On the MIMO Channel Capacity for the Nakagami-m Channel

Gustavo Fraidenraich, *Student Member, IEEE*, Olivier Lévêque, *Member, IEEE*, and John M. Cioffi, *Fellow, IEEE*

***Abstract*—This paper presents the multiple-input multiple-output (MIMO) channel capacity over the Nakagami-m fading channel. The joint eigenvalue density function of $W = HH^\dagger$, where H is the channel matrix, is derived in closed form for $\boldsymbol{H}$ (2×2) and any integer values of m, as well as for $\boldsymbol{H}$ (2×3) with $m = 2$ and $m = 3$. The marginal eigenvalue distribution of W is also derived in a closed-form expression. All the results are validated by numerical Monte Carlo simulations and are in excellent agreement.**

***Index Terms*—Eigenvalue distribution, fading distributions, multiple-input multiple-output (MIMO) channels, Nakagami-m distribution, Rayleigh distribution.**

I. Introduction

It has been acknowledged in recent years that the use of multiple-inputs multiple-outputs (MIMOs) can potentially provide large spectral efficiency for wireless communications in the presence of multipath fading environments. In the papers presented by Winters [1], Foschini [2], and Telatar [3], the capacity was in particular shown to scale linearly with the number of antennas.

In most previous research on MIMO capacity, the channel fading is assumed to be Rayleigh distributed. Of course, the Rayleigh-fading model is known to be a reasonable assumption for the fading encountered in many wireless communications systems. Nevertheless, many measurements campaigns [4], [5] show that the Nakagami-m distribution provides a much better fitting for the fading channel distribution. In fact, since the Nakagami-m distribution has one more free parameter, it allows for more flexibility. It moreover contains the Rayleigh distribution $(m = 1)$, the one-sided Gaussian distribution $(m \to 0.5)$, and the uniform distribution on the unit circle[1] $(m \to \infty)$ as special (extreme) cases.

The Nakagami-m distribution is a general, but approximate solution to the random phase problem [6]. The exact solution to this problem involves the knowledge of the distribution and the correlations of all of the partial waves composing the total signal and becomes infeasible due to its complexity [7]. This has been circumvented by Nakagami [6] who, through empirical methods based on field measurements followed by a curve-fitting process, obtained the approximate distribution.

Since the publication of [3], other distributions have been considered, as in [8] for the Ricean case, and recently [9] addressed the Hoyt distribution case. A more realistic MIMO Rayleigh channel including correlation has been addressed in [10]. The Nakagami-m distribution has also been addressed for some special cases as in [11] for the single-input multiple-output (SIMO) and multiple-input single-output (MISO) cases, and in [12] for the keyhole channel.

The MIMO channel capacity can be computed by means of the joint eigenvalue density function (JEDF) of the matrix $\boldsymbol{W} = \boldsymbol{H}\boldsymbol{H}^\dagger$, where † denotes the complex conjugate transpose. In the classical Rayleigh fading model, the entries of $\boldsymbol{H}$ are assumed to be i.i.d. zero mean complex Gaussian, resulting in a matrix $\boldsymbol{W}$ which is Wishart distributed. This model takes advantages of the numerous results provided by the literature [13], [14]. Unfortunately, as one departs from this model, there is not too much that can be said.

The result presented here follows the sequence of decompositions $\boldsymbol{H} = \boldsymbol{L}Q, \boldsymbol{W} = \boldsymbol{L}L^\dagger$, and $\boldsymbol{W} = \boldsymbol{S\Delta S}^\dagger$. The decomposition through the unitary matrices $\boldsymbol{Q}$ provides a way to integrate over the volume dQ, which $\boldsymbol{W} = \boldsymbol{H}\boldsymbol{H}^\dagger$ does not provide. Denoting dH as $dH = J\, dL\, dQ$ where J is the Jacobian of the transformation, the distribution of $p(\boldsymbol{H})$ does not depend on Q for the Rayleigh case. Hence, the integration on the Stiefel manifold can be computed in closed form [15]. Unfortunately, for the Nakagami-m case, $p(\boldsymbol{H})$ depends on Q, so a larger number of variables has to be integrated. To be more accurate, $2rt - r^2$ integrations over real variables are necessary (the matrix $\boldsymbol{H}$ has $2rt$ real variables and $\boldsymbol{L}$ has r^2 real variables, so $\boldsymbol{Q}$ has $2rt - r^2$ variables). For either $m = 1$ or $m = 0.5$, the resulting integral is simply the volume of the Stiefel manifold.

This paper presents the MIMO channel capacity over the Nakagami-m fading channel for $\boldsymbol{H}$ with dimensions 2×2 and 2×3. Channel state information is assumed at the receiver side only. In this case, it is shown that the uniform power distribution across the transmitting antennas achieves the channel capacity. Assuming that the entries of $\boldsymbol{H}$ are i.i.d. with Nakagami-m distributed envelopes and uniform phases, an elegant and simple expression for the JEDF of $\boldsymbol{W}$ is derived.

The paper is organized as follows. Section II. introduces the Nakagami distribution in more detail. Section III defines channel capacity, and Section IV presents the main result for $\boldsymbol{H}$ 2×2 and any integer m. Section V provides the result for $\boldsymbol{H}$ 2×3, and $m = 2$ or $m = 3$. Section VI presents the asymptotic result. Section VII presents some numerical simulations and Section VIII draws the conclusion.

