# The Mutual Information of a MIMO Channel: A Survey

Walid Hachem, Philippe Loubaton and Jamal Najim

***Abstract*— In this paper, we survey recent mathematical results devoted to the study of the mutual information of MIMO channels in the case where transmit and receive antennas converge to $\infty$ at the same rate.**

**We express the different results in a unified framework and the emphasis is put on non-asymptotic deterministic approximations of the mutual information, asymptotic limits (when existing) and Ricean correlated channels.**

## I. Introduction

It is well-known that the mutual information of a MIMO channel is given by

$$C(\varsigma^2) = \mathbb{E} \log\det \left( I + \frac{H_n H_n^*}{\varsigma^2} \right)$$

where $\varsigma^2$ is the variance of an additive corrupting noise and the $N \times n$ matrix $H_n = (H_{ij}^n)$ represents (up to a factor $\sqrt{n}$) the complex gain between transmit and receive antennas. In his seminal paper [13], Telatar has proved that in the case where the entries of the matrix are i.i.d. centered Gaussian random variables with variance $\frac{\sigma^2}{n}$, the mutual information properly normalized, i.e. $C_n(\varsigma^2) = \frac{C(\varsigma^2)}{N}$ converges toward a deterministic quantity involving Marčenko-Pastur probability distribution in the case where $\frac{N}{n} \to c > 0$. Telatar relied on Marčenko-Pastur's theorem from the theory of Large Random Matrices. Of importance is the fact that the mutual information of the channel grows proportionally to the number of transmit antennas (or receive ones since their ratio is assumed to be constant).

The question soon arised to extend these results to more realistic models, especially to those models where the entries of the matrix are no longer independent and have a covariance function of the form:

$$\mathrm{cov}(H_{ij}^n, H_{i'j'}^n) = \frac{a(i-i')b(j-j')}{n}$$

where $f$ and $g$ are two given functions. Such results, based on an extensive use of the Stieltjes transform $\mathbf{f}$ of a probability measure $\mu$:

$$\mathbf{f}(z) = \int_{\mathbb{R}^+} \frac{\mu(d\lambda)}{\lambda - z},$$

This work was partially supported by the "Fonds National de la Science" via the ACI program "Nouvelles Interfaces des Mathématiques", project MALCOM number 205

W. Hachem is with Ecole Supérieure d'Electricité (Supélec), Plateau de Moulon, 3 rue Joliot-Curie 91192 Gif Sur Yvette Cedex, France. walid.hachem@supelec.fr

P. Loubaton is with IGM LabInfo, UMR 8049, Institut Gaspard Monge, Université de Marne La Vallée, France. loubaton@univ-mlv.fr

J. Najim is with CNRS and Télécom Paris 46, rue Barrault, 75013 Paris, France. najim@tsi.enst.fr

have been developed by Chuah et al. in [2], relying on results from Girko [6]. Relying on replica methods, Moustakas et al. [11] have been able to compute an equivalent of the mean of $\frac{1}{N} \log\det \left( I + \frac{H_n H_n^*}{\varsigma^2} \right)$ (the variance has also been computed but second-order computations are out of the scope of the present survey).

We shall survey all this line of results and present recent results [10] where an equivalent of the mutual information is computed in the case where the covariance of $H_n$ is of a general form

$$\mathrm{cov}(H_{ij}^n, H_{i'j'}^n) = \frac{\kappa(i-i', j-j')}{n}.$$

and in the case where $H_n$ is no longer centered, i.e. $H_n = Z_n + B_n$ where $B_n$ is deterministic and $\mathbb{E}(Z_n) = 0$. Such a case is known as the Ricean channel and has led to partial studies under different sets of assumptions: See [3], [12] (using the replica method in the uncorrelated case), [5] (receive correlated Ricean channels).

In the sequel, we deal with the following model of non-centered random matrices with a variance profile:

$$\Sigma_n = Y_n + A_n$$

where $\Sigma_n, Y_n$ and $A_n$ are $N \times n$ random matrices. Matrix $Y_n$ has a variance profile, i.e. the entries of $Y_n = (Y_{ij}^n)$ have the form $Y_{ij}^n = \frac{\sigma_{ij}^{(n)}}{\sqrt{n}} X_{ij}^n$, the $X_{ij}^n$ being independent and identically distributed $(0,1)$ complex circular gaussian (denoted $\mathcal{CN}(0,1)$) random variables. Matrix $A_n$ is assumed to be deterministic. Otherwise stated, $\mathbb{E} Y_n = 0$ and $\mathbb{E}\Sigma_n = A_n$. Particular attention will be devoted to the case of a Ricean channel, i.e. $A_n \neq 0$.

