ON THE FLUCTUATIONS OF THE MUTUAL INFORMATION FOR NON CENTERED MIMO CHANNELS: THE NON GAUSSIAN CASE

Abla Kammoun$^{(1)}$, *Malika Kharouf*$^{(1,2)}$, *Walid Hachem*$^{(3)}$, *Jamal Najim*$^{(3)}$ *and Ahmed El Kharroubi*$^{(2)}$

$^{(1)}$TELECOM ParisTech (Paris, France),
$^{(2)}$Hassan II University (Casablanca, Morocco),
$^{(3)}$CNRS LTCI ; TELECOM ParisTech,
kammoun,kharouf,hachem,najim@telecom-paristech.fr
elkharroubi@facsc-achok.ac.ma

ABSTRACT

The use of Multiple Input Multiple Output (MIMO) systems has been widely recognized as an efficient solution to increase the data rate of wireless communications. In this regard, several contributions investigate the performance improvement of MIMO systems in terms of Shannon's mutual information. In most of these contributions, elements of the MIMO channel matrix are assumed to belong to a multivariate Gaussian distribution. The non Gaussian case, which is realistic in many practical environments, has been much less studied. This contribution is devoted to the study of the mutual information of MIMO channels when the channel matrix elements are Ricean with the non-Ricean component being iid but non-Gaussian. In this context, the mutual information behavior is studied in the large dimensional regime where both channel matrix dimensions converge to infinity at the same pace. In this regime, a Central Limit Theorem on the mutual information is provided. In particular, the mutual information variance is determined in terms of the parameters of the channel statistical model. Since non Gaussian entries are allowed, a new term proportional to the fourth cumulant of their distribution arises in the expression of the asymptotic variance. In addition, a bias term proportional to this fourth order cumulant appears.

1. INTRODUCTION

It is widely known that high spectral efficiencies are attained when multiple antennas are used at both the transmitter and the receiver of a wireless communication system. Consider the classical transmission model $\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{v}$, where $\mathbf{y}$ is the received signal, $\mathbf{s}$ is the vector of transmitted symbols, $\mathbf{v}$ is a complex white Gaussian noise, and $\mathbf{H}$ is the $N \times n$ Multiple Input Multiple Output (MIMO) channel matrix with N antennas at the receiver's site and n antennas at the transmitter's. Due to the mobility and to the presence of a large number of reflected and scattered signal paths, the elements of the channel matrix $\mathbf{H}$ are often modeled as random variables. Assuming a random model for this matrix, Telatar [1] and Foschini [2] realized in the mid-nineties that Shannon's mutual information of such channels increases at the rate of $\min(N, n)$. The authors of [1] and [2] assumed that the elements of the channel matrix $\mathbf{H}$ are centered, independent and identically distributed (i.i.d.) elements. In this context, a well-known result in Large Random Matrix Theory (LRMT) [3] states that the eigenvalue distribution of the Gram matrix $\mathbf{H}\mathbf{H}^*$ where $\mathbf{H}^*$ is the Hermitian adjoint of $\mathbf{H}$ converges to a deterministic probability distribution as n goes to infinity and N/n converges to a positive constant. Recalling that Shannon's mutual information per receiver antenna is $N^{-1}\mathbb{E} \log\det(\mathbf{I} + \rho^{-1}\mathbf{H}\mathbf{H}^*)$ where $\rho > 0$ is the additive noise variance, one consequence of [3] is that this mutual information converges to a constant. This fact already observed in [1] sustains the assertion of a linear increase of mutual information with the number of antennas. In addition, this convergence proves to be sufficiently fast. As a matter of fact, the asymptotic results predicted by LRMT remain relevant for systems with a moderate number of antennas.

The next step was to apply this theory for channel models that include correlations between the elements of $\mathbf{H}$. In this context, centered matrix channels have been first studied. The first order behaviour (LRMT asymptotic approximation of the mutual information) as well as the second order (Central Limit Theorem) have been established for channels with Gaussian elements whose variances are described by the so called Kronecker model [4,5]. Recently, the general variance profile has been considered and the Gaussian assumption has been raised [6]. It should be noted that these results deal with the case where the channel has no line of sight nor a reflected component, implying that the channel elements are centered random variables. The non centered case has been considered in *e.g.* [7, 8]. In these contributions, the non centered Gaussian distribution with a Kronecker variance profile has been treated, and the first order approximation as well as the variance of the $\log\det$ functional have been derived. The approach of [7, 8] is based on the replica method, which can be used as long as Gaussian distributions are considered. The results obtained by this method are relevant, but its assumptions are not always rigorous.

