On the asymptotic analysis of mutual information of MIMO Rician correlated channels

J. Dumont, W. Hachem, P. Loubaton, J. Najim

***Abstract*—In this paper, we show how to address the asymptotic behavior of the mutual information of correlated MIMO Ricean channels when the number of transmit and receive antennas converge to $+\infty$ at the same rate. Our approach is based on the extensive use of the Stieljès transform of the Gram matrix of the channel matrix, and is inspired by previous works of Girko in the context of simpler models. We give a closed form expression of the first order approximation of the average mutual information, and evaluate the convergence rate of the corresponding error.**

I. Introduction

The analysis of mutual information and capacity of block fading MIMO static channels has generated considerable interest since the seminal work of Telatar in 1995. The vast majority of these works have addressed the case of Rayleigh MIMO channels. Despite its importance, the context of Rician MIMO channels seems to have been studied much less extensively. We however mention [14], [10], [11], [12]. In particular, important problems such as the study of the impact of the statistical properties of the channel on the mutual information and the capacity, or the derivation of reliable algorithms for finding the optimum input covariance matrix achieving the ergodic capacity of general correlated Rician channels have not yet been solved completely. The above conceptual difficulties are mainly due to the quite intricate structure of the probability distribution of the mutual information of correlated Rician channels.

A possible approach to overcome these problems consists in studying the mutual information in the case where the number of transmit and receive antennas converge to $+\infty$ at the same rate. In effect, this approach has been found quite useful in the context of Rayleigh MIMO channels in the sense that the asymptotic approximations of, e.g. the expectation and the variance of the mutual information, have rather simple and easy to interpret expressions. Moreover, these asymptotic predictions were shown to be quite reliable, even for a quite moderate number of antennas (see e.g. [15]).

In our knowledge, the asymptotic analysis of Rician channels has been considered in [3] (using a result of Girko [7]) and [16] (using the replica method) in the uncorrelated case and in [5] in the case of receive correlated Ricean channels. In this paper, we provide a comprehensive review of the mathematical aspects of this problem, and indicate how the asymptotic behaviour of the mutual information of Rician channels can be addressed by rigorous, and simpler than expected methods. We first review the results of [9] related to the asymptotic behavior of the Stieljès transform of the eigenvalue distribution of the Gram matrix of the channel, and provide a sketch of a simpler proof. Next, we discuss on the asymptotic gaussianity of the mutual information, and on the convergence speed of the average mutual information. We hope that this paper will help the reader to have a better understanding of the random matrix methods used in the asymptotic analysis of MIMO channels.

This work was supported by Newcom

J. Dumont and P. Loubaton are with IGM LabInfo, UMR CNRS 8049, Université de Marne la Vallée, France, dumont,loubaton@univ-mlv.fr

W. Hachem is with Dept. Telecommunications, SUPELEC, France walid.hachem@supelec.fr

J. Najim is with LTCI, UMR CNRS/ENST 5141, France, najim@tsi.enst.f

II. Presentation of the model.

We consider a block fading MIMO static channel and denote by n and N the number of transmit and receive antennas respectively. The $N \times n$ channel matrix, denoted Σ, is supposed to be given by $\Sigma = A + Y$. Y is a zero mean $N \times n$ complex Gaussian random matrix given by $Y = \frac{1}{\sqrt{n}} R^{1/2} X \tilde{R}^{1/2}$ where R and $\tilde{R}$ are the receive and transmit correlation matrices, and where X is a zero mean independent identically distributed complex Gaussian matrix such that $\mathbb{E}|X_{kl}|^2 = 1$. We also assume that the real and imaginary parts of the entries of X are independent, and have the same variance, i.e. $\frac{1}{\sqrt{2}}$. A represents a deterministic $N \times n$ matrix. Very often, A is assumed to be a rank one matrix (see e.g. [8], [13]). However, in important contexts, this hypothesis is not valid. Macro diversity is a typical example in which A is likely to be full rank. In this context, transmit antennas are very far one from each other, while the distance between the receive antennas are of the order of the wavelenght of the transmitted signals. In such a context, the line of sight components between each transmit antenna and the receive antenna arrays are different, so that A is likely to be full rank. If the receive antennas array is linear and uniform, a typical example for A is

$$A = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_n)]\,\Lambda \tag{1}$$

where $\mathbf{a}(\theta) = \frac{1}{\sqrt{N}}(1, e^{i\theta}, \ldots, e^{i(N-1)\theta})^T$ and Λ is a diagonal matrix, the entries of which represent the complex amplitudes of the n line of sight components. Therefore, we do not formulate any assumption on the rank of A.

We denote by $D = \mathrm{diag}(d_i, i = 1, \ldots, N)$ and $\tilde{D} = \mathrm{diag}(\tilde{d}_j, j = 1, \ldots, n)$ the diagonal eigenvalues matrices of R and $\tilde{R}$. Using the unitary invariance of the mutual information and of the Gaussian distribution of X, we can assume without restrictions that matrices R and $\tilde{R}$ are replaced by D and $\tilde{D}$ respectively. From now on, we therefore assume that

$$\Sigma = A + Y \tag{2}$$

where

$$Y = \frac{1}{\sqrt{n}} D^{1/2} X \tilde{D}^{1/2} \tag{3}$$

Remark in particular that the entries of Y are independent, but have different variances. In the following, we denote by $(\xi_j, \mathbf{a}_j, \mathbf{y}_j, \mathbf{x}_j)$ the j-th columns of Σ, A, Y, X respectively.