In Section II, we survey mutual informations results in the case where $\Sigma_n$ is centered, that is in the case where $A_n = 0$. Non-asymptotic formulas are given for a general variance profile $\sigma_{ij}^{(n)}$ and asymptotic formulas are provided in the case where the variance profile is the sampling of a continuous function, i.e. $\sigma_{ij}^{(n)} = \sigma(i/N, j/n)$.

In Section III, the general non-centered case is addressed. Non-asymptotic formulas for the mutual information are provided.

In Section IV, we revisit the general case when the variance profile is separable, that is when $\sigma_{ij}^{(n)} = d_i \tilde{d}_j$. This assumption induces major simplifications in the computation of the deterministic equivalent of the mutual information.

As will be shown in Section V, the case of a Gaussian matrix $Z_n$ with correlated entries is very close to the case of a matrix $Y_n$ with a variance profile. The intuitive equivalence

$Y_n \approx F_N Z_n F_n^*$ where $F_p = (F^p_{j_1,j_2})_{0\le j_1,j_2<p}$ is the $p\times p$ Fourier matrix:

$$F^p_{j_1,j_2} = \frac{1}{\sqrt{p}} \exp 2\mathbf{i}\pi \left(\frac{j_1 j_2}{p}\right) \tag{1}$$

is fully explained.

## II. THE CENTERED CASE

In this section, we take $A_n = 0$ that is $\Sigma_n = Y_n$. As a major consequence of this assumption, the normalized mutual information $\frac{1}{N}\mathbb{E}\log\det\left(I_N + \frac{\Sigma\Sigma^*}{\varsigma^2}\right)$ converges toward a deterministic limit in the case where the variance profile is the sampling of a continuous function (see Assumption (A-1) and Theorem 2.3).

We introduce the following notations:

$$\begin{array}{ll} D_j = \mathrm{diag}\left(\sigma^2_{ij},\ 1\le i\le N\right), & T = \mathrm{diag}(T_i,\ 1\le i\le N), \\ \tilde{D}_i = \mathrm{diag}\left(\sigma^2_{ij},\ 1\le j\le n\right), & \tilde{T} = \mathrm{diag}(\tilde{T}_j,\ 1\le j\le n) \end{array}$$

where both $T$ and $\tilde{T}$ are defined by the following system of $N+n$ equations.

*Theorem 2.1 (see [10]):* Consider the following system of $N+n$ equations:

$$\begin{aligned} T_i(z) &= \frac{-1}{z(1+\frac{1}{n}\mathrm{Tr}\tilde{D}_i\tilde{T}(z))}, \quad 1\le i\le N, \\ \tilde{T}_j(z) &= \frac{-1}{z(1+\frac{1}{n}\mathrm{Tr}D_j T(z))}, \quad 1\le j\le n \end{aligned}$$

then this system admits a unique solution $(T,\tilde{T})$ among the class of diagonal matrices such that $T_i(z)$ and $\tilde{T}_j(z)$ are Stieltjes transforms of probability measures.

*Theorem 2.2 (see [10]):* Denote by $\bar{C}_n(\varsigma^2)$ the quantity

$$\begin{aligned} \bar{C}_n(\varsigma^2) = &-\frac{1}{N}\sum_{i=1}^N \log\varsigma^2 T_i(-\varsigma^2) - \frac{1}{N}\sum_{j=1}^n \log\varsigma^2\tilde{T}_j(-\varsigma^2) \\ &- \frac{\varsigma^2}{Nn}\sum_{\substack{i=1:N\\ j=1:n}} \sigma^2_{ij} T_i(-\varsigma^2)\tilde{T}_j(-\varsigma^2) \end{aligned}$$

Assume that the variance profile is bounded:

$$\sup_{i,j,n}\sigma^{(n)}_{ij} \le \sigma_{\max}.$$

Then the following holds true:

$$\frac{1}{N}\mathbb{E}\log\det\left(I_N + \frac{\Sigma\Sigma^*}{\varsigma^2}\right) - \bar{C}_n(\varsigma^2) \xrightarrow[n\to\infty]{} 0.$$

Of interest is the case where the convergence of $\bar{C}_n(\varsigma^2)$ occurs. This is the aim of next assumption and next theorem.