In this paper, we consider non centered channels which are not necessarily Gaussian. Our study, based on the martingale method, leads to an improved fit of the LRMT deterministic approximation with the true mutual information in many practical situations, especially met in severe fading cases. Among the most commonly used non Gaussian distributions, we mention the m-Nakagami, the Weibull and the Lognormal distributions. Several works support the pertinence of these distributions to model real propagation channels. In particular, the Nakagami distribution is particularly suited to some urban multipath environments [9]. The Weibull distribution was shown to exhibit a good fit for both indoor and outdoor environments, [10, 11], whereas the Lognormal distribution is known to approximate shadowing effects.

Despite their importance for modelling line-of-sight radio chan-

This work was partially supported by the French program ANR-07-MDCO-012-01 "SESAME"

nels, these distributions have not been considered so far in the analysis of the mutual information. This motivates our work. In particular, we prove that in the LRMT asymptotic regime, the $\log\det$ functional fluctuates around its deterministic approximation as a Gaussian random variable. Since we allow for non Gaussian channels, a new term proportional to the fourth cumulant κ of the channel matrix elements arises in the expression of the asymptotic variance of the $\log\det$. Interestingly, this new correction of the variance not only depends on the singular value distribution of the MIMO channel mean matrix but also on its singular vectors. We also identify the singular vector matrices that maximize and minimize the asymptotic variance, and show in particular that the sign of the fourth cumulant plays a key role in this respect. In addition, a bias term also proportional to κ appears, and this term also depends on the singular values and the singular vectors of the channel mean matrix. Finally we test our findings on m Nakagami, Weibull and Lognormal distributions.

2. SYSTEM MODEL AND PROBLEM SETTING

We consider a MIMO system with N receiving antennas and n transmitting antennas. We denote by $\mathbf{H}$ the $N \times n$ channel matrix which is assumed to be given by:

$$\mathbf{H} = \frac{1}{\sqrt{n}} \left(\sqrt{\frac{1}{K+1}} \mathbf{X} + \sqrt{\frac{K}{K+1}} \mathbf{A} \right)$$

where the constant K is sometimes called the Rice factor, $\mathbf{X}$ is a random matrix with i.i.d entries with zero mean and unit variance, and $\mathbf{A}$ is a deterministic matrix accounting for the line of sight or reflected components. Denoting by ρ the additive Gaussian noise variance, the mutual information normalized by the number of receiver antennas is $\mathbb{E}\boldsymbol{\mathcal{I}}_n(\rho)$ where

$$\boldsymbol{\mathcal{I}}_n(\rho) = \frac{1}{N} \log\det \left(\mathbf{I}_N + \frac{1}{\rho} \mathbf{H}\mathbf{H}^* \right).$$

For fixed size dimensions, the study of the mutual information is rather difficult. Instead, we will consider the asymptotic regime defined as $N, n \to \infty$ in such a way that

$$0 < \liminf \frac{N}{n} \le \limsup \frac{N}{n} < \infty,$$

which we refer to as $n \to \infty$ for notation simplicity. However, even after relaxing the assumption of finite size dimensions, the mutual information still does not generally have a closed-form expression. Instead, the asymptotic approximation of the mutual information can be defined as the solution of a system of equations:

Theorem 1 ([12]). *For any $\rho > 0$, the deterministic system:*

$$\begin{cases} \delta(\rho) = \frac{1}{n} \mathrm{Tr} \mathbf{T}_n(\rho) & (1) \\ \widetilde{\delta}(\rho) = \frac{1}{n} \mathrm{Tr} \widetilde{\mathbf{T}}_n(\rho) & (2) \end{cases}$$

where $\mathbf{T}_n$ and $\widetilde{\mathbf{T}}_n$ are the matrices

$$\begin{cases} \mathbf{T}_n(\rho) = \left(\rho(K+1)(1+\widetilde{\delta}(\rho)) \mathbf{I}_N + \frac{K\mathbf{A}\mathbf{A}^*}{n(1+\delta(\rho))} \right)^{-1} & (3) \\ \widetilde{\mathbf{T}}_n(\rho) = \left(\rho(K+1)(1+\delta(\rho)) \mathbf{I}_n + \frac{K\mathbf{A}^*\mathbf{A}}{n(1+\widetilde{\delta}(\rho))} \right)^{-1} & (4) \end{cases}$$

admits a unique solution $(\delta, \widetilde{\delta})$ in $(0, \infty)^2$.