II. The Nakagami-m Distribution

The entries of the $r \times t$ channel matrix $\boldsymbol{H}$ are assumed to be i.i.d. and distributed as

$$Z = R\, \exp(j\Theta) \tag{1}$$

where the phase Θ is uniformly distributed and independent of the envelope R. R is in turn given by

$$R^2 = \sum_{i=1}^{m} X_i^2 + Y_i^2 \tag{2}$$

where X_i and Y_i are i.i.d. zero mean Gaussian distributed with variance $\Omega/2m$. The distribution of R is therefore the Nakagami-m distribution [6] given by

$$p(r) = \frac{2m^m r^{2m-1}}{\Omega^m \Gamma(m)} \exp\left(-\frac{mr^2}{\Omega}\right) \tag{3}$$

where $\Omega = \mathbb{E}[R^2]$, $m = \mathbb{E}[R^2]^2/\mathrm{Var}[R^2]$, and $\Gamma(\cdot)$ denotes the Euler Gamma function. In addition, $p(r,\theta) = p(r)\frac{1}{2\pi}$, yielding

$$p(r,\theta) = \frac{m^m r^{2m-1}}{\pi\, \Omega^m \Gamma(m)} \exp\left(-\frac{mr^2}{\Omega}\right) \tag{4}$$

Manuscript received April 24, 2007; revised December 13, 2007. The work of G. Fraidenraich was supported by the Cnpq (Brazil) under Grant 200869/2005-1. The work of O. Lévêque was supported in part by the Swiss NSF under Grant PA002-108987.

G. Fraidenraich and J. M. Cioffi are with the Electrical Engineering Department, Stanford University, Stanford, CA 94305 USA (e-mail: gustavofraidenra@gmail.com; cioffi@stanford.edu).

O. Lévêque was with the Electrical Engineering Department, Stanford University, Stanford, CA 94305 USA. He is now with the Ecole Polytechnique Fédérale de Lausanne, Switzerland (e-mail: oliver.leveque@epfl.ch).

Communicated by A. J. Grant, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2008.926467

[1]When the uniform phase distribution is considered.

0018-9448/$25.00 © 2008 IEEE

Note that $p(r,\theta)$ reduces to the Rayleigh distribution for $m = 1$ and to the uniform distribution on the circle of radius $\sqrt{\Omega}$ for $m \to \infty$. The above family of distributions therefore allows to interpolate between the classical Rayleigh distribution and the "pure random phase" distribution.

Using the standard polar-rectangular transformation, the joint distribution of the real and imaginary parts of Z is given by $p(x,y) = \frac{p(r,\theta)}{r}$, thus

$$p(x,y) = \frac{m^m (x^2+y^2)^{m-1}}{\pi\,\Omega^m \Gamma(m)}\, e^{-\frac{m(x^2+y^2)}{\Omega}}. \tag{5}$$

III. MIMO Channel Capacity

The following MIMO single-user Gaussian channel is considered, with t antennas at the transmitter and r antennas at the receiver

$$\boldsymbol{y} = \boldsymbol{H}\,\boldsymbol{x} + \boldsymbol{n} \tag{6}$$

$\boldsymbol{H}$ is the $r \times t$ channel matrix with i.i.d. entries h_{ij}, each distributed as the random variable Z defined in (1).

The vector $\boldsymbol{y} \in \mathcal{C}^r$, $\boldsymbol{x} \in \mathcal{C}^t$, and $\boldsymbol{n}$ is zero-mean complex Gaussian noise with $\mathbb{E}[\boldsymbol{n}\,\boldsymbol{n}^\dagger] = \boldsymbol{I}$. In addition, a total transmit power constraint $\mathbb{E}[\boldsymbol{x}^\dagger \boldsymbol{x}] \le P$ is assumed.

For a given input covariance matrix $\boldsymbol{\Sigma}$, the MIMO instantaneous capacity is given by

$$C(\boldsymbol{\Sigma}) = \log_2 \det(\boldsymbol{I} + \boldsymbol{H}\boldsymbol{\Sigma}\boldsymbol{H}^\dagger) \tag{7}$$

and the MIMO channel capacity in the absence of channel knowledge at the transmitter is given by [3]

$$C = \sup_{\boldsymbol{\Sigma}\ge 0 : \mathrm{tr}[\boldsymbol{\Sigma}] \le P} \mathbb{E}[C(\boldsymbol{\Sigma})]. \tag{8}$$

Since the entries h_{ij} are i.i.d. and the distribution of h_{ij} is the same as that of $-h_{ij}$ for all i,j, one obtains from [16, Corollary 1b] that the uniform power allocation over the t transmit antennas achieves the channel capacity. The capacity is therefore given by

$$C = \mathbb{E}\left[\log_2 \det\left(\boldsymbol{I} + \frac{P}{t}\boldsymbol{W}\right)\right] \tag{9}$$

where

$$\boldsymbol{W} = \begin{cases} \boldsymbol{H}\boldsymbol{H}^\dagger & r \le t \\ \boldsymbol{H}^\dagger \boldsymbol{H} & r > t \end{cases} \tag{10}$$

IV. Main Result for $\boldsymbol{H}$ 2×2

Theorem 1: The JEDF of the 2×2 matrix $\boldsymbol{W} = \boldsymbol{H}\boldsymbol{H}^\dagger$, where the entries of the 2×2 matrix $\boldsymbol{H}$ are i.i.d. with Nakagami-m envelope and uniform phase, is given by

$$p(\lambda_1,\lambda_2) = K_{22}\, e^{-\frac{m(\lambda_1+\lambda_2)}{\Omega}} (\lambda_1-\lambda_2)^2\, F(\lambda_1,\lambda_2) \tag{11}$$

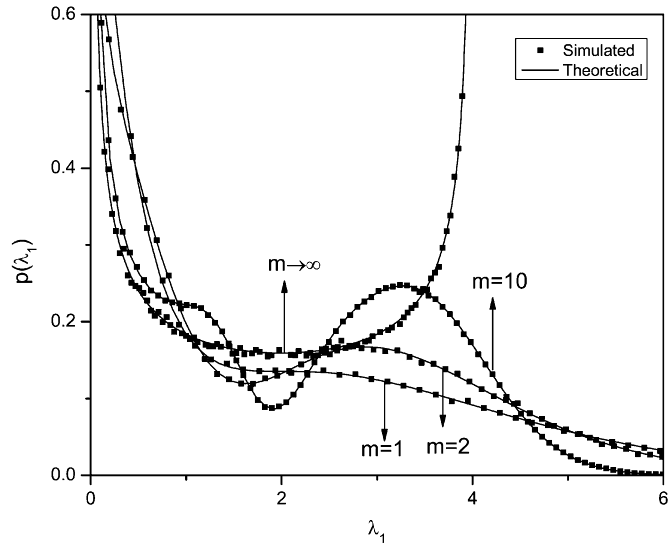

Fig. 1. The Nakagami eigenvalue distribution function for $\boldsymbol{H}$ (2×2) and $m = 1, 2, 10, 100$ $(\Omega = 1)$.

