Output Statistics of MIMO Channels with General Input Distribution

Original
Output Statistics of MIMO Channels with General Input Distribution / Alfano G.; Chiasserini C.F.; Nordio A.; Zhou S.. - STAMPA. - (2013), pp. 829-833. ((Intervento presentato al convegno IEEE International Symposium on Information Theory (ISIT 2013) tenutosi a Istanbul (Turkey) nel July 2013.

Availability:
This version is available at: 11583/2513885 since: 2016-11-19T14:34:57Z

Publisher:
IEEE

Published
DOI:10.1109/ISIT.2013.6620342

Output Statistics of MIMO Channels with General Input Distribution

G. Alfano
DET
Politecnico di Torino
Torino, Italy, 10129
alfano@tlc.polito.it

C-F. Chiasserini
DET
Politecnico di Torino
Torino, Italy, 10122
chiasserini@polito.it

A. Nordio
CNR IEIIT
Torino, Italy, 10129
alessandro.nordio@ieiit.cnr.it

S. Zhou
DET
Politecnico di Torino
Torino, Italy, 10129
siyuan.zhou@polito.it

***Abstract*—The information that can be conveyed through a wireless channel, with multiple-antenna equipped transmitter and receiver, crucially depends on the channel behavior as well as on the input structure. In this paper, we derive analytical results, concerning the probability density function (pdf) of the output of a single-user, multiple-antenna communication. The analysis is carried out under the assumption of an optimized input structure, and assuming Gaussian noise and a Rayleigh block-fading channel. Our analysis therefore provides a quite general and compact expression for the conditional output pdf. We also highlight the relation between such an expression and the results already available in the literature for some specific input structures.**

I. INTRODUCTION

The availability of an explicit statistical characterization for the output of a wireless channel impaired by additive and multiplicative random disturbance is of paramount importance to information-theoretic purposes. Specifically, such characterization is required for the evaluation of the mutual information between the input and the output signals of the channel. The output probability density function (pdf) is particularly relevant as the communication takes place in absence of perfect Channel State Information (CSI) at both ends of the link (see the seminal paper thereabout [1] and its relevant consequences in [2]). Indeed, this case is of particular interest as the availability of CSI would imply a high energy and time consumption at both the transmitter and the receiver.

In this paper, we therefore refer to the above scenario and consider a single-user, multiple-antenna channel, affected by AWGN and block Rayleigh fading. In our scenario, a MIMO channel can be adequately modeled as a matrix of jointly Gaussian entries. This yields a Gaussian behavior of the output signal, conditionally on the input structure. Nowadays, several results are available in the literature about finite-dimensional random matrices with jointly Gaussian entries, and the most popular decompositions of such matrices (e.g., Bartlett, Singular Value, and Cholesky [3]) are completely characterized from a statistical point of view [4]. In such characterizations, we often encounter special functions of matrix arguments, which are largely adopted in theoretical physics for the study of particles behavior. Based on some results on the integration of these special functions over complex unitary and positive definite matrices, we present hereafter a new closed form expression for the output pdf of the single-user MIMO channel in Rayleigh block fading channels, conditionally on the input power distribution and on the channel eigenvalues. We also highlight the relation between our result and the ones that are available in the literature for some special cases, such as the i.i.d. Gaussian or isotropically random input structures [5], [6], [2].

The paper is organized as follows. Section II presents the notations used throughout the paper and provides relevant mathematical background on matrix-variate distributions and their eigenvalue law. Section III introduces the model of the wireless system. Then, Section IV presents the derivation of our new expression of the channel output pdf, as well as the particularization of such an expression to some relevant input structure cases. Directions for future work are discussed in Section V. The integration rules we use in our analysis are reported in the Appendix.

II. MATHEMATICAL BACKGROUND

A. Notations

Throughout the paper, matrices are denoted by uppercase boldface letters, vectors by lowercase boldface. The pdf of a random matrix $\mathbf{Z}$, $p_{\mathbf{Z}}(\hat{\mathbf{Z}})$, is simply denoted by $p(\mathbf{Z})$. $(\cdot)^\dagger$ indicates the conjugate transpose operator, $|\cdot|$ and $\mathrm{Tr}(\cdot)$ denote, respectively, the determinant and the trace of a square matrix, and $||\cdot||$ stands for the Euclidean norm[1]. $\Gamma_p(q)$, with $p\leq q$, is the complex multivariate Gamma function [7]

$$\Gamma_p(q) = \pi^{\frac{p(p-1)}{2}} \prod_{\ell=1}^{p} (q-\ell)!$$

and

$${}_pF_q(a_1, \dots, a_p; b_1 \dots, b_q; \cdot, \dots, \cdot)\,,$$

with p and q non-negative integers, denotes the generalized hypergeometric function [8]. The arguments of such a function can be either scalars or square matrices; there is in general no limit to the number of arguments, and hypergeometric functions of multiple matrix arguments are defined also for set of square matrices of different size. We denote by $\mathbf{I}_m$ the $m \times m$ identity matrix. Then, let $\mathbf{A}$ be an $n \times n$ Hermitian