3. FIRST AND SECOND ORDER RESULTS

Having introduced matrices $\mathbf{T}_n(\rho)$ and $\widetilde{\mathbf{T}}_n(\rho)$, we are now in position to provide the first and second order results of the mutual information. The first order result is a by-product of theorem 4.1 in [12] and can be given as:

Theorem 2. *Assume $\sup_n \| \frac{1}{\sqrt{n}} \mathbf{A} \| < \infty$ where $\|.\|$ is the spectral norm. In the asymptotic regime, the following holds true:*

$$\boldsymbol{\mathcal{I}}_n(\rho) - V_n(\rho) \xrightarrow[n\to\infty]{} 0 \quad \textit{almost surely}$$

where

$$V_n(\rho) = \frac{1}{N} \log\det(\rho(K+1)(1+\widetilde{\delta})\mathbf{I}_N + \frac{K\mathbf{A}\mathbf{A}^*}{n(1+\delta)}) + \frac{n}{N}\log(1+\delta) - \frac{n\rho(K+1)}{N}\delta\widetilde{\delta} - \log(\rho(K+1)).$$

Remark 1. *It can be seen from Theorems 1 and 2 that the first order approximation $V_n(\rho)$ of $\boldsymbol{\mathcal{I}}_n(\rho)$ depends on the channel mean matrix $\mathbf{A}$ through its singular values only. This approximation would be the same if $\mathbf{H}$ were Gaussian.*

In addition to the first order result, another question regarding the fluctuations of the $\log\det$ around $V_n(\rho)$ is worth studying. Determining the nature of these fluctuations gives us insights about the variance of the $\log\det$ and the outage probability, which is the pertinent information theoretic performance measure when the channel is a slow fading channel. Under some mild technical assumptions, we can prove the following Central Limit Theorem:

Theorem 3. *Let $\gamma = \frac{1}{n}\mathrm{Tr}(\mathbf{T}^2)$, $\widetilde{\gamma} = \frac{1}{n}\mathrm{Tr}(\widetilde{\mathbf{T}}^2)$, $\mathbf{S} = \mathrm{diag}(\mathbf{T})$, and $\widetilde{\mathbf{S}} = \mathrm{diag}(\widetilde{\mathbf{T}})$. Let κ be the fourth cumulant of the entries of $\mathbf{X}$ given by $\kappa = \mathbb{E}|X_{1,1}|^4 - 2$. Define Δ_n as*

$$\Delta_n = \left(1 - \frac{K}{n^2(1+\delta)^2} \mathrm{Tr}(\mathbf{A}\mathbf{A}^*\mathbf{T}^2) \right)^2 - \rho^2(K+1)^2\gamma\widetilde{\gamma}$$

Then, the following holds true :

1. *The sequence of real numbers*

$$\Theta_n^2 = -\log\Delta_n + \kappa\rho^2(K+1)^2 \frac{1}{n}\mathrm{Tr}\mathbf{S}^2 \frac{1}{n}\mathrm{Tr}\widetilde{\mathbf{S}}^2$$

satisfies $0 < \liminf_n \Theta_n^2 \le \limsup_n \Theta_n^2 < \infty$.

2. *The mutual information satisfies:*

$$\frac{N}{\Theta_n} \left(\boldsymbol{\mathcal{I}}_n(\rho) - \mathbb{E}\boldsymbol{\mathcal{I}}_n(\rho) \right) \xrightarrow[n\to\infty]{} \mathcal{N}(0,1) \quad \textit{in distribution}.$$

Remark 2. *It should be noted that the variance Θ_n^2 is the sum of two terms. The first term Δ_n would be the same if $\mathbf{H}$ were Gaussian, and it depends on $\mathbf{A}$ through its singular values only. The second term accounts for the impact of the fourth cumulant and depends also on the left and right singular vectors of $\mathbf{A}$. Also, setting $\kappa = 0$, we get back the expression of the variance established in [7] for the Gaussian case.*

In case the cumulant is nonzero, it might be of interest determine the singular vector matrices of $\mathbf{A}$ which provide the extremal values of Θ_n^2. Since κ can be positive or negative, the extrema might refer to the minimum or the maximum depending on $\mathrm{sign}(\kappa)$:

