Fluctuations of the Mutual Information in Large Distributed Antenna Systems with Colored Noise

Jakob Hoydis*†, Jamal Najim‡, Romain Couillet†§ and Mérouane Debbah†

*Department of Telecommunications, Supélec, 91192 Gif-sur-Yvette, France

†Alcatel-Lucent Chair on Flexible Radio, Supélec, 91192 Gif-sur-Yvette, France

‡CNRS and Télécom Paristech, 75013 Paris, France

§ST-Ericsson, 06560 Sophia Antipolis, France

{jakob.hoydis,romain.couillet,merouane.debbah}@supelec.fr

najim@telecom-paristech.fr

Abstract

This paper studies the fluctuations of the mutual information of a class of multiple-input multiple-output (MIMO) channels with arbitrary correlated noise in the large system limit. Under the assumption that the channel dimensions grow infinitely large at the same rate, we find a deterministic approximation of the ergodic mutual information and study its fluctuations around this value in form of a central limit theorem (CLT). This result can be used to predict the outage probability for slow fading channels. The channel model considered in this contribution has a particular application in the context of distributed antenna or network MIMO systems where the path loss between any pair of transmit and receive antennas has a different value. As shown by simulations, the asymptotic approximations translate well into systems of small dimensions.

I. Introduction

Consider a wireless communication channel from n single antenna transmitters to a receiver equipped with N antennas. Let $\mathbf{H} \in \mathbb{C}^{N \times n}$ be the channel matrix representing the complex channel gains from the transmitters to the receiving antennas. The receive vector $\mathbf{y} \in \mathbb{C}^N$ at a given time instant reads

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^n$ is the vector of the transmitted signals and $\mathbf{z} \in \mathbb{C}^N$ is a vector of complex Gaussian noise with covariance matrix $\mathbb{E}\mathbf{z}\mathbf{z}^{\mathsf{H}} = \rho \mathbf{I}_N + \mathbf{A}\mathbf{A}^{\mathsf{H}}$. The elements $(h_{ij},\, 1 \leq i \leq N,\, 1 \leq j \leq n)$ of the channel matrix $\mathbf{H}$ are modeled as

$$h_{ij} = \frac{\sigma_{ij} w_{ij}}{\sqrt{n}} \tag{2}$$

where $\left(\sigma_{ij}^2\right)$ is a sequence of positive real numbers called a variance profile and the w_{ij} are independent complex Gaussian random variables with zero mean and unit variance. For a complex Gaussian channel input vector $\mathbf{x}$ with covariance matrix $\mathbb{E}\mathbf{x}\mathbf{x}^{\mathsf{H}} = \mathbf{I}_n$ and full channel knowledge at the receiver, the normalized ergodic capacity of the channel is given by $I(\rho) = \mathbb{E}\mathcal{I}(\rho)$, where

$$\mathcal{I}(\rho) = \frac{1}{N} \log\det \left(\mathbf{I}_N + \left(\rho \mathbf{I} + \mathbf{A}\mathbf{A}^{\mathsf{H}} \right)^{-1} \mathbf{H}\mathbf{H}^{\mathsf{H}} \right). \tag{3}$$

The aim of this paper is to derive a deterministic approximation $V(\rho)$ of $I(\rho)$ and to study the fluctuations of the random variable $N(\mathcal{I}(\rho) - V(\rho))$ in the large system limit when $n \to \infty$ while $N/n \to c > 0$. A well known result of random matrix theory states that the empirical eigenvalue distribution of the Gram matrix $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ converges weakly to a limit distribution function when the elements of $\mathbf{H}$ are independent and identically distributed (i.i.d) [1]. This fact leads also to the convergence of $\mathcal{I}(\rho)$ to a deterministic limit which can be given in closed form [2]. Similar results could be established for more complicated models such as the Kronecker model $\mathbf{H} = \mathbf{\Phi}_R \mathbf{W} \mathbf{\Phi}_T$ [3], [4], where $\mathbf{W}$ is a $N \times n$ standard Gaussian matrix and $\mathbf{\Phi}_R$ and $\mathbf{\Phi}_T$ are $N \times N$ and $n \times n$ matrices capturing the effects of transmit and receive antenna correlation, sums of matrices $\mathbf{H}_k$ each having a Kronecker variance structure [5] and also non-centered channel matrices with a variance profile [6]. These works provide deterministic approximations $V(\rho)$ of $I(\rho)$, only depending on N, n and the distribution of $\mathbf{H}$, in the sense that $I(\rho) - V(\rho) \to 0$ for $n \to \infty$ while $N/n \to c > 0$. Apart from some special cases, the function $V(\rho)$ is rarely available in closed-form and requires the solution of a set of implicit equations. However, their computation is in general less complex than the capacity evaluation by Monte Carlo simulations. Moreover, the deterministic approximations have been shown by simulations to yield very accurate results for small channel dimensions with as little as two transmit and receive antennas. Recently, also the fluctuations of the mutual information have been studied in terms of central limit theorems (CLTs). One is generally interested in obtaining results of the form:

$$\frac{N^{\alpha}}{\Theta_n} \left(\mathcal{I}(\rho) - V(\rho) \right) \to \mathcal{N}(0,1) \tag{4}$$

in distribution, where α is a measure of the convergence speed and Θ_n^2 determines the variance. In a slow fading scenario, these results allow to approximate the outage probability, i.e., $\Pr\left(N\mathcal{I}(\rho) \leq R \right)$, for a given desired target rate R. A CLT for channel matrices with left-sided correlation was established in [7] and the more general case of a variance profile addressed in [8]. Also the fluctuations of the mutual information in the presence of correlated interference under the Kronecker model were studied in [4] relying on the replica method.