The mutual information of the channel at the noise level σ^2 is given by

$$I(\sigma^2) = \log\det\left[I + \frac{\Sigma\Sigma^H}{\sigma^2}\right] \quad (4)$$

and can also be written as

$$I(\sigma^2) = N\int_{\sigma^2}^{+\infty}\left[\frac{1}{\omega^2} - \frac{1}{N}\mathrm{Tr}(\Sigma\Sigma^H + \omega^2 I)^{-1}\right]d\omega^2 \quad (5)$$

In the following, we study the asymptotic behaviour of $I(\sigma^2)$ when N and n converge to $+\infty$ in such a way that $\frac{n}{N} \to \alpha$, $0 < \alpha < +\infty$. For this, Eq. (5) shows that it is sufficient to study $\frac{1}{N}\mathrm{Tr}(\Sigma\Sigma^H + \sigma^2 I)^{-1}$, i.e. the Stieljès transform of the empirical eigenvalue distribution of matrix $\Sigma\Sigma^H$. We denote by $Q(\sigma^2)$ and $\tilde{Q}(\sigma^2)$ the so-called resolvents of $\Sigma\Sigma^H$ and $\Sigma^H\Sigma$ respectively, i.e.

$$Q(\sigma^2) = (\Sigma\Sigma^H + \sigma^2 I)^{-1},\ \tilde{Q}(\sigma^2) = (\Sigma^H\Sigma + \sigma^2 I)^{-1} \quad (6)$$

$f(\sigma^2)$ and $\tilde{f}(\sigma^2)$ represent the normalized traces of $Q(\sigma^2)$ and $\tilde{Q}(\sigma^2)$, i.e. $f(\sigma^2) = \frac{1}{N}\mathrm{Tr}(Q(\sigma^2))$ and $\tilde{f}(\sigma^2) = \frac{1}{n}\mathrm{Tr}(\tilde{Q}(\sigma^2))$. In order to simplify the notations of this paper, we omit to mention that all the above terms depend on n and N. However, here and there, it will be important to indicate this dependence, and will index $\Sigma, A, Y, Q, f, \ldots$ by n only. In particular, the symbol $n \to +\infty$ should be understood as n and N converge to $+\infty$ in such a way $\frac{n}{N} \to \alpha$.

III. Asymptotic behaviour of $f_n(\sigma^2)$

In this section, we present some convergence results concerning $f_n(\sigma^2) = \frac{1}{N}\mathrm{Tr}(Q_n(\sigma^2))$. As Σ_n in non zero mean, only few mathematical results are available to address the problem. Girko [6] and Dozier-Silverstein [4] have considered the case where D_n and $\tilde{D}_n$ are reduced to I, which corresponds to an uncorrelated Rayleigh part. Under the assumption that the eigenvalue distribution of $A_nA_n^H$ converges to a limit distribution, [4] showed that the eigenvalue distribution of $\Sigma_n\Sigma_n^H$ converges to a limit deterministic distribution μ_*. This in particular implies that for each σ^2, $\lim_{n\to+\infty} f_n(\sigma^2) = t_*(\sigma^2)$ where t_* represents the Stieljès transform of this distribution, i.e. $t_*(\sigma^2) = \int_0^{+\infty}\frac{1}{\lambda+\sigma^2}\,d\mu_*(\lambda)$. The case where D_n and $\tilde{D}_n$ do not coincide with I can be addressed using the results of [7], chap. 7, only valid if $\sup_{n,i}\sum_{j=1}^n |A_{n,ij}| < +\infty$ and $\sup_{n,j}\sum_{i=1}^N |A_{n,ij}| < +\infty$. The eigenvalue distribution of $\Sigma_n\Sigma_n^H$ does not converge to a limit, but it can be approximated by a deterministic distribution, which, in general, does not converge. This means that for each n, there exists a deterministic probability measure μ_n carried by $\mathbb{R}^+$ for which, almost surely

$$\lim_{n\to+\infty} f_n(\sigma^2) - t_n(\sigma^2) = 0 \quad (7)$$

for each σ^2, where t_n represents the Stieljès transform of μ_n, that can be calculated by solving certain equations (see below). However, $\lim_{n\to+\infty} t_n(\sigma^2)$ does not exist in general. In the following, $t_n(\sigma^2)$ will be refered to as the "deterministic equivalent" of $f_n(\sigma^2) = \frac{1}{N}\mathrm{Tr}(Q(\sigma^2))$.