where

$$K_{ij} = \left(\frac{m^m}{\pi\Omega^m\Gamma(m)}\right)^{i\,j} \tag{12}$$

and $F(\lambda_1,\lambda_2)$ is given by (13) shown at the bottom of the page, and $f(i_1,i_2,k_1,k_2)$ is given by (14) shown at the bottom of the page.

In order to gain some intuition on the above result, note that the function $F(\lambda_1,\lambda_2)$ is a homogeneous polynomial of order $4(m-1)$. In the Rayleigh case $(m=1)$, F is a constant and one recovers the classical JEDF of a Wishart matrix [13]:

$$p(\lambda_1,\lambda_2) = \frac{1}{2\Omega^4}\, e^{-\frac{(\lambda_1+\lambda_2)}{\Omega}} (\lambda_1-\lambda_2)^2. \tag{15}$$

The effect of the Nakagami-m envelope distribution on the JEDF is therefore expressed by the polynomial $F(\lambda_1,\lambda_2)$.

To illustrate the effect of the parameter m, Fig. 1 shows the resulting marginal eigenvalue distribution $p(\lambda)$ for $m = 1, 2, 10, \infty$ (see also Corollary 2). It can be observed that as the parameter m increases, the eigenvalues concentrate on the interval $[0, 4\Omega]$, with higher probability on the boundary of the interval. It can be also seen in Fig. 1 that the result is in perfect agreement with Monte-Carlo simulations.

Proof: In order to get the JEDF of $\boldsymbol{W}$, the following steps need to be performed: 1) the joint distribution of $\boldsymbol{H}$ can be easily found, since its entries are i.i.d.; 2) the matrix $\boldsymbol{H}$ is decomposed as $\boldsymbol{H} = \boldsymbol{L}Q$, where $\boldsymbol{L}$ is a complex lower triangular matrix with real positive diagonals and $\boldsymbol{Q}$ is a complex unitary matrix $(\boldsymbol{Q}Q^\dagger = \boldsymbol{I})$; 3) therefore, $\boldsymbol{W} = \boldsymbol{L}\boldsymbol{L}^\dagger$;

$$F(\lambda_1,\lambda_2) = \frac{\pi^4}{\Gamma(4m-2)} \sum_{i_1=0}^{m-1}\sum_{i_2=0}^{m-1}\sum_{k_1=0}^{i_1}\sum_{k_2=0}^{i_2} f(i_1,i_2,k_1,k_2)(\lambda_1\lambda_2)^{\frac{1}{2}(i_1+i_2-2(k_1+k_2))+m-1}(\lambda_1-\lambda_2)^{-i_1-i_2+2(k_1+k_2+m-1)} \tag{13}$$

$$f(i_1,i_2,k_1,k_2) = \frac{\binom{i_1}{k_1}\binom{i_2}{k_2}\binom{m-1}{i_1}\binom{m-1}{i_2}(-1)^{-i_1-2i_2+3m+1}\, 2^{-2(k_1+k_2+1)}\Gamma\left(-\frac{i_1}{2}-\frac{i_2}{2}+m-\frac{1}{2}\right)}{\Gamma\left(-\frac{i_1}{2}-\frac{i_2}{2}+m\right)\Gamma\left(-\frac{i_1}{2}-\frac{i_2}{2}+k_1+k_2+m+\frac{1}{2}\right)} \times \Gamma\left(-\frac{i_1}{2}-\frac{i_2}{2}+k_1+k_2+m\right)\Gamma\left(-\frac{i_1}{2}+\frac{i_2}{2}+k_1-k_2+2m-1\right)\Gamma\left(\frac{i_1}{2}-\frac{i_2}{2}-k_1+k_2+2m-1\right). \tag{14}$$

4) finally, performing the eigenvalue decomposition $\boldsymbol{W} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\dagger$, one obtains the JEDF of $\boldsymbol{W}$.

Since the entries of the channel matrix $\boldsymbol{H}$ are independent, their joint distribution is given by

$$p(\boldsymbol{H}) = K_{22} \exp\left(-\frac{m\,\mathrm{tr}(\boldsymbol{H}\boldsymbol{H}^\dagger)}{\Omega}\right) \prod_{i,j=1}^{2} |h_{ij}|^{2(m-1)}. \tag{16}$$

Using the LQ decomposition, the matrix $\boldsymbol{H}$ can be written as $\boldsymbol{H} = \boldsymbol{L}Q$, where the matrix $\boldsymbol{Q}$ is given as [17]

$$\boldsymbol{Q} = \begin{pmatrix} e^{j\phi_1}\cos(\theta) & e^{j\phi_2}\sin(\theta) \\ -e^{j(\phi_3-\phi_2)}\sin(\theta) & e^{j(\phi_3-\phi_1)}\cos(\theta) \end{pmatrix} \tag{17}$$

and the variables are defined in the following range $0 \le \phi_1, \phi_2, \phi_3 \le 2\pi, 0 \le \theta \le \pi/2$. The matrix $\boldsymbol{L}$ is given by

$$\boldsymbol{L} = \begin{pmatrix} l_{11} & 0 \\ l_{21R} + j\,l_{21I} & l_{22} \end{pmatrix}. \tag{18}$$