The novelty of the results derived in this paper in contrast to [8] and [4] is the consideration of arbitrary correlated Gaussian noise whose covariance matrix can be written in the form $\rho \mathbf{I}_N + \mathbf{A}\mathbf{A}^{\mathsf{H}}$. Typically, $\rho \mathbf{I}_N$ represents an uncorrelated thermal noise component with power ρ while $\mathbf{A}\mathbf{A}^{\mathsf{H}}$ accounts for a source of correlated interference whose covariance matrix has the non-negative square root $\mathbf{A} \in \mathbb{C}^{N \times m}$. This model is more general than the particular case where the interference term can be written in the form $\mathbf{H}_I \mathbf{x}_I$, where $\mathbf{H}_I$ is a random matrix which follows the same statistical model as the channel matrix $\mathbf{H}$ and $\mathbf{x}_I$ is a standard complex Gaussian vector [4]. Here, the mutual information can be decomposed into two terms without interference, i.e., $\mathcal{I}(\rho) = \frac{1}{N} \log\det \left(\rho \mathbf{I}_N + \mathbf{H}_I \mathbf{H}_I^{\mathsf{H}} + \mathbf{H}\mathbf{H}^{\mathsf{H}} \right) - \frac{1}{N} \log\det \left(\rho \mathbf{I}_N + \mathbf{H}_I \mathbf{H}_I^{\mathsf{H}} \right)$, where the first can be seen as the mutual information of the compound channel $[\mathbf{H}\mathbf{H}_I]$ and the second as the mutual information of the interference channel $\mathbf{H}_I$. Note that both matrices $\mathbf{H}$ *and* $\mathbf{H}_I$ are considered random while $\mathbf{A}$ in our model is assumed to be deterministic.

The channel model considered in this work has a particular application in the context of distributed antenna or network MIMO systems where the signals received at several spatially separated antennas are jointly processed to provide *macro diversity*. For more details on this topic we refer the reader to the comprehensive surveys [9], [10]. More precisely, the channel in (1) can be seen as a multiple access channel (MAC) with macro diversity where the value of a particular σ_{ij}^2 represents the inverse path loss between the j^{th} transmitter and the i^{th} receiving antenna. For example, assuming a log-distance path loss model, we have $\sigma_{ij}^2 = \left(\frac{1}{d_{ij}}\right)^{\beta}$, where d_{ij} is the distance between transmitter j and receive antenna i and β is the path loss exponent whose value lies usually in the range from 2 to 5 dependent on the radio environment. The application of random matrix theory to this field of research is not new. The sum-capacity scaling of large cooperative cellular networks has been studied in [11] and the downlink of large multi-cell systems with optimal power allocation and user scheduling was considered in [12]. However, both works build upon the assumption that the inter-cell interference is of the form $\mathbf{H}_I\mathbf{x}_I$ as discussed above.

The remaining part of the paper is structured as follows. A first order result in form of a deterministic equivalent of the normalized mutual information is derived in Section II. That is, we find a deterministic function $V(\rho)$ such that $I(\rho) - V(\rho) \to 0$ for $n \to \infty$ while $N/n \to c > 0$. The fluctuations of the random variable $N(\mathcal{I}(\rho) - V(\rho))$ are studied in Section III where we establish a CLT of the form (4) and provide an explicit expression for the variance Θ^2. Numerical results are presented in Section IV which corroborate the analysis and demonstrate their applicability to channels with even small system dimensions. Section V concludes the paper.

II. Deterministic Approximation of $I(\rho)$

Recall that $\mathbf{H}$ is a $N \times n$ matrix, $\mathbf{A}$ is $N \times m$. The aim of this section is to propose a deterministic equivalent to the normalized ergodic mutual information

$$I(\rho) = \frac{1}{N}\mathbb{E}\log\det\left(\rho\mathbf{I}_N + \mathbf{A}\mathbf{A}^{\mathsf{H}} + \mathbf{H}\mathbf{H}^{\mathsf{H}}\right) - \log\det\left(\rho\mathbf{I}_N + \mathbf{A}\mathbf{A}^{\mathsf{H}}\right),$$

as N and n grow to infinity at the same pace:

$$0 \quad \leq \quad \liminf_{n\to\infty}\frac{N}{n} \quad \leq \quad \limsup_{n\to\infty}\frac{N}{n} \quad < \quad \infty \,. \tag{5}$$

The requirement related to dimension m is:

$$0 \quad \leq \quad \limsup_{n\to\infty}\frac{m}{n} \quad < \quad \infty \,. \tag{6}$$

In the sequel, the notation $N, n \to \infty$ will refer to (5) and (6). Consider the following technical assumptions:

A 1: Consider a family of non-negative real numbers $(\sigma_{ij}^{(n)};\ 1 \leq i \leq N;\ 1 \leq j \leq n)$. Then there exists a non-negative real number $\sigma_{\max}$ such that:

$$\sup_{i\leq N, j\leq n, n\geq 1} \sigma_{ij}^{(n)} \quad \leq \quad \sigma_{\max} \quad < \quad \infty.$$

A 2: Consider a family $(\mathbf{A}_{N,m}; N \geq 1,\ m \geq 1)$ of $N \times m$ matrices and denote by $\|\mathbf{A}_{n,m}\|$ the spectral norm of matrix $\mathbf{A}_{N,m}$, then there exists a non-negative real number $\mathbf{a}_{\max}$ such that:

$$\sup_{N,m}\|\mathbf{A}_{N,m}\| \quad \leq \quad \mathbf{a}_{\max} \quad < \quad \infty \,.$$

A 3: Consider a family of non-negative real numbers $(\sigma_{ij}^{(n)};\ 1 \le i \le N;\ 1 \le j \le n)$. Then there exists a non-negative real number $\sigma_{\min}$ such that:

$$\liminf_{n\ge 1} \min_{1\le j\le n} \frac{1}{n}\sum_{i=1}^{N} \sigma_{ij}^2 \quad \ge \quad \sigma_{\min}^2 .$$

In the sequel, we will drop the dependencies in N, n and m and simply write σ_{ij} and $\mathbf{A}$ instead of $\sigma_{ij}^{(n)}$ and $\mathbf{A}_{N,m}$.