[1]As applied to a matrix, we mean $||\mathbf{A}||^2 = \mathrm{Tr}(\mathbf{A}^\dagger \mathbf{A})$

matrix with ordered eigenvalues [2] $\lambda_1, \ldots, \lambda_n$. We denote by $\mathcal{V}(\mathbf{A})$ the Vandermonde determinant of $\mathbf{A}$ [3], i.e.,

$$\mathcal{V}(\mathbf{A}) = \prod_{1 \le i < \ell \le n} (\lambda_i - \lambda_\ell).$$

The differential of $n \times m$ complex matrix variables are always defined, unless otherwise stated, as $d\mathbf{A} = \pi^{-nm} \prod_{i,j=1}^{n,m} d\mathrm{Re} A_{i,j} d\mathrm{Im} A_{i,j}$.

B. Matrix-variate distributions

Definition 1. *Let $\mathbf{H}$ be an $m \times n$ matrix whose columns are zero-mean independent complex Gaussian vectors with covariance matrix $\mathbf{\Theta}_m$. Then, the $m \times m$ random matrix $\mathbf{W} = \mathbf{H}\mathbf{H}^\dagger$ is a (central) complex Wishart matrix, with n degrees of freedom and covariance matrix $\mathbf{\Theta}_m$ ($\mathbf{W} \sim \mathcal{W}_m(n, \mathbf{\Theta}_m)$). For $n \ge m$, the pdf of $\mathbf{W}$ is [7]*

$$p(\mathbf{W}) = \frac{|\mathbf{W}|^{n-m}}{\Gamma_m(n)|\mathbf{\Theta}_m|^n} \exp\left\{-\mathrm{Tr}\left({\mathbf{\Theta}_m}^{-1}\mathbf{W}\right)\right\}.$$

The joint distribution of the ordered eigenvalues of $\mathbf{W}$ coincides with the law of the squared non-zero singular values of $\mathbf{H}$. Specifically, denoting by $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}^{1/2}\mathbf{V}^\dagger$ the Singular Value Decomposition of $\mathbf{H}$, it turns out that $\mathbf{W} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^\dagger$. Notice that $\mathbf{\Sigma}$ is positive definite and, by $\mathbf{\Sigma}^{1/2}$ we mean the $m \times n$ matrix with the m non-zero singular values of $\mathbf{H}$ on its main diagonal. If $\mathbf{\Theta}_m = \mathbf{I}_m$, then $\mathbf{U}$ and $\mathbf{V}$ are *isotropic*[3] matrices independent of $\mathbf{\Sigma}$, and the joint eigenvalue distribution can be written as

$$p(\mathbf{\Sigma}) = \frac{\prod_{k=1}^{m} \sigma^{n-m} e^{-\sigma}}{\prod_{\ell=1}^{s}(n-\ell)!(m-\ell)!} \mathcal{V}^2(\mathbf{\Sigma}).$$

Definition 2. *Let $\mathbf{H}$ be an $m \times n$ matrix whose rows are zero-mean independent complex Gaussian vectors with covariance matrix $\mathbf{\Theta}_m$. Then, the $m \times m$ random matrix $\widetilde{\mathbf{W}} = \mathbf{H}\mathbf{H}^\dagger$ is a (central)* complex pseudo-Wishart *matrix, with n degrees of freedom and covariance matrix $\mathbf{\Theta}_n$, ($\widetilde{\mathbf{W}} \sim \widetilde{\mathcal{W}}_m(n, \mathbf{\Theta}_n)$). For $n \ge m$, the pdf of $\widetilde{\mathbf{W}}$ is [9]*

$$p(\widetilde{\mathbf{W}}) = \frac{|\widetilde{\mathbf{W}}|^{n-m} {}_0F_0(\mathbf{\Theta}_n^{-1}, -\widetilde{\mathbf{W}})}{\prod_{i=1}^{m}(n-i)!|\mathbf{\Theta}_n|^m}$$

where ${}_0F_0(\cdot,\cdot)$ is the hypergeometric function of exponential type of two square matrix arguments of different size [9].

Definition 3. *The $m \times m$ random matrix $\mathbf{B}$ is Beta-distributed with positive integer parameters p and n, ($\mathbf{B} \sim \mathrm{Beta}_m(p,n)$) if it can be written as $\mathbf{B} = (\mathbf{T}^\dagger)^{-1}\mathbf{C}\mathbf{T}$ where $\mathbf{C} \sim \mathcal{W}_m(p, \mathbf{\Theta}_m)$, and, given $\mathbf{A} \sim \mathcal{W}_m(n, \mathbf{\Theta}_m)$, $\mathbf{A} + \mathbf{C} = \mathbf{T}^\dagger\mathbf{T}$, with $\mathbf{T}$ upper triangular with positive diagonal elements. Notice that, if either or both $p < m$ and $n < m$, the distribution is referred to as pseudo-Beta since it involves pseudo-Wishart matrices [2, and references therein].*

[2]In the following, all the joint eigenvalue distributions assume the eigenvalues to be ordered.

[3]In the multi-antenna literature, a matrix is often referred to as isotropic if it belongs to the set of Haar matrices, i.e., unitary matrices uniformly distributed on their group.