The Jacobian of this transformation is given by $|J| = l_{11}^3 l_{22}\,\sin(\theta)\cos(\theta)$, so the joint probability density function (PDF) of $\boldsymbol{L}$ and $\boldsymbol{Q}$ is given by

$$\begin{aligned} p(\boldsymbol{L},\boldsymbol{Q}) &= \frac{K_{22}}{2^{2m-1}} e^{(-\frac{m\,\mathrm{tr}(\boldsymbol{L}\boldsymbol{L}^\dagger)}{\Omega})} l_{11}^{4m-1} l_{22} \sin^{2m-1}(2\theta) \\ &\quad \times \left(l_{22}^2\cos^2(\theta) + T_{1L}\sin^2(\theta) - T_{2L}\sin(2\theta)\right)^{m-1} \\ &\quad \times \left(T_{1L}\cos^2(\theta) + l_{22}^2\sin^2(\theta) + T_{2L}\sin(2\theta)\right)^{m-1} \end{aligned} \tag{19}$$

where

$$T_{1L} = \left(l_{21I}^2 + l_{21R}^2\right) \tag{20}$$

and

$$\begin{aligned} T_{2L} &= l_{22}(l_{21I}\sin(\phi_1+\phi_2-\phi_3) \\ &\quad - l_{21R}\cos(\phi_1+\phi_2-\phi_3)) \end{aligned} \tag{21}$$

Since m is integer, it is possible to use the classical binomial expansion in (19). Therefore

$$\begin{aligned} p(\boldsymbol{L},\boldsymbol{Q}) &= \frac{K_{22}}{2^{2m-1}} e^{(-\frac{m\,\mathrm{tr}(\boldsymbol{L}\boldsymbol{L}^\dagger)}{\Omega})} l_{11}^{4m-1} l_{22} \\ &\quad \times \sum_{i_1=0}^{m-1}\sum_{k_1=0}^{i_1}\sum_{i_2=0}^{m-1}\sum_{k_2=0}^{i_2} \binom{m-1}{i_1}\binom{i_1}{k_1}\binom{m-1}{i_2} \\ &\quad \times \binom{i_2}{k_2}(-1)^{m-1-i_2} T_{1L}^{k_1+k_2} T_{2L}^{2m-2-i_1-i_2} \\ &\quad \times l_{22}^{2(i_2-k_2+i_1-k_1)} \\ &\quad \times \cos^{2(k_1+i_2-k_2)}(\theta)\,\sin^{2(i_1-k_1+k_2)} \\ &\quad \times (\theta)\,\sin(2\theta)^{4m-3-i_1-i_2}. \end{aligned} \tag{22}$$

Integrating now over θ, ϕ_1, ϕ_2, and ϕ_3, the distribution of $p(\boldsymbol{L})$ is obtained as (23) shown at the bottom of the page. The next transformation is given by

$$\boldsymbol{W} = \boldsymbol{L}L^\dagger = \begin{pmatrix} w_1 & w_3 - jw_4 \\ w_3 + jw_4 & w_2 \end{pmatrix} \tag{24}$$

where the Jacobian of this change is given by $|J| = 4l_{11}^3 l_{22}$. Using this transformation, the distribution of $p(\boldsymbol{W})$ can be easily obtained.

The next step it to apply the eigenvalue decomposition $\boldsymbol{W} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\dagger$, where the matrix $\boldsymbol{S}$ is given by

$$\boldsymbol{S} = \begin{pmatrix} \cos(\kappa) & -e^{i\psi}\sin(\kappa) \\ e^{-j\psi}\sin(\kappa) & \cos(\kappa) \end{pmatrix} \tag{25}$$

$0 \le \psi \le 2\pi, 0 \le \kappa \le \pi/2$, and $\boldsymbol{\Lambda}$ is the eigenvalue matrix given by

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{26}$$

The Jacobian of this last transformation can be obtained as $|J| = \frac{1}{2}(\lambda_1 - \lambda_2)^2 \sin(2\kappa)$. Finally, applying this transformation and integrating over ψ and κ, the JEDF (11) is obtained. □

The following corollary is a direct consequence of Theorem 1.

Corollary 2: The marginal distribution $p(\lambda)$ is given by

$$\begin{aligned} p(\lambda) &= \frac{K_{22}\pi^4}{\Gamma(4m-2)} \sum_{i_1=0}^{m-1}\sum_{i_2=0}^{m-1}\sum_{k_1=0}^{i_1}\sum_{k_2=0}^{i_2}\sum_{i_3=0}^{2(k_1+k_2+m)-i_1-i_2} \\ &\quad \times f_1(i_1,i_2,k_1,k_2) f_2(i_1,i_2,k_1,k_2,i_3)\, e^{-\frac{m\lambda}{\Omega}} \\ &\quad \times \lambda^{\frac{1}{2}(i_1+i_2+2(m+i_3-k_1-k_2-1))} \end{aligned} \tag{27}$$

where $f_2(i_1,i_2,k_1,k_2,i_3)$ is given by (28)

$$\begin{aligned} f_2&(i_1,i_2,k_1,k_2,i_3) \\ &= \binom{-i_1-i_2+2(m+k_1+k_2)}{i_3} \\ &\quad \times (-1)^{2m-i_1-i_2-i_3+2k_1+2k_2} \\ &\quad \times m^{\frac{1}{2}(-6m+i_1+i_2+2i_3-2k_1-2k_2)}\, \Omega^{3m-\frac{i_1}{2}-\frac{i_2}{2}-i_3+k_1+k_2} \\ &\quad \times \Gamma\left(3m - \frac{i_1}{2} - \frac{i_2}{2} - i_3 + k_1 + k_2\right). \end{aligned} \tag{28}$$