Usual line of sight components models (see e.g. (1)) do not satisfy $\sup_{n,i}\sum_{j=1}^n |A_{n,ij}| < +\infty$ and $\sup_{n,j}\sum_{i=1}^N |A_{n,ij}| < +\infty$. Therefore, we have addressed the problem considered in [7] in the case where

Assumption 1: $\sup_{n,i}\sum_{j=1}^n |A_{n,ij}|^2 < +\infty$, $\sup_{n,j}\sum_{i=1}^N |A_{n,ij}|^2 < +\infty$

a much more realistic condition verified in particular by example (1). For this, we have considered in [9] a different approach. In the rest of this section, we sketch a simpler method of proof of the main results of [9].

Theorem 1: For σ^2 fixed, consider the system of equations

$$\begin{cases} \kappa = \frac{1}{n}\mathrm{Tr}\left[D\left(\sigma^2(I + D\tilde{\kappa}) + A(I + \tilde{D}\kappa)^{-1}A^H\right)^{-1}\right] \\ \tilde{\kappa} = \frac{1}{n}\mathrm{Tr}\left[\tilde{D}\left(\sigma^2(I + \tilde{D}\kappa) + A^H(I + D\tilde{\kappa})^{-1}A\right)^{-1}\right] \end{cases}. \quad (8)$$

Then, equations (8) have unique positive solutions $(\delta(\sigma^2), \tilde{\delta}(\sigma^2))$. We denote by $T(\sigma^2)$ and $\tilde{T}(\sigma^2)$ the following matrix valued functions:

$$\begin{aligned} T(\sigma^2) &= \left[\sigma^2(I + \tilde{\delta}D) + A(I + \delta\tilde{D})^{-1}A^H\right]^{-1} \\ \tilde{T}(\sigma^2) &= \left[\sigma^2(I + \delta\tilde{D}) + A^H(I + \tilde{\delta}D)^{-1}A\right]^{-1} \end{aligned} \quad (9)$$

Then,

$$\begin{aligned} \delta(\sigma^2) &= \frac{1}{n}\mathrm{Tr}\left[DT(\sigma^2)\right] \\ \tilde{\delta}(\sigma^2) &= \frac{1}{n}\mathrm{Tr}\left[\tilde{D}\tilde{T}(\sigma^2)\right] \end{aligned} \quad (10)$$

Moreover, $t(\sigma^2) = \frac{1}{N}\mathrm{Tr}(T(\sigma^2))$ coincides with the Stieljès transform of a probability measure μ, and, under Assumption 1, almost surely, $f(\sigma^2) - t(\sigma^2)$ converges to 0 for each σ^2. •

We omit to prove that (8) has unique solutions, and that t is the Stieljès transform of a probability measure. We first observe that using (9), Eq. (10) is equivalent to $(\delta(\sigma^2), \tilde{\delta}(\sigma^2))$ solution of (8). Next, we justify that almost surely, $f(\sigma^2) - t(\sigma^2)$ converges to 0. For this, 3 different steps are necessary.

The first step, introduced by Girko in other contexts, consists in showing that $f(\sigma^2) - \mathbb{E}(f(\sigma^2))$ converges to 0 almost surely. For this, Girko observed that the above term can be written as the sum of n martingale increments. More precisely, we denote by $\mathbb{E}_j$ the conditional expectation operator over the σ algebra generated by $\xi_j, l \geq j$; operator $\mathbb{E}_{n+1}$ is defined as the conventional mathemitical expectation. Matrix $\Sigma^{(j)}$ is the $N \times (n-1)$ matrix obtained by deleting the j-th column ξ_j from Σ, and $Q^{(j)}(\sigma^2)$ and $f^{(j)}(\sigma^2)$ are defined by $Q^{(j)}(\sigma^2) = (\Sigma^{(j)}\Sigma^{(j)H} + \sigma^2 I)^{-1}$ and $f^{(j)}(\sigma^2) = \frac{1}{N}\mathrm{Tr}(Q^{(j)}(\sigma^2))$. Then, it is clear that

$$f(\sigma^2) - \mathbb{E}(f(\sigma^2)) = \sum_{j=1}^n (\mathbb{E}_j - \mathbb{E}_{j+1})f(\sigma^2)$$

We denote by $\gamma_j(\sigma^2)$ the term $(\mathbb{E}_j - \mathbb{E}_{j+1})f(\sigma^2)$. Then, the $(\gamma_j(\sigma^2))_{j=1,\ldots n}$ are martingale increments in the sense that $\gamma_j(\sigma^2)$ depends on the $(\xi_k)_{k\geq j}$ and $\mathbb{E}_{j+1}(\gamma_j(\sigma^2)) = 0$. This in particular implies that

$$\mathbb{E}\left|f(\sigma^2) - \mathbb{E}(f(\sigma^2))\right|^2 = \sum_{j=1}^n \mathbb{E}|\gamma_j(\sigma^2)|^2$$

In order to show that $f(\sigma^2) - \mathbb{E}(f(\sigma^2))$ converges almost surely to 0, it is sufficient to show that $\mathbb{E}\left|f(\sigma^2) - \mathbb{E}(f(\sigma^2))\right|^2$ converges to 0 faster than $\frac{1}{n}$. In fact, it turns out that