A Beta-distributed matrix admits, like the Wishart, an eigendecomposition where the matrix of the eigenvectors is independent of the matrix of the eigenvalues. The joint eigenvalues distribution of $\mathbf{B} \sim \mathrm{Beta}_m(p,n)$ can be written as

$$p(\mathbf{\Lambda}) = \frac{\pi^{m(m-1)}\Gamma_m(p+n)}{\Gamma_m(m)\Gamma_m(p)\Gamma_m(n)} \cdot |\mathbf{I} - \mathbf{\Lambda}|^{n-m}|\mathbf{\Lambda}|^{p-m}\mathcal{V}^2(\mathbf{\Lambda}), \tag{1}$$

for $n \ge m$, or, rather, as

$$p(\mathbf{\Lambda}_{m-n+1}) = \frac{\pi^{n(m-1)}\Gamma_n(p+n)}{\Gamma_n(m)\Gamma_n(p+n-m)\Gamma_n(n)} \cdot |\mathbf{I} - \mathbf{\Lambda}_{m-n+1}|^{m-n}|\mathbf{\Lambda}_{m-n+1}|^{p-m}\mathcal{V}^2(\mathbf{\Lambda}_{m-n+1}),$$

if $m > n$. Notice that, for $n \ge m$, the matrix has m nonzero eigenvalues, whose joint law is given by (1). Instead, for $m > n$, the first $m-n$ eigenvalues are equal to 1 with probability (w.p.) 1, and thus $\mathbf{\Lambda}_{m-n+1} = \{\lambda_{m-n+1}, \ldots, \lambda_m\}$.

III. System Model

We consider a single-user multiple-antenna communication, with n_R and n_T denoting the number of receive and transmit antennas. Assuming block-fading Rayleigh with block length n_b, the channel output can be described by the following linear relationship:

$$\mathbf{Y} = \sqrt{\gamma}\mathbf{H}\,\mathbf{X} + \mathbf{N}. \tag{2}$$

In (2), $\mathbf{Y}$ is the $n_R \times n_b$ output, $\mathbf{X}$ is the complex $n_T \times n_b$ input matrix, and $\mathbf{N}$ is the $n_R \times n_b$ matrix of additive complex circularly symmetric Gaussian noise. $\mathbf{H}$ is the $n_R \times n_T$ complex channel matrix, whose entries represent the fading coefficients between each transmit and each receive antenna. Finally, $\gamma = \mathrm{SNR}/n_T$ represents the normalized per-transmit antenna Signal-to-Noise Ratio (SNR).

For ease of notation, let us denote

$$s = \min\{n_T, n_R\}, \quad r = \max\{n_T, n_R\}, \quad \tau = r - s,$$
$$m = \min\{n_T, n_b\}, \quad n = \max\{n_T, n_b\}, \quad \delta = n - m$$
$$p = \min\{m, s\}, \qquad q = \max\{n, r\}.$$

The input matrix $\mathbf{X}$, unless otherwise stated, is assumed to have a product structure, i.e., $\mathbf{X} = \mathbf{D}^{1/2}\mathbf{\Phi}$, where $\mathbf{D}$ is a random, n_T-dimensional, diagonal matrix, which is positive definite w.p. 1. The entries of $\mathbf{D}$ represent the amount of transmit power allocated to each of the n_T transmit antennas, while $\mathbf{\Phi}$ is an $n_T \times n_b$ isotropic matrix. As usually done in the literature [10], [6], we will refer to square isotropic matrices as Haar and to rectangular isotropic matrices as Stiefel. We stress that the above structure of the input matrix $\mathbf{X}$ allows to achieve the capacity limit in absence of CSI at both the link ends [6, Thm2].

IV. Statistical characterization of the channel output

We now present our main result, i.e., the expression of the output pdf, conditioned on the input power allocation and on the channel.

Theorem 1. *Given a channel as in (2), the pdf of its matrix-variate output, conditionally on the n_T-dimensional (diagonal) input power allocation matrix $\mathbf{D}$ and on the s-dimensional matrix of the non-zero squared singular values of the channel, $\mathbf{\Sigma}$, can be expressed as*

$$p(\mathbf{Y}|\mathbf{D},\mathbf{\Sigma}) = \frac{e^{-||\mathbf{Y}||^2}}{\widetilde{v}\pi^{n_R n_b}} {}_0F_2(n_T, n_R; -\mathbf{\Sigma}^2, \gamma^2\mathbf{D}^2, \mathbf{Y}^\dagger\mathbf{Y}) \,. \tag{3}$$

In (3), $\widetilde{v} = 2^\delta \pi^{\delta^2}/\Gamma_\delta(\delta)$ is the volume of the unitary group of dimension δ, and the hypergeometric function of three matrix arguments can be expanded in zonal polynomials [7] $\mathcal{C}_\kappa(\cdot)$, following [11, Appendix B] as

$$\begin{aligned} &{}_0F_2(n_T, n_R; -\mathbf{\Sigma}^2, \gamma^2\mathbf{D}^2, \mathbf{Y}^\dagger\mathbf{Y}) = \\ &(2\pi)^{2q-m-n_T} \sum_{k=0}^{+\infty} \sum_{\#\kappa=k} (n_T)_\kappa (n_R)_\kappa \\ &\frac{\mathcal{C}_\kappa(-\mathbf{\Sigma}^2)\mathcal{C}_\kappa(\gamma^2\mathbf{D}^2)\mathcal{C}_\kappa(\mathbf{Y}^\dagger\mathbf{Y})}{k!\mathcal{C}_\kappa^2(\mathbf{I}_q)} \,, \end{aligned} \tag{4}$$

with $(\cdot)_\kappa$ being the multivariate Pochhammer symbol of order κ [7], and $\#\kappa$ the cardinality of κ.