Consider the following diagonal $N \times N$ matrices:

$$\mathbf{D}_j = \mathrm{diag}\left(\sigma_{1j}^2, \dots, \sigma_{Nj}^2\right), \quad 1 \le j \le n \ . \tag{7}$$

Denote by $\mathbb{C}_+ = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$, and by $\mathcal{S}$ the class of functions f analytic over $\mathbb{C}_+$, such that $f : \mathbb{C}_+ \to \mathbb{C}_+$ and $\lim_{y\to\infty} -\mathbf{i}yf(\mathbf{i}y) = 1$ [1]. We are now in position to state the first result of the paper:

Theorem 1 (Deterministic Equivalent): Assume that Assumptions ($\mathbf{A}1$) and ($\mathbf{A}2$) hold true, then:

(i) The following equation:

$$\mathbf{T}(z) = \left(-z\mathbf{I}_N + \mathbf{A}\mathbf{A}^{\mathsf{H}} + \frac{1}{n}\sum_{j=1}^{n} \frac{\mathbf{D}_j}{1 + \frac{1}{n}\mathrm{tr}\mathbf{D}_j\mathbf{T}(z)} \right)^{-1} \tag{8}$$

admits a unique solution $\mathbf{T}(z)$ among the $N \times N$ matrices such that there exists a $N \times N$ matrix-valued measure[2] μ such that:

$$\mathbf{T}(z) = \int_{\mathbb{R}^+} \frac{\mu(d\lambda)}{\lambda - z} \quad \text{where} \quad \mu(\mathbb{R}^+) = \mathbf{I}_N \ . \tag{9}$$

In particular, $\frac{1}{N}\mathrm{tr}\,\mathbf{T}(z) \in \mathcal{S}$.

(ii) Let $\rho > 0$, denote $\mathbf{T}_\rho = \mathbf{T}(-\rho)$ and consider the quantity:

$$V(\rho) = -\frac{1}{N}\log\det\left(\mathbf{T}_\rho\left(\rho\mathbf{I}_N + \mathbf{A}\mathbf{A}^{\mathsf{H}}\right)\right) + \frac{1}{N}\sum_{j=1}^{n}\log\left(1 + \frac{1}{n}\mathrm{tr}\mathbf{D}_j\mathbf{T}_\rho\right) - \frac{1}{Nn}\sum_{\substack{i=1,\dots,N\\ j=1,\dots,n}} \frac{\sigma_{ij}^2\mathbf{T}_{ii}(-\rho)}{1 + \frac{1}{n}\mathrm{tr}\mathbf{D}_j\mathbf{T}_\rho} \ .$$

Then, the following limit holds true:

$$I(\rho) - V(\rho) \xrightarrow[N,n\to\infty]{} 0.$$

Theorem 1 is essentially a consequence of Theorems 2.4 and 4.1 in [6]. Details are provided below.

Proof: The main idea is to cast the model $\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{A}\mathbf{A}^{\mathsf{H}}$ into an extended model which fits into the framework of [6]. Consider the $N \times (n+m)$ matrices $\mathbf{Z} = [\mathbf{H}\ \mathbf{0}_{N\times m}]$ and $\mathbf{\Gamma} = [\mathbf{0}_{N\times n}\ \mathbf{A}]$; then $(\mathbf{Z}+\mathbf{\Gamma})(\mathbf{Z}+\mathbf{\Gamma})^{\mathsf{H}} = \mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{A}\mathbf{A}^{\mathsf{H}}$, which is precisely the model under investigation. Introduce the following notations, for $1 \le i \le N$ and $1 \le j \le n+m$:

$$\rho_{ij} = \begin{cases} \sqrt{\frac{n+m}{n}} \times \sigma_{ij} & \text{if } j \le n \\ 0 & \text{if } j \ge n+1 \end{cases} , \quad \tilde{\mathbf{\Delta}}_i = \mathrm{diag}(\rho_{ij}^2;\ 1 \le j \le n+m) \ , \quad \mathbf{\Delta}_j = \mathrm{diag}(\rho_{ij}^2;\ 1 \le i \le N) \ .$$

[1]Such functions are known to be Stieltjes transforms of probability measures over $\mathbb{R}$ - see for instance [6, Proposition 2.2].

[2]For details, see for instance [6, Theorem 2.4].

Note that $\mathbf{\Delta}_j = \mathbf{0}$ if $j \geq n+1$. We can now write down the equations associated to the model $(\mathbf{Z}+\mathbf{\Gamma})(\mathbf{Z}+\mathbf{\Gamma})^{\mathsf{H}}$ as given in [6, Theorem 2.4]. Let $\mathbf{\Psi} = \mathrm{diag}(\Psi_i, 1 \leq i \leq N)$ and $\tilde{\mathbf{\Psi}} = \mathrm{diag}(\tilde{\Psi}_j, 1 \leq j \leq n+m)$ where:

$$\Psi_i = -\frac{1}{z\left(1+\frac{1}{n+m}\mathrm{tr}\tilde{\mathbf{\Delta}}_i\tilde{\mathbf{T}}\right)} \ , \quad \tilde{\Psi}_j = -\frac{1}{z\left(1+\frac{1}{n+m}\mathrm{tr}\mathbf{\Delta}_j\mathbf{T}\right)} \tag{10}$$

and

$$\mathbf{T} = \left(\mathbf{\Psi}^{-1} - z\mathbf{\Gamma}\tilde{\mathbf{\Psi}}\mathbf{\Gamma}^{\mathsf{H}}\right)^{-1} \ , \quad \tilde{\mathbf{T}} = \left(\tilde{\mathbf{\Psi}}^{-1} - z\mathbf{\Gamma}^{\mathsf{H}}\mathbf{\Psi}\mathbf{\Gamma}\right)^{-1} \ . \tag{11}$$

Then this system admits a unique solution $(\Psi_i, \tilde{\Psi}_j) \in \mathcal{S}^{N+n+m}$. In particular, $\mathbf{T}$ satisfies (9).