$$\begin{aligned} p(\boldsymbol{L}) &= \frac{K_{22}\,8\,\pi^{5/2}}{2^{2m-1}} e^{-\frac{m\,\mathrm{tr}(\boldsymbol{L}L^\dagger)}{\Omega}} l_{11}^{4m-1} \sum_{i_1=0}^{m-1}\sum_{k_1=0}^{i_1}\sum_{i_2=0}^{m-1}\sum_{k_2=0}^{i_2} \binom{m-1}{i_1}\binom{i_1}{k_1}\binom{m-1}{i_2}\binom{i_2}{k_2} \\ &\quad \times (-1)^{m-1-i_2}\left(l_{21I}^2 + l_{21R}^2\right)^{k_1+k_2+\frac{1}{2}(2m-i_1-i_2-2)} l_{22}^{i_2-2k_2+i_1-2k_1+2m-1}\,\Gamma\left(\frac{1}{2}(2m-i_1-i_2-1)\right) \\ &\quad \times \frac{2^{4m-4-i_1-i_2}\Gamma\left(\frac{1}{2}(4m+2k_1+i_2-2k_2-2-i_1)\right)}{\Gamma(4(m-1))} \frac{\Gamma\left(\frac{1}{2}(i_1-2k_1+2k_2+4m-2-i_2)\right)}{\Gamma\left(\frac{1}{2}(2m-i_1-i_2)\right)}. \end{aligned} \tag{23}$$

A. *Special Cases:* $M = 1, 2, 3$ *and* ∞

As already mentioned, the JEDF (11) specializes to the classical Wishart distribution (15) in the case $m = 1$. For $m = 2$, the JEDF (11) specializes to

$$p(\lambda_1,\lambda_2) = \frac{4e^{-\frac{2(\lambda_1+\lambda_2)}{\Omega}}}{225\pi^3\Omega^8}(\lambda_1-\lambda_2)^2 \times \left(\lambda_1^4+\lambda_2\lambda_1^3+26\lambda_2^2\lambda_1^2+\lambda_2^3\lambda_1+\lambda_2^4\right). \quad (29)$$

In the same way for $m = 3$, the following is obtained

$$\begin{aligned}p(\lambda_1,\lambda_2) &= \frac{6561e^{-\frac{3(\lambda_1+\lambda_2)}{\Omega}}(\lambda_1-\lambda_2)^2}{1254400\pi^3\Omega^{12}}\\ &\times\left(\lambda_1^8+\lambda_2\lambda_1^7+37\lambda_2^2\lambda_1^6+37\lambda_2^3\lambda_1^5+478\lambda_2^4\lambda_1^4\right.\\ &\left.+37\lambda_2^5\lambda_1^3+37\lambda_2^6\lambda_1^2+\lambda_2^7\lambda_1+\lambda_2^8\right). \quad (30)\end{aligned}$$

The case $m \to \infty$ requires a slightly different analysis, which leads to the result below.

Theorem 3: The marginal eigenvalue distribution of the 2×2 matrix $\boldsymbol{W} = \boldsymbol{H}\boldsymbol{H}^\dagger$, where $\boldsymbol{H}$ (2×2) has i.i.d. entries with Nakagami-∞ envelope and uniform phase, is given by

$$p(\lambda) = \frac{1}{\pi\sqrt{4\Omega\lambda-\lambda^2}}1_{0<\lambda<4\Omega}. \quad (31)$$

Proof: As $m \to \infty$, the joint distribution of (R,Θ) given in (4) tends to the uniform distribution on the circle of radius $\sqrt{\Omega}$, therefore

$$\boldsymbol{H} = \sqrt{\Omega}\begin{pmatrix} e^{j\theta_{11}} & e^{j\theta_{12}}\\ e^{j\theta_{21}} & e^{j\theta_{22}}\end{pmatrix} \quad (32)$$

where θ_{ij} are uniformly distributed between 0 and 2π. In this case, the matrix $\boldsymbol{W}$ is given by

$$\boldsymbol{W} = \Omega\begin{pmatrix} 2 & w_{12}\\ \overline{w_{12}} & 2\end{pmatrix} \quad (33)$$

where $w_{12} = e^{j(\theta_{11}-\theta_{21})}+e^{j(\theta_{12}-\theta_{22})}$ and $\overline{w_{12}}$ denotes its conjugate. Using the decomposition $\boldsymbol{W} = \boldsymbol{S\Lambda S}^\dagger$, it is possible to write $\lambda_1+\lambda_2 = \mathrm{tr}(\boldsymbol{W}) = 4\Omega$ as well as $\lambda_1\lambda_2 = \Omega^2(4-|w_{12}|^2)$, so

$$\lambda_1 = \Omega(2\pm|w_{12}|) \quad (34)$$

and $\lambda_2 = 4\Omega - \lambda_1$. Note therefore that in this case, the correlation between the eigenvalues λ_1 and λ_2 is much stronger than in the finite m case. In particular, the joint eigenvalue density $p(\lambda_1,\lambda_2)$ does not exist.

The modulus of w_{12} can be written as $|w_{12}| = \sqrt{2(1+\cos(\theta_{11}-\theta_{21}-\theta_{12}+\theta_{22}))}$. Since the distribution of $\cos(\theta_{ij})$ is the same as the distribution of $y = \cos(\theta_{11}-\theta_{21}-\theta_{12}+\theta_{22})$, the distribution of this term is $p(y) = \frac{1}{\pi\sqrt{1-y^2}}1_{|y|<1}$. Making the transformation of variable and computing the distribution of $|w_{12}|$, the following distribution is obtained

$$p(|w_{12}|) = \frac{1}{\pi\sqrt{1-\frac{|w_{12}|^2}{4}}}1_{|w_{12}|<1} \quad (35)$$

From (34) and (35), (31) follows directly. □

In Fig. 1, the expression (31) is compared with simulations for $m = 100$, and a nearly perfect match is observable.