Proof. The output pdf of the MIMO channel (2) can be expressed as follows:

$$p(\mathbf{Y}) = \int p(\mathbf{Y}|\mathbf{X},\mathbf{H}) p(\mathbf{X}) p(\mathbf{H}) d\mathbf{X} d\mathbf{H} \,, \tag{5}$$

with the integral being over the appropriate matrix spaces, which will be specified step by step throughout the proof. The Gaussianity of both channel and noise leads to

$$p(\mathbf{Y}|\mathbf{X},\mathbf{H}) = \frac{e^{-\mathrm{Tr}\left(\left(\mathbf{Y}-\sqrt{\gamma}\mathbf{H}\mathbf{X}\right)\left(\mathbf{Y}-\sqrt{\gamma}\mathbf{H}\mathbf{X}\right)^\dagger\right)}}{\pi^{n_R n_b}} \,.$$

Notice that, expanding the product in the exponent and decomposing in its singular values/vectors the channel matrix $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}^{1/2}\mathbf{V}^\dagger$, one obtains, term by term, $e^{-||\mathbf{Y}||^2}$, which is independent of $\mathbf{H}$ and $\mathbf{X}$ and, thus, it can be factored out of the integral (5);

$$\exp\left\{-\gamma \mathrm{Tr}\left(\mathbf{H}\mathbf{X}\mathbf{X}^\dagger\mathbf{H}^\dagger\right)\right\} = \exp\left\{-\gamma \mathrm{Tr}\left(\mathbf{\Sigma}\mathbf{V}^\dagger\mathbf{D}\mathbf{V}\right)\right\} \,,$$

which depends only on $\mathbf{V}$, and finally

$$\exp\{\sqrt{\gamma}\mathrm{Tr}\left(\mathbf{Y}\mathbf{X}^\dagger\mathbf{H}^\dagger + \mathbf{H}\mathbf{X}\mathbf{Y}^\dagger\right)\} \,,$$

which is the only term dependent on $\mathbf{U}$. Hence, we can write

$$\begin{aligned} p(\mathbf{Y}) = &\frac{\exp\{-||\mathbf{Y}||^2\}}{\pi^{n_R n_b}} \int p(\mathbf{D}) d\mathbf{D} p(\mathbf{\Sigma}) d\mathbf{\Sigma} \int_{\mathbf{\Phi}} d\mathbf{\Phi} \\ &\int_{\mathcal{U}(n_T)} \exp\left\{-\gamma \mathrm{Tr}\left(\mathbf{\Sigma}\mathbf{V}^\dagger\mathbf{D}\mathbf{V}\right)\right\} d_*\mathbf{V} \\ &\int_{\mathcal{U}(n_R)} e^{\sqrt{\gamma}\mathrm{Tr}(\mathbf{Y}\mathbf{\Phi}^\dagger\mathbf{D}^{\frac{1}{2}}\mathbf{V}\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{U}^\dagger + \mathbf{U}\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{V}^\dagger\mathbf{D}^{\frac{1}{2}}\mathbf{\Phi}\mathbf{Y}^\dagger)} d_*\mathbf{U} \end{aligned}$$

where $d_*\mathbf{V}$ and $d_*\mathbf{U}$ are normalized measures[4]. The inner integral over the unitary group can be expressed in terms of the Bessel hypergeometric functions of matrix argument, by virtue of [7, Formula (91)],

$$\begin{aligned} &\int_{\mathcal{U}(n_R)} e^{\sqrt{\gamma}\mathrm{Tr}(\mathbf{Y}\mathbf{\Phi}^\dagger\mathbf{D}^{\frac{1}{2}}\mathbf{V}\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{U}^\dagger + \mathbf{U}\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{V}^\dagger\mathbf{D}^{\frac{1}{2}}\mathbf{\Phi}\mathbf{Y}^\dagger)} d_*\mathbf{U} \\ &= {}_0F_1(n_R; \gamma\mathbf{V}\mathbf{\Sigma}\mathbf{V}^\dagger\mathbf{D}^{1/2}\mathbf{\Phi}\mathbf{Y}^\dagger\mathbf{Y}\mathbf{\Phi}^\dagger\mathbf{D}^{1/2}) \,. \end{aligned} \tag{6}$$