Taking advantage of the particular forms of ρ_{ij} and $\mathbf{\Delta}_j$, one can prove that $\mathbf{T}$ as defined in the previous equation satisfies the following equation

$$\mathbf{T}(z) = \left(-z\mathbf{I}_N + \mathbf{A}\mathbf{A}^{\mathsf{H}} + \frac{1}{n}\sum_{j=1}^{n}\frac{\mathbf{D}_j}{1+\frac{1}{n}\mathrm{tr}\mathbf{D}_j\mathbf{T}(z)}\right)^{-1} \ .$$

Hence, the existence of a solution $\mathbf{T}$ to (8) is established; moreover $\mathbf{T}$ admits representation (9).

To complete the proof of (i), it remains to check that such a $\mathbf{T}$ is unique. Assume that there exists $\mathbf{T}$ satisfying (8) with representation (9). Define $\tilde{\mathbf{\Psi}}$ with the help of the second part of (10), $\mathbf{\Psi}$ with the help of the first part of (11) and $\tilde{\mathbf{T}}$ with the help of the second part of (11). It is then a matter of routine to check that $\mathbf{\Psi}$ and $\tilde{\mathbf{\Psi}}$ satisfy the system (10)-(11) (it remains basically to check that the first part of (10) is satisfied). As $\mathbf{T}$ admits representation (9), Ψ_i and $\tilde{\Psi}_j$ belong to $\mathcal{S}$. Hence $\mathbf{\Psi}$ and $\tilde{\mathbf{\Psi}}$ are uniquely defined and so is $\mathbf{T}$.

Part (ii) of the theorem is a direct application of [6, Theorem4.1]; details are therefore omitted. ■

III. Fluctuations of $\mathcal{I}(\rho)$: A Central Limit Theorem

A number of studies has been devoted to the fluctuations of the mutual information, with various statistical assumptions for the channel $\mathbf{H}$. Let us cite [4], [13], and in a more mathematical flavour [14] (separable variance profile) [8] (general variance profile) and [15] (Rician channel with separable variance profile). A common feature of these works, although perhaps not much known, is the nice and concise closed-form expression of the variance which always writes

$$\Theta^2 = -\log\det\left(\mathbf{I} - \mathbf{J}\right) \ ,$$

where $\mathbf{J}$ is a Jacobian matrix associated to the set of fundamental equations of the matrix model under study. The fluctuations of the model $\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{A}\mathbf{A}^{\mathsf{H}}$ have not been studied yet, but relying on the previous observation, it is easy to infer the formula for the variance. Let $\delta_j = \frac{1}{n}\mathrm{tr}\,\mathbf{D}_j\mathbf{T}_\rho$. Multiplying $\mathbf{T}_\rho$ in (8) by $\mathbf{D}_j$ and taking the normalized trace yields the following system of n equations:

$$\begin{aligned} \delta_j &= \frac{1}{n}\mathrm{tr}\,\mathbf{D}_j\left(\rho\mathbf{I}_N + \mathbf{A}\mathbf{A}^{\mathsf{H}} + \frac{1}{n}\sum_{j=1}^{n}\frac{\mathbf{D}_j}{1+\delta_j}\right)^{-1} \\ &\stackrel{\triangle}{=} \Gamma_j(\delta_1,\ldots,\delta_n) \ . \end{aligned}$$

The computation of the Jacobian matrix $\mathbf{J}_n$ of function $\Gamma = (\Gamma_1, \cdots, \Gamma_n)$ is then straightforward:

$$[\mathbf{J}_n]_{k\ell} \;=\; \frac{\partial \Gamma_k}{\partial \delta_\ell} \;=\; \frac{1}{n} \frac{\frac{1}{n} \mathrm{tr}\, \mathbf{D}_k \mathbf{T}_\rho \mathbf{D}_\ell \mathbf{T}_\rho}{(1+\delta_\ell)^2} \ .$$

Based on the previous remarks, we are now in position to state the claim related to the fluctuations of the mutual information for the channel model under investigation.

Claim 1 (The CLT): Assume that Assumptions ($\mathbf{A}1$), ($\mathbf{A}2$) and ($\mathbf{A}3$) hold true. Recall the definition of $\mathbf{T}_\rho = \mathbf{T}(-\rho)$ and consider the following $n \times n$ matrix $\mathbf{J}_n$ defined by:

$$[\mathbf{J}_n]_{k,\ell} = \frac{1}{n} \frac{\frac{1}{n} \mathrm{tr}\, \mathbf{D}_k \mathbf{T}_\rho \mathbf{D}_\ell \mathbf{T}_\rho}{\left(1 + \frac{1}{n} \mathrm{tr}\, \mathbf{D}_\ell \mathbf{T}_\rho\right)^2} \ . \tag{12}$$

Then:

(i) The real number $\Theta_n^2 = -\log\det\left(\mathbf{I}_n - \mathbf{J}_n\right)$ is well-defined and satisfies

$$0 \quad < \quad \liminf_{N,n\to\infty} \Theta_n^2 \quad \le \quad \limsup_{N,n\to\infty} \Theta_n^2 \quad < \quad \infty.$$

(ii) The following convergence holds true

$$\frac{N}{\Theta_n}\left(\mathcal{I}(\rho) - V(\rho)\right) \xrightarrow[N,n\to\infty]{\mathcal{D}} \mathcal{N}(0,1) \ ,$$

where $\mathcal{D}$ stands for the convergence in distribution.

