{S}\|_\mathrm{F}^2\}$, the power constraint in (2) can be expressed in terms of $\mathbf{S}$ as

$$\frac{1}{T}\mathrm{E}\{\|\mathbf{S}\|_\mathrm{F}^2\} \leq \mathsf{SNR} \tag{18}$$

where SNR is the average SNR per receive antenna. Henceforth, we restrict our attention to spatial multiplexing MIMO systems with CSI-T, as formalized in the following assumptions:

Assumption 1: The tth transmitted vector, denoted by $\mathbf{s}_t \in \mathbb{C}^{M\times 1}$, within a given block $\mathbf{S} = [\mathbf{s}_1\,\mathbf{s}_2\,\ldots\,\mathbf{s}_T]$ is generated independently from all other transmitted vectors in $\mathbf{S}$, $\{\mathbf{s}_i\}_{i=1,\ldots,T,\,i\neq t}$, or, equivalently, coding is performed only across space. In this sense, we can restrict our attention to the case $T = 1$, although our results hold for any given block length $T < \infty$.

Assumption 2: $K \leq \min\{M,N\}$ i.i.d., zero-mean, and unit energy data symbols, denoted by $\{z_k\}_{k=1,\ldots,K}$ have to be transmitted per channel use, such that (omitting the index t within a block):

$$\mathbf{s} = \sqrt{\mathbf{P}(\mathbf{H})}\mathbf{z} \tag{19}$$

[3]We call this matrix diagonal even though it may be not square.

where $\mathbf{z} \in \mathbb{C}^{K\times 1}$ gathers the K data symbols, and $\mathbf{P}(\mathbf{H}) \in \mathbb{R}^{M\times K}$ is a nonnegative matrix whose off-diagonal entries are zero and the K nonzero diagonal entries, $\{p_k \triangleq [\mathbf{P}(\mathbf{H})]_{kk}\}_{k=1,\ldots,K}$, contain the power allocated among the K established substreams and satisfy

$$\sum_{k=1}^{K} p_k \leq \mathsf{SNR} \tag{20}$$

due to the power constraint in (18).

The signal model corresponding to spatial multiplexing MIMO systems with CSI-T (see Assumptions 1 and 2) is then given by

$$\hat{\mathbf{s}} = \boldsymbol{\Lambda}\sqrt{\mathbf{P}(\mathbf{H})}\mathbf{z} + \mathbf{n} \tag{21}$$

or, component-wise,[4]

$$\hat{s}_k = \sqrt{\lambda_k p_k}z_k + n_k \qquad k = 1,\ldots,K \tag{22}$$

where λ_k is the kth ordered $(\lambda_1 \geq \cdots \geq \lambda_K)$ eigenvalue of $\mathbf{H}\mathbf{H}^\dagger$ (squared modulus of the kth channel singular value) and n_k is the kth component of the noise vector $\mathbf{n}$.

B. Motivation of Spatial Multiplexing MIMO System With CSI-T

The analysis of spatial multiplexing MIMO systems with CSI-T is mainly motivated by their optimality in terms of the channel capacity. Telatar showed in [3] that the ergodic MIMO channel capacity with perfect CSI-T can be achieved by splitting the incoming data stream into $\min\{M,N\}$ substreams, coding these substreams separately using i.i.d. Gaussian codes, and waterfilling the available transmit power as

$$p_k = \left(\mu - \lambda_k^{-1}\right)^+ \quad k = 1,\ldots,\min\{M,N\} \tag{23}$$

where the water level μ is selected to satisfy the power constraint in (20) with equality. Note that this scheme can be described using the general signal model of spatial multiplexing MIMO systems in (22) with $K = \min\{M,N\}$ and the power allocation given in (23). It is worth pointing out that, sacrificing the low-complexity of the previous coding strategy (and the low-complexity of the corresponding optimum decoding), one can approach capacity with lower error probabilities using multidimensional codes, as can be inferred from the theory of error exponents of parallel channels [22, Sec. 7.5]. As it will be seen in Section 6, this fact has important consequences in terms of the diversity and multiplexing tradeoff[5] achievable by spatial multiplexing MIMO schemes with perfect CSI-T.

However, the motivation behind the analysis of spatial multiplexing MIMO systems with CSI-T is not only supported by their optimality from the channel capacity point-of-view. Let us assume that the ideal Gaussian codes are substituted with practical constellations (e.g., QAM) and that the MIMO system is equipped with a linear transmitter $\mathbf{B} \in \mathbb{C}^{M\times K}$ and a linear

[4]Observe that $\{\hat{s}_k\}_{k=K+1,\ldots,N}$ only contain noise.

[5]The similitudes and differences between the theory of error exponents and the diversity and multiplexing tradeoff framework are investigated in [9, Sec. VI].

receiver $\mathbf{A} \in \mathbb{C}^{N \times K}$ (linear transceiver). The resulting signal model is given by

$$\hat{\mathbf{s}} = \mathbf{A}^{\dagger}(\mathbf{HBz} + \mathbf{w}) \tag{24}$$

where $\hat{\mathbf{s}} \in \mathbb{C}^{K \times 1}$ is the estimated data vector and $\mathbf{z} \in \mathbb{C}^{K \times 1}$ contains the K data symbols as in (19). The uncoded linear transceiver optimization with perfect CSI-T, i.e., the joint optimization of $\mathbf{A}$ and $\mathbf{B}$ in (24) subject to the short term power constraint (equivalent to (20))

$$\mathrm{tr}(\mathbf{BB}^{\dagger}) \leq \mathsf{SNR} \tag{25}$$

has been largely studied in the literature under practical design criteria based on performance measures such as the SNR, the mean-square error (MSE), or the bit error rate (BER), e.g., [6], [11]–[17]. In most cases (see details in Section IV-C), the optimum strategy results in transmitting the K independent data substreams through the K strongest eigenmodes with a particular power allocation that depends on the specific design criterion. Hence, most of the linear MIMO transceiver designs proposed in the literature can be also expressed as in (22).

IV. Diversity and Multiplexing Tradeoff of the Individual Substreams

In this section we analyze the diversity and multiplexing tradeoff behavior of each individual substream transmitted in parallel through the channel eigenmodes, given the spatial multiplexing signal model presented in the previous section. We focus first on the capacity-achieving solution and derive the fundamental tradeoff of each individual MIMO eigenchannel. Then, we extend our analysis to include a more general class of power allocation policies and show that the given individual diversity and multiplexing tradeoff curves also holds for some interesting practical linear transceiver designs.

A. Capacity-Achieving Spatial Multiplexing MIMO System With CSI-T

Let us consider the spatial multiplexing MIMO system in (22) with the power allocation $\{p_k\}_{k=1,\ldots,K}$ given by the capacity-achieving waterfilling in (23). Following the same procedure as in [9], the diversity and multiplexing tradeoff can be derived by bounding the individual error probability of the kth substream, $\mathrm{P}_e^{(k)}(R_k)$, as

$$\mathrm{P}_{\mathrm{out}}^{(k)}(R_k) \,\dot{\leq}\, \mathrm{P}_e^{(k)}(R_k) \,\dot{\leq}\, \mathrm{PEP}^{(k)}(R_k) \tag{26}$$

where $P_{\mathrm{out}}^{(k)}(R_k)$ denotes the outage probability and $\mathrm{PEP}^{(k)}(R_k)$ denotes the pairwise error probability averaged over the no-outage channel states (cf. (9)) of the kth substream. The outage probability is given by [8, Sec. 5.4]

$$\mathrm{P}_{\mathrm{out}}^{(k)}(R_k) = \Pr(\log(1 + p_k\lambda_k) \leq r_k \log \mathsf{SNR}) \tag{27}$$

where $0 \leq r_k \leq 1$ can not possibly exceed the value of 1, since it represents the rate of the kth eigenchannel in which capacity scales as $\log(1+p_k\lambda_k)$. The short-term power constraint in (20) implies that $p_k \leq \mathsf{SNR}$ and, hence, $\mathrm{P}_{\mathrm{out}}^{(k)}(R_k)$ can be lower-bounded with the outage probability obtained when allocating all available power to the kth substream independently of the channel state:

$$\mathrm{P}_{\mathrm{out}}^{(k)}(R_k) \geq \Pr(\log(1 + \mathsf{SNR}\lambda_k) \leq r_k \log \mathsf{SNR}) \tag{28}$$
$$\dot{\geq} \Pr(\lambda_k \leq \mathsf{SNR}^{r_k - 1}). \tag{29}$$

The pairwise error probability $\mathrm{PEP}^{(k)}(R_k)$ can be upper-bounded by substituting the ideal Gaussian codes with some suboptimal coding strategy. Let us assume, for instance, that the data symbols are drawn from a QAM constellation of size SNR^{r_k} in order to sustain a data rate of $R_k = r_k \log \mathsf{SNR}$. Hence, the minimum distance between symbols is $\mathsf{SNR}^{-r_k/2}$. The pairwise error probability between two arbitrary symbols transmitted in the kth substream satisfies[6] [9, Sec. VII-A]

$$\mathrm{PEP}^{(k)}(R_k) \doteq \Pr\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{SNR}^{-r_k/2} < 1\right). \tag{30}$$

In order to take into account the waterfilling mechanism, we rewrite (30) as

$$\mathrm{PEP}^{(k)}(R_k) \,\dot{\leq}\, \Pr\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{SNR}^{-r_k/2} < 1 \,\middle|\, p_k > 0\right) \tag{31}$$
$$\times\ (1 - \Pr(p_k = 0)) + \Pr(p_k = 0)$$
$$\leq \Pr\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{SNR}^{-r_k/2} < 1 \,\middle|\, p_k > 0\right) \tag{32}$$
$$+ \Pr(p_k = 0)$$

and the upper bound in (32) is shown in Appendix A to satisfy

$$\Pr\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{SNR}^{-r_k/2} < 1 \,\middle|\, p_k > 0\right)$$
$$+\Pr(p_k = 0) \,\dot{\leq}\, \Pr(\lambda_k \leq \mathsf{SNR}^{r_k - 1}). \tag{33}$$

Finally, combining (29) and (33) we can conclude that

$$\Pr(\lambda_k \leq \mathsf{SNR}^{r_k-1}) \,\dot{\leq}\, \mathrm{P}_e^{(k)}(R_k)$$
$$\dot{\leq} \Pr(\lambda_k \leq \mathsf{SNR}^{r_k-1}) \tag{34}$$

and, hence

$$\mathrm{P}_e^{(k)}(R_k) \doteq \Pr(\lambda_k \leq \mathsf{SNR}^{r_k-1}). \tag{35}$$

The resulting individual diversity and multiplexing tradeoff curves are characterized in the following theorem and plotted in Fig. 2.