Lemma 1:

$$\mathbb{E}\left|f(\sigma^2) - \mathbb{E}(f(\sigma^2))\right|^2 = O(\frac{1}{n^2}) \quad (11)$$

To show this, it is sufficient to establish that $\sup_j \mathbb{E}|\gamma_j(\sigma^2)|^2 = O(\frac{1}{n^3})$. The trick is to observe that, as $f^{(j)}(\sigma^2)$ is independent of ξ_j, then $(\mathbb{E}_j - \mathbb{E}_{j+1})f^{(j)}(\sigma^2) = 0$. Therefore,

$$\gamma_j(\sigma^2) = (\mathbb{E}_j - \mathbb{E}_{j+1})(f(\sigma^2) - f^{(j)}(\sigma^2))$$

Using the matrix inversion lemma, it is easily seen that $f(\sigma^2) - f^{(j)}(\sigma^2) = \frac{1}{N}\frac{\xi_j^H (Q^{(j)})^2 \xi_j}{1+\xi_j^H Q^{(j)}\xi_j}$. It follows that

$$\gamma_j(\sigma^2) = (\mathbb{E}_j - \mathbb{E}_{j+1})\frac{1}{N}\frac{\xi_j^H (Q^{(j)})^2 \xi_j}{1+\xi_j^H Q^{(j)}\xi_j} \tag{12}$$

In order to evaluate γ_j, we introduce

$$\varepsilon_j = \xi_j^H Q^{(j)}\xi_j - \tilde{d}_j \frac{1}{n}\mathrm{Tr}(DQ^{(j)}) - \mathbf{a}_j^H Q^{(j)}\mathbf{a}_j \tag{13}$$

Writing $\xi_j = \mathbf{a}_j + \mathbf{y}_j$ and $\mathbf{y}_j = \frac{1}{\sqrt{n}}\tilde{d}_j^{1/2} D^{1/2}\mathbf{x}_j$, we get that the second term of the righthandside of (13) coincides with the mathematical expectation of $\xi_j^H Q^{(j)}\xi_j$ w.r.t. the probability distribution of $\mathbf{y}_j$. It is clear that $1+\xi_j^H Q^{(j)}\xi_j = 1+\tilde{d}_j\frac{1}{n}\mathrm{Tr}(DQ^{(j)}) + \mathbf{a}_j^H Q^{(j)}\mathbf{a}_j + \varepsilon_j$. We claim that $\sup_j \mathbb{E}|\varepsilon_j|^2 = O(\frac{1}{n})$. To see this, we note that

$$\varepsilon_j = \mathbf{y}_j^H Q^{(j)}\mathbf{y}_j - \frac{1}{n}\tilde{d}_j \mathrm{Tr}(DQ^{(j)}) + \mathbf{y}_j^H Q^{(j)}\mathbf{a}_j + \mathbf{a}_j^H Q^{(j)}\mathbf{y}_j \tag{14}$$

We study the 3 terms of the righthandside of (14), and begin by the second one. As matrix $Q^{(j)}$ satisfies $Q^{(j)} \leq \frac{I}{\sigma^2}$, it is clear that $\mathbb{E}|\mathbf{y}_j^H Q^{(j)}\mathbf{a}_j|^2 \leq \frac{1}{n}\frac{\tilde{d}_j}{\sigma^2}\mathbf{a}_j^H D\mathbf{a_j}$. As the entries of D and the norms of $\mathbf{a}_j$ are uniformly bounded w.r.t. n, it is clear that $\sup_j \mathbb{E}|\mathbf{y}_j^H Q^{(j)}\mathbf{a}_j|^2 = O(\frac{1}{n})$. The third term has a similar behaviour. The first term can be written as $\frac{1}{n}\tilde{d}_j \mathbf{x}_j^H D^{1/2} Q^{(j)} D^{1/2}\mathbf{x}_j - \frac{1}{n}\tilde{d}_j \mathrm{Tr}(DQ^{(j)})$. The components of $\mathbf{x}_j$ are iid with variance 1, matrix $\tilde{d}_j D^{1/2} Q^{(j)} D^{1/2}$ is independent of the entries of $\mathbf{x}_j$ and uniformly bounded w.r.t. n and j. Using a well known result of [1], it turns out that the supremum over j of the variance of the first term is $O(\frac{1}{n})$. This in turn establishes that $\sup_j \mathbb{E}|\varepsilon_j|^2 = O(\frac{1}{n})$.