*Assumption A-1:* The variance profile is the sampling of a continuous function:

$$\sigma^{(n)}_{ij} = \sigma\left(\frac{i}{N},\frac{j}{n}\right) \tag{2}$$

where $\sigma(x,y)$ is continuous.

*Theorem 2.3:* Assume now that (A-1) holds and consider the following functional equation:

$$k(u,z) = \frac{1}{-z + \int_0^1 \frac{\sigma^2(u,t)}{1+c\int_0^1\sigma^2(x,t)k(x,z)dx}dt}.$$

This equation admits a unique solution in the class of functions $k$ such that

1) $z\mapsto k(u,z)$ is the Stieltjes transform of a probability measure,
2) $[0,1]\ni u\mapsto k(u,z)$ is continuous.

We denote $k_\varsigma(u) = k(u,-\varsigma^2)$. The following convergence holds true:

$$\bar{C}_n(\varsigma^2) \xrightarrow[n\to\infty]{} C^*(\varsigma^2)$$

where $C^*(\varsigma^2)$ is given by the following formula

$$\begin{aligned} C^*(\varsigma^2) = &-\int_0^1 \log(\varsigma^2 k_\varsigma(x))\,dx \\ &- \frac{1}{c}\int_0^1 \log\left(\frac{1}{1+c\int_0^1\sigma^2(u,x)k_\varsigma(u)du}\right)dx \\ &- \int_{[0,1]^2}\frac{\sigma^2(x,y)k_\varsigma(x)}{1+c\int_0^1\sigma^2(u,y)k_\varsigma(u)du}dx\,dy \end{aligned}$$

Mathematical details are provided in [9] and [10].

## III. THE RICEAN CASE

In the general case, that is when $A_n \ne 0$ one cannot expect the convergence of the empirical distribution of the eigenvalues of $\Sigma_n\Sigma_n^*$ in the case where $A_n\ne 0$. Only very specific cases can be studied ([4], [9]) in a fully asymptotic perspective. However, one can still compute a deterministic approximation as in Theorem 2.2.

*Assumption A-2:* We assume that the $N\times n$ matrix $A_n = (A^n_{ij})$ whose columns $(\mathbf{a}^n_k)_{1\le k\le n}$ and rows $(\tilde{\mathbf{a}}^n_\ell)_{1\le\ell\le N}$ satisfies

$$\sup_{n\ge 1}\max_{k,\ell}(\|\mathbf{a}^n_k\|,\|\tilde{\mathbf{a}}^n_\ell\|) < +\infty \tag{3}$$

where $\|\cdot\|$ stands for the Euclidean norm.

Assumption (A-2) is very relevant in the context of digital communication and encompasses usual line of sight components models for instance .

*Theorem 3.1 (see [10], see also [7]):* Assume that (A-2) holds and let $A_n$ be a $N\times n$ deterministic matrix. The deterministic system of $N+n$ equations:

$$\psi_i(z) = \frac{-1}{z\left(1+\frac{1}{n}\mathrm{Tr}\,\tilde{D}_i\tilde{T}(z)\right)} \quad \text{for } 1\le i\le N, \tag{4}$$

$$\tilde{\psi}_j(z) = \frac{-1}{z\left(1+\frac{1}{n}\mathrm{Tr}\,D_j T(z)\right)} \quad \text{for } 1\le j\le n, \tag{5}$$

where

$$\Psi(z) = \mathrm{diag}(\psi_i(z), 1\le i\le N), \tag{6}$$

$$\tilde{\Psi}(z) = \mathrm{diag}(\tilde{\psi}_j(z), 1\le j\le n), \tag{7}$$

$$T(z) = \left(\Psi^{-1}(z) - zA\tilde{\Psi}(z)A^*\right)^{-1}, \tag{8}$$

$$\tilde{T}(z) = \left(\tilde{\Psi}^{-1}(z) - zA^*\Psi(z)A\right)^{-1}. \tag{9}$$

admits a unique solution $(\psi_1,\ldots,\psi_N,\tilde{\psi}_1,\ldots,\tilde{\psi}_n)$ in the class of the functions which are Stieltjes transforms.
In the sequel, we denote by $\Psi_\varsigma = \Psi(-\varsigma^2)$ and by $\tilde{\Psi}_\varsigma = \tilde{\Psi}(-\varsigma^2)$.