Theorem 1: Consider a spatial multiplexing MIMO system with CSI-T (see Assumptions 1 and 2) and consider that the power allocated to the substream transmitted through the kth ordered channel eigenmode is given by the capacity-achieving waterfilling, i.e., $p_k = (\mu - \lambda_k^{-1})^+$. The individual diversity and multiplexing tradeoff $d_S^{(k)}(r)$ of the kth substream in an uncorrelated Rayleigh flat fading $N \times M$ MIMO channel is given by

$$d_S^{(k)}(r_k) = d_k(1 - r_k) \quad 0 \leq r_k \leq 1 \tag{36}$$

[6]The only difference with [9] is that in our case the channel power gain is λ_k instead of $\|\mathbf{H}\|_F^2$.

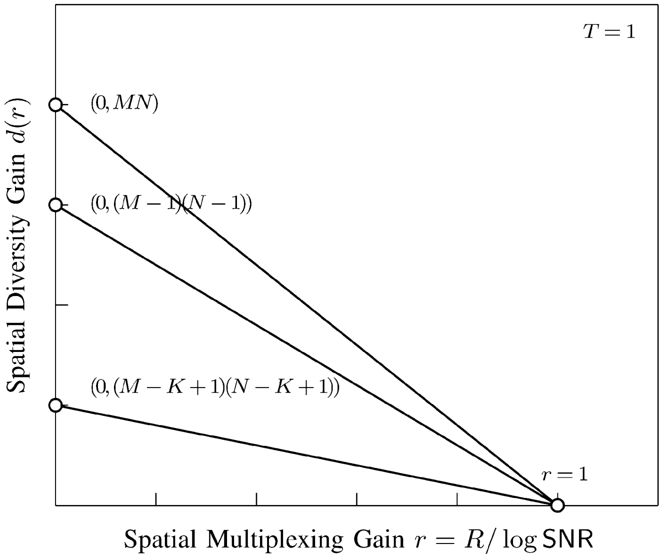

Fig. 2. Fundamental diversity and multiplexing tradeoff of the MIMO channel eigenmodes.

where d_k is defined as

$$d_k = (M-k+1)(N-k+1). \tag{37}$$

Proof: See Appendix B. □

Note that the individual diversity and multiplexing tradeoff curves presented in Theorem 1 are, in fact, the fundamental tradeoff curves of the eigenmodes of an uncorrelated Rayleigh flat fading $N \times M$ MIMO channel.

B. General Spatial Multiplexing MIMO Systems With CSI-T

Let us now consider a spatial multiplexing MIMO system with CSI-T such that the power allocated to the kth substream is a function of the SNR and possibly a function of the K strongest channel eigenvalues, i.e., $p_k = g_k(\lambda_1, \ldots, \lambda_K, \mathsf{SNR})$, and satisfies the short-term power constraint in (20). We also assume that there exits a deterministic strictly positive quantity, ϕ_k, such that

$$\Pr(p_k \leq \phi_k \mathsf{SNR}) \dot{\leq} \mathrm{P}_{\mathrm{out}}^{(k)}(R_k). \tag{38}$$

This condition ensures that the power allocation does not inherently limit the achievable individual tradeoff either by allocating zero or a small amount of power in terms of the SNR with large probability. Observe that, due to the short-term power constraint, it is not possible to design a power allocation which privileges the individual tradeoffs of some substreams by penalizing the tradeoffs of the remaining substreams. The individual diversity and multiplexing tradeoff curves presented in Theorem 1 for the capacity-achieving waterfilling also hold for this general power allocation policy as presented in the following corollary.

Corollary 1: Consider a spatial multiplexing MIMO system with CSI-T (see Assumptions 1 and 2) and consider that the power allocated to the substream transmitted through the kth ordered channel eigenmode is given by $p_k = g_k(\lambda_1, \ldots, \lambda_K, \mathsf{SNR})$ under the condition in (38) and the short-term power constraint in (20). The individual diversity and multiplexing tradeoff $d_S^{(k)}(r)$ of the kth substream in an uncorrelated Rayleigh flat fading $N \times M$ MIMO channel is given by

$$d_S^{(k)}(r_k) = d_k(1-r_k) \quad 0 \leq r_k \leq 1 \tag{39}$$

where d_k is defined in Theorem 1.

Proof: See Appendix C. □

Corollary 1 shows that the power allocated to an individual substream (under a short-term power constraint) cannot improve the SNR exponent of the error probability, which is on the contrary fixed by the order of the channel eigenmode used for transmission.

C. Practical Spatial Multiplexing MIMO Systems With CSI-T

In Section III-B, we introduced linear MIMO transceivers as a practical implementation of spatial multiplexing systems with CSI-T. Palomar developed in [17] a unifying framework, in which the design problem of linear MIMO transceivers under perfect CSI-T is formulated as the minimization of some cost function of the MSEs of the individual substreams, since the other common practical system quality measures such as the SNR, or the BER can be easily related to the MSEs. More exactly, the optimum transmit matrix $\mathbf{B}$ and receive matrix $\mathbf{A}$ (see (24)) are obtained as

$$\{\mathbf{A}, \mathbf{B}\} = \arg \min_{\mathbf{A},\mathbf{B}} f_0(\{\mathsf{MSE}_k\}_{k=1,\ldots,K}) \tag{40}$$

$$\text{subject to } \mathrm{tr}(\mathbf{B}\mathbf{B}^\dagger) \leq \mathsf{SNR} \tag{41}$$

where $f_0(\cdot)$ denotes the design cost function and MSE_k is the kth diagonal element of the MSE matrix $\mathbf{E} = \mathrm{E}\{(\hat{\mathbf{s}} - \mathbf{z})(\hat{\mathbf{s}} - \mathbf{z})^\dagger\}$. In particular, [17] shows that the optimum linear receiver is always given by the MMSE solution [17, eq. (7)]

$$\mathbf{A} = (\mathbf{H}\mathbf{B}\mathbf{B}^\dagger\mathbf{H}^\dagger + \mathbf{I}_N)^{-1}\mathbf{H}\mathbf{B} \tag{42}$$

and provides the optimum linear transmitter for the class of Schur-concave and Schur-convex cost functions. Interestingly, this framework embraces most of the linear transceiver schemes previously proposed in the literature, e.g., [6], [11]–[15].

In the case of Schur-concave cost functions, the optimum transmit strategy establishes K independent data streams through the K strongest channel eigenmodes with a power allocation that depends on the particular cost function [17, eq. (14)]

$$\mathbf{B} = \mathbf{U}\sqrt{\mathbf{P}(\mathbf{H})} \tag{43}$$

where $\mathbf{U} \in \mathbb{C}^{M \times K}$ has as columns the eigenvectors of $\mathbf{H}\mathbf{H}^\dagger$ corresponding to the K largest eigenvalues and $\mathbf{P}(\mathbf{H}) \in \mathbb{C}^{K \times K}$ is a diagonal matrix containing the power allocation policy $\{p_k\}_{k=1,\ldots,K}$. For instance, when the design criterion is the minimization of the product of MSEs, the optimum power allocation is given by the capacity-achieving waterfilling in (23) [17, eq. (24)], when it is the minimization of the weighted sum of MSEs, the optimum power allocation is given by [17, eq. (22)]

$$p_k = \left(\mu w_k^{1/2} \lambda_k^{-1/2} - \lambda_k^{-1}\right)^+ \quad k = 1, \ldots, K \tag{44}$$

where w_k is a strictly positive constant (weight) and μ is chosen to fulfill the power constraint in (41), and when it is the maximization of the weighted product of SNRs, the optimum power allocation is given by [17, eq. (40)]

$$p_k = \frac{w_k}{\sum_{i=1}^{K} w_i}\mathsf{SNR} \quad k = 1, \ldots, K \tag{45}$$

where w_k is a strictly positive constant (weight). In the following propositions we show that the power allocation policies in (44) and (45) satisfy the conditions of Corollary 1.

Proposition 1: The power allocation that minimizes the weighted sum of MSEs given in (44) satisfies the conditions of Corollary 1.

Proof: See Appendix D. □

Proposition 2: The power allocation that maximizes the weighted product of SNRs given by (45) satisfies the conditions of Corollary 1.

Proof: The power allocation in (45) is channel nondependent and, thus, the condition in (38) is directly satisfied. □

Finally, since the linear transceivers of (24) with $\mathbf{A}$ and $\mathbf{B}$ as given in (42) and (43), respectively, can be described by the general signal model presented in (22) for spatial multiplexing MIMO system with CSI-T [23], it follows that the tradeoff curves given in Corollary 1 also characterize the individual substreams of practical linear transceiver designed under Schur-concave cost functions.

In the case of Schur-convex cost functions, however, the optimum transmit strategy transmits a unitary transformation of the K independent data streams through the K strongest channel eigenmodes [17, eq. (14)]

$$\mathbf{B} = \mathbf{U}\sqrt{\mathbf{P}(\mathbf{H})}\mathbf{Q} \tag{46}$$

where $\mathbf{U}$ and $\sqrt{\mathbf{P}(\mathbf{H})}$ are defined as in (43), and $\mathbf{Q} \in \mathbb{C}^{K\times K}$ is a unitary matrix such that all K substreams experience the same equivalent channel (see [17] for details). Furthermore, the optimum power allocation is always given by (44). Schur-convex cost functions appear, for instance, in the minimization of the maximum of the MSEs, or in the minimization of the sum of BERs (under equal constellations). These schemes cannot be described by the signal model in (22), since the individual data symbols are not transmitted in parallel through the channel eigenmodes and, hence, Corollary 1 cannot be applied.