We pursue the proof of Theorem 1, and remark that

$$\frac{1}{1+\xi_j^H Q^{(j)}\xi_j} = \frac{1}{1+\tilde{d}_j\frac{1}{n}\mathrm{Tr}(DQ^{(j)}) + \mathbf{a}_j^H Q^{(j)}\mathbf{a}_j} - \frac{\varepsilon_j}{(1+\tilde{d}_j\frac{1}{n}\mathrm{Tr}(DQ^{(j)}) + \mathbf{a}_j^H Q^{(j)}\mathbf{a}_j)(1+\xi_j^H Q^{(j)}\xi_j)} \tag{15}$$

Hence, using (12), we get that

$$N\gamma_j = (\mathbb{E}_j - \mathbb{E}_{j+1})\left[\frac{\xi_j^H (Q^{(j)})^2\xi_j}{1+\tilde{d}_j\frac{1}{n}\mathrm{Tr}(DQ^{(j)}) + \mathbf{a}_j^H Q^{(j)}\mathbf{a}_j}\right] - (\mathbb{E}_j - \mathbb{E}_{j+1})\left[\frac{\varepsilon_j \xi_j^H (Q^{(j)})^2\xi_j}{(1+\tilde{d}_j\frac{1}{n}\mathrm{Tr}(DQ^{(j)}) + \mathbf{a}_j^H Q^{(j)}\mathbf{a}_j)(1+\xi_j^H Q^{(j)}\xi_j)}\right] \tag{16}$$

Using that $(1+\tilde{d}_j\frac{1}{n}\mathrm{Tr}(DQ^{(j)}) \geq 1$, $(1+\xi_j^H Q^{(j)}\xi_j) \geq 1$, and $\sup_j \mathbb{E}|\varepsilon_j|^2 = O(\frac{1}{n})$, it is possible to show that the variances of the terms of the righthandside of the above equation are bounded by a term such as $\frac{K}{n}$, where K is independent of j and n (but depend on σ^2). This, in turn, shows that $\sup_j \mathbb{E}|N\gamma_j|^2 = O(\frac{1}{n})$, i.e. that $\sup_j \mathbb{E}|\gamma_j(\sigma^2)| = O(\frac{1}{n^3})$ as expected. •

We now sketch the second step of the proof. At this stage of the proof, it is not yet possible to study directly $\frac{1}{N}\mathrm{Tr}(Q(\sigma^2)) - \frac{1}{N}\mathrm{Tr}(T(\sigma^2))$. We introduce functions $c(\sigma^2), \tilde{c}(\sigma^2)$, and $R(\sigma^2), \tilde{R}(\sigma^2)$ which play an intermediate role between $\frac{1}{n}\mathrm{Tr}(DQ(\sigma^2))$ and $\delta(\sigma^2)$, $\frac{1}{n}\mathrm{Tr}(\tilde{D}\tilde{Q}(\sigma^2))$ and $\tilde{\delta}(\sigma^2)$, $Q(\sigma^2)$ and $T(\sigma^2)$, $\tilde{Q}(\sigma^2)$ and $\tilde{T}(\sigma^2)$ respectively. We define:

$$\begin{aligned} c(\sigma^2) &= \frac{1}{n}\mathrm{Tr}(D\mathbb{E}(Q(\sigma^2))) \\ \tilde{c}(\sigma^2) &= \frac{1}{n}\mathrm{Tr}(\tilde{D}\mathbb{E}(\tilde{Q}(\sigma^2)) \end{aligned} \tag{17}$$

$$\begin{aligned} R(\sigma^2) &= \left[\sigma^2(I+\tilde{c}(\sigma^2)D) + A(I+c(\sigma^2)\tilde{D})^{-1}A^H\right]^{-1} \\ \tilde{R}(\sigma^2) &= \left[\sigma^2(I+c(\sigma^2)\tilde{D}) + A^H(I+\tilde{c}(\sigma^2)D)^{-1}A\right]^{-1} \end{aligned} \tag{18}$$

Then, the following result holds.

Proposition 1: If U and $\tilde{U}$ are uniformly bounded (w.r.t. n) matrices, then, for each σ^2,

$$\frac{1}{n}\mathrm{Tr}\left[(\mathbb{E}(Q(\sigma^2)) - R(\sigma^2))U\right] \rightarrow 0 \tag{19}$$

$$\frac{1}{n}\mathrm{Tr}\left[(\mathbb{E}(\tilde{Q}(\sigma^2)) - \tilde{R}(\sigma^2))\tilde{U}\right] \rightarrow 0 \tag{20}$$

We just sketch the proof of this result. We first remark that $R-Q$ can be written as

$$(R-Q)U = Q(Q^{-1}-R^{-1})RU = Q(\Sigma\Sigma^H - \sigma^2\tilde{c}D - \sum_{j=1}^{n}\frac{1}{1+c\tilde{d}_j}\mathbf{a}_j\mathbf{a}_j^H)RU \tag{21}$$