*Theorem 3.2:* Denote by $\bar{C}_n(\varsigma^2)$ the quantity

$$\bar{C}_n(\varsigma^2) = \frac{1}{N}\log\det\left[\frac{\Psi_\varsigma^{-1}}{\varsigma^2} + A\tilde{\Psi}_\varsigma A^*\right] + \frac{1}{N}\log\det\frac{\tilde{\Psi}_\varsigma^{-1}}{\varsigma^2} - \frac{\varsigma^2}{Nn}\sum_{\substack{i=1:N\\ j=1:n}} \sigma_{ij}^2 T_i(-\varsigma^2)\tilde{T}_j(-\varsigma^2) \quad (10)$$

Assume that the variance profile is bounded. Then the following holds true:

$$\frac{1}{N}\mathbb{E}\log\det\left(I_N + \frac{\Sigma\Sigma^*}{\varsigma^2}\right) - \bar{C}_n(\varsigma^2) \xrightarrow[n\to\infty]{} 0.$$

Mathematical details are provided in [10].

## IV. THE RICEAN CASE (REVISITED)

In this section, we assume that the variance profile $\sigma_{ij}^{(n)}$ is separable:

*Assumption A-3:* The variance profile $\sigma_{ij}^{(n)}$ is assumed to be separable, i.e.:

$$\sigma_{ij}^{(n)} = d_i\tilde{d}_j; \quad 1\le i\le N,\ 1\le j\le n.$$

As we shall see, Assumption (A-3) induces major simplification over the system of $N+n$ equations of Theorem 3.1 since the system reduces to 2 equations in this case (in accordance with [11] for instance). Denote by

$$\begin{aligned} D &= \operatorname{diag}(d_i,\ 1\le i\le N) \\ \tilde{D} &= \operatorname{diag}(\tilde{d}_j,\ 1\le j\le n) \end{aligned}$$

*Theorem 4.1 (see [10]):* Assume that (A-3) holds and consider the following system of equations

$$\begin{cases} \delta(z) = \frac{1}{n}\mathrm{Tr}\left[D\left(-z(I+D\tilde{\delta}) + A(I+\tilde{D}\delta)^{-1}A^T\right)^{-1}\right] \\ \tilde{\delta}(z) = \frac{1}{n}\mathrm{Tr}\left[\tilde{D}\left(-z(I+\tilde{D}\delta) + A^T(I+D\tilde{\delta})^{-1}A\right)^{-1}\right] \end{cases}.$$

Then this system admits a unique solution in the class of Stieltjes transforms of positive measures $\mu$ and $\tilde{\mu}$ such that $\mu(\mathbb{R}^+) = \frac{1}{n}\mathrm{Tr}D$ and $\tilde{\mu}(\mathbb{R}^+) = \frac{1}{n}\mathrm{Tr}\tilde{D}$.
We can now define properly the related quantities $T, \tilde{T}, \Psi$ and $\tilde{\Psi}$ as:

$$\begin{aligned} \Psi(z) &= -\frac{(I+\tilde{\delta}D)^{-1}}{z},\ \tilde{\Psi}(z) = -\frac{(I+\delta\tilde{D})^{-1}}{z} && (11) \\ T(-z) &= \left(-z(1+\tilde{\delta}D) + A(I+\delta\tilde{D})^{-1}A^*\right)^{-1} && (12) \\ \tilde{T}(-z) &= \left(-z(1+\delta\tilde{D}) + A(I+\tilde{\delta}D)^{-1}A^*\right)^{-1} && (13) \end{aligned}$$

and accordingly their evaluations at the point $z=-\varsigma^2$: $\Psi_\varsigma$, $\tilde{\Psi}_\varsigma$, $T_\varsigma$ and $\tilde{T}_\varsigma$.

*Theorem 4.2:* The statement of Theorem 3.2 remains valid with $T, \tilde{T}, \Psi$ and $\tilde{\Psi}$ given by (11), (12) and (13).

## V. FROM INDEPENDENCE TO STATIONARITY: THE CASE OF GAUSSIAN MATRICES

We now turn to the relation between random matrices based on a Gaussian stationary field and matrices with independent entries and a variance profile.