Proposition 1. *Let $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$ be a spectral decomposition of $\mathbf{A}$ where $\mathbf{U} = [u_{ik}]$ and $\mathbf{V} = [v_{jl}]$ are the matrices of singular vectors.*

- *When $|u_{ik}|^2 = 1/N$ and $|v_{jl}|^2 = 1/n$, Θ_n^2 attains its minimum with respect to $(\mathbf{U}, \mathbf{V})$ if $\kappa > 0$ and its maximum if $\kappa < 0$.*
  Such a situation arises e.g. *when $\mathbf{U}$ and $\mathbf{V}$ are Fourier matrices.*
- *When $\mathbf{U} = \mathbf{I}_N$ and $\mathbf{V} = \mathbf{I}_n$, Θ_n^2 attains its minimum with respect to $(\mathbf{U}, \mathbf{V})$ if $\kappa < 0$ and its maximum if $\kappa > 0$.*

Proof. We fix $\mathbf{\Lambda}$ and look for matrices $(\mathbf{U}, \mathbf{V})$ which minimize or maximize Θ_n^2. One can notice that the term Δ_n in the expression of Θ_n^2 does not depend on $(\mathbf{U}, \mathbf{V})$. Hence we only need to consider the cumulant term and minimize or maximize

$$\kappa \sum_{i=1}^{N} t_{i,i}^2 \sum_{j=1}^{n} \tilde{t}_{j,j}^2$$

Using (3) and (4), the spectral decompositions of $\mathbf{T}$ and $\widetilde{\mathbf{T}}$ are given by: $\mathbf{T} = \mathbf{U}\mathbf{\Omega}\mathbf{U}^*$ and $\widetilde{\mathbf{T}} = \mathbf{V}\widetilde{\mathbf{\Omega}}\mathbf{V}^*$ where $\mathbf{\Omega}$ and $\widetilde{\mathbf{\Omega}}$ depend only on $\mathbf{\Lambda}$. Letting $\mathbf{\Omega} = \mathrm{diag}(\omega_1, \cdots, \omega_N)$, we will show that for any matrix $\mathbf{U}$, we have:

$$\frac{1}{N}\left(\sum_{i=1}^{N} \omega_i\right)^2 \overset{(a)}{\leq} \sum_{i=1}^{N} t_{i,i}^2 \overset{(b)}{\leq} \sum_{i=1}^{N} \omega_i^2$$

and that inequality (a) becomes an equality when the elements of $\mathbf{U}$ satisfy $|u_{i,k}|^2 = 1/N$, whereas inequality (b) is an equality when $\mathbf{U} = \mathbf{I}_N$. Since $\mathbf{T} = \mathbf{U}\mathbf{\Omega}\mathbf{U}^*$ where $\mathbf{\Omega} = \mathrm{diag}\,(\omega_1, \cdots, \omega_N)$, we get:

$$\sum_{i=1}^{N} t_{i,i}^2 = \sum_{i=1}^{N}\left(\sum_{k=1}^{N} |u_{i,k}|^2 \omega_k\right)^2 \tag{5}$$

Denote by $\mathbf{P}$ the doubly stochastic matrix $\mathbf{P} = \left[|u_{i,k}|^2\right]_{i,k=1}^{N}$ and by $\boldsymbol{\omega}$ the vector $\boldsymbol{\omega} = [\omega_1, \cdots, \omega_N]^{\mathrm{T}}$. Therefore:

$$\sum_{i=1}^{N} t_{i,i}^2 = \boldsymbol{\omega}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{P}\boldsymbol{\omega}$$

It is clear that since $\mathbf{P}$ is a doubly stochastic matrix, the vector $\frac{1}{\sqrt{N}}\mathbf{1}$ is the eigenvector of $\mathbf{P}^{\mathrm{T}}\mathbf{P}$ corresponding to the eigenvalue equal to one where $\mathbf{1}$ is the vector of all ones. Thus, for any vector $\boldsymbol{\omega}$, we get

$$\boldsymbol{\omega}^{\mathrm{T}}\left(\mathbf{P}^{\mathrm{T}}\mathbf{P} - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)\boldsymbol{\omega} = \boldsymbol{\omega}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{P}\boldsymbol{\omega} - \frac{1}{N}\left(\sum_{i=1}^{N}\omega_i\right)^2 \geq 0$$