Writing $\Sigma\Sigma^H = \sum_{j=1}^n (\mathbf{a}_j+\mathbf{y}_j)(\mathbf{a}_j+\mathbf{y}_j)^H$, evaluating the trace of the righthandside of (21), using the identity $Q = Q^{(j)} - \frac{Q^{(j)}\xi_j\xi_j^H Q^{(j)}}{1+\xi_j^H Q^{(j)}\xi_j}$, we get, after tedious manipulations, that

$$\frac{1}{N}\mathrm{Tr}\left[(\mathbb{E}(Q(\sigma^2)) - R(\sigma^2))U\right] = \frac{1}{N}\mathbb{E}(Z_1+Z_2+Z_3+Z_4+Z_5) \tag{22}$$

where

$$\begin{aligned} Z_1 &= \textstyle\sum_{j=1}^n \frac{\mathbf{y}_j^H RUQ^{(j)}\mathbf{a}_j}{1+\xi_j^H Q^{(j)}\xi_j} \\ Z_2 &= \textstyle\sum_{j=1}^n \frac{\mathbf{a}_j^H RUQ^{(j)}\mathbf{y}_j}{1+\xi_j^H Q^{(j)}\xi_j}\left(1+\frac{1}{1+c\tilde{d}_j}(\mathbf{a}_j^H Q^{(j)}\mathbf{a}_j + \mathbf{y}_j^H Q^{(j)}\mathbf{a}_j)\right) \\ Z_3 &= -\textstyle\sum_{j=1}^n \frac{\mathbf{a}_j^H RUQ^{(j)}\mathbf{a}_j}{1+\xi_j^H Q^{(j)}\xi_j}\frac{1}{1+c\tilde{d}_j}\mathbf{a}_j^H Q^{(j)}\mathbf{y}_j \\ Z_4 &= \textstyle\sum_{j=1}^n \frac{\mathbf{a}_j^H RUQ^{(j)}\mathbf{a}_j}{1+\xi_j^H Q^{(j)}\xi_j}\left(1-\frac{1}{1+c\tilde{d}_j}(1+\mathbf{y}_j^H Q^{(j)}\mathbf{y}_j)\right) \\ Z_5 &= \textstyle\sum_{j=1}^n \frac{\mathbf{y}_j^H RUQ^{(j)}\mathbf{y}_j}{1+\xi_j^H Q^{(j)}\xi_j} - \sigma^2\mathrm{Tr}(\tilde{c}D) \end{aligned} \tag{23}$$

To establish that the righthandside of (22) converges to 0, it is sufficient to show that $\frac{1}{N}\mathbb{E}(Z_l) \rightarrow 0$ for $l = 1,\ldots,5$. This is tedious, but rather easy. We just check that $\frac{1}{N}\mathbb{E}(Z_4)$ converges to 0. For this, we remark that $Z_4 = \sum_{j=1}^n Z_{4,j}$ where the definition of $Z_{4,j}$ is obvious, and verify that $\sup_j(\mathbb{E}|Z_{4,j}|^2)^{1/2} \rightarrow 0$. If this holds, the Minkowski inequality gives immediately that $\left(\mathbb{E}|\frac{1}{N}\sum_{j=1}^n Z_{4,j}|^2\right)^{1/2} \rightarrow 0$, and $\frac{1}{N}\mathbb{E}(Z_4) \rightarrow 0$. To check that $\sup_j(\mathbb{E}|Z_{4,j}|^2)^{1/2} \rightarrow 0$, we remark that $\frac{1}{1+\xi_j^H Q^{(j)}\xi_j} \leq 1$. Moreover, it is clear from the definition of R that $R \leq \frac{I}{\sigma^2}$. As $Q^{(j)}$ also satisfies this inequality, we get that $|\frac{\mathbf{a}_j^H RUQ^{(j)}\mathbf{a}_j}{1+\xi_j^H Q^{(j)}\xi_j}| \leq \frac{1}{\sigma^4}||U||\mathbf{a}_j^H\mathbf{a}_j$. This implies that this term is uniformly bounded w.r.t. j and n. It is therefore sufficient to verify that $\sup_j(\mathbb{E}|W_{4,j}|^2)^{1/2} \rightarrow 0$, where

$$W_{4,j} = 1 - \frac{1}{1+c\tilde{d}_j}(1+\mathbf{y}_j^H Q^{(j)}\mathbf{y}_j) = -\frac{\mathbf{y}_j^H Q^{(j)}\mathbf{y}_j - c\tilde{d}_j}{1+c\tilde{d}_j}$$

Using [1], we obtain that $\sup_j \mathbb{E}|\mathbf{y}_j^H Q^{(j)}\mathbf{y}_j - \frac{1}{n}\tilde{d}_j\mathrm{Tr}(DQ^{(j)})|^2 \rightarrow 0$. It remains to remark that $\sup_j \mathbb{E}|\frac{1}{n}\mathrm{Tr}(DQ^{(j)}) - \frac{1}{n}\mathrm{Tr}(DQ)|^2 = O(\frac{1}{n})$ and that $\mathbb{E}|\frac{1}{n}\mathrm{Tr}(DQ) - c|^2 = O(\frac{1}{n^2})$ by a straightforward extension of Lemma (1). As $\frac{1}{1+c\tilde{d}_j} \leq 1$, this proves that $\sup_j(\mathbb{E}|W_{4,j}|^2)^{1/2} \rightarrow 0$. •