The integration over $\mathbf{V}$ is a bit trickier, and follows as a consequence of [12, Thms I and III]. Indeed, exploiting the linear independence of the differential operators of the complex variable $\mathbf{V}$ and of its conjugate (transpose) $\mathbf{V}^\dagger$, we can integrate separately with respect to each of the variables. We stress that both the *product property* and the *splitting property* (see Appendix) of integration over unitary groups are exploited. Indeed [12, Formula (70)],

$$\begin{aligned} &\int_{\mathcal{U}(n_T)} d_*\mathbf{V} \int_{\mathcal{U}(n_T)} d_*\mathbf{V}^\dagger \exp\{-\mathrm{Tr}\left(\gamma\mathbf{\Sigma}\mathbf{V}^\dagger\mathbf{D}\mathbf{V}\right)\} \\ &{}_0F_1(N_R; \gamma\mathbf{V}\mathbf{\Sigma}\mathbf{V}^\dagger\mathbf{D}^{1/2}\mathbf{\Phi}\mathbf{Y}^\dagger\mathbf{Y}\mathbf{\Phi}^\dagger\mathbf{D}^{1/2}) = \\ &{}_0F_2(n_T, n_R; -\mathbf{\Sigma}^2, \gamma^2\mathbf{D}^2\mathbf{\Phi}\mathbf{Y}^\dagger\mathbf{Y}\mathbf{\Phi}^\dagger) \,. \end{aligned}$$

The last integration is taken with respect to a Stiefel matrix and results in [2, Formula (54)]

$$\begin{aligned} &\int_{\mathbf{\Phi}} {}_0F_2(n_T, n_R; -\mathbf{\Sigma}^2, \gamma^2\mathbf{D}^2\mathbf{\Phi}\mathbf{Y}^\dagger\mathbf{Y}\mathbf{\Phi}^\dagger) d\mathbf{\Phi} = \\ &\frac{1}{\widetilde{v}} {}_0F_2(n_T, n_R; -\mathbf{\Sigma}^2, \gamma^2\mathbf{D}^2, \mathbf{Y}^\dagger\mathbf{Y}) \,, \end{aligned}$$

from where the theorem statement follows.

A. Special cases

1) i.i.d. Gaussian input: In the case of an ergodic channel[5], the capacity-achieving input distribution, as the transmitter has neither instantaneous or statistical CSI, is a Gaussian vector with i.i.d. components. For non-ergodic channels, under the same assumptions about the CSI availability, there are no results that prove the optimality of isotropic input. This is mainly due to the fact that optimality depends on the operating SNR regime and on the relationships among r, s and n_b.

Nevertheless, also in absence of CSI, the i.i.d. input distribution has been taken as a baseline [5, and references therein]. Indeed, in such a case, the input matrix $\mathbf{X}$ distribution is invariant under both left and right multiplications times unitary matrices, and integration is noticeably simplified with respect to the general case. Under such an assumption, we can write the output pdf as

$$\begin{aligned} p(\mathbf{Y}|\mathbf{X},\mathbf{\Sigma},\mathbf{U}) = &\frac{\exp\left\{-||\mathbf{Y}||^2\right\}}{\pi^{n_R n_b}} \exp\left\{-\gamma \mathrm{Tr}\left(\mathbf{\Sigma}\widetilde{\mathbf{X}}\widetilde{\mathbf{X}}^\dagger\right)\right\} \\ &\exp\left\{\sqrt{\gamma}\mathrm{Tr}\left(\mathbf{Y}\widetilde{\mathbf{X}}^\dagger\mathbf{\Sigma}^{1/2}\mathbf{U}^\dagger + \mathbf{U}\mathbf{\Sigma}^{1/2}\widetilde{\mathbf{X}}\mathbf{Y}^\dagger\right)\right\} \end{aligned}$$

[4]By normalized measures we mean $d_*\mathbf{U} = d\mathbf{U}/\mathrm{Vol}(\mathcal{U}(n_R))$ and, respectively, $d_*\mathbf{V} = d\mathbf{V}/\mathrm{Vol}(\mathcal{U}(n_T))$.

[5]A channel is ergodic when a codeword spans many realizations of the fading coefficients.

with $\widetilde{\mathbf{X}} = \mathbf{V}^\dagger \mathbf{X}$ still being an i.i.d. Gaussian matrix. This way, integration over $\mathbf{U}$ yields, as in (6),

$$p(\mathbf{Y}|\widetilde{\mathbf{X}}, \boldsymbol{\Sigma}) = \frac{e^{-||\mathbf{Y}||^2} e^{-\gamma \mathrm{Tr}\left(\boldsymbol{\Sigma}\widetilde{\mathbf{X}}\widetilde{\mathbf{X}}^\dagger\right)} {}_0F_1\left(n_R; \gamma \widetilde{\mathbf{X}}^\dagger \boldsymbol{\Sigma} \widetilde{\mathbf{X}} \mathbf{Y}^\dagger \mathbf{Y}\right)}{\pi^{n_R n_b}}$$