If we choose $\mathbf{U}$ in such a way to have $|u_{i,j}|^2 = \frac{1}{N}$ then $\mathbf{P}^{\mathrm{T}}\mathbf{P} = \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}$, and as such (a) becomes an equality. Inequality (b) can be deduced from (5) and from the convexity of the function $f(x) = x^2$. In the same way, we can easily see that:

$$\frac{1}{n}\left(\sum_{j=1}^{n} \widetilde{\omega}_j\right)^2 \overset{(c)}{\leq} \sum_{j=1}^{n} \widetilde{t}_{j,j}^2 \overset{(d)}{\leq} \sum_{j=1}^{n} \widetilde{\omega}_j^2$$

where inequalities (c) (resp. (d)) become equalities when $|\widetilde{v}_{i,j}|^2 = 1/n$ (resp $\mathbf{V} = \mathbf{I}_n$).
The proposition follows from these results. $\square$

Remark 3. *Note that we can obtain the same result by using a Schur convexity argument*[1]. *Actually, if $\kappa > 0$, one can easily see that all real positive vectors whose entries have sum $\frac{1}{N}\mathrm{Tr}\mathbf{T}$ weakly majorize the vector $\mathbf{u}_{unif} = \left(\frac{\frac{1}{n}\mathrm{Tr}\mathbf{T}}{N}, \ldots, \frac{\frac{1}{n}\mathrm{Tr}\mathbf{T}}{N}\right)^T$. Consider the Schur Convex function $f : \mathbb{R}^N \to \mathbb{R}$, $\mathbf{u} \mapsto \sum_{i=1}^{N} |u_i|^2$, then f attains its minimum when $\mathbf{u} = \mathbf{u}_{unif}$. In particular, we obtain: $\frac{1}{n}\mathrm{Tr}\mathbf{S}^2 = f(\bar{\mathbf{u}}) \geq f(\mathbf{u}_{unif})$, where,*
$\bar{\mathbf{u}} = \left(\frac{1}{n}\sum_{i=1}^{N}\omega_i|u_{1i}|^2, \ldots, \frac{1}{n}\sum_{i=1}^{N}\omega_i|u_{Ni}|^2\right)^{\mathrm{T}}$. The same kind of argument can be used to prove that $\frac{1}{n}\mathrm{Tr}\tilde{\mathbf{S}}^2$ is minimized if the entries of matrix $\mathbf{V}$ verify: $|v_{jl}|^2 = 1/n$.

The bias term is characterized by the following theorem:

Theorem 4. *With the notations of Theorem 3, let*

$$\beta_n = \kappa \frac{A(\omega)}{B(\omega)}$$

with

$$\begin{aligned}
A(\omega) &= \omega \frac{1}{n}\mathrm{Tr}S(\omega)^2 \frac{1}{n}\mathrm{Tr}\widetilde{S}(\omega)^2 \\
&\quad - \omega^2(1+\tilde{\delta}(\omega))\frac{1}{n}\mathrm{Tr}\widetilde{S}(\omega)^2 \frac{1}{n}\mathrm{Tr}S(\omega)T(\omega)^2 \\
&\quad - \omega^2(1+\delta(\omega))\frac{1}{n}\mathrm{Tr}S(\omega)^2 \frac{1}{n}\mathrm{Tr}\widetilde{S}(\omega)\widetilde{T}(\omega)^2, \\
B(\omega) &= 1 + \omega(1+\delta(\omega))\tilde{\gamma}(\omega) + \omega(1+\tilde{\delta}(\omega))\gamma(\omega).
\end{aligned}$$

Then

$$\sup_n \int_{\rho(K+1)}^{\infty} |\beta_n(\omega)| d\omega \;<\; \infty$$

and furthermore

$$N\left(\mathbb{E}\boldsymbol{\mathcal{I}}_n(\rho) - V_n(\rho)\right) - \int_{\rho(K+1)}^{\infty} \beta(\omega)\, d\omega \xrightarrow[n\to\infty]{} 0\,.$$