The step 3 establishes connections between c and δ, R and T, and $\tilde{c}$ and $\tilde{\delta}$, $\tilde{R}$ and $\tilde{T}$. For this, we use Proposition (1) for $U = D$ and $\tilde{U} = \tilde{D}$. It follows that $c(\sigma^2) = \frac{1}{n}\mathrm{Tr}(DR(\sigma^2)) + \varepsilon(\sigma^2)$ and $\tilde{c}(\sigma^2) = \frac{1}{n}\mathrm{Tr}(\tilde{D}\tilde{R}(\sigma^2)) + \tilde{\varepsilon}(\sigma^2)$, where $\varepsilon(\sigma^2)$ and $\tilde{\varepsilon}(\sigma^2)$ converge to 0. Matrices T and $\tilde{T}$ can be interpreted as the solutions of (9) and (10). The definitions of R and $\tilde{R}$ show that, up to the terms

$\varepsilon(\sigma^2)$ and $\tilde{\varepsilon}(\sigma^2)$, R and $\tilde{R}$ satisfy the same equations than T and $\tilde{T}$. This is the key observation to prove that

$$\frac{1}{N}\mathrm{Tr}\left[R(\sigma^2) - T(\sigma^2)\right] \quad \rightarrow \quad 0 \tag{24}$$

$$\frac{1}{n}\mathrm{Tr}\left[\tilde{R}(\sigma^2) - \tilde{T}(\sigma^2)\right] \quad \rightarrow \quad 0 \tag{25}$$

In sum, we have proved in step 1 that $\frac{1}{N}\mathrm{Tr}(Q(\sigma^2) - \mathbb{E}(Q(\sigma^2)) \rightarrow 0$ (see Lemma 1). In step 2, we have established Proposition 1, which for $U = I$, gives $\frac{1}{N}\mathrm{Tr}(\mathbb{E}(Q(\sigma^2)) - R(\sigma^2)) \rightarrow 0$. Finally, we have indicated in step 3 why $\frac{1}{N}\mathrm{Tr}\left[R(\sigma^2) - T(\sigma^2)\right] \rightarrow 0$ should hold. This, in turn, shows that $\frac{1}{N}\mathrm{Tr}(Q(\sigma^2)) - \frac{1}{N}\mathrm{Tr}(T(\sigma^2))$ converges to 0.

IV. Study of the mutual information.

We now briefly indicate how it is possible to address the asymptotic behavior of $I(\sigma^2)$. In the context of simpler models, Girko's approach still uses a martingale difference representation. More precisely, $I(\sigma^2) - \mathbb{E}(I(\sigma^2))$ can be written as

$$I(\sigma^2) - \mathbb{E}(I(\sigma^2)) = \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j+1}) I(\sigma^2)$$

Denote by $I^{(j)}(\sigma^2)$ the mutual information associated to the "channel" $\Sigma^{(j)}$ defined by $I^{(j)}(\sigma^2) = \log\det(I + \frac{\Sigma^{(j)}\Sigma^{(j)H}}{\sigma^2})$. It is straightforward that $(\mathbb{E}_j - \mathbb{E}_{j+1}) I^{(j)}(\sigma^2) = 0$, so that

$$I(\sigma^2) - \mathbb{E}(I(\sigma^2)) = \sum_{j=1}^{n} (\mathbb{E}_j - \mathbb{E}_{j+1})(I(\sigma^2) - I^{(j)}(\sigma^2))$$

Using the Schur formula, we obtain immediately that $I(\sigma^2) - I^{(j)}(\sigma^2) = \log\sigma^2 + \log(1 + \xi_j^H Q^{(j)} \xi_j)$. This can probably be used to evaluate the variance of $I(\sigma^2)$, and, using central limit theorem for sum of martingale increments, to derive Gaussian approximation like results for $I(\sigma^2)$.

It is also interesting to study the average mutual information $\mathbb{E}(I(\sigma^2))$, which by (5), is given by

$$\mathbb{E}(I(\sigma^2)) = N \int_{\sigma^2}^{+\infty} (\frac{1}{\omega^2} - f(\omega^2))\, d\omega^2$$

As $f(\omega^2) - t(\omega^2)$ converges to 0, it is clear that $\mathbb{E}(I(\sigma^2))$ can be approximated by $\bar{I}(\sigma^2)$ defined by

$$\bar{I}(\sigma^2) = N \int_{\sigma^2}^{+\infty} (\frac{1}{\omega^2} - t(\omega^2))\, d\omega^2$$

We first mention that $\bar{I}(\sigma^2)$ can be evaluated more explicitely. In effect, it can be shown (see [9]) that

$$\begin{aligned} \bar{I}(\sigma^2) = & \log\det\left[I + \tilde{\delta}(\sigma^2) D + \tfrac{1}{\sigma^2} A (I + \delta(\sigma^2)\tilde{D})^{-1} A^H\right] \\ & + \log\det\left[I + \delta(\sigma^2)\tilde{D}\right] - \sigma^2 \tfrac{n}{N} \delta(\sigma^2)\tilde{\delta}(\sigma^2) \end{aligned} \tag{26}$$

Second, it is of course important to evaluate the behaviour of the error $\mathbb{E}(I(\sigma^2)) - \bar{I}(\sigma^2)$. For this, we again use (5), and get that

$$\bar{I}(\sigma^2) - \mathbb{E}(I(\sigma^2)) = \int_{\sigma^2}^{+\infty} \left[\mathrm{Tr}(\mathbb{E}(Q(\omega^2)) - \mathrm{Tr}(T(\omega^2))\right] d\omega^2$$