V. 2×3 Channel Matrix $\boldsymbol{H}$, With $m = 2$ and $m = 3$

We now address the case where $\boldsymbol{H}$ is a 2×3 matrix. Unfortunately, because the complexity of the computation, only the cases $m = 2$ and $m = 3$ will be presented[2], but the approach is valid for any integer m. Since the entries of the channel matrix $\boldsymbol{H}$ are independent, their joint distribution is given by

$$p(\boldsymbol{H}) = K_{23}\exp\left(-\frac{m\,\mathrm{tr}(\boldsymbol{H}\boldsymbol{H}^\dagger)}{\Omega}\right)\prod_{i=1}^{2}\prod_{j=1}^{3}|h_{ij}|^{2(m-1)} \quad (36)$$

where K_{23} is given in (12). Using the LQ decomposition, the matrix $\boldsymbol{H}$ can be written as $\boldsymbol{H} = \boldsymbol{L}Q$, where the matrix $\boldsymbol{Q}$, in this case, will be given by (37) shown at the bottom of the page. This unitary matrix was generated using [17] for the 3×3 case, and then two out of three columns were chosen. The work in [17] provides a way to generate a $n \times n$ unitary matrix starting from a 2×2 unitary matrix. Unfortunately, this method produces a null jacobian J in some cases. For this reason, the matrix in (17) was used as the seed for the generation of the 3×3 matrix. Using this method, the Jacobian of this transformation is given by $|J| = \frac{l_{11}^5 l_{22}^3}{4}\cos(\theta_2)^3\sin(\theta_2)\sin(2\theta_1)\sin(2\theta_3)$.

Following the same steps, i.e., $\boldsymbol{H} = \boldsymbol{L}\boldsymbol{Q}, \boldsymbol{W} = \boldsymbol{L}\boldsymbol{L}^\dagger$, and then $\boldsymbol{W} = \boldsymbol{S\Delta S}^\dagger$, the following distributions are obtained for $m = 2$ and $m = 3$, respectively

$$\begin{aligned}p(\lambda_1,\lambda_2) &= \frac{32e^{-\frac{2(\lambda_1+\lambda_2)}{\Omega}}\lambda_1\lambda_2(\lambda_1-\lambda_2)^2}{33075\Omega^{12}}\\ &\times\left(3\lambda_1^6+3\lambda_2\lambda_1^5+101\lambda_2^2\lambda_1^4+101\lambda_2^3\lambda_1^3\right.\\ &\left.+101\lambda_2^4\lambda_1^2+3\lambda_2^5\lambda_1+3\lambda_2^6\right) \quad (38)\end{aligned}$$

$$\begin{aligned}p(\lambda_1,\lambda_2) &= \frac{1594323e^{-\frac{3(\lambda_1+\lambda_2)}{\Omega}}\lambda_1\lambda_2(\lambda_1-\lambda_2)^2}{17955857920000\Omega^{18}}\\ &\times\left(450\lambda_1^{12}+450\lambda_2\lambda_1^{11}+24110\lambda_2^2\lambda_1^{10}+24110\lambda_2^3\lambda_1^9\right.\\ &+453539\lambda_2^4\lambda_1^8+453539\lambda_2^5\lambda_1^7+1067009\lambda_2^6\lambda_1^6\\ &+453539\lambda_2^7\lambda_1^5+453539\lambda_2^8\lambda_1^4+24110\lambda_2^9\lambda_1^3\\ &\left.+24110\lambda_2^{10}\lambda_1^2+450\lambda_2^{11}\lambda+450\lambda_2^{12}\right). \quad (39)\end{aligned}$$

[2]In fact, the result can be found for any m, but the number of summations is extremely large.

$$\boldsymbol{Q} = \begin{pmatrix} e^{j(\phi_1+\phi_5)}\cos(\theta_1)\cos(\theta_3) - e^{j(\phi_2+\phi_4)}\sin(\theta_1)\sin(\theta_2)\sin(\theta_3) & e^{j(\phi_2+\phi_3)}\cos(\theta_2)\sin(\theta_1)\\ -e^{j(\phi_5-\phi_2)}\cos(\theta_3)\sin(\theta_1) - e^{j(\phi_4-\phi_1)}\cos(\theta_1)\sin(\theta_2)\sin(\theta_3) & e^{j(\phi_3-\phi_1)}\cos(\theta_1)\cos(\theta_2)\\ -e^{-j\phi_3}\cos(\theta_2)\sin(\theta_3) & -e^{-j\phi_4}\sin(\theta_2)\end{pmatrix}^\dagger. \quad (37)$$

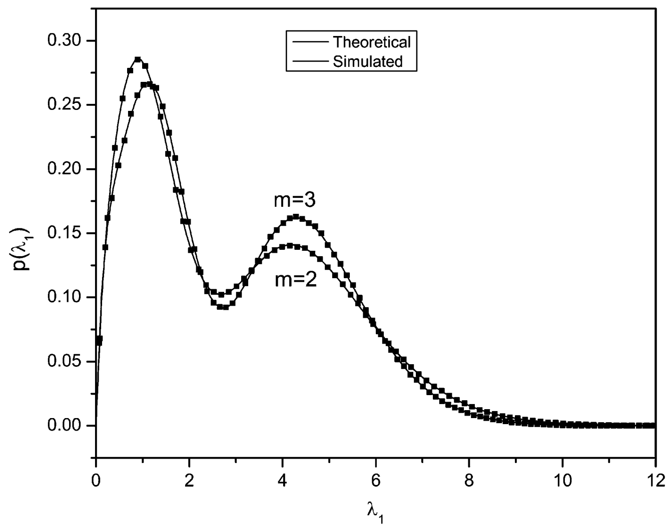

Fig. 2. The Nakagami eigenvalue distribution function for $\boldsymbol{H}$ (2×3) and $m = 2, 3$ ($\Omega = 1$).

In order to validate these expressions, the theoretical and simulated marginal eigenvalue distributions are plotted in Fig. 2. Note the perfect agreement between them.