4. SIMULATIONS

In this section, we verify by simulations the accuracy of our results. We assume a non centered channel with $\mathbf{A} = [\mathbf{a}(\alpha_1), \cdots, \mathbf{a}(\alpha_n)]$ where $\mathbf{a}(\alpha) = \left[1, e^{\jmath\alpha}, \cdots, e^{\jmath(N-1)\alpha}\right]^{\mathrm{T}}$ is a directional vector, the α_i being some given phase variables. The entries of the non line of sight matrix $\mathbf{X}$ are assumed to verify $[\mathbf{X}]_{i,j} = r_{i,j}\exp(\jmath\theta_{i,j})$, where $\theta_{i,j}$ are i.i.d uniform phase variables over $[0, 2\pi]$ and $r_{i,j}$ are i.i.d real positive random variables. According to the distribution of $r_{i,j}$, we distinguish three type of channels, whose main properties are summarized in Table 1. Each distribution in Table 1 has originally two degrees of freedom, but adding the constraint $\mathbb{E}|X_{1,1}|^2 = 1$, the number of degrees of freedom is reduced to one, and this is captured by the equations in the third row of Table 1. The parametric Nakagami-m and Weibull density functions cover a wide range of distributions. For instance, the Rayleigh distribution (corresponding to a Gaussian channel) can be obtained by setting the parameter μ to 1 and k to 2 for respectively, the Nakagami-m and the Weibull density functions. Moreover, other less common distributions can be also generated, like the truncated Gaussian distribution on the real positive axis, for $\mu = \frac{1}{2}$, and the exponential distribution when $k = 1$. The coefficient of variation shown in Table 1 quantifies the severity of the fading [13]. It is given by:

$$\mathrm{CV} = \frac{\sqrt{\mathrm{var}(r_{i,j})}}{\mathbb{E}r_{i,j}} = \sqrt{\frac{1}{(\mathbb{E}r_{i,j})^2} - 1}$$

[1]The authors thank the reviewer who suggested this argument.

	Nakagami-m (μ,ω)	Weibull (k,λ)	Log-Normal (μ,σ)
distribution	$f_{\mu,\omega}(x) = \frac{2\mu^{\mu}}{\Gamma(\mu)\omega^{\mu}} x^{2\mu-1} e^{-\frac{\mu}{\omega}x^2}, \ x > 0$	$f_{\lambda,k}(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}, \ x > 0$	$f_{\mu,\sigma}(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \ x > 0$
Parameter setting	$\omega = 1$	$\lambda = \sqrt{\frac{1}{\Gamma(1+\frac{2}{k})}}$	$\mu = -\sigma^2$
cumulant	$\kappa = \left(1 + \frac{1}{\mu}\right) - 2$	$\kappa = \frac{\Gamma(1+\frac{4}{k})}{\left(\Gamma(1+\frac{2}{k})\right)^2} - 2$	$\kappa = e^{4\sigma^2} - 2$
Coefficient of Variation	$\mathrm{CV} = \sqrt{\frac{\mu(\Gamma(\mu))^2}{\left(\Gamma(\mu+\frac{1}{2})\right)^2} - 1}$	$\mathrm{CV} = \sqrt{\frac{\Gamma(1+\frac{2}{k})}{\left(\Gamma(1+\frac{1}{k})\right)^2} - 1}$	$\mathrm{CV} = \sqrt{\exp(\sigma^2) - 1}$

Table 1. Main characteristics of some non Gaussian distributions

The Nakagami-m and Weibull distributions exhibit severe fadings when their respective parameters μ and k are very small, whereas the lognormal distribution exhibit severe fadings for large values of σ^2. In this section, we will first verify the accuracy of our derived results for the case of Gaussian entries. The effect of the coefficient of variation and the eigenvectors of $\mathbf{A}$ will be discussed afterwards.

4.1. Gaussian entries

In this section, we consider Gaussian $\mathbf{H}$ and assess the accuracy of the results of Theorems 2 and 3 for finite N and n. We also investigate the effect of the Rice factor K on the variance.

Fig. 1 displays the empirical estimation of $\mathbb{E}\boldsymbol{\mathcal{I}}_n$ as well as V_n with respect to $c = \frac{N}{n}$ when $n = 16$, N ranging from 2 to 15, and $\rho = 0.2$. We also set the Rice factor K to 1. We notice a very good fit on the whole range of c.

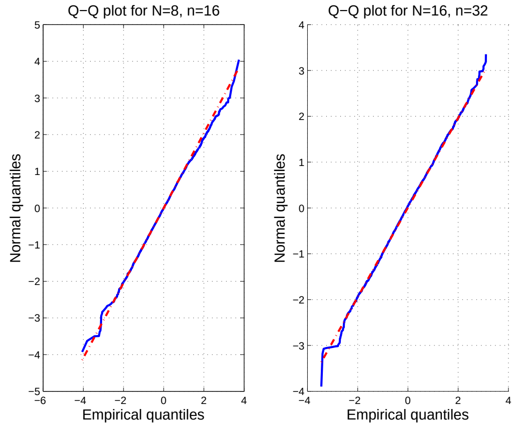

Fig. 2. $Q - Q$ plot for $\frac{N}{\Theta_n}\left(I_n(\rho) - \mathbb{E}I_n(\rho)\right)$.