V. Diversity and Multiplexing Tradeoff of Spatial Multiplexing Systems With CSI-T

In the previous section, we derived the SNR exponent of the error probability associated to the individual substreams transmitted in parallel through the channel eigenmodes when using the capacity-achieving waterfilling or a general power allocation which satisfies the condition in (38). Based on this result, we obtained the individual diversity and multiplexing tradeoff curves of the established substreams in Theorem 1 and Corollary 1, respectively. In contrast, in this section we are interested in the global diversity and multiplexing tradeoff of the spatial multiplexing MIMO system with CSI-T that results from the combination of these individual substreams.

Recall that the K-dimensional data symbol vector $\mathbf{z} \in \mathbb{C}^{K\times 1}$ is formed by the K i.i.d. individual data symbols $\{z_k\}_{k=1,\ldots,K}$ transmitted through each channel eigenmode (see Assumption 2). We define the global error probability of spatial multiplexing MIMO systems with CSI-T, denoted by $\mathrm{P}_e(R)$, $R = \sum_{k=1}^{K} r_k \log \mathsf{SNR}$, as the probability of having an error in the detection of at least one of the K individual symbols in the data symbol vector $\mathbf{z}$. Thus, $\mathrm{P}_e(R)$ can be bounded as

$$\max_{1\le k\le K} \mathrm{P}_e^{(k)}(R_k) \le \mathrm{P}_e(R) \le \sum_{k=1}^{K} \mathrm{P}_e^{(k)}(R_k). \tag{47}$$

The SNR exponent of $\sum_{k=1}^{K} \mathrm{P}_e^{(k)}(R_k)$ is dominated by the term with the lowest exponent and this term corresponds precisely to the substream with the worst error probability, i.e., $\max_{1\le k\le K} \mathrm{P}_e^{(k)}(R_k)$. Assuming that the power allocation among the K substreams is either given by the capacity-achieving waterfilling (or satisfies the condition in (38)), we can use Theorem 1 (or Corollary 1) and (47) to obtain that

$$\mathrm{P}_e(R) \doteq \mathsf{SNR}^{-\min_K d_S^{(K)}(r_K)} \tag{48}$$

where $d_S^{(K)}(r_K)$ is defined in (36) and[7] $K \ge \lceil r \rceil$, since the individual multiplexing gains $\{r_k\}_{k=1,\ldots,K}$ cannot exceed the value 1. Hence, as it becomes apparent later on in this section, the diversity and multiplexing tradeoff curve depends on both the number of active substreams K and the rate allocation policy adopted by the transmitter, $\{r_k\}_{k=1,\ldots,K}$. In the following, we first present the tradeoff curve obtained when using a uniform rate allocation among the active substreams. Then, we derive the optimum rate allocation, which results in the fundamental diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI-T as formalized in Assumptions 1 and 2.

A. *Diversity and Multiplexing Tradeoff With Uniform Rate Allocation*

Let us denote by $K(r)$ the number of active substreams as a function of the spatial multiplexing gain $r = \sum_{k=1}^{K(r)} r_k$ with $0 < r_k \le 1$ and let us assume a uniform rate allocation policy among these $K(r)$ substreams, i.e.

$$r_k = r/K(r) \quad k = 1, \ldots, K(r). \tag{49}$$

Then, as shown in (48), the diversity and multiplexing tradeoff is obtained by deriving the optimum number of active substreams, i.e.

$$K^*(r) = \arg \max_{K\ge\lceil r\rceil} \min_{1\le k\le K} d_S^{(k)}(r/K). \tag{50}$$

This problem is implicitly solved in the following theorem and the resulting diversity and multiplexing tradeoff curve $d_{SU}^*(r)$ is plotted in Fig. 3.

[7] $\lceil a \rceil$ denotes the smallest integer bigger than or equal to a.

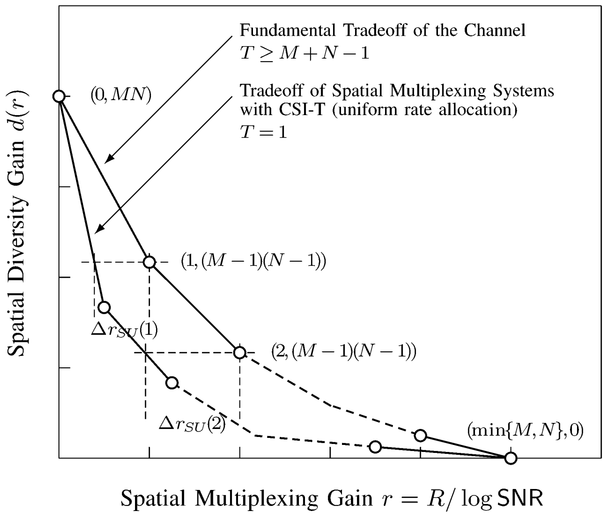

Fig. 3. Diversity and multiplexing tradeoff of spatial multiplexing systems with CSI-T (uniform rate allocation).

Theorem 2: Consider a spatial multiplexing MIMO system with CSI-T (see Assumptions 1 and 2) and consider that the power allocation among the active substreams is either given by the capacity-achieving waterfilling in (23) or satisfies the condition in (38). The diversity and multiplexing tradeoff $d^*_{SU}(r)$ achievable with a uniform rate allocation policy in an uncorrelated Rayleigh flat fading $N \times M$ MIMO channel is given by the piecewise-linear function connecting the points $(r(K), d^*_{SU}(K))$ and $(\min\{M,N\}, 0)$, where

$$\begin{aligned} r(K) &= K - \frac{K d_{K+1}}{(K+1)d_K - K d_{K+1}} \\ d^*_{SU}(K) &= \frac{d_K d_{K+1}}{(K+1)d_K - K d_{K+1}} \\ &\qquad \text{for } K = 0, \ldots, \min\{M,N\} - 1 \end{aligned} \tag{51}$$

with $d_K = (M-K+1)(N-K+1)$ and $r(K)$ denoting the values of r at which the number of active substreams is increased from K to $K+1$.

Proof: Assuming a uniform rate allocation among the K active eigenchannels, the SNR exponent of the kth substream is equal to

$$d_S^{(k)}(r/K) = d_k(1 - r/K) \quad k = 1, \ldots, K \tag{52}$$

where $d_k = (M-k+1)(N-k+1)$. Since it holds that $d_k > d_{k+1}$ for $k = 1, \ldots, K-1$, the global performance of the spatial multiplexing scheme with K active channel eigenmodes is dictated by the individual performance of the Kth substream. Hence, for a given multiplexing gain r, we can maximize the diversity gain by optimizing the number of active substreams K, i.e.

$$K^*(r) = \arg \max_{K \geq \lceil r \rceil} d_S^{(K)}(r/K) \tag{53}$$

and the global diversity and multiplexing tradeoff is given by the supremum of the individual tradeoff curves, $\{d_S^{(K)}(r/K)\}_{K=1,\ldots,\min\{M,N\}}$, i.e.,

$$d^*_{SU}(r) = \max_{K \geq \lceil r \rceil} d_S^{(K)}(r/K). \tag{54}$$

Clearly, the resulting tradeoff curve is a piecewise-linear function connecting the points $r(K)$, defined as the values of r at which the optimum number of active substreams is increased from K to $K+1$, i.e., the discontinuity points of $K^*(r)$. Forcing

$$d_S^{(K)}(r/K) = d_S^{(K+1)}(r/(K+1)) \tag{55}$$

it follows that

$$\begin{aligned} r(K) &= K - \frac{K d_{K+1}}{(K+1)d_K - K d_{K+1}} \\ &\qquad K = 1, \ldots, \min\{M,N\} - 1 \end{aligned} \tag{56}$$

and, finally, substituting (56) in (52), we obtain the corresponding diversity gains

$$\begin{aligned} d^*_{SU}(K) &= d_S^{(K)}(r(K)/K) \\ &= \frac{d_K d_{K+1}}{((K+1)d_K - K d_{K+1})} \\ &\qquad K = 1, \ldots, \min\{M,N\} - 1 \end{aligned} \tag{57}$$

which completes the proof. □

B. Diversity and Multiplexing Tradeoff With Optimal Rate Allocation

The derivation of the fundamental tradeoff of spatial multiplexing systems with CSI-T requires maximizing the minimum SNR exponent not only over the number of active channel eigenmodes, $K(r)$, but also over the rate allocation policy among the established substreams, $\{r_k\}_{k=1,\ldots,K(r)}$, i.e.

$$d^*_S(r) = \max_{K, \{r_k\}} \min_{1 \leq k \leq K} d_S^{(k)}(r_k) \tag{58}$$

$$\text{subject to } K \geq \lceil r \rceil \tag{59}$$

$$\sum_{k=1}^{K} r_k = r \tag{60}$$

$$0 < r_k \leq 1. \tag{61}$$

This problem is implicitly solved in the following theorem and the resulting diversity and multiplexing tradeoff curve $d^*_S(r)$ is plotted in Fig. 4.