VI. Asymptotic Case

In the case where the number of antennas grows to infinity, the result given in [18] can be used, since the entries of $\boldsymbol{H}$ are i.i.d. The limiting eigenvalue distribution of the matrix $\frac{1}{n}\boldsymbol{W}$ is given by

$$p(\lambda) = \frac{1}{2\pi\lambda\beta\Omega}\sqrt{(b-\lambda)(\lambda-a)} \quad a \le \lambda \le b \tag{40}$$

where $a = \Omega(1-\sqrt{\beta})^2, b = \Omega(1+\sqrt{\beta})^2, \beta = r/t$. Using (40) in (9), the following result is obtained for the asymptotic Nakagami channel capacity

$$\frac{C}{n} r, t \rightrightarrows \infty \int_a^b \log_2(1+P\lambda)\, p(\lambda)\, d\lambda. \tag{41}$$

Using the result presented in [19], one obtains

$$\frac{C}{n} r, t \rightrightarrows \infty \frac{1}{\beta \ln 2}(\beta \ln(1+P-P\,v(\beta,P)) + \ln(1+P\,\beta - P\,v(\beta,P)) - v(\beta,P)) \tag{42}$$

where

$$v(\beta,P) = \frac{1}{2}\left(1+\beta+\frac{1}{P}-\sqrt{(1+\beta+P^{-1})^2-4\,\beta}\right). \tag{43}$$

Although this formula is asymptotic, it is shown below by simulation that it is quite accurate, even for a small number of antennas.

VII. Numerical Results

As already presented, Fig. 1 validates with Monte-Carlo simulations the expression given in (11) for the JEDF for the cases $m = 1, 2, 10, \infty$. Note that there is an excellent agreement between the simulations and the theoretical results.

Fig. 3 compares the simulated channel capacity to the theoretical result (9) for the 2×2 case and $m = 0.5, 1, 20$. As can be seen in the figure, when m increases, the channel capacity also increases.

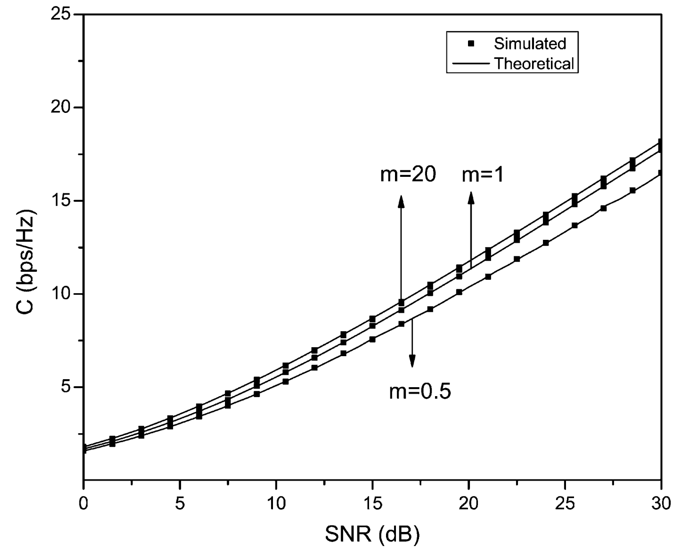

Fig. 3. The Nakagami channel capacity for the 2×2 case ($\Omega = 1$).

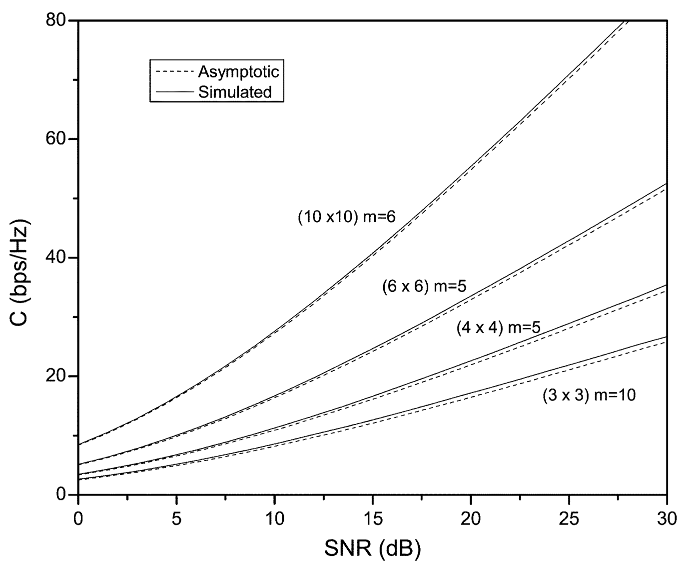

Fig. 4. The Asymptotic Nakagami channel capacity ($\Omega = 1$).

The difference between the case $m = 0.5$ and $m = 20$ is small for low values of the power P, and increases as the power increases. The effect of the Nakagami parameter m depends on the power P, but in the worst case, it degrades up to 12.5% of the capacity for a 2×2 channel matrix. In Fig. 4, for the $3 \times 3, 4 \times 4, 6 \times 6$, and 10×10 cases for different values of m, simulations are performed and compared with the asymptotic result given in (42).

VIII. Conclusion

In this paper, a general JEDF for the Nakagami-m channel is presented in a closed-form expression considering a 2×2 channel matrix $\boldsymbol{H}$ and any integer m. The marginal distribution is also derived in a closed-form expression, and for the $m \to \infty$ case, a simple and elegant expression is obtained. The JEDF was also found for the case where $\boldsymbol{H}$ is a 2×3 matrix, with $m = 2$ or $m = 3$. The ergodic MIMO channel capacity is computed for the Nakagami-m channel, and for the case $\boldsymbol{H}$ (2×2), the effect of the parameter m is shown to degrade up to 12.5% of the channel capacity. In all cases, the results are validated by Monte-Carlo simulations.

REFERENCES

[1] J. H. Winters, "On the capacity of radio communications systems with diversity in rayleigh fading environments," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 5, pp. 871–4878, Jun. 1987.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Feb. 1998.

[3] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[4] H. Suzuki, "A statistical model for urban radio propagation," *IEEE Trans. Commun.*, vol. 25, no. 7, pp. 673–679, 1977.