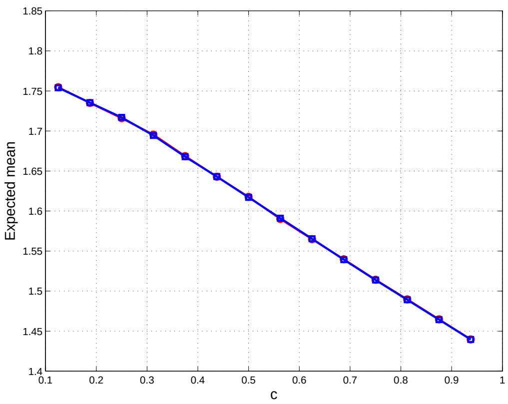

Fig. 1. Theoretical and Empirical expected mean with respect to c.

Fig. 2 presents the Quantile-Quantile (Q-Q) plot for the distribution of the Shannon capacity when $\rho = 0.2$ and $K = 1$. This figure shows that the normal approximation is suitable in the range of a few standard deviations and could be used to approximate the outage probability under these conditions.

Fig. 3 shows the empirical and the LRMT theoretical variance with respect to the Rice factor when $n = 32$, $N = 16$, and $\rho = 0.2$. We note that when K increases, the line of sight component prevails over the non-line of sight one, to the point that the channel is almost constant. Thus, its fluctuations around its mean become small.

4.2. Non Gaussian entries

In this experiment, we investigate the effect of the coefficient of variation on the variance of the mutual information. For each distribution (Nakagami-m, Weibull or Lognormal), we make the coefficient of variation CV vary from 0 to 1. Fig. 4 displays the empirical and theoretical variances with respect to CV, when $N = 32, n = 64, \rho = 0.2$ and $K = 1$. We note that when the channel exhibit severe fadings (high CV), the Lognormal channel is the one that undergoes the highest variance.

We now investigate by simulations the asymptotic behaviour of the bias. We consider the m-Nakagami distribution with $\mu = 0.05$. We also set $\rho = 0.2$, $K = 1$ and show on the same figure the empirical bias $|\left(\mathbb{E}\boldsymbol{\mathcal{I}}_n(\rho) - V_n(\rho)\right)|$ computed over 5000 channel realizations and the theoretical bias, $|\frac{1}{N}\int_{\rho(K+1)}^{+\infty}\beta(\omega)d\omega|$, for different values of N. The ratio $c = N/n$ is set to 0.5. Fig. 5 illustrates the obtained results. As expected, the empirical and theoretical biases coincide and are of order $\mathcal{O}(\frac{1}{N})$.

Finally, we apply Proposition 1 to the Nakagami-m, Weibull and Lognormal distributions. A close look at these distributions shows that their fourth cumulant is positive if and only if their coefficient

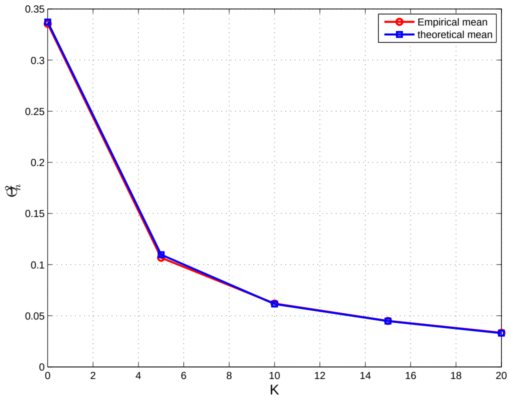

Fig. 3. Theoretical and Empirical variance with respect to K.