Recalling that the i.i.d. input assumption implies

$$p(\mathbf{X}) = \frac{e^{-||\mathbf{X}||^2}}{\pi^{n_T n_b}},$$

the average over the input is performed by the help of [12, Thm. II and Consequence II], namely

$$\int_{\mathbf{X}} e^{-||\mathbf{X}||^2} e^{-\gamma \mathrm{Tr}\left(\boldsymbol{\Sigma}\mathbf{X}\mathbf{X}^\dagger\right)} {}_0F_1\left(n_R; \gamma \mathbf{X}^\dagger \boldsymbol{\Sigma} \mathbf{X} \mathbf{Y}^\dagger \mathbf{Y}\right) d\mathbf{X} = {}_0F_1\left(n_R; -\gamma^2 \boldsymbol{\Sigma}^2, \mathbf{Y}^\dagger \mathbf{Y}\right) .$$

As a consequence,

$$p(\mathbf{Y}|\boldsymbol{\Sigma}) = \frac{\exp\{-||\mathbf{Y}||^2\}}{\pi^{n_R n_b}} {}_0F_1\left(n_R; -\gamma^2 \boldsymbol{\Sigma}^2, \mathbf{Y}^\dagger \mathbf{Y}\right) ,$$

where, as in [5], the average over $\boldsymbol{\Sigma}$ has to be performed numerically.

As mentioned above, in absence of CSI the capacity-achieving input structure depends on the relationships among r, s and n_b. The existing results on this aspect can be essentially grouped into two categories, depending on whether $n_b \geq n_R + s$, or the other way round. Below, we consider the two cases separately.

B. A limited number of antennas: $n_b \geq n_R + s$

In this case, it turns out that the capacity-achieving input structure is such that $\mathbf{D} = c\mathbf{I}$ w.p. 1, and $c = \frac{n_b n_R}{\min\{n_R, n_b - n_T\}}$ [2]. Thus, the conditional law of the output can be evaluated following the footsteps of the proof of Theorem I. Indeed, in this case

$$\exp\{-\gamma ||\mathbf{H}\mathbf{X}||^2\} = \exp\{-c\gamma \mathrm{Tr}\left(\boldsymbol{\Sigma}\right)\} ,$$

while

$$\int_{\mathcal{U}(n_R)} e^{\sqrt{\gamma}\mathrm{Tr}\left(\mathbf{Y}\boldsymbol{\Phi}^\dagger \mathbf{U}^\dagger \boldsymbol{\Sigma}^{1/2}\mathbf{V} + \mathbf{U}\boldsymbol{\Sigma}^{1/2}\mathbf{V}^\dagger \boldsymbol{\Phi}\mathbf{Y}^\dagger\right)} d_*\mathbf{U}$$
$$= {}_0F_1(n_R; \gamma \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^\dagger \mathbf{D}^{1/2} \boldsymbol{\Phi}\mathbf{Y}^\dagger \mathbf{Y} \boldsymbol{\Phi}^\dagger \mathbf{D}^{1/2}) . \qquad (7)$$

Replacing the capacity-achieving expression of $\mathbf{D}$ in (7) and noticing that $\boldsymbol{\Phi}^\dagger \mathbf{V} = \widetilde{\mathbf{V}}$ is an $n_b \times n_T$ Stiefel matrix, we can exploit again the splitting property to obtain

$$p(\mathbf{Y}|\boldsymbol{\Sigma}) = \frac{\exp\{-||\mathbf{Y}||^2\}}{\widetilde{v}\pi^{n_R n_b}} {}_0F_1(n_R; \boldsymbol{\Sigma}, \gamma c \mathbf{Y}^\dagger \mathbf{Y}) e^{-\gamma c \mathrm{Tr}(\boldsymbol{\Sigma})} .$$

This case is the only one for which we are able to give a closed form expression of the unconditional law of the output signal, $p(\mathbf{Y})$. The result has been already obtained in a slightly different way in [13]. We first observe that, due to our assumptions, $\boldsymbol{\Sigma}$ and $\mathbf{Y}\mathbf{Y}^\dagger$ are not warranted to have the same size. In particular, if $n_R = s$, they are square matrices of the same size, and then the determinant representation of ${}_0F_1(n_R; \boldsymbol{\Sigma}, \gamma c \mathbf{Y}^\dagger \mathbf{Y})$ is given by [12, Formula (35)]. Once the hypergeometric function is expressed as a ratio of determinants, we resort to [14, Lemma III] in order to average over $\boldsymbol{\Sigma}$ whose law is given by (1). We obtain

$$p(\mathbf{Y}) = K \frac{\exp\{-||\mathbf{Y}||^2\}}{\mathcal{V}(\gamma c \mathbf{Y}^\dagger \mathbf{Y})} |\mathbf{F}| ,$$

with

$$K = \frac{c_{n_R}}{s! \tilde{v} \pi^{n_R n_b} \prod_{\ell=1}^{s} (n_R - \ell)!(n_T - \ell)!}$$

being a normalizing constant [6] and[7]

$$\mathbf{F}_{i,j} = \Gamma(\tau + j)_1F_1\left(\tau + j, 1, \frac{y_i^2 \gamma c}{1 + \gamma c}\right) , \qquad (8)$$

where y_i^2 denotes the i-th eigenvalue of $\mathbf{Y}^\dagger \mathbf{Y}$. Notice that, for $n_T = s$, the determinant representation of ${}_0F_1(n_R; \boldsymbol{\Sigma}, \gamma c \mathbf{Y}^\dagger \mathbf{Y})$ is a bit more involved and can be found in [15, Lemma II]. Apart from that, the evaluation of the closed-form pdf of $\mathbf{Y}$ follows the same steps as in the simplest case of $n_R = s$.

