



HAL
open science

On the mutual information of 3D massive MIMO systems: An asymptotic approach

Qurrat-Ul-Ain Nadeem, Abla Kammoun, Mérouane Debbah, Mohamed-Slim Alouini

► **To cite this version:**

Qurrat-Ul-Ain Nadeem, Abla Kammoun, Mérouane Debbah, Mohamed-Slim Alouini. On the mutual information of 3D massive MIMO systems: An asymptotic approach. 2015 IEEE International Symposium on Information Theory (ISIT), Jun 2015, hong kong, Hong Kong SAR China. 10.1109/ISIT.2015.7282924 . hal-01242453

HAL Id: hal-01242453

<https://hal.science/hal-01242453>

Submitted on 14 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the Mutual Information of 3D Massive MIMO Systems: An Asymptotic Approach

Qurrat-Ul-Ain Nadeem*, Abla Kammoun*, Mérouane Debbah[†], and Mohamed-Slim Alouini*

* CEMSE Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.

Email: {qurratulain.nadeem, abla.kammoun, slim.alouini}@kaust.edu.sa

[†] Mathematical and Algorithmic Sciences Lab, Huawei France R&D, Paris, France. Email: merouane.debbah@huawei.com

Abstract—Motivated by the recent interest in 3D beamforming to enhance system performance, we present an information-theoretic channel model for multiple-input multiple-output (MIMO) systems, that can support the elevation dimension. The principle of maximum entropy is used to determine the distribution of the channel matrix consistent with the prior angular information. We provide an explicit expression for the cumulative density function (CDF) of the mutual information in the large number of transmit antennas and paths regime. The derived Gaussian approximation is quite accurate even for realistic system dimensions. The simulation results study the achievable performance through the meticulous selection of the transmit antenna downtilt angles. The results are directly applicable to the analysis of 5G 3D massive MIMO systems.

Index Terms—Massive MIMO, channel modeling, maximum entropy, 3D beamforming, mutual information.

I. INTRODUCTION

Channel modeling and mutual information (MI) analysis is a fundamental step towards the performance evaluation of several transmission techniques. Over the years, channel models have evolved to address the challenges of wireless communication technologies. Recently, the trend is to exploit the channel's degrees of freedom in the elevation leading to several 3D beamforming techniques. The existing multiple-input multiple-output (MIMO) systems were designed to support antenna configurations at the base station (BS) capable of adaptation in the azimuth only. Several measurement campaigns have demonstrated the significant impact of elevation on the system performance [1], [2] and the 3GPP is currently defining future mobile communication standards to help evaluate the potential of 3D beamforming [3]. An important feature of these 3D channels is the downtilt angle, that can be optimized to change the vertical beam pattern dynamically and yield the promised gains [4]–[6].

To allow for a better understanding of the limits of wireless systems through the MI analysis, a large effort is required to accurately model the 3D channels. Given the available knowledge on channel parameters like angle of departure (AoD), angle of arrival (AoA), delay, amplitude, finding the best way to attribute a joint distribution to the channel matrix is of vital importance. It has been shown that the choice of

distribution with the greatest entropy creates a model out of available information without making arbitrary assumptions on information not available [7], [8]. In the context of wireless communications, the authors in [9] made an extensive use of the principle of maximum entropy together with the consistency argument to address the question of MIMO channel modeling from an information-theoretic point of view for the 2D case. To the best of authors' knowledge, there has not been any comprehensive study to develop an information-theoretic 3D channel model and characterize its MI distribution.

The aims of this paper are twofold. First is to provide an information-theoretic channel model for 3D MIMO systems, inspired from models presented in standards [10]–[12]. We follow the guidelines provided in [9] to derive an entropy maximizing 3D channel model consistent with the state of available knowledge of AoDs and AoAs, known a priori to the modeler of the channel. The resulting model has a systematic structure, which is exploited to effectively characterize its MI. The second aim is to present an asymptotic analysis of the statistical distribution of the MI and show that it is well approximated by a Gaussian distribution for any number of Rx antennas, as the number of paths and Tx antennas grow large. The simulation results illustrate an excellent fit between the theoretical and simulated CDFs for finite-sized systems and provide a flavor of the performance gains realizable through the meticulous selection of the downtilt angles. The results are applicable to the design of 3D 5G massive MIMO systems.