*Assumption A-4:* Consider the $N\times n$ matrix whose entries are given by

$$Z_{j_1j_2}^n = \frac{1}{\sqrt{n}}\sum_{(k_1,k_2)\in\mathbb{Z}^2} h(k_1,k_2)U(j_1-k_1, j_2-k_2),$$

where $h$ is a deterministic complex summable sequence (over $\mathbb{Z}^2$) and $(U(j_1,j_2); (j_1,j_2)\in\mathbb{Z}^2)$ is a sequence of $\mathcal{CN}(0,1)$ random variables.

Such a matrix is a good model for a Gaussian stationary field since every entry $Z_{j_1j_2}^n$ is complex gaussian, centered and

$$\operatorname{cov}(Z_{j_1j_2}^n, Z_{j_1'j_2'}^n) = \frac{\kappa(j_1-j_1', j_2-j_2')}{n}$$

where

$$\kappa(j_1,j_2) = \sum_{(k_1,k_2)\in\mathbb{Z}^2} h(k_1,k_2)h^*(k_1-j_1, k_2-j_2)$$

Consider on the other hand the $N\times n$ matrix $Y_n = (Y_{j_1,j_2}^n)$ where

$$Y_{j_1,j_2}^n = \frac{\Phi\left(\frac{j_1}{N}, \frac{j_2}{n}\right)}{\sqrt{n}}X_{j_1,j_2} \quad (14)$$

where the $(X_{j_1,j_2})$ are i.i.d. $\mathcal{CN}(0,1)$ random variables and

$$\Phi(t_1,t_2) = \sum_{(\ell_1,\ell_2)\in\mathbb{Z}^2} h(\ell_1,\ell_2)e^{2\pi\mathbf{i}(\ell_1t_1-\ell_2t_2)} \quad (15)$$

The similar asymptotic behavior of the spectral measure of $Z_nZ_n^*$ and $Y_nY_n^*$ are part of the folklore in the MIMO capacity literature. We give here a formal justification to this fact, based on [8], and extend Theorem 3.2 to the case of matrices with Gaussian stationary entries. The following holds true:

*Theorem 5.1 (see [8]):* Let $H_n = Z_n + B_n$ where $B_n$ satisfies (A-2) and $Z_n$ satisfies (A-4). Then the conclusions of Theorems 3.1 and 3.2 remain valid with the following slight modifications:

$$\begin{aligned} D_j &= \operatorname{diag}\left\{|\Phi|^2\left(\frac{i}{N},\frac{j}{n}\right);\ 1\le i\le N\right\}; \\ \tilde{D}_i &= \operatorname{diag}\left\{|\Phi|^2\left(\frac{i}{N},\frac{j}{n}\right);\ 1\le j\le n\right\}; \\ A &= F_N^* B F_n. \end{aligned}$$

where $\Phi$ is given by (15) and $F_N$ and $F_n$ are Fourier matrices defined by (1). Moreover,

$$\frac{1}{N}\mathbb{E}\log\det\left(I + \frac{HH^*}{\varsigma^2}\right) - \bar{C}_n(\varsigma^2) \xrightarrow[n\to\infty]{} 0,$$

where $\bar{C}_n(\varsigma^2)$ is given by (10).

*Elements of proof*

The proof of Theorem 5.1 relies on two main components.

1) A periodization scheme popular in signal processing. We introduce the matrix $\tilde{Z}_n = (\tilde{Z}^n_{j_1 j_2})$ where

$$\tilde{Z}^n_{j_1 j_2} = \frac{1}{\sqrt{n}} \sum_{(k_1,k_2)\in\mathbb{Z}^2} h(k_1,k_2) \times U\left((j_1-k_1) \bmod N, (j_2-k_2) \bmod n\right),$$

and mod denotes modulo. The main interest of matrix $\tilde{Z}_n$ comes from the fact that it can be fully decorrelated by Fourier multiplication:

$$F_N^* \tilde{Z}_n F_n = Y_n,$$

where $Y_n$ is defined by (14).

2) The second element is an inequality due to Bai [1] involving the Lévy distance $\mathcal{L}$ between distribution functions:

$$\mathcal{L}^4(F^{AA^*}, F^{BB^*}) \leq \frac{2}{N^2}\mathrm{Tr}(A-B)(A-B)^*\mathrm{Tr}(AA^*+BB^*),$$

where $F^{AA^*}$ denotes the empirical distribution function of the eigenvalues of the matrix $AA^*$. This inequality turns out to be perfectly suited to evaluate the difference between the spectrum of matrices $Z_n Z_n^*$ (resp. $(Z_n+B_n)(Z_n+B_n)^*$) and $\tilde{Z}_n\tilde{Z}_n^*$ (resp. $(\tilde{Z}_n + B_n)(\tilde{Z}_n+B_n)^*$)

Mathematical details are provided in [8].