Proof of part (i) closely follows [8, Theorem 3.1] and is therefore omitted. Due to the term $\mathbf{A}\mathbf{A}^{\mathsf{H}}$ in the model $\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{A}\mathbf{A}^{\mathsf{H}}$, the proof of the fluctuations (ii) is not a simple consequence of Theorems 3.2 and 3.3 in [8] and necessitates special mathematical developments. We however provide the proof of the fluctuations in two specific cases, namely:

1) The case where $\mathbf{A}\mathbf{A}^{\mathsf{H}} = \mathbf{\Lambda}$ is a $N \times N$ diagonal matrix.
2) The case where the variance profile is separable, i.e. $w_{ij} = \sqrt{d_i \tilde{d}_j}$.

Beyond the proof of the fluctuations for these cases, simulations are provided that validate the variance formula in the general case.

Proof of Claim 1 in case 1): Let $\mathbf{A}\mathbf{A}^{\mathsf{H}} = \mathbf{\Lambda} = \mathrm{diag}(\lambda_i^2;\ 1 \le i \le N)$. Denote $\mathbf{\Delta} = (\mathbf{\Lambda} + \rho \mathbf{I}_N)^{-1}$, then:

$$\log\det(\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{\Lambda} + \rho\mathbf{I}_N) \quad = \quad -\log\det\mathbf{\Delta} + \log\det\left(\mathbf{\Delta}^{1/2}\mathbf{H}\mathbf{H}^{\mathsf{H}}\mathbf{\Delta}^{1/2} + \mathbf{I}_N\right) \ .$$

Consider $\tilde{\mathbf{H}} = \mathbf{\Delta}^{1/2}\mathbf{H}$, then $\tilde{\mathbf{H}}$ is a centered matrix with a variance profile given by: $\kappa_{ij} = \frac{\sigma_{ij}}{\sqrt{\lambda_i^2 + \rho}}$. Hence, the fluctuations of $\log\det\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\mathsf{H}} + \mathbf{I}_N\right)$ fall into the framework of Theorems 3.2 and 3.3 in [8]. In particular, $N(I(\rho) - V(\rho)) \to 0$ as $N, n \to \infty$ and $\tilde{\Theta}_n^{-1} N(\mathcal{I}(\rho) - I(\rho)) \to \mathcal{N}(0,1)$ in distribution, where:

$$\tilde{\Theta}_n = -\log\det(\mathbf{I}_n - \tilde{\mathbf{J}}_n) \ , \quad [\tilde{\mathbf{J}}_n]_{k,\ell} = \frac{1}{n} \frac{\frac{1}{n}\mathrm{tr}\, \mathbf{D}_k \mathbf{\Upsilon} \mathbf{D}_\ell \mathbf{\Upsilon}}{\left(1 + \frac{1}{n}\mathrm{tr}\, \mathbf{D}_\ell \mathbf{\Upsilon}\right)^2} \ .$$

and where $\mathbf{\Upsilon}$ satisfies the following equation:

$$\mathbf{\Upsilon} = \left(\mathbf{I}_N + \frac{1}{n}\sum_{j=1}^{n} \frac{\mathbf{\Delta}_j}{1 + \frac{1}{n}\mathrm{tr}\, \mathbf{\Delta}_j \mathbf{\Upsilon}}\right)^{-1} \qquad \text{with} \quad \mathbf{\Delta}_j = \mathbf{D}_j \mathbf{\Delta} \ . \tag{13}$$

In order to establish Claim 1 in this case, it remains to prove that $\tilde{\Theta}_n$ as just defined is equal to Θ_n. From (13), it is straightforward to prove that $\mathbf{\Delta\Upsilon}$ satisfies (8) with $z = -\rho$, and thus is equal to $\mathbf{T}$ due to the uniqueness of the solution of (8). It readily follows that $\mathbf{J}_n = \tilde{\mathbf{J}}_n$, which implies $\Theta_n = \tilde{\Theta}_n$. Claim 1 is proved in the case where $\mathbf{A}\mathbf{A}^{\mathsf{H}} = \mathbf{\Lambda}$. ■

Proof of Claim 1 in case 2): In the case where the variance profile is separable, i.e., $w_{ij} = \sqrt{d_i \tilde{d}_j}$, $\mathbf{H}$ writes $\mathbf{H} = n^{-1/2}\mathbf{D}^{1/2}\mathbf{W}\tilde{\mathbf{D}}^{1/2}$, where $\mathbf{D} = \mathrm{diag}(d_i,\ 1 \le i \le N)$ and $\tilde{\mathbf{D}} = \mathrm{diag}(\tilde{d}_j,\ 1 \le j \le n)$ and where $\mathbf{W}$ has i.i.d. circular Gaussian entries. Consider the following extended model: $\tilde{\mathbf{W}} = [\mathbf{W}\,\mathbf{W}_1]$, where $\mathbf{W}_1$ is a $N \times m$ matrix with i.i.d. circular Gaussian entries; $\mathbf{\Delta} = \alpha\mathbf{D}$, where $\alpha = \sqrt{\frac{n+m}{n}}$; $\mathbf{\Gamma} = [\mathbf{0}_{N\times n}\ \mathbf{A}]$; and finally $\tilde{\mathbf{\Delta}} = \alpha\,\mathrm{diag}(\tilde{\mathbf{D}}, \mathbf{0}_{m\times m})$.