Theorem 3: Consider a spatial multiplexing MIMO system with CSI-T (see Assumptions 1 and 2) and consider that the power allocation among the active substreams is either given by the capacity-achieving waterfilling in (23) or satisfies the condition in (38). The fundamental diversity and multiplexing tradeoff $d^*_S(r)$ in an uncorrelated Rayleigh flat fading $N \times M$ MIMO channel is given by the piecewise-linear

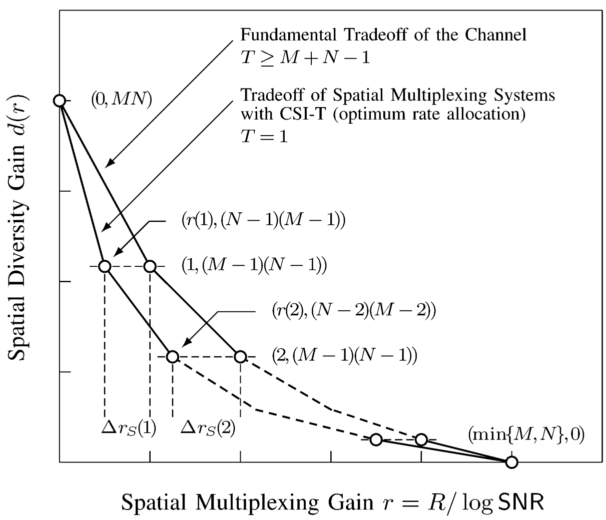

Fig. 4. Diversity and multiplexing tradeoff of spatial multiplexing systems with CSI-T (optimum rate allocation).

function connecting the points $(0, MN)$, $(r(K), d_S^*(K))$, and $(\min\{M,N\}, 0)$, where

$$r(K) = K - d_{K+1}\left(\sum_{k=1}^{K} 1/d_k\right)$$
$$d_S^*(K) = (M-K)(N-K)$$
$$\text{for } K = 1, \ldots, \min\{M,N\} - 1 \quad (62)$$

with $d_K = (M-K+1)(N-K+1)$ and $r(K)$ denoting the values of r at which the number of active substreams is increased from K to $K+1$. The fundamental diversity and multiplexing tradeoff is achievable if and only if the rate is allocated among the optimum number of active substreams $K^*(r)$ as

$$r_k = 1 - \frac{1/d_k}{\sum_{k=1}^{K^*(r)} 1/d_k}(K^*(r) - r) \qquad k = 1, \ldots, K^*(r). \quad (63)$$

Proof: The constrained optimization problem in (58) is equivalent to first imposing a rate allocation that assures the same SNR exponent $d(K,r)$ for the K active substreams, i.e.

$$d(K,r) = d_S^{(k)}(r_k) = d_k(1 - r_k) \qquad k = 1, \ldots, K \quad (64)$$

and, then, maximizing the resulting SNR exponent $d(K,r)$ over K subject to the constraint in (59), i.e.

$$d_S^*(r) = \max_{K \geq \lceil r \rceil} d(K,r) \quad (65)$$
$$K^*(r) = \arg \max_{K \geq \lceil r \rceil} d(K,r). \quad (66)$$

From (64), we obtain the individual rates that force all K active substreams to have the same SNR exponent

$$r_k = 1 - \frac{d(K,r)}{d_k} \quad k = 1, \ldots, K. \quad (67)$$

and, combining (67) and (60), it follows that

$$d(K,r) = \frac{K - r}{\sum_{k=1}^{K} 1/d_k}. \quad (68)$$

Then, the fundamental diversity and multiplexing tradeoff $d_S^*(r)$ is given by the supremum of the individual curves, $\{d(K,r)\}_{K=1,\ldots,\min\{M,N\}}$. As in Theorem 2, the resulting tradeoff curve is a piecewise-linear function connecting the points $r(K)$, defined as the values of r at which the optimum number of active substreams is increased from K to $K+1$, i.e., the discontinuity points of $K^*(r)$. Forcing

$$d(K, r(K)) = d(K+1, r(K)) \quad (69)$$

it follows that

$$r(K) = K - d_{K+1}\left(\sum_{k=1}^{K} 1/d_k\right)$$
$$K = 1, \ldots, \min\{M,N\} - 1 \quad (70)$$

and, substituting (70) back in (68), we obtain the corresponding diversity gains

$$d_S^*(K) = d(K, r(K)) = (M-K)(N-K)$$
$$K = 1, \ldots, \min\{M,N\} - 1. \quad (71)$$

Finally, the optimal rate allocation among substreams comes simply from combinig (67) and (68) and this completes the proof. □

C. Achievability of the Diversity and Multiplexing Tradeoff

In the previous section we obtained the diversity and multiplexing tradeoff of spatial multiplexing systems with CSI-T based on the individual tradeoff curves derived in Theorem 1. Thus, the diversity and multiplexing tradeoff curves provided in Theorems 2 and 3 for the uniform and the optimal rate allocation policy, respectively, are achieved whenever the individual tradeoffs in Theorem 1 are achieved. In addition to the capacity-achieving spatial multiplexing scheme analyzed in Section IV-A, we considered spatial multiplexing systems with a general power allocation in Section IV-B and practical linear MIMO transceivers in Section IV-B. Consequently, the fundamental diversity and multiplexing tradeoff of spatial multiplexing systems with CSI-T can be achieved with practical linear MIMO transceiver designs even with a channel nondependent power allocation and using QAM constellations.

VI. Analysis of the Results

In this section, we analyze the obtained diversity and multiplexing tradeoff of spatial multiplexing systems with CSI-T ($T = 1$ or Assumption 1 is satisfied) with respect to the fundamental tradeoff of the channel ($T \geq M+N-1$) and the achievable tradeoffs under different assumptions when only CSI-R is available.

A. Tradeoff of Spatial Multiplexing With CSI-T ($T = 1$) Versus Fundamental Tradeoff of the Channel ($T \geq M + N - 1$)

In this section, we compare the diversity and multiplexing tradeoff of spatial multiplexing systems with CSI-T with respect to the fundamental tradeoff offered by the channel for $T \geq M + N - 1$ as illustrated in Figs. 3 and 4 for the uniform and the optimum rate allocation policies, respectively. It can be observed that, for a given diversity gain, spatial multiplexing schemes suffer from a degradation with respect to the diversity and multiplexing limits of the channel, due to the lack of coding between substreams (see Assumption 2). In fact, the established substreams are coded independently and transmitted through the K strongest channel eigenmodes and, hence, the outage event must be understood under the perspective of the individual eigenmodes, since spatial multiplexing systems experiences an outage whenever at least one of the K established substreams is in outage.

For instance, let us consider the fundamental tradeoff of the channel and the fundamental tradeoff of spatial multiplexing systems with CSI-T (see Fig. 4) and let us focus on values of r close to 0, such that only the first linear part of both curves is taken into consideration. This requires $0 \leq r \leq 1$ in the fundamental tradeoff of the channel and $0 \leq r \leq 1 - \frac{(M-1)(N-1)}{MN}$ in the spatial multiplexing case (see Fig. 5). In this region, the tradeoff exponent of the spatial multiplexing system, $d_S^*(r)$, is derived evaluating the probability of outage for the transmission along the eigenmode associated with the maximum eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$, denoted by λ_1. More exactly, applying Theorem 1 and noting that for $0 \leq r \leq 1 - \frac{(M-1)(N-1)}{MN}$ only one substream is used, it follows that $r_1 = r$ and

$$\begin{aligned} \mathrm{P_{out}}(r) &\doteq \Pr(\lambda_1 \leq \mathsf{SNR}^{-(1-r)}) \\ &\doteq \mathsf{SNR}^{-d_S^*(r)} \\ &= \mathsf{SNR}^{-MN(1-r)}. \end{aligned} \tag{72}$$

On the other hand, the fundamental tradeoff curve is obtained evaluating the dominant term of the SNR exponent of the outage probability, which in the interval $0 \leq r \leq 1$ can be expressed as [9]

$$\begin{aligned} \mathrm{P_{out}}(r) \doteq \Pr(&\lambda_1 \leq \mathsf{SNR}^{-(1-r)}, \lambda_2 \leq \mathsf{SNR}^{-1}, \dots, \\ &\lambda_{\min\{M,N\}} \leq \mathsf{SNR}^{-1}) \doteq \mathsf{SNR}^{-d^*(r)} \end{aligned} \tag{73}$$

where $d^*(r) = (M-1)(N-1) + (M+N-1)(1-r)$. Comparing expressions (72) and (73), it is now apparent that the probability of outage of the strongest eigenmode is higher than the dominant outage event probability of the MIMO channel, which in addition requires all other eigenmodes ($\min\{M,N\} - 1$) to be fully ineffective. A similar argument can be used in different regions of the tradeoff curve (see Fig. 4) and this explains the diversity and multiplexing loss.

This performance degradation can be characterized more exactly by defining the spatial multiplexing loss $\Delta r(k)$ as

$$\begin{aligned} \Delta r(K) &= K - r(d^*(K)) \\ K &= 1, \dots, \min\{M,N\} - 1 \end{aligned} \tag{74}$$

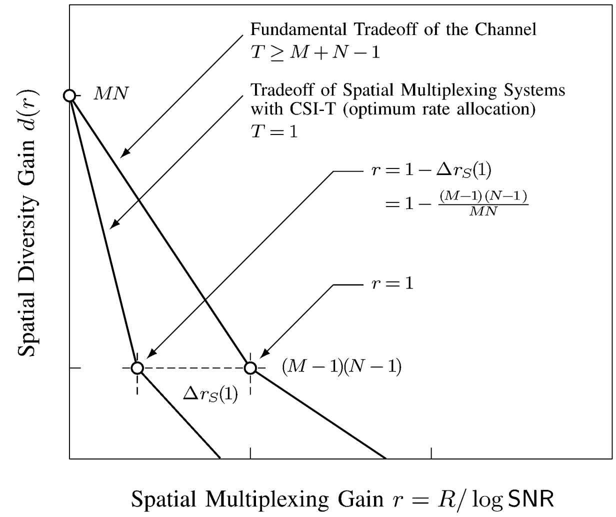

Fig. 5. Diversity and multiplexing tradeoff of the channel and of spatial multiplexing systems with CSI-T (low spatial multiplexing regime).

where $r(d^*(K))$ is the spatial multiplexing gain for which the spatial multiplexing system achieves a diversity gain of $(M - K)(N - K)$. Using Theorems 2 and 3, we obtain

$$\Delta r_{SU}(k) = \min_{0 \leq k < K} k + d_{K+1}\frac{K-k}{d_{K-k}} \tag{75}$$

for the uniform rate allocation and

$$\Delta r_S(K) = d_{K+1}\sum_{k=1}^{K}\frac{1}{d_k}. \tag{76}$$

for the optimum rate allocation. Observe that in (75) the multiplexing loss is dominated by $K-k$ times the inverse of the maximum diversity of the worst active eigenmode, $(K-k)/d_{K-k}$, whereas in (76) the multiplexing loss depends on the sum of the inverse of the maximum diversities associated the K active channel eigenmodes, $\sum_{k=1}^{K} 1/d_k$. By allocating the same rate among all substreams, the performance of the system is limited by the performance of the $(K-k)$th channel eigenmode, since it is the worst active subchannel and is the first one to become in outage. Therefore, in this case, not even the active subchannels are fully exploited. On the contrary, by using the optimal rate allocation, we force all active channel eigenmodes to become in outage simultaneously or, equivalently, we transmit always (for any given diversity) at the maximum individual rate supported by each channel eigenmode. Moreover, with the optimal rate allocation the number of active substreams is optimal in the sense that it coincides with the number of channel eigenmodes which are typically not in outage (see geometrical interpretation in [9]) and this is not always the case when using the uniform rate allocation, where the term k in (75) can further increase the multiplexing loss. As we illustrate in the following section, however, spatial multiplexing with CSI-T often provides an advantage with respect to space-only codes and spatial-multiplexing with CSI-R only.