## VI. Conclusion

In this survey, we have presented up-to-date mathematical results related to the study of the normalized mutual information of a MIMO channel:

$$\bar{C}_n(\varsigma^2) = \frac{1}{N}\mathbb{E}\log\det\left(I + \frac{\Sigma\Sigma^*}{\varsigma^2}\right)$$

in the case where the number of receive and transmit antennas go to $\infty$, their ratio being constant. The model under study is $\Sigma = Y + A$ where $Y$ is a random matrix with a variance profile and $A$ is a deterministic matrix.

In the case where $\mathbb{E}\Sigma = A = 0$, we provide both asymptotic and non-asymptotic results while in the case of a Ricean channel ($\mathbb{E}\Sigma = A \neq 0$), the mutual information $\bar{C}_n(\varsigma^2)$ might not converge, however we provide non-asymptotic results.

We show how the variance profile $\sigma^{(n)}_{ij}$ of $Y$'s entries has an impact on the complexity of the results. In particular, if the variance profile is separable, i.e. $\sigma^{(n)}_{ij} = d_i\tilde{d}_j$ the non-asymptotic deterministic equivalent of the mutual information relies on a system of 2 equations instead of $N+n$ equations.

Finally, we describe precisely the links between Gaussian random matrices with a variance profile and stationary Gaussian random matrices. This enables us to study the mutual information of general stationary Gaussian models where the matrix $H$ has the form

$$H = Z + B$$

where $Z$ is random and $B$, deterministic and where $H$'s entries have the following correlation structure:

$$\mathrm{cov}(H^n_{ij}, H^n_{i'j'}) = \frac{\kappa(i-i', j-j')}{n}.$$

## References


[1] Z. D. Bai. Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices. *Ann. Probab.*, 21(2):649–672, 1993.

[2] C.-N. Chuah, D.N.C. Tse, J.M. Kahn, and R.A. Valenzuela. Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Trans. Inform. Theory*, 48(3):637–650, 2002.

[3] L. Cottatellucci, and M. Debbah. The Effect of Line of Sight Components on the Asymptotic Capacity of MIMO Systems in *Proceedings ISIT 2004*, Chicago, 2004.

[4] R. B. Dozier and J. W. Silverstein. On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices. preprint 2004.

[5] J. Dumont, P. Loubaton, S. Lasaulce, and M. Debbah. On the asymptotic performance of mimo correlated ricean channels. In *ICASSP Proceedings*, 2005.

[6] V. L. Girko. *Theory of random determinants*, volume 45 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1990.

[7] V. L. Girko. *Theory of stochastic canonical equations. Vol. I*, volume 535 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2001.

[8] W. Hachem, P. Loubaton, and J. Najim. The empirical eigenvalue distribution of a gram matrix: From independence to stationarity. *Markov Process, Related Fields.*, 11(4):629–648, 2005.

[9] W. Hachem, P. Loubaton, and J. Najim. The empirical distribution of the eigenvalues of a Gram matrix with a given variance profile. available at http://arxiv.org/abs/math.PR/0411333, 2004, to be published in Annales de l'I.H.P., Probab. & Stat.

[10] W. Hachem, P. Loubaton, and J. Najim. Deterministic equivalents for certain functionals of large random matrices. available at http://arxiv.org/abs/math.PR/0507172, 2005.

[11] A. L. Moustakas, S. H. Simon, and A. M. Sengupta. MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large $N$ analysis. *IEEE Trans. Inform. Theory*, 49(10):2545–2561, 2003.

[12] A. L. Moustakas, and S. H. Simon. Random Matrix theory of multi-antenna communications: the Ricean case *J. Phys. A:Math. Gen.* , 38:10859–10872, 2005.

[13] I.E. Telatar. Capacity of multi-antenna gaussian channel. *European Transactions on Telecommunications*, 1999.

[14] A. M. Tulino, A. Lozano and S. Verdú. Impact of Antenna Correlation on the Capacity of Multiantenna Channels *IEEE Trans. Inform. Theory*, 51(7):2491–2509, 2005.

[15] A. Tulino and S. Verdú. Random matrix theory and wireless communications. In *Foundations and Trends in Communications and Information Theory*, volume 1, pages 1–182. Now Publishers, june 2004.