We remark that, since $\tau + j \geq 1$, $\forall \tau, j$, (8) can be written as a Laguerre polynomial [8] rather than as an infinite, tough convergent, series.

C. The Massive MIMO *regime:* $n_b < n_R + s$

In this case, the high-SNR capacity-achieving input has been proven in [2] to be given by $\mathbf{X} = \sqrt{c}\mathbf{D}^{1/2}\boldsymbol{\Phi}$, with $\mathbf{D}$ being a diagonal n_T-dimensional matrix whose squared entries are jointly distributed as the eigenvalues of a Beta-distributed matrix. Evaluation of the output pdf in this case has been carried out in [18], where the mutual information conveyed by the MIMO channel (2) is analytically characterized, leading to

$$p(\mathbf{Y}) = T \cdot L \cdot \frac{e^{-\|\mathbf{Y}\|^2}}{\mathcal{V}(\mathbf{Y}^\dagger \mathbf{Y})} \cdot |\boldsymbol{\Psi}| \cdot |\widetilde{\mathbf{M}}| \qquad (9)$$

with

$$T = \frac{\pi^{-n_b n_R} \prod_{i=n_b-n_R+1}^{n_b} \Gamma(i)}{(c\gamma)^{n_T(\frac{n_T-3}{2}+n_b)}} ,$$

$$L = \frac{\pi^{n_T(n_T-1)} \Gamma_{n_T}(n_R)}{\Gamma_{n_T}(n_T)\Gamma_{n_T}(n_b - n_T)\Gamma_{n_T}(n_R + n_T - n_b)} ,$$

and

$$\widetilde{\mathbf{M}}_{i,j} = \int_0^1 \frac{(1-x)^{n_R - n_b} x^{i-1-n_T}}{(1 + c\gamma x)^{n_R - n_b + 1}} \cdot \left[\exp\left(\frac{c\gamma x y_j^2}{1 + c\gamma x}\right) - \sum_{l,k=1}^{n_b - n_T} (\boldsymbol{\Psi}^{-1})_{l,k} \exp\left(\frac{c\gamma x y_{l+n_T}^2}{1 + c\gamma x}\right) y_j^{2(n_b - k - n_T)}\right] \mathrm{d}x ,$$

where y_i^2, $i = 1, \ldots, n_b$ are the eigenvalues of the square matrix $\mathbf{Y}\mathbf{Y}^\dagger$ and $\boldsymbol{\Psi}$ is a square matrix of size $n_b - n_T$ whose elements are given by $(\boldsymbol{\Psi})_{i,j} = y_{n_T+j}^{2(n_b - n_T - i)}$ $i, j = 1, \ldots, n_b - n_T$.

[6] In K, $c_{n_R} = \prod_{t=1}^{s-1}(t(n_R + t))^{s-t}$ and for the general case c_m is defined in the Appendix.

[7] We have herein skipped analytical details due to space limitation, however the expression for $\mathbf{F}_{i,j}$ can be obtained through [16, Formulae (6.643.2) and (9.220.2)].

V. Discussion and future work

We obtained a new expression for the law of the output signal of a block-Rayleigh fading MIMO channel, relying on sophisticated results in the field of finite-dimensional random matrix theory. We focused on channels where neither the transmitter nor the receiver, each equipped with multiple antennas, are aware of the CSI. This assumption is relevant in energy-efficient wireless systems, in that CSI sharing is time and resource-consuming. From the mathematical point of view, this implies that a relevant structure for the input is the product of a positive-definite (diagonal) matrix times a Stiefel matrix. The main expression we obtained for the output law is conditioned to the channel eigenvalues and to the input power allocation, while we were able to average over the input and the channel eigenvectors distributions. We also highlighted the relation between our result and previous studies.

The following further observations on our result are worthwhile being underlined. Although very compact, in the general case the expression in (3) presents several difficulties as one aims at further averaging it over either $\mathbf{\Sigma}$, or $\mathbf{D}$, or over both matrices. First of all, yet no determinant representation is known for hypergeometric functions of more than two matrix arguments. This prevents us from expressing the output pdf as a ratio of determinants, and forces us to exploit the zonal polynomial expansion in (4). Henceforth, one should adopt the procedure in [17, Appendix], expanding the zonal polynomials as ratio of determinants and choosing the best way to recover a closed-form result summing over partitions of properly defined sets. The combinatorial complexity of this procedure is noticeable, and its application to our expressions is subject of ongoing work. Furthermore, it would be worthwhile extending the analysis to block-fading channels with Line of Sight component.

Acknowledgment

C.F. Chiasserini and S. Zhou were supported by NPRP grant $\sharp 5-782-2-322$ from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors. This work was also supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM$\sharp$ (Grant agreement no. 318306).