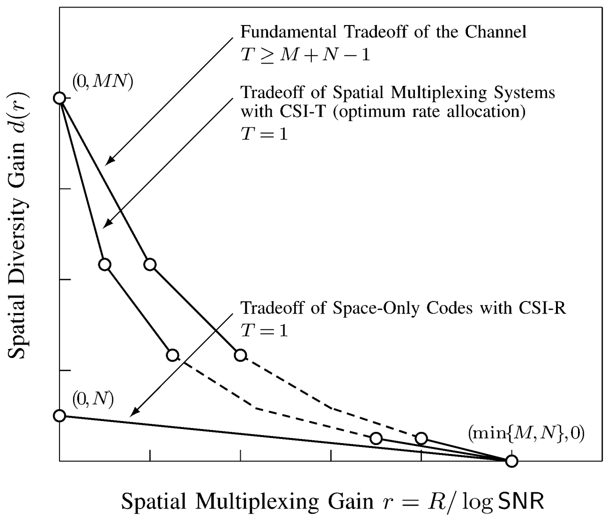

Fig. 6. Comparison of the tradeoff curves of spatial multiplexing systems with CSI-T and space-only codes with CSI-R only.

B. Tradeoff of Spatial Multiplexing With CSI-T ($T = 1$) *Versus Tradeoff of Space-Only Codes With CSI-R* ($T = 1$)

In this section we focus on the tradeoff achieved by any space-only coding scheme ($T = 1$) and CSI-R only. The fundamental diversity and multiplexing tradeoff of space-only codes was derived in [9] as given in the following lemma.

Lemma 2 ([9, Sec. IV-D]): Consider the MIMO system in (1) with $T = 1$ and perfect CSI-R only. The fundamental diversity and multiplexing tradeoff $d^*_{S_0}(r)$ in an uncorrelated Rayleigh flat fading $N \times M$ MIMO channel with $M \leq N$ is given by

$$d^*_{S_0}(r) = N(1 - r/\min\{M,N\}) \qquad 0 \leq r \leq \min\{M,N\}. \tag{77}$$

Observe that the fundamental tradeoff of spatial multiplexing MIMO systems with CSI-T obtained in Theorem 3 can expressed as

$$d^*_S(r) = \max_{K, K \geq r} \frac{K}{\sum_{k=1}^{K} 1/d_k}(1-r/K) \quad 0 \leq r \leq \min\{M,N\}. \tag{78}$$

Clearly, it holds that $d^*_S(r) \geq d^*_{S_0}(r)$ as long as $M \neq N$ (see Fig. 6). Hence, CSI-T exploitation, or more exactly, the diversity gain increase obtained by selecting only the strongest channel eigenmodes and optimally allocating the rate overcomes the limitations inherent to the independent coding procedure of spatial multiplexing systems. Only when $M = N$ and for

$$r \geq \max_{K, K \geq r} \frac{K - N\sum_{k=1}^{K} 1/d_k}{1 - \sum_{k=1}^{K} 1/d_k} \tag{79}$$

higher diversity gains can be achieved with space-only codes. Intuitively, in the high multiplexing gain regime, all channel eigenmodes are used to transmit independent data streams, rendering CSI-T less useful. On the contrary, an arbitrary space coding scheme can potentially benefit from coding among different streams. This turns out to be more beneficial than spatial multiplexing with CSI-T when $M = N$, i.e., the weakest eigenmode has diversity 1. In fact, when $M = N$ and $r > \min\{M,N\} - 1$, $d^*_{S_0}(r)$ coincides with the fundamental tradeoff of the channel $d^*(r)$.

Observe that it is false to conclude from this result that there are circumstances under which is better to ignore the perfect CSI-T information at the transmitter. Indeed, the comparison is not completely fair, since space-only codes can transmit any data vector while spatial multiplexing MIMO systems are restricted to transmit data vectors with i.i.d. components due to Assumption 2. This low-complexity spatial multiplexing strategy, although capacity-achieving, is suboptimal in terms of error probability (see e.g., [3, Ex. 1]).

C. Tradeoff of Spatial Multiplexing With CSI-T ($T = 1$) *Versus Tradeoff of V-Blast* ($T = 1$)

Finally, in this section we compare the fundamental tradeoff of spatial multiplexing systems with CSI-T and with CSI-R only, in order to evaluate the benefits of having perfect channel knowledge at the transmitter. Let us consider the spatial multiplexing system with CSI-R only, in which an independent substream is transmitted through each transmit antenna ($M \leq N$) with the same rate and the receiver uses the nulling and canceling algorithm. This scheme is commonly known as V-BLAST [10] and the corresponding diversity and multiplexing tradeoff curve is presented in the following theorem.

Theorem 4: Consider the spatial multiplexing MIMO system with CSI-R in [10]. The fundamental diversity and multiplexing tradeoff curve $d_V(r)$ in an uncorrelated Rayleigh flat fading $N \times M$ MIMO channel with $M \leq N$ is given by

$$d_V(r) = (N - M + 1)(1 - r/\min\{M,N\}) \qquad 0 \leq r \leq \min\{M,N\}. \tag{80}$$

Proof: The diversity and multiplexing tradeoff of V-BLAST was obtained in [9, Sec. VII-B] for the case $N = M$. Using the procedure presented in [9] along with the high-SNR performance analysis of V-BLAST in [24], the tradeoff curve given in Theorem 4 follows. □

Using (78), it is not difficult to show that $d^*_S(r) > d_V(r)$ for all values or r. This result is not surprising, since the spatial multiplexing scheme with CSI-T exploits the channel information at the transmitter to select only the best eigenmodes and perform the optimal rate allocation among them, whereas the V-BLAST architecture must rely blindly on all eigenchannels. In fact, for this reason, the tradeoff curves for spatial multiplexing with CSI-T and uniform rate allocation and for V-BLAST coincide in the high multiplexing gain regime, i.e., when r is such that $K^*(r) = \min\{M,N\}$, since this particular spatial multiplexing scheme with CSI, does not use the CSI-T information for either dicarding substreams or allocating the target rate optimally.

VII. CONCLUSION

This paper addresses the study of the fundamental tradeoff of MIMO systems when not only the receiver but also the transmitter has access to the channel matrix. First we show how the fundamental tradeoff is not altered by channel knowledge at the transmit side, as long as the duration of the encoding blocks satisfies $T \geq M + N - 1$. The paper then concentrates on the analysis of spatial multiplexing MIMO systems, a family of transceiver schemes which has been subject to thorough study in the literature owing to its simplicity and capacity achieving capabilities. These systems transmit independent symbols along the MIMO channel eigenmodes. The fundamental tradeoff of each of these substreams was determined and allowed the formulation of a fundamental tradeoff optimal rate allocation strategy for linearly combining different parallel channels as done by most of the spatial multiplexing schemes. The fact that channel knowledge at the transmitter does not increase the diversity versus multiplexing tradeoff of the MIMO channel was highly predictable. However, what was not known so far, is how much loss would simple spatial multiplexing schemes with CSI-T incur with respect to the fundamental limits of the channel and how much would they be able to benefit from channel knowledge at the transmitter. Precise answer to this two questions has been given in this paper comparing the fundamental tradeoff of spatial multiplexing schemes with the fundamental tradeoff of the channel, the tradeoff of space-only codes and the one of V-BLAST.