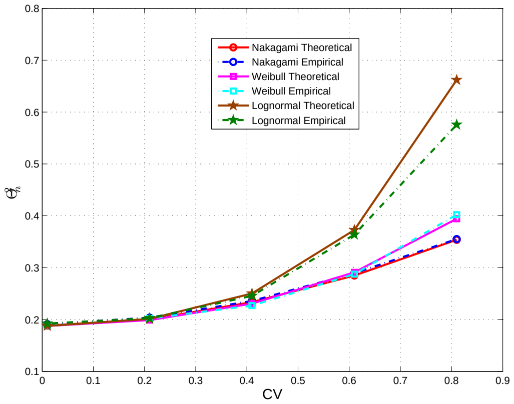

Fig. 4. Theoretical and empirical variances with respect to the coefficient of variation

of variation is larger than the coefficient of variation of the Rayleigh distribution. Hence, assuming $\mathbf{D}$ is a diagonal matrix and $\mathbf{F}_\ell$ is the Fourier $\ell \times \ell$ matrix, we have the results shown in Table 2.

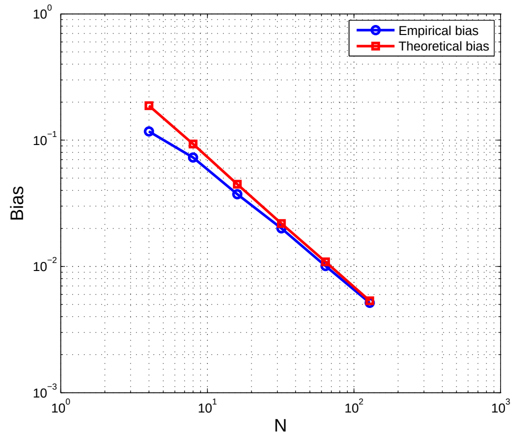

Fig. 5. Theoretical and empirical biases with respect to N

	$\mathbf{A} = \mathbf{D}$	$\mathbf{A} = \mathbf{F}_N \mathbf{D} \mathbf{F}_n^*$
CV $>$ CV(Rayleigh) (severe fading)	Θ_n^2 is max.	Θ_n^2 is min.
CV $<$ CV(Rayleigh) (non severe fading)	Θ_n^2 is min.	Θ_n^2 is max.

Table 2. Effect of the singular vectors of $\mathbf{A}$ on Θ_n^2

5. REFERENCES

[1] I.E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," published in *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov/Dec 1999, Technical Memorandum, Bell Laboratories, Lucent Technologies, October 1995.

[2] Foschini G.J., "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[3] V.A. Marchenko and L.A. Pastur, "Distribution of Eigenvalues for Some Sets of Random Matrices," *Math. USSR – Sbornik*, vol. 1, no. 4, pp. 457–483, 1967.

[4] 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. on Information Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.

[5] W. Hachem, O. Khorunzhiy, Ph. Loubaton, J. Najim, and L. Pastur, "A New Approach for Capacity Analysis of Large Dimensional Multi-Antenna Channels," *IEEE Trans. on Information Theory*, vol. 54, no. 9, Sept. 2008.

[6] W. Hachem, P. Loubaton, and J. Najim, "A CLT for Information Theoretic Statistics of Gram Random Matrices With a given Variance Profile," *Annals of Applied Probability*, vol. 18, no. 6, pp. 2071–2130, 2008.

[7] G. Taricco, "On the Capacity of Separately-Correlated MIMO Rician Fading Channels," in *Proceedings of the GLOBECOM conference*, San Francisco USA, Nov. 2006.

[8] G. Taricco, "Asymptotic Ergodic Capacity of Wideband MIMO Channels with Separately-Correlated Rician Fading," in *Proceedings of the GLOBECOM conference*, New Orelans, LA, USA, Dec. 2008.

[9] K. Yip and T. Ng, "A Simulation Model for Nakagami-m Fading Channels $m < 1$," *IEEE Trans. on Communications*, vol. 48, no. 2, pp. 214–221, Feb. 2000.

[10] H. Hashemi, "The indoor radio propagation channel," *Proceedings of the IEEE*, vol. 81, July 1993.

[11] N. S. Adawi etal, "Coverage Prediction for Mobile Radio Systems Operating in the 800/900Mhz ," *IEEE Trans. on Vehicular Technology*, vol. 37, pp. 3–62, Feb. 1998.

[12] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *Ann. Appl. Probab.*, vol. 17, no. 3, pp. 875–930, 2007.

[13] S. Ho Choi, P. Smith, B. Allen, W. Q. Malik, and M. Shafi, "Severely Fading MIMO Channels: Models and Mutual Information," in *Proceedings of the Int. Conf. on Communications*, Glasgow, Scotland, June 2007, pp. 4628–4633.