VI. Appendix: Properties of hypergeometric functions

Hereinafter, we list some essential properties of hypergeometric functions of (multiple) matrix arguments, which follow from matrix-variate complex zonal polynomials properties. In the following, both $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices.

Splitting property of the integral:

$$\int_{\mathcal{U}(n)} {}_pF_q\left(a_1,\dots,a_p;b_1\dots,b_q;\mathbf{AUBU}^\dagger\right) d_*\mathbf{U} = {}_pF_q\left(a_1,\dots,a_p;b_1\dots,b_q;\mathbf{A},\mathbf{B}\right).$$

Product property:

$$\int_{\mathcal{U}(n)} {}_pF_q\left(a_1,\dots,a_p;b_1\dots,b_q;\mathbf{AU}\right) {}_{\tilde{p}}F_{\tilde{q}}\left(a_1,\dots,a_{\tilde{p}};b_1\dots,b_{\tilde{q}};\mathbf{BU}^\dagger\right) d_*\mathbf{U} = {}_{p+\tilde{p}}F_{q+\tilde{q}+1}\left(\begin{array}{c} a_1,\dots,a_p,a_1,\dots,a_{\tilde{p}} \\ n,b_1\dots,b_q,b_1\dots,b_{\tilde{q}}\end{array};\mathbf{AB}\right). \quad (10)$$

Notice that (10) and (10) hold, *mutatis mutandis*, also for the integration over generic complex matrices [12].

Determinant representations. For single and double matrix arguments, hypergeometric functions can be efficiently expressed as ratio of determinants involving scalar hypergeometric functions and Vandermonde determinants of the matrix arguments. Indeed, for the case of a single matrix argument it holds [12, Formula 34], while for two matrix arguments one should refer to [12, Formula 35]. Unfortunately, no such representations are available for more than two matrix arguments.

References

[1] A. Lapidoth, S. M. Moser, "Capacity Bounds via Duality with Applications to Multiple-Antenna Systems on Flat Fading Channels," *IEEE Trans. on Information Theory,* Vol. 49, No. 10, pp. 2426–2467, Oct. 2003.
[2] W. Yang, G. Durisi, E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Sel. Areas Commun.*, Feb. 2013, to appear.
[3] R. A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
[4] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, MIT PhD Dissertation, 1989.
[5] F. Rusek, A. Lozano, N. Jindal "Mutual Information of IID Complex Gaussian Signals on Block Rayleigh-Faded Channels", *IEEE Trans. on Information Theory* Vol. 58, No. 1, pp. 331–340 Jan. 2012.
[6] T. L. Marzetta, B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. on Information Theory,* Vol. 45, No. 1, pp. 139–157, Jan. 1999.
[7] A. T. James, "Distribution of Matrix Variates and Latent Roots Derived from Normal Samples," *The Annals of Mathematical Statistics,* Vol. 35, No. 2, pp. 474–501, June 1964.
[8] M. Abramowitz, and I. A. Stegun, "Handbook of Mathematical Functions", *New York: Dover Publications*, 1972.
[9] C. G. Khatri, "On Certain Distribution Problems Based on Positive Definite Quadratic Functions in Normal Vectors," *The Annals of Mathematical Statistics,* Vol. 37, No. 2, pp. 467–479, Apr. 1966.
[10] L. Zheng, D. Tse, "Communicating on the Grassmann Manifold: A Geometric Approach to the Non-coherent Multiple Antenna Channel," *IEEE Trans. on Information Theory*, Vol. 48, No. 2, pp. 359–383, Feb. 2002.
[11] L. Hanlen, A. Grant, "Capacity Analysis of Correlated MIMO Channels," *IEEE Trans. Information Theory*, Vol. 58, No. 11, pp. 6773–6787, Nov. 2012.
[12] A. Y. Orlov "New solvable matrix integrals," *Int. Journ. of Modern Physics A,* Vol. 19, pp. 279–293, 2004.
[13] B. Hassibi, T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: the received signal density in closed form," *IEEE Trans. on Information Theory,* Vol. 48, No. 6, pp. 1473–1484, June 2002.
[14] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of Spatially Correlated MIMO Rayleigh-Fading Channels," *IEEE Transactions on Information Theory,* Vol. 49, No. 10, pp. 2363–2371, Oct. 2003.
[15] G. Alfano, A. Lozano, A. Tulino, and S. Verdú, "Mutual Information and Eigenvalue Distribution of MIMO Ricean Channels," *Proc. of ISITA 2004,* Parma, Italy, Oct. 2004.
[16] I.S. Gradshteyn and I.M. Ryzhik, Table of Integrals, Series, and Products. *Academic Press*, New York, 1980.
[17] M. R. McKay, A. J. Grant, I.B. Collings, "Performance Analysis of MIMO-MRC in Double-Correlated Rayleigh Environments," *IEEE Trans. on Communications*, Vol. 55, No. 3, pp. 497–507, Mar. 2007.
[18] G. Alfano, C.-F. Chiasserini, A. Nordio, S. Zhou, "High-SNR Block-faded Rayleigh MIMO Channels," submitted to *IEEE Wireless Communications Letters*, Apr. 2013.