Appendix A
Exponent of the Individual Pairwise Error Probability

In this Appendix, we derive the exponent of the bound on the pairwise error probability of the kth substream given in (32)

$$\mathrm{PEP}^{(k)}(R_k) \dot{\leq} \Pr\left(\sqrt{\frac{p_k\lambda_k}{2}}\mathsf{SNR}^{-r_k/2} < 1 \,|\, p_k > 0\right) + \Pr(p_k = 0) \tag{81}$$

where $\Pr(p_k = 0)$ denotes the probability of not transmitting power through the kth channel eigenmode with the capacity-achieving waterfilling. Recall that p_k is given by (see (23))

$$p_k = \left(\mu - \lambda_k^{-1}\right)^+ \tag{82}$$

where the water level μ is chosen such that

$$\sum_{k=1}^{K} p_k = \sum_{i=1}^{K}\left(\mu - \lambda_i^{-1}\right)^+ = \mathsf{SNR}. \tag{83}$$

Hence, $p_k = 0$, if $(\mu - \lambda_k^{-1}) \leq 0$ for the water level μ calculated as in (83) with $K = k$, and $\Pr(p_k = 0)$ is then given by

$$\Pr(p_k = 0) = \Pr\left((k-1)\lambda_k^{-1} - \sum_{i=1}^{k}\lambda_i^{-1} \geq \mathsf{SNR}\right) \tag{84}$$

This probability can be upper-bounded as

$$\begin{aligned}\Pr(p_k = 0) &\leq \Pr\left(\lambda_k^{-1} \geq \frac{\mathsf{SNR}}{k-1}\right) \\ &\doteq \Pr(\lambda_k \leq \mathsf{SNR}^{-1}).\end{aligned} \tag{85}$$

Now, let us assume that $\tilde{K}$ substreams are transmitted with nonzero power, such that $k < \tilde{K} \leq K$. Then, the power allocated to the kth substream is

$$p_k = \frac{\mathsf{SNR}}{\tilde{K}} + \frac{1}{\tilde{K}}\sum_{i=1}^{\tilde{K}}\lambda_i^{-1} - \lambda_k^{-1} \tag{86}$$

and the first term in (81) satifies

$$\begin{aligned}\mathrm{PEP}^{(k)}(R_k \,|\, p_k > 0) &\dot{\leq} \Pr\left(\frac{p_k\lambda_k}{2}\mathsf{SNR}^{-r_k} < 1 \,\Big|\, p_k > 0\right) \\ &= \Pr\left(\left(\frac{\mathsf{SNR}}{\tilde{K}}\lambda_k + \frac{1}{\tilde{K}}\sum_{i=1}^{\tilde{K}}\frac{\lambda_k}{\lambda_i} - 1\right)\right. \\ &\qquad \left.\times\ \mathsf{SNR}^{-r_k} < 1\right) \\ &\leq \Pr\left(\frac{\mathsf{SNR}}{\tilde{K}}\lambda_k < \mathsf{SNR}^{r_k}\right) \\ &\doteq \Pr(\lambda_k \leq \mathsf{SNR}^{r_k - 1}).\end{aligned} \tag{87}$$

Finally, combining (85) and (87), it follows that

$$\mathrm{PEP}^{(k)}(R_k) \dot{\leq} \Pr(\lambda_k \leq \mathsf{SNR}^{r_k - 1}) \tag{88}$$

since $0 \leq r_k \leq 1$ and this completes the proof. □

Appendix B
Proof of Theorem 1

Theorem 1 presents the diversity and multiplexing tradeoff of the the kth channel eigenmode, assuming an uncorrelated flat-fading Rayleigh channel. In Section IV, we showed that

$$\mathrm{P}_e^{(k)}(R_k) \doteq \Pr(\lambda_k \leq \mathsf{SNR}^{r_k - 1}) \tag{89}$$

where λ_k is the kth ordered eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$. Thus, the diversity and multiplexing tradeoff can be obtained as

$$\begin{aligned}d_S^{(k)}(r) &= -\lim_{\mathsf{SNR}\to\infty}\frac{\log\Pr(\lambda_k \leq \mathsf{SNR}^{r_k - 1})}{\log\mathsf{SNR}} \\ &= \lim_{\lambda\to 0}\frac{\log\Pr(\lambda_k \leq \lambda^{1-r_k})}{\log\lambda}\end{aligned} \tag{90}$$

since $0 \leq r_k \leq 1$, and the proof of Theorem 1 reduces to obtaining the exponent[8] of the marginal pdf of the kth ordered eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$, whose distribution is given in the following lemma.

[8]Henceforth the exponent is defined as $\lambda_k \to 0$ and $g(\lambda_k) \doteq h(\lambda_k)$ denotes

$$\lim_{\lambda_k\to 0}\frac{g(\lambda_k)}{\log\lambda_k} = \lim_{\lambda_k\to 0}\frac{h(\lambda_k)}{\log\lambda_k}. \tag{91}$$

Lemma 3 ([25, eq. (95)]): Let the entries of the $N \times M$ matrix $\mathbf{H}$ be i.i.d. complex Gaussian with zero mean and unit variance and let $n = \min\{N, M\}$ and $m = \max\{N, M\}$. The joint pdf of the ordered strictly positive eigenvalues of the complex central Wishart distributed matrix $\mathbf{H}^\dagger \mathbf{H}$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, equals

$$f(\boldsymbol{\lambda}) \triangleq f(\lambda_1, \lambda_2, \ldots, \lambda_n) = K_{m,n}^{-1} \prod_{i=1}^{n} e^{-\lambda_i} \lambda_i^{m-n} \prod_{i<j}^{n} (\lambda_i - \lambda_j)^2 \quad (92)$$

where the normalizing constant $K_{m,n}$ is given by

$$K_{m,n} = \prod_{i=1}^{n} (n-i)!(m-i)!. \quad (93)$$

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

$$f_{\lambda_k}(\lambda_k) = \int_{\lambda_k}^{\infty} \int_{\lambda_k}^{\lambda_1} \cdots \int_{\lambda_k}^{\lambda_{k-2}} \int_{0}^{\lambda_k} \cdots \int_{0}^{\lambda_{n-1}} f(\boldsymbol{\lambda}) \, d\lambda_n \cdots d\lambda_{k+1} d\lambda_{k-1} \cdots d\lambda_2 d\lambda_1. \quad (94)$$

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 [21] the authors of [26] 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 4 ([26, Corollary 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:

$$\int_{\mathcal{D}_{\text{ord}}} |\mathbf{F}(\mathbf{x})||\mathbf{G}(\mathbf{x})| \prod_{i=1}^{n} h(x_i) d\mathbf{x} = |\mathbf{A}| \quad (95)$$

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

$$[\mathbf{A}]_{ij} = \int_a^b f_i(x) g_j(x) h(x) dx. \quad (96)$$

We define the Vandermonde matrix of order v in $\boldsymbol{\phi}$, $\mathbf{V}(\boldsymbol{\phi}) \triangleq \mathbf{V}(\phi_1, \phi_2, \ldots, \phi_v)$, as the $v \times v$ matrix with elements [27, eq. (6.1.32)]

$$[\mathbf{V}(\boldsymbol{\phi})]_{i,j} = \phi_i^{j-1} \quad i,j = 1, \ldots, v \quad (97)$$

and whose determinant satisfies [27, eq. (6.1.33)]

$$|\mathbf{V}(\boldsymbol{\phi})| = \prod_{i<j}^{v} (\phi_j - \phi_i). \quad (98)$$

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., $\{\lambda_i\}_{i=k+1,\ldots,n}$ and observe that[9]

$$\begin{aligned}
|\mathbf{V}(\boldsymbol{\lambda})|^2 &= \prod_{i<j}^{n} (\lambda_i - \lambda_j)^2 = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} (\lambda_i - \lambda_j)^2 \\
&= \left[\prod_{i=1}^{k-2} \prod_{j=i+1}^{k-1} (\lambda_i - \lambda_j)^2 \right] \\
&\quad \times \left[\prod_{i=1}^{k-2} \prod_{j=k}^{n} (\lambda_i - \lambda_j)^2 \right] \\
&\quad \times \left[\prod_{i=k+1}^{n-1} \prod_{j=i+1}^{n} (\lambda_i - \lambda_j)^2 \right] \\
&\quad \times \left[\prod_{i=k-1}^{k} \prod_{j=i+1}^{n} (\lambda_i - \lambda_j)^2 \right] \\
&= |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 |\mathbf{V}(\boldsymbol{\lambda}_k^-)|^2 \left[\prod_{i=1}^{k-1} (\lambda_i - \lambda_k)^2 \right] \\
&\quad \times \left[\prod_{i=1}^{k} \prod_{j=k+1}^{n} (\lambda_k - \lambda_j)^2 \right]
\end{aligned} \quad (99)$$

where $\mathbf{V}(\boldsymbol{\lambda})$, $\mathbf{V}(\boldsymbol{\lambda}_k^-)$, and $\mathbf{V}(\boldsymbol{\lambda}_k^+)$ are Vandermonde matrices of order n, $k-1$, and $n-k$ in variables $\boldsymbol{\lambda}$, $\boldsymbol{\lambda}_k^+$, and $\boldsymbol{\lambda}_k^-$, respectively. Then, the pdf in (92) can be expressed as

$$\begin{aligned}
f(\boldsymbol{\lambda}) &= K_{m,n}^{-1} e^{-\lambda_k} \lambda_k^{m-n} \left| \mathbf{V}(\boldsymbol{\lambda}_k^+) \right|^2 \\
&\quad \times \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2 \right] \\
&\quad \times |\mathbf{V}(\boldsymbol{\lambda}_k^-)|^2 \\
&\quad \times \left[\prod_{i=k+1}^{n} e^{-\lambda_i} \lambda_i^{m-n} \prod_{j=1}^{k} (\lambda_i - \lambda_j)^2 \right] \\
&= K(\boldsymbol{\lambda}_k^+, \lambda_k) |\mathbf{V}(\boldsymbol{\lambda}_k^-)|^2 \prod_{i=k+1}^{n} \xi_k^-(\lambda_i)
\end{aligned} \quad (100)$$

where

$$K(\boldsymbol{\lambda}_k^+, \lambda_k) = K_{m,n}^{-1} e^{-\lambda_k} \lambda_k^{m-n} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \times \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2 \right] \quad (101)$$

$$\xi_k^-(\lambda_i) \triangleq \xi_k^-(\lambda_i, \boldsymbol{\lambda}_k^+, \lambda_k) = e^{-\lambda_i} \lambda_i^{m-n} \prod_{j=1}^{k} (\lambda_i - \lambda_j)^2. \quad (102)$$

[9]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.