[5] T. Aulin, "Characteristics of a digital mobile channel type," *IEEE Trans. Veh. Technol.*, vol. 30, no. 1, pp. 45–53, Feb. 1981.

[6] M. Nakagami, "The m-distribution—A General formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*, W. C. Hoffman, Ed. Elmsford, NY: Pergamon, 1960.

[7] M. D. Yacoub, J. E. V. Bautista, and L. G. de Rezende Guedes, "On higher order statistics of the Nakagami-m distribution," *IEEE Trans. Veh. Technol.* , vol. 48, no. 3, pp. 790–794, May 1999.

[8] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in rician fading," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1102–1111, May 2005.

[9] G. Fraidenraich, O. Leveque, and J. M. Cioffi, "On the MIMO channel capacity for the dual and asymptotic cases over Hoyt channels," *IEEE Commun. Lett*, vol. 11, no. 1, pp. 31–33, Jan. 2007.

[10] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[11] F. Zheng and T. Kaiser, "On the channel capacity of multi-antenna systems with nakagami fading," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1–11, 2006.

[12] A. Müller and J. Speidel, "Ergodic capacity and information outage probability of mimo nakagami-m keyhole channels with general branch parameters," in *Proc. Wireless Commun. Networking Conf.—WCNC 2007*, Mar. 2007, pp. 2184–2189.

[13] A. T. James, "Distributions of matrix variates and latents roots derived from normal samples," *Ann. Math. Stat.*, vol. 35, no. 2, pp. 475–501, Jun. 1964.

[14] M. L. Mehta, *Random Matrices*. New York: Academic, 1990.

[15] A. Edelman, "Eigenvlaues and Conditions Numbers of Random Matrices," Ph.D. dissertation, Massachusetts Institute of Technology—MIT, Cambridge, MA, May 1989.

[16] E. Abbe, I. E. Telatar, and L. Zheng, "The algebra of MIMO channels," in *Proc. Allerton Annu. Conf. Commun., Contr. Comput.*, Oct. 2005.

[17] K. Zyczkowski and M. Kus, "Random unitary matrices," *J. Phys., A.: Math Gen.*, vol. 27, 1994.

[18] Z. D. Bai, "Methodologies in spectral analysis of large dimensional random matrices, a review," *Statistica Sinica*, vol. 9, pp. 611–677, 1999.

[19] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

Estimation of the Rate–Distortion Function

Matthew T. Harrison, *Member, IEEE*, and Ioannis Kontoyiannis, *Senior Member, IEEE*

***Abstract*—Motivated by questions in lossy data compression and by theoretical considerations, the problem of estimating the rate–distortion function of an unknown (not necessarily discrete-valued) source from empirical data is examined. The focus is the behavior of the so-called "plug-in" estimator, which is simply the rate–distortion function of the empirical distribution of the observed data. Sufficient conditions are given for its consistency, and examples are provided demonstrating that in certain cases it fails to converge to the true rate–distortion function. The analysis of its performance is complicated by the fact that the rate–distortion function is not continuous in the source distribution; the underlying mathematical problem is closely related to the classical problem of establishing the consistency of maximum-likelihood estimators (MLEs). General consistency results are given for the plug-in estimator applied to a broad class of sources, including all stationary and ergodic ones. A more general class of estimation problems is also considered, arising in the context of lossy data compression when the allowed class of coding distributions is restricted; analogous results are developed for the plug-in estimator in that case. Finally, consistency theorems are formulated for modified (e.g., penalized) versions of the plug-in, and for estimating the optimal reproduction distribution.**

***Index Terms*—Consistency, entropy, estimation, maximum-likelihood estimation (MLE), plug-in estimator, rate–distortion function.**

I. INTRODUCTION

Suppose a data string $x_1^n := (x_1, x_2, \dots, x_n)$ is generated by a stationary memoryless source $(X_n : n \geq 1)$ with unknown marginal distribution P on a discrete alphabet A. In many theoretical and practical problems arising in a wide variety of scientific contexts, it is desirable—and often important—to obtain accurate estimates of the entropy $H(P)$ of the source, based on the observed data x_1^n; see, e.g., [5]–[10]. Perhaps the simplest method is via the so-called **plug-in estimator**, where the entropy of P is estimated by $H(P_{x_1^n})$, the entropy of the empirical distribution $P_{x_1^n}$ of x_1^n. The plug-in estimator satisfies the basic statistical requirement of consistency: $H(P_{X_1^n}) \to H(P)$ in probability as $n \to \infty$. In fact, it is strongly consistent; the convergence holds with probability 1 (w.p. 1) [11].

A natural generalization is the problem of estimating the rate–distortion function $R(P, D)$ of a (not necessarily discrete-valued) source. Motivation for this comes in part from lossy data compression, where we may need an estimate of how well a given data set could potentially be compressed, cf. [12], and also from cases where we want to quantify the "information content" of a particular signal, but the data under examination take values in a continuous (or more general) alphabet, cf. [13].

The rate–distortion function estimation question appears to have received little attention in the literature. Here, we present some basic

Manuscript received August 31, 2006; revised April 11, 2008. This work was supported in part by a Sloan Research Fellowship from the Sloan Foundation, the National Science Foundation under Grant 0073378-CCR, and by the USDA-IFAFS under Grant 00-52100-9615. Additionally, the work of M. T. Harrison was supported in part by a National Defense Science and Engineering Graduate Fellowship. The material in this correspondence was presented at the IEEE International Symposium on Information Theory, Seattle, WA, July 2006.

M. T. Harrison is with the Department of Statistics, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: mtharris@cmu.edu).

I. Kontoyiannis is with the Department of Informatics, Athens University of Economics and Business, Athens 10434, Greece (e-mail: yiannis@aueb.gr).

Communicated by M. Effros, Associate Editor for Source Coding.

Digital Object Identifier 10.1109/TIT.2008.926387

0018-9448/$25.00 © 2008 IEEE