It is therefore useful to study $\mathrm{Tr}(\mathbb{E}(Q(\sigma^2)) - \mathrm{Tr}(T(\sigma^2))$. Using an approach similar to the step 3 of the proof of Theorem 1, one can first show that $\mathrm{Tr}(\mathbb{E}(Q(\sigma^2) - T(\sigma^2))$ has the same behaviour than $\mathrm{Tr}(\mathbb{E}(Q(\sigma^2)) - R(\sigma^2))$. The study of this term can be addressed by refining the step 2 of the proof of Theorem 1. We first we recall that

$$\mathrm{Tr}(\mathbb{E}(Q(\sigma^2)) - R(\sigma^2)) = \mathbb{E}(Z_1 + Z_2 + Z_3 + Z_4 + Z_5)$$

In contrast to the analysis of section III, $\frac{1}{1+\xi_j^H Q^{(j)} \xi_j}$ has to be expanded at the first order (see 15) or at the second order to capture the order of magnitude of $\mathrm{Tr}(\mathbb{E}(Q(\sigma^2)) - R(\sigma^2))$. In the context $A = 0$ and $\tilde{D} = I$, [2] showed this term is in general $O(1)$, but converges to 0 in the complex Gaussian case. Under the extra assumption that A is uniformly bounded, we have been able to generalize this result to the present context. Moreover, we have shown that $\mathrm{Tr}(\mathbb{E}(Q(\sigma^2)) - R(\sigma^2)) = O(\frac{1}{n^{1/2}})$, but believe that it is possible to show with some efforts that $\mathrm{Tr}(\mathbb{E}(Q(\sigma^2)) - R(\sigma^2)) = O(\frac{1}{n})$. This would be in accordance with the evaluations obtained using the replica method in the contexte $A = 0$ ([15]), and would explain why the approximant is reliable even for quite moderate values of n and N.

V. Conclusion

In this paper, we have briefly shown how it is possible to study rigorously the asymptotic behaviour of the mutual information of Rician correlated MIMO channels. We believe that the proposed approach is simpler than expected, and perhaps easier to follow than the replica method.

References

[1] Z. Bai, J. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices", Ann. Probab., 26(1):316-345, 1998.

[2] Z. Bai, J. Silverstein, "CLT for linear statistics of large-dimensional sample covariance matrices", Ann. Probab., 32(1A):553-605, 2004.

[3] L. Cottatellucci, M. Debbah, "The Effect of Line of Sight Components on the Asymptotic Capacity of MIMO Systems", in Proc. ISIT 04, Chicago, June 27-July 2 2004.

[4] B. Dozier, J. Silverstein, "On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices", Preprint.

[5] J.Dumont, Ph.Loubaton, S.Lasaulce, M.Debbah, "On the Asymptotic Performance of MIMO Correlated Ricean Channels", Proc ICASSP05, vol. 5, pp. 813-816, March 2005.

[6] V.L. Girko, "Introduction to Statistical Analysis of Random Arrays", Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[7] V.L. Girko, "Theory of Stochastic Canonical Equations, Volume I", Chap. 7, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001.

[8] A. Goldsmith, S.A. Jafar, N. Jindal, S. Vishwanath, "Capacity Limits of MIMO Channels", IEEE J. Sel. Areas in Comm.,vol.21,no 5,June 2003

[9] W. Hachem, Ph. Loubaton, J. Najim, "Deterministic equivalents for certain functionals of large random matrices", Preprint.

[10] D. Hoesli, K. Young-Han, A. Lapidoth, "IEEE Trans. on Information Theory, vol. 51, no. 12, pp. 4334-4339, December 2005.

[11] S.K. Jayaweera, H.V. Poor, "MIMO capacity results for Rician fading channels", Proc. Globecom03, vol. 4, pp. 1806-1810, December 2003.

[12] G. Lebrun, M. Faulkner, M. Shafi , P.J. Smith, "MIMO Ricean channel capacity", Proc. ICC 2004, vol. 5, pp. 2939-2943, June 2004.

[13] A. Lozano, A.M. Tulino, S. Verdú, "Multiple-Antenna Capacity in the Low-Power Regime", Trans. on Inf. Theo., vol.49,no 10,pp 2527-2544, Oct. 03

[14] M.R. McKay, I.B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels", IEEE Trans. on Information Theory, vol. 51, no. 9, pp. 3121-3145, September 2005.

[15] A.L. Moustakas, S.H. Simon, A.M. Sengupta, "MIMO Capacity Through Correlated Channels in the Presence of Correlated Interference and Noise : A (Not so) Large N Analysis", Trans. on Inf. Theo., vol.49,no 10,pp 2545-2561, Oct. 03.

[16] A.L. Moustakas, S.H. Simon, "Random matrix theory of multi-antenna communications: the Ricean case", J. Phys. A: Math. Gen. 38 (2005) 10859-10872.