First, we integrate the joint pdf of the ordered eigenvalues over the domain of $\boldsymbol{\lambda}_k^-$, defined as $\mathcal{D}_k^- \triangleq \mathcal{D}_k^-(\boldsymbol{\lambda}_k^-, \lambda_k) = \{0 < \lambda_n \leq \cdots \leq \lambda_{k-1} \leq \lambda_k\}$, applying Lemma 4

$$f(\boldsymbol{\lambda}_k^+, \lambda_k) = K(\boldsymbol{\lambda}_k^+, \lambda_k) \int_{\mathcal{D}_k^-} |V(\boldsymbol{\lambda}_k^-)|^2 \times \prod_{i=k+1}^{n} \xi_k^-(\lambda_i) d\boldsymbol{\lambda}_k^- \quad (103)$$

$$= K(\boldsymbol{\lambda}_k^+, \lambda_k) |\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)| \quad (104)$$

where

$$[\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)]_{ij} = \int_0^{\lambda_k} e^{-x} x^{m-n+j+i-2} \prod_{s=1}^{k} (x - \lambda_s)^2 dx \quad i,j = 1, \dots, n-k. \quad (105)$$

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 $\mathcal{D}_k^+ \triangleq \mathcal{D}_k^+(\boldsymbol{\lambda}_k^+, \lambda_k) = \{\lambda_k \leq \lambda_{k-1} \leq \cdots \leq \lambda_1 < \infty\}$

$$f_{\lambda_k}(\lambda_k) = \int_{\mathcal{D}_k^+} K(\boldsymbol{\lambda}_k^-, \lambda_k) |\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)| d\boldsymbol{\lambda}_k^+$$

$$= K(\lambda_k) \int_{\mathcal{D}_k^+} |\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)| |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \times \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2\right] d\boldsymbol{\lambda}_k^+ \quad (106)$$

where

$$K(\lambda_k) = K_{m,n}^{-1} e^{-\lambda_k} \lambda_k^{m-n}. \quad (107)$$

Expanding the determinant of $\mathbf{U}(\boldsymbol{\lambda}_k^+, \lambda_k)$, we obtain that

$$f_{\lambda_k}(\lambda_k) = K(\lambda_k) \sum_{\boldsymbol{\sigma}} \mathrm{sgn}(\boldsymbol{\sigma}) \int_{\mathcal{D}_k^+} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \times \left[\prod_{i=1}^{n-k} \int_0^{\lambda_k} e^{-x} x^{m-n+i+\sigma_i-2} \times \prod_{j=1}^{k} (x - \lambda_j)^2 dx\right] \times \left[\prod_{i=1}^{k-1} e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2\right] d\boldsymbol{\lambda}_k^+ \quad (108)$$

where the summation over $\boldsymbol{\sigma} = [\sigma_1, \dots, \sigma_n]$ is for all permutations of the integers $1, \dots, n$ and $\mathrm{sgn}(\boldsymbol{\sigma})$ denotes the sign of the permutation. 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 $\mathcal{D}_{\mathbf{x}} = \{0 < x_s \leq \lambda_k\}_{s=1,\dots,n-k}$ as

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

$$= K(\lambda_k) \sum_{\boldsymbol{\sigma}} \mathrm{sgn}(\boldsymbol{\sigma}) \times \int_{\mathcal{D}_{\mathbf{x}}} \left[\prod_{s=1}^{n-k} e^{-x_s} x_s^{m-n+s+\sigma_s-2} (x_s - \lambda_k)^2\right] \times \left[\int_{\mathcal{D}_k^+} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \prod_{i=1}^{k-1} \xi_k^+(\lambda_i) d\boldsymbol{\lambda}_k^+\right] d\mathbf{x} \quad (109)$$

where

$$\xi_k^+(\lambda_i) \triangleq \xi_k^+(\lambda_i, \lambda_k, \mathbf{x}) = e^{-\lambda_i} \lambda_i^{m-n} (\lambda_i - \lambda_k)^2 \prod_{s=1}^{n-k} (\lambda_i - x_s)^2. \quad (110)$$

The integral over $\boldsymbol{\lambda}_k^+$ can be solved applying again Lemma 4

$$\int_{\mathcal{D}_k^+} |\mathbf{V}(\boldsymbol{\lambda}_k^+)|^2 \prod_{i=1}^{k-1} \xi_k^+(\lambda_i) d\boldsymbol{\lambda}_k^+ = |\mathbf{W}(\lambda_k, \mathbf{x})| \quad (111)$$

where

$$[\mathbf{W}(\lambda_k, \mathbf{x})]_{ij} = \int_{\lambda_k}^{\infty} e^{-y} y^{m-n+i+j-2} (y - \lambda_k)^2 \times \prod_{s=1}^{n-k} (y - x_s)^2 dy \quad i,j = 1, \dots, k-1. \quad (112)$$

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

$$\prod_{s=1}^{n-k} (y - x_s)^2 = \sum_{v=0}^{2(n-k)} C_v(\mathbf{x}) y^v \quad (113)$$

where $C_v(\mathbf{x})$ are the unique polynomials in $\mathbf{x}$ such that (113) is satisfied, we have that

$$[\mathbf{W}(\lambda_k, \mathbf{x})]_{ij} = \sum_{v=0}^{2(n-k)} C_v(\mathbf{x}) \times \int_{\lambda_k}^{\infty} e^{-y} y^{b(i+j+v-2)} (y - \lambda_k)^2 dy \quad (114)$$

where $b(i) = m - n + i$. The integral in (114) has the same structure as the upper incomplete gamma function, defined as [28, eq. (6.5.3)]

$$\Gamma(a, \lambda) = \int_{\lambda}^{\infty} e^{-t} t^{a-1} dt. \quad (115)$$

Since we are only interested in the exponent of $f_{\lambda_k}(\lambda_k)$ in λ_k, noting that, for $a \in \mathbb{N}$

$$\Gamma(a, \lambda) = (a-1)! e^{-\lambda} \sum_{i=0}^{a-1} \frac{\lambda^i}{i!} \doteq \lambda^0 \quad (116)$$

we can expand $(y - \lambda_k)^2$ in (114) and neglect the terms with λ_k and λ_k^2. Hence, it follows that

$$[\mathbf{W}(\lambda_k, \mathbf{x})]_{ij} \doteq \sum_{v=0}^{2(n-k)} C_v(\mathbf{x}) \quad (117)$$

and, therefore,

$$|\mathbf{W}(\lambda_k, \mathbf{x})| \doteq \sum_{v=0}^{2(n-k)} C_v(\mathbf{x}). \tag{118}$$

Then, the marginal pdf of λ_k satisfies

$$\begin{aligned} f_{\lambda_k}(\lambda_k) \doteq K(\lambda_k) \sum_{\boldsymbol{\sigma}} \operatorname{sgn}(\boldsymbol{\sigma}) \\ \times \int_{\mathcal{D}_{\mathbf{x}}} \left[\prod_{s=1}^{n-k} e^{-x_s} x_s^{b(s+\sigma_s-2)} \right. \\ \left. \times\ (x_s - \lambda_k)^2 \sum_{v=0}^{2(n-k)} C_v(\mathbf{x}) \right]. \end{aligned} \tag{119}$$

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 [28, eq. (6.5.2)]

$$\gamma(a, \lambda) = \int_0^{\lambda} e^{-t} t^{a-1} dt. \tag{120}$$

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

$$\begin{aligned} \gamma(a, \lambda) &= (a-1)! \left(1 - e^{-\lambda} \sum_{i=0}^{a-1} \frac{\lambda^i}{i!} \right) \\ &= (a-1)! \left(1 - e^{-\lambda} \left(e^{\lambda} - \sum_{i=a}^{\infty} \frac{\lambda^i}{i!} \right) \right) \\ &\doteq \lambda^a \end{aligned} \tag{121}$$

the term with the lowest λ_k exponent of $f_{\lambda_k}(\lambda_k)$ in (119) 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 (118). Thus

$$\begin{aligned} f_{\lambda_k}(\lambda_k) \doteq K(\lambda_k) \sum_{\boldsymbol{\sigma}} \operatorname{sgn}(\boldsymbol{\sigma}) \\ \times \left[\int_{\mathcal{D}_{\mathbf{x}}} \prod_{s=1}^{n-k} e^{-x_s} x_s^{b(s+\sigma_s-2)} (x_s - \lambda_k)^2 d\mathbf{x} \right]. \end{aligned} \tag{122}$$

and rewriting the multiple integral over $\mathbf{x}$ as a product of $n-k$ single integrals, it follows that

$$\begin{aligned} f_{\lambda_k}(\lambda_k) &= K(\lambda_k) \sum_{\boldsymbol{\sigma}} \operatorname{sgn}(\boldsymbol{\sigma}) \\ &\quad \times \left[\prod_{s=1}^{n-k} \int_0^{\lambda_k} e^{-x} x^{b(s+\sigma_s-2)} (x - \lambda_k)^2 dx \right] \\ &\doteq e^{-\lambda_k} \lambda_k^{m-n} \sum_{\boldsymbol{\sigma}} \operatorname{sgn}(\boldsymbol{\sigma}) \\ &\quad \times \left[\prod_{s=1}^{n-k} \int_0^{\lambda_k} e^{-x} x^{m-n+s+\sigma_s} dx \right] \\ &\doteq \lambda_k^{(m-n)+(n-k)(m-n+1)+\sum_{s=1}^{n-k}(\sigma_s+s)} \end{aligned} \tag{123}$$

where we have used that $e^{-\lambda_k} \doteq \lambda_k^0$ and the equivalence in (121). Then, since it holds that [29, eq. (0.121.1)]

$$\sum_{s=1}^{n-k} (\sigma_s + s) = 2 \sum_{s=1}^{n-k} s = (n-k)(n-k+1) \tag{124}$$

the exponent of the marginal pdf of λ_k is given by

$$f_{\lambda_k}(\lambda_k) \doteq \lambda_k^{(n-k)(m-k+1)+(m-k+1)}. \tag{125}$$

Finally, the marginal distribution of the kth eigenvalue is obtained from its marginal pdf as

$$\Pr(\lambda_k \le \lambda) = \int_0^{\lambda} f_{\lambda_k}(x) dx \tag{126}$$

and consequently the exponent of $\Pr(\lambda_k \le \lambda)$ is given by

$$\begin{aligned} \Pr(\lambda_k \le \lambda) &\doteq \int_0^{\lambda} \lambda_k^{(n-k)(m-k+1)+(m-k)} d\lambda_k \\ &\doteq \lambda^{(n-k+1)(m-k+1)}. \end{aligned} \tag{127}$$

Now, returning to the equivalence presented in (90), we have that

$$\mathrm{P}_e^{(k)}(R_k) \doteq \mathsf{SNR}^{-(M-k+1)(N-k+1)(1-r_k)} \tag{128}$$

which completes the proof of the theorem. □

APPENDIX C
PROOF OF COROLLARY 1

In this proof, we derive the exponent of the error probability of spatial multiplexing MIMO system with a general power allocation satisfying the condition in (38). Since this proof is strongly based on the procedure used for the capacity-achieving waterfilling power allocation in Section IV-A, some repetitive parts are omitted.