The rest of the paper is organized as follows. In section II, we introduce the information-theoretic 3D channel model. In Section III, we provide an asymptotic analysis of the MI distribution. The derived results are corroborated using simulations in section IV and finally in section V, we conclude.

II. INFORMATION-THEORETIC CHANNEL MODEL

The mobile communication standards like 3GPP SCM [10], ITU [12] and WINNER [13] follow a system level, stochastic channel modeling approach. The extensions of these 2D models to the 3D case have started to emerge recently [11], [3]. We base the evaluation of our work on these models.

A. 3D MIMO Channel Model

Based on the aforementioned standards and assuming vertically polarized antennas, the effective channel between BS antenna port s and mobile station (MS) antenna port u is given

The work of Q.-U.-A. Nadeem, A. Kammoun and M. -S. Alouini was supported by a CRG4 grant from the Office of Competitive Research Funding (OCRF) at KAUST. The work of Mérouane Debbah was supported by ERC Starting Grant 305123 MORE (Advanced Mathematical Tools for Complex Network Engineering).

$$\mathbf{H}]_{su} = \frac{1}{N} \sum_{n=1}^N \alpha_n \sqrt{g_t(\phi_n, \theta_n, \theta_{tilt})} \exp(ik(s-1)d_t \sin \phi_n \sin \theta_n) \sqrt{g_r(\varphi_n, \vartheta_n)} \exp(ik(u-1)d_r \sin \varphi_n \sin \vartheta_n), \quad (1)$$

by (1), where ϕ_n and θ_n are the azimuth and elevation AoD of the n^{th} path respectively, φ_n and ϑ_n are the azimuth and elevation AoA of the n^{th} path respectively. The elevation angle of the antenna boresight is denoted by θ_{tilt} and α_n is the complex random amplitude of the n^{th} path. Also $\sqrt{g_t(\phi_n, \theta_n, \theta_{tilt})}$ and $\sqrt{g_r(\varphi_n, \vartheta_n)}$ are the global patterns of Tx and Rx antennas respectively. d_t and d_r are the separations between Tx antenna ports and Rx antenna ports respectively and k is the wave number that equals $\frac{2\pi}{\lambda}$, λ being the wavelength. The two exponential terms in (1) are the array responses of the s^{th} Tx and the u^{th} Rx antenna, that can be denoted as $[\mathbf{a}_t(\phi, \theta)]_s$ and $[\mathbf{a}_r(\varphi, \vartheta)]_u$ respectively. Fig. 1 shows the 3D channel model.

To enable an abstraction of the role played by the antennas in performing downtilt, the 3GPP and ITU approximate the global pattern of each port $\sqrt{g_t(\phi_n, \theta_n, \theta_{tilt})}$ by a narrow beam in the elevation as [12],

$$\sqrt{[17\text{dBi} - \min\{-(A_H(\phi) + A_V(\theta, \theta_{tilt})), 20\text{dB}\}]_{\text{lin}}}, \quad (2)$$

$$\text{where, } A_H(\phi) = -\min \left[12 \left(\frac{\phi}{\phi_{3dB}} \right)^2, 20 \right] \text{ dB}, \quad (3)$$

$$A_V(\theta, \theta_{tilt}) = -\min \left[12 \left(\frac{\theta - \theta_{tilt}}{\theta_{3dB}} \right)^2, 20 \right] \text{ dB}, \quad (4)$$

where ϕ_{3dB} is the horizontal 3 dB beamwidth and θ_{3dB} is the vertical 3 dB beamwidth. $g_r(\varphi_n, \vartheta_n)$ is taken to be 0 dB.

B. Maximum Entropy Channel

The theoretical analysis of the ray-tracing 3D channel model just presented is difficult. The multiple propagation paths result in a large number of random variables in the model. Secondly, the model exhibits non-linearity with respect to the AoDs and AoAs. To circumvent these problems encountered in the theoretical analysis, we use the principle of maximum entropy to develop an equivalent information-theoretic model.

It was rigorously proved in [7] that the principle of maximum entropy yields models that express the constraints of our knowledge of model parameters and avoid making any arbitrary assumptions on information not available. The authors in

[9] utilized this framework to devise theoretical grounds for constructing channel models for 2D MIMO systems, consistent with the state of available knowledge of channel parameters.

The geometry-based MIMO channel models presented in standards assume single-bounce scattering between the transmitter and the receiver, i.e., there is a one-to-one mapping between the AoDs and AoAs [14]. Therefore $\phi_n, \theta_n, \varphi_n, \vartheta_n$ are assumed to be known apriori and fixed over channel realizations, during the modeling phase. The only random component in the 3D channel model is now α_n , so a suitable distribution needs to be assigned to it such that the obtained model is consistent with the information available.

It was proved through an extensive analysis in [9] that in the presence of prior information on the number of paths, AoDs, AoAs and Tx and Rx power of the propagation paths, the maximum entropy channel model is given by,

$$\mathbf{H} = \mathbf{\Psi} \mathbf{P}_{Rx}^{\frac{1}{2}} (\mathbf{\Omega} \circ \mathbf{G}) \mathbf{P}_{Tx}^{\frac{1}{2}} \mathbf{\Phi}^H, \quad (5)$$

where \mathbf{P}_{Rx} and \mathbf{P}_{Tx} contain the Rx and Tx powers of the multipaths respectively, $\mathbf{\Omega}$ is the mask matrix that captures path gains, $\mathbf{\Psi}$ and $\mathbf{\Phi}$ capture the antenna array responses and \circ is the Hadamard product. The solution of the consistency argument that maximizes entropy is to take \mathbf{G} to be i.i.d zero mean Gaussian with unit variance [9], [14].

To represent the channel model in (1) with a structure similar to that of (5), we define \mathbf{A} and \mathbf{B} as $N_{BS} \times N$ and $N_{MS} \times N$ deterministic matrices given by,

$$\mathbf{A} = \frac{1}{\sqrt{N}} [\mathbf{a}_t(\phi_1, \theta_1), \mathbf{a}_t(\phi_2, \theta_2), \dots, \mathbf{a}_t(\phi_N, \theta_N)] \circ \begin{bmatrix} \sqrt{g(\phi_1, \theta_1, \theta_{tilt_1})} & \dots & \sqrt{g(\phi_N, \theta_N, \theta_{tilt_1})} \\ \vdots & \ddots & \vdots \\ \sqrt{g(\phi_1, \theta_1, \theta_{tilt_{N_{BS}}})} & \dots & \sqrt{g(\phi_N, \theta_N, \theta_{tilt_{N_{BS}}})} \end{bmatrix}, \quad (6)$$

where $\mathbf{a}_t(\phi_n, \theta_n) = [[\mathbf{a}_t(\phi_n, \theta_n)]_{s=1}, \dots, [\mathbf{a}_t(\phi_n, \theta_n)]_{s=N_{BS}}]^H$.

$$\mathbf{B} = [\mathbf{a}_r(\varphi_1, \vartheta_1), \dots, \mathbf{a}_r(\varphi_N, \vartheta_N)], \quad (7)$$

where $\mathbf{a}_r(\varphi_n, \vartheta_n) = [[\mathbf{a}_r(\varphi_n, \vartheta_n)]_{u=1}, \dots, [\mathbf{a}_r(\varphi_n, \vartheta_n)]_{u=N_{MS}}]^T$.

Note that the array responses and antenna patterns are captured in \mathbf{A} and \mathbf{B} . Since the Tx and Rx powers are incorporated in the antenna patterns so \mathbf{P}_{Rx} and \mathbf{P}_{Tx} in (5) are identity matrices. For the single-bounce scattering model, the mask matrix $\mathbf{\Omega}$ will be diagonal. Taking $\mathbf{\Psi}$ and $\mathbf{\Phi}$ to be \mathbf{B} and \mathbf{A} respectively, the solution of the maximum entropy problem for the $N_{MS} \times N_{BS}$ 3D MIMO channel matrix with entries given by (1) will have a systematic structure as,

$$\mathbf{H} = \frac{1}{\sqrt{N}} \mathbf{B} \text{diag}(\boldsymbol{\alpha}) \mathbf{A}^H, \quad (8)$$

where $\boldsymbol{\alpha}$ is an N dimensional vector with entries that are i.i.d zero mean, unit variance Gaussian RVs.

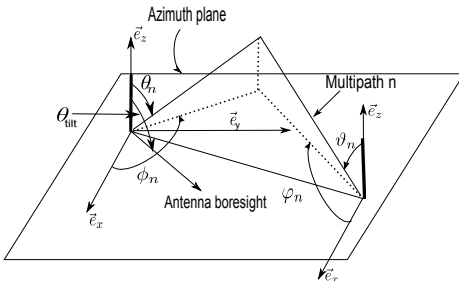


Fig. 1. 3D channel model.

III. ASYMPTOTIC ANALYSIS OF THE MI

We consider a point-to-point communication link where \mathbf{H} is a $N_{MS} \times N_{BS}$ MIMO channel matrix from (8). The channel is linear and time-invariant. A time-division duplex (TDD) protocol is considered. The complex received baseband signal $\mathbf{y} \in \mathbb{C}^{N_{MS} \times 1}$ is given by,

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \quad (9)$$

where $\mathbf{s} \in \mathbb{C}^{N_{BS} \times 1}$ is the Tx signal from the BS and $\mathbf{n} \in \mathbb{C}^{N_{MS} \times 1}$ is the complex additive noise such that $\mathbb{E}[\mathbf{n}\mathbf{n}^H] = \mathbf{R} + \sigma^2\mathbf{I}$, where \mathbf{R} is the covariance matrix of interference experienced and σ^2 is the noise variance at the MS. In this context, the MI of the 3D MIMO system is given by,

$$I(\sigma^2) = \log \det(\mathbf{I}_{N_{MS}} + (\mathbf{R} + \sigma^2\mathbf{I}_{N_{MS}})^{-1}\mathbf{H}\mathbf{H}^H). \quad (10)$$

The reduced degree of randomness in our model makes it difficult to characterize the behaviour of the MI. The asymptotic analysis provided here differs from the previous analysis because of the non-Wishart nature of $\mathbf{H}\mathbf{H}^H$ and the incorporation of the elevation dimension and antenna tilt angles in the channel model. In order to study the asymptotic approximation to the MI distribution, the following assumption is required over the number of BS antennas and multipaths.

Assumption A-1. In the large (N_{BS}, N) regime, N_{BS} and N tend to infinity such that

$$0 < \liminf \frac{N_{BS}}{N} \leq \limsup \frac{N_{BS}}{N} < +\infty, \quad (11)$$

a condition we shall refer to by writing $N_{BS}, N \rightarrow \infty$.

A. Distribution of $\mathbf{H}\mathbf{H}^H$

The analysis starts by characterizing the distribution of $\mathbf{H}\mathbf{H}^H$, which would later be followed by transformations to complete the characterization of MI. Defining $\bar{\mathbf{b}}_i$ as the i^{th} row vector of \mathbf{B} in (7), the $(k, l)^{th}$ entry of $\mathbf{H}\mathbf{H}^H$ is given by,

$$[\mathbf{H}\mathbf{H}^H]_{kl} = \frac{1}{N} \alpha^H (\bar{\mathbf{b}}_k^H \bar{\mathbf{b}}_l) \circ (\mathbf{A}^H \mathbf{A})^T \alpha. \quad (12)$$

Denoting N_{MS} by M , $\mathbf{H}\mathbf{H}^H$ is a $M \times M$ matrix of the following general form,

$$\frac{1}{N} \begin{bmatrix} \alpha^H \mathbf{C}^{1,1} \alpha & \alpha^H \mathbf{C}^{1,2} \alpha & \dots & \alpha^H \mathbf{C}^{1,M} \alpha \\ \alpha^H \mathbf{C}^{2,1} \alpha & \alpha^H \mathbf{C}^{2,2} \alpha & \dots & \alpha^H \mathbf{C}^{2,M} \alpha \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^H \mathbf{C}^{M,1} \alpha & \alpha^H \mathbf{C}^{M,2} \alpha & \dots & \alpha^H \mathbf{C}^{M,M} \alpha \end{bmatrix}, \quad (13)$$

where $\mathbf{C}^{