Then $\mathbf{H}\mathbf{H}^{\mathsf{H}} + \mathbf{A}\mathbf{A}^{\mathsf{H}}$ writes

$$\left(\mathbf{\Delta}^{1/2}\frac{\tilde{W}}{\sqrt{n+m}}\tilde{\mathbf{\Delta}}^{1/2} + \mathbf{\Gamma}\right)\left(\mathbf{\Delta}^{1/2}\frac{\tilde{W}}{\sqrt{n+m}}\tilde{\mathbf{\Delta}}^{1/2} + \mathbf{\Gamma}\right)^* .$$

The CLT of the mutual information associated to this model has recently been established in [15]:

$$\frac{N}{\tilde{\Theta}_n}\left(\mathcal{I}(\rho) - I(\rho)\right) \xrightarrow[N,n\to\infty]{\mathcal{D}} \mathcal{N}(0,1) .$$

Moreover, it has been proved in [16, Theorem 2] that $N(I(\rho) - V(\rho)) \to 0$. Let us first provide the equations associated to this model in order to describe the variance $\tilde{\Theta}_n^2$. The following system in $(\delta, \tilde{\delta})$ admits a unique pair of nonnegative solutions $(\delta > 0, \tilde{\delta} > 0)$ (see for instance [16, Theorem 1]):

$$\begin{aligned}
\delta &= \frac{1}{n+m}\mathrm{tr}\,\mathbf{\Delta}\left(\rho(\mathbf{I}_N + \tilde{\delta}\mathbf{\Delta}) + \mathbf{\Gamma}(\mathbf{I}_{n+m} + \delta\tilde{\mathbf{\Delta}})^{-1}\mathbf{\Gamma}^{\mathsf{H}}\right)^{-1} \\
\tilde{\delta} &= \frac{1}{n+m}\mathrm{tr}\,\tilde{\mathbf{\Delta}}\left(\rho(\mathbf{I}_{n+m} + \delta\tilde{\mathbf{\Delta}}) + \mathbf{\Gamma}^{\mathsf{H}}(\mathbf{I}_N + \tilde{\delta}\mathbf{\Delta})^{-1}\mathbf{\Gamma}\right)^{-1} .
\end{aligned}$$

Introduce the matrices

$$\mathbf{\Upsilon} = \left(\rho(\mathbf{I}_N + \tilde{\delta}\mathbf{\Delta}) + \mathbf{\Gamma}(\mathbf{I}_{n+m} + \delta\tilde{\mathbf{\Delta}})^{-1}\mathbf{\Gamma}^{\mathsf{H}}\right)^{-1} \quad \text{and} \quad \tilde{\mathbf{\Upsilon}} = \left(\rho(\mathbf{I}_{n+m} + \delta\tilde{\mathbf{\Delta}}) + \mathbf{\Gamma}^{\mathsf{H}}(\mathbf{I}_N + \tilde{\delta}\mathbf{\Delta})^{-1}\mathbf{\Gamma}\right)^{-1} ,$$

and the quantities $\gamma = (n+m)^{-1}\mathrm{tr}\,\mathbf{\Delta}^2\mathbf{\Upsilon}^2$ and $\tilde{\gamma} = (n+m)^{-1}\mathrm{tr}\,\tilde{\mathbf{\Delta}}^2\tilde{\mathbf{\Upsilon}}^2$. These quantities enable us to express the variance associated to the CLT:

$$\tilde{\Theta}_n^2 = -\log\left(\left(1 - \frac{1}{m+n}\mathrm{tr}\,\mathbf{\Delta}^{1/2}\mathbf{\Upsilon}\mathbf{\Gamma}(\mathbf{I} + \delta\tilde{\mathbf{\Delta}})^{-1}\tilde{\mathbf{\Delta}}(\mathbf{I} + \delta\tilde{\mathbf{\Delta}})^{-1}\mathbf{\Gamma}^{\mathsf{H}}\mathbf{\Upsilon}\mathbf{\Delta}^{1/2}\right)^2 - \rho^2\gamma\tilde{\gamma}\right) .$$

Due to the particular form of the matrices associated to the extended model, one can readily prove that the variance takes the simpler form $\tilde{\Theta}_n^2 = -\log(1 - \rho^2\gamma\tilde{\gamma})$.

It remains now to prove that $\tilde{\Theta}_n = \Theta_n$.

Easy matrix computations yield to the fact that

$$\mathbf{\Upsilon} \;=\; \left[\rho\left(\mathbf{I} + \tilde{\delta}\mathbf{\Delta}\right) + \mathbf{A}\mathbf{A}^*\right]^{-1} \quad \text{and} \quad \tilde{\mathbf{\Upsilon}} = \begin{bmatrix} \rho\left(\mathbf{I} + \alpha\delta\tilde{\mathbf{D}}\right) & \mathbf{0} \\ \mathbf{0} & \rho\mathbf{I} + \mathbf{A}^*\left(\mathbf{I} + \alpha\tilde{\delta}\mathbf{D}\right)\mathbf{A} \end{bmatrix}^{-1} . \tag{14}$$

Hence, considering $\tilde{\mathbf{\Upsilon}}$ as a block-diagonal matrix of inverses, we get:

$$\tilde{\delta} \;=\; \frac{\alpha}{\rho(n+m)}\mathrm{tr}\,\tilde{\mathbf{D}}\left[\mathbf{I}+\alpha\delta\tilde{\mathbf{D}}\right]^{-1} \quad \text{and} \quad \tilde{\gamma} \;=\; \frac{1}{n+m}\mathrm{tr}\,\tilde{\mathbf{\Delta}}^2\tilde{\mathbf{\Upsilon}}^2 \;=\; \frac{1}{\rho^2 n}\mathrm{tr}\,\tilde{\mathbf{D}}^2\left[\mathbf{I}+\delta\tilde{\mathbf{D}}\alpha\right]^{-2} \;. \tag{15}$$

Consider Eq. (8), which defines $\mathbf{T}$, for $z=-\rho$; note that $\mathbf{D}_j = \tilde{d}_j\mathbf{D}$ and introduce $\kappa = \frac{1}{n}\mathrm{tr}\,\mathbf{D}\mathbf{T}_\rho$ so that $\mathbf{T}_\rho$ satisfies the equation:

$$\mathbf{T}_\rho \;=\; \left[\rho\mathbf{I}+\mathbf{A}\mathbf{A}^* + \frac{1}{n}\sum_{j=1}^{n}\frac{\tilde{d}_j\mathbf{D}}{1+\kappa\tilde{d}_j}\right]^{-1} \;=\; \left[\rho\mathbf{I}+\mathbf{A}\mathbf{A}^* + \left(\frac{1}{n}\mathrm{tr}\,\tilde{\mathbf{D}}(\mathbf{I}+\kappa\tilde{\mathbf{D}})^{-1}\right)\mathbf{D}\right]^{-1}$$

Considering the definitions of $\mathbf{\Upsilon}$ and $\tilde{\delta}$ as given in Eq. (14) and (15), one can prove that $\mathbf{T}_\rho$ and $\mathbf{\Upsilon}$ satisfy the same equation and hence are equal. In particular, $\kappa = \alpha\delta$ and $\tilde{\Theta}_n^2$ writes:

$$\tilde{\Theta}_n^2 = -\log\left(1-\frac{1}{n}\mathrm{tr}\,\mathbf{D}^2\mathbf{T}_\rho^2 \times \frac{1}{n}\mathrm{tr}\,\mathbf{D}^2\left[\mathbf{I}+\kappa\tilde{\mathbf{D}}\right]^{-2}\right) \;.$$

We now rewrite Θ_n^2 as given in Claim 1.

$$\mathbf{J}_{k\ell} \;=\; \frac{1}{n}\frac{\frac{1}{n}\tilde{d}_k\tilde{d}_\ell\mathrm{tr}\,\mathbf{D}^2\mathbf{T}_\rho^2}{(1+\tilde{d}_\ell)^2} \;.$$

Hence,

$$\mathbf{J} = \left(\frac{1}{n^2}\,\mathrm{tr}\,\mathbf{D}^2\mathbf{T}_\rho^2\right)\mathbf{u}^{\mathsf{H}}\mathbf{v} \;,$$

where $\mathbf{u} = [\tilde{d}_1,\ldots,\tilde{d}_n]$ and $\mathbf{v} = [\tilde{d}_1(1+\kappa\tilde{d}_1)^{-2},\ldots,\tilde{d}_n(1+\kappa\tilde{d}_n)^{-2}]$. Now,

$$\begin{aligned} -\log\det\left(\mathbf{I}-\left(\frac{1}{n^2}\,\mathrm{tr}\,\mathbf{D}^2\mathbf{T}_\rho^2\right)\mathbf{u}^{\mathsf{H}}\mathbf{v}\right) \;&=\; -\log\left(1-\left(\frac{1}{n^2}\,\mathrm{tr}\,\mathbf{D}^2\mathbf{T}_\rho^2\right)\mathbf{v}\mathbf{u}^{\mathsf{H}}\right) \\ &=\; -\log\left(1-\frac{1}{n}\mathrm{tr}\,\mathbf{D}^2\mathbf{T}_\rho^2\times\frac{1}{n}\mathrm{tr}\,\mathbf{D}^2\left[\mathbf{I}+\kappa\tilde{\mathbf{D}}\right]^{-2}\right) \;, \end{aligned}$$

which is exactly the expression of $\tilde{\Theta}_n^2$ and the proof is completed. ■

IV. Numerical Results

In order to verify the accuracy of the analysis in the preceding sections, we provide now some simulation results. We consider a variance profile where each σ_{ij}^2 is drawn randomly from the interval $[0,10]$. The interference covariance matrix $\mathbf{A}\mathbf{A}^{\mathsf{H}}$ is also generated in a random fashion by letting $\mathbf{A} = \frac{1}{\sqrt{N}}\mathbf{X}$, where $\mathbf{X}$ is a standard complex Gaussian $N\times m$ matrix. Both the variance profile (σ_{ij}^2) and $\mathbf{A}$ are chosen at random at the beginning of the simulations and then kept constant. We define the signal-to-noise-ratio as SNR $= \frac{1}{\rho}$ and let $m=3$. Figure 1 shows the normalized ergodic mutual information $I(\rho)$ versus SNR for several different values of N and n. Solid lines represent the deterministic equivalent approximation $V(\rho)$ as given by Theorem 1. Markers are obtained by Monte Carlo simulations for 10000 different realizations of $\mathbf{H}$. We observe a very good fit between both results which demonstrates that the asymptotic analysis yields accurate approximations for small channel dimensions. Figure 2 depicts the histogram of the random variable $\frac{N}{\Theta_n}(\mathcal{I}(\rho)-V(\rho))$ in comparison with the normal distribution for two different pairs of parameters N,n. While we observe some mismatch for the case $N=2$ receive antennas and $n=1$ transmitter, the overlap is almost perfect for a slightly larger system with $N=16$ and $n=8$. These plots further validate the CLT as stated in Claim 1.

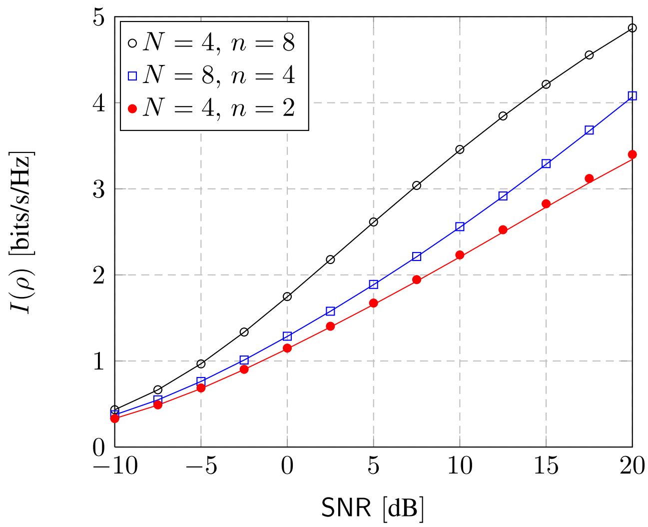
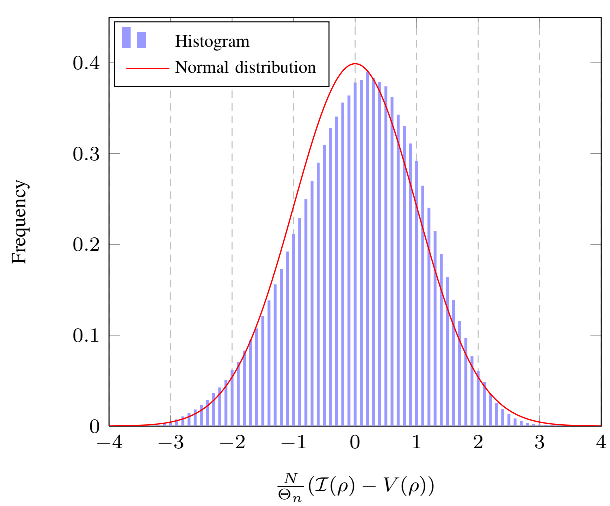
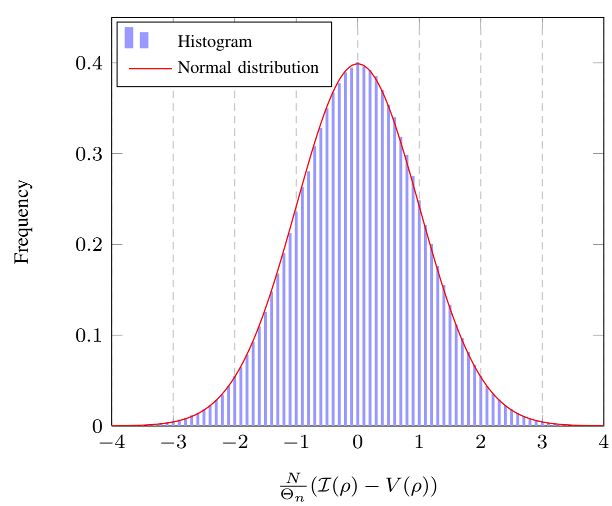


Fig. 1: Normalized ergodic mutual information $I(\rho)$ versus SNR for different channel dimensions N, n. Solid lines correspond to the deterministic equivalent approximation $V(\rho)$. Markers are obtained by Monte Carlo simulations.



(a) $N = 2$, $n = 1$



(b) $N = 16$, $n = 8$

Fig. 2: Histogram of $\frac{N}{\Theta_n}(\mathcal{I}(\rho) - V(\rho))$ in comparison with the normal distribution $\mathcal{N}(0, 1)$.

V. Conclusions

We have studied the fluctuations of the mutual information of a class of large-dimensional MIMO channels with arbitrary colored noise. First, we have provided a deterministic approximation of the mutual information in the asymptotic limit. Second, we have established the fluctuations of the mutual information around this approximation in form of a CLT. Both analytical results have then been confirmed by simulations and it was shown that the asymptotic results yield accurate approximations for even small channel dimensions.

References

[1] 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.

[2] S. Verdu and S. Shamai, "Spectral Efficiency of CDMA with Random Spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622 –640, Mar. 1999.

[3] C.-N. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity Scaling in MIMO Wireless Systems Under Correlated Fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637 –650, Mar. 2002.

[4] A. Moustakas, S. Simon, and A. Sengupta, "MIMO Capacity Through Correlated Channels in the Presence of Correlated Interferers and Noise: A (Not So) Large N Analysis," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.

[5] R. Couillet, M. Debbah, and J. Silverstein, "A Deterministic Equivalent for the Capacity Analysis of Correlated Multi-User MIMO Channels," *IEEE Trans. Inf. Theory*, 2010, Second round review. [Online]. Available: http://www.supelec.fr/offres/file_inline_src/342/342_P_11867_15.pdf

[6] W. Hachem, P. Loubaton, and J. Najim, "Deterministic Equivalents for Certain Functionals of Large Random Matrices," *Annals of Applied Probability*, vol. 17, no. 3, pp. 875–930, 2007.

[7] Z. D. Bai and J. W. Silverstein, "CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices," *Ann. Prob.*, vol. 32, no. 1A, pp. 553–605, 2004.

[8] 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.

[9] O. Somekh, O. Simeone, Y. Bar-Ness, A. Haimovich, U. Spagnolini, and S. Shamai (Shitz), *An Information Theoretic View of Distributed Antenna Processing in Cellular Systems: Open Architecture for Future Wireless Communications*. Auerbach Publications, CRC Press, New York, NY, USA, 2007.

[10] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference," *IEEE J. Sel. Areas Commun., Special issue on "Cooperative Communications in MIMO Cellular Networks"*, 2010, to appear.

[11] D. Aktas, M. N. Bacha, J. E. Evans, and S. Hanly, "Scaling Results on the Sum Capacity of Cellular Networks with MIMO Links," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3264–3274, Jul. 2006.

[12] H. Huh, G. Caire, S.-H. Moon, and I. Lee, "Multi-Cell MIMO Downlink with Fairness Criteria: The Large-System Limit," in *Proc. International Symposium on Information Theory (ISIT'10)*, Austin, Texas, Jun. 2010. [Online]. Available: http://arxiv.org/abs/1001.1133

[13] G. Taricco, "On the Capacity of Separately-Correlated MIMO Rician Fading Channels," in *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, Nov. 2006, pp. 1 –5.

[14] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A New Approach for Mutual Information Analysis of Large Dimensional Multi-Antenna Channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987 –4004, Sep. 2008.

[15] M. Kharouf, "Fluctuations de fonctionnelles spectrales de grandes matrices aléatoires et applications aux communications numériques," Ph.D. dissertation, Télécom Paristech (Paris) and Université Hassan II (Casablanca), Jun. 2010, *available upon request*.

[16] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim, "On the Capacity Achieving Covariance Matrix for Rician MIMO Channels: An Asymptotic Approach," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1048–1069, Mar. 2010.