The individual error probability of the kth substream, $\mathrm{P}_e^{(k)}(R_k)$ with the general power allocation described in Corollary 1 can be lower-bounded by the outage probability obtained when allocating all available power to the kth substream

$$\mathrm{P}_e^{(k)}(R_k) \ge \mathrm{P}_{\text{out}}^{(k)}(R_k) \dot{\ge} \Pr(\lambda_k \le \mathsf{SNR}^{r_k - 1}) \tag{129}$$

since $p_k \le \mathsf{SNR}$ due to the short-term power constraint. In addition, $\mathrm{P}_e^{(k)}(R_k)$ can be upper-bounded as (see Section IV-A)

$$\begin{aligned} \mathrm{PEP}^{(k)}(R_k) &\le \Pr\left(\sqrt{\frac{p_k \lambda_k}{2}} \mathsf{SNR}^{-r_k/2} < 1 \right) \\ &= \Pr\left(\frac{p_k \lambda_k}{2} \mathsf{SNR}^{-r_k} < 1 \middle| p_k > \phi_k \mathsf{SNR} \right) \\ &\quad \times (1 - \Pr(p_k \le \phi_k \mathsf{SNR})) \\ &\quad + \Pr\left(\frac{p_k \lambda_k}{2} \mathsf{SNR}^{-r_k} < 1 \middle| p_k \le \phi_k \mathsf{SNR} \right) \\ &\quad \times \Pr(p_k \le \phi_k \mathsf{SNR}) \\ &\le \Pr\left(\frac{p_k \lambda_k}{2} \mathsf{SNR}^{-r_k} < 1 \middle| p_k = \phi_k \mathsf{SNR} \right) \\ &\quad + \Pr(p_k \le \phi_k \mathsf{SNR}) \end{aligned}$$

$$\leq \Pr(\phi_k \lambda_k \mathsf{SNR}^{-r_k+1} < 1) + \Pr(p_k \leq \phi_k \mathsf{SNR}) \dot{\leq} \Pr(\lambda_k \leq \mathsf{SNR}^{r_k-1}) \quad (130)$$

where we have used that the power allocation satisfies the condition in (38)

$$\Pr(p_k \leq \phi_k \mathsf{SNR}) \dot{\leq} \mathrm{P}_{\mathrm{out}}^{(k)}(R_k) \dot{\leq} \Pr(\lambda_k \leq \mathsf{SNR}^{r_k-1}). \quad (131)$$

Finally, combining (129) and (130), it follows that

$$\mathrm{P}_e^{(k)}(R_k) \doteq \Pr(\lambda_k \leq \mathsf{SNR}^{r_k-1}) \quad (132)$$

which coincides with the exponent derived for the capacity-achieving spatial multiplexing MIMO system in Theorem 1 and this completes the proof. □

Appendix D
Proof of Proposition 1

We want to prove that power allocation in (44) satisfies the condition in (38). Since the exponent of the error probability is given by (see Corollary 1)

$$d_S^{(k)}(r_k) = d_k(1 - r_k) \quad (133)$$

where $d_k = (M-k+1)(N-k+1)$ and $0 \leq r_k \leq 1$, we have to show that

$$\delta_k = \lim_{\mathsf{SNR}\to\infty} -\frac{\log \Pr(p_{\mathrm{mse},k} \leq \phi_k \mathsf{SNR})}{\log \mathsf{SNR}} > d_k \quad (134)$$

where ϕ_k is a strictly positive constant. The individual power p_k is maximum when the weakest $K-k$ substreams are discarded, i.e., $\{p_k\}_{i=k+1,\dots,K} = 0$, and, thus, we can assume without loss of generality that only $k \leq K$ substreams are transmitted with nonzero power. Then, $\Pr(p_k \leq \phi_k \mathsf{snr})$ is given by

$$\begin{aligned}\Pr(p_k \leq \phi_k \mathsf{SNR}) &= \Pr\left(\frac{(w_k/\lambda_k)^{1/2}}{\sum_{i=1}^k (w_i/\lambda_i)^{1/2}} \times \left(\mathsf{SNR} + \sum_{i=1}^k \lambda_i^{-1}\right) - \lambda_k^{-1} \leq \phi_k \mathsf{SNR}\right)\\ &= \Pr\left(\left((w_k\lambda_k)^{-1/2} \sum_{i=1}^k (w_i/\lambda_i)^{1/2} - \sum_{i=1}^k \lambda_i^{-1}\right) \geq \mathsf{SNR}\left(1 - \phi_k (w_k/\lambda_k)^{-1/2} \sum_{i=1}^k (w_i/\lambda_i)^{1/2}\right)\right).\end{aligned} \quad (135)$$

Observing that

$$\phi_k (w_k/\lambda_k)^{-1/2} \sum_{i=1}^k (w_i/\lambda_i)^{1/2} \leq \phi_k \sum_{i=1}^k (w_i/w_k)^{1/2} \quad (136)$$

since $\lambda_k/\lambda_i \leq 1$ for $i = 1, \dots, k$, $\Pr(p_k \leq \phi_k \mathsf{SNR})$ can be upper-bounded as

$$\begin{aligned}\Pr(p_k \leq \phi_k \mathsf{SNR}) &\leq \Pr\left(\lambda_k^{-1/2} \sum_{i=1}^{k-1} (w_i/\lambda_i)^{1/2} \geq \mathsf{SNR}\left(1 - k\phi_k \sum_{i=1}^k (w_i/w_k)^{1/2}\right)\right)\\ &\leq \Pr\left(\lambda_k \lambda_{k-1} \leq \left(\frac{k-1}{1 - k\phi_k \sum_{i=1}^k (w_i/w_k)^{1/2}}\right)^2 \mathsf{SNR}^{-2}\right)\end{aligned} \quad (137)$$

whenever $\phi_k < 1/(k\sum_{i=1}^k (w_i/w_k)^{1/2})$. Hence, it follows that

$$\delta_k \geq \lim_{\mathsf{snr}\to\infty} \frac{\log \Pr(\lambda_K \lambda_{K-1} \leq \mathsf{snr}^{-2})}{\log \mathsf{snr}^{-1}}. \quad (138)$$

Then, by defining $\eta_k = \lambda_{k-1}\lambda_k$, we can rewrite (138) as

$$\begin{aligned}d_{\mathrm{mse},k} &\geq \lim_{x\to 0} \frac{\log \Pr\left(\eta_k \leq x^2\right)}{\log x}\\ &= \lim_{x\to 0} \frac{\log(\int_0^{x^2} f_{\eta_k}(\eta) d\eta)}{\log x}\end{aligned} \quad (139)$$

where $f_{\eta_k}(\eta)$ is the pdf of a product of two random variables and is given by [30, Sec. 4.4, Th. 7]

$$f_{\eta_k}(\eta) = \int_0^\infty \frac{1}{x} f_{\lambda_{k-1},\lambda_k}(x, \eta/x) dx. \quad (140)$$

We are interested in $f_{\eta_k}(\eta)$ as $\eta \to 0$ and, thus, 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 B), it can be shown that[10]

$$f_{\lambda_{k-1},\lambda_k}(\lambda_{k-1}, \lambda_k) = g(\lambda_{k-1}) \lambda_k^{d_k-1} + o\left(\lambda_k^{d_k-1}\right) \quad (141)$$

where $g(\lambda_{k-1})$ is a function of λ_{k-1}. Then, substituting back this result in the expression of $f_{\eta_k}(\eta)$ in (140), it follows that

$$f_{\eta_k}(\eta) = a\eta^{d_k-1} + o(\eta^{d_k-1}) \quad (142)$$

where a is a fixed constant in terms of η. Finally, the exponent of $\Pr(p_{\mathrm{mse},k} \leq \phi_k \mathsf{SNR})$ can be bounded as

$$\begin{aligned}\delta_k &\geq \lim_{x\to 0} \frac{\log\left(a\int_0^{x^2} \eta^{d_k-1} d\eta\right)}{\log x}\\ &= \lim_{\eta\to 0} \frac{\log(\eta^{2d_k})}{\log \eta}\\ &= 2d_k > d_k\end{aligned} \quad (143)$$

and this completes the proof. □

[10]We say that $f(x) = o(g(x))$ if $f(x)/g(x) \to 0$ as $x \to 0$.

REFERENCES

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Syst. 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," *Europ. 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 Pers. 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. 5, 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, 2002, pp. 90.1–90.14.

[8] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[9] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[10] G. Foschini, G. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.

[11] 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.

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

[13] 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.

[14] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[15] 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.

[16] 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.

[17] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[18] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

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

[20] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 1273–1289, May 2001.

[21] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. New York: Now, 2004.

[22] R. B. Gallager, *Information Theory and Reliable Communication*. New York: Wiley , 1968.

[23] L. G. Ordóñez, D. P. Palomar, A. Pagès-Zamora, and J. R. Fonollosa, "High SNR analytical performance of spatial multiplexing MIMO systems with CSI," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5447–5463, Nov. 2001.

[24] Y. Jiang, X. Zheng, and J. Li, "Asymptotic performance analysis of V-BLAST," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, 2005.

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

[26] 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, Oct. 2003.

[27] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

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

[29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Orlando, FL: Academic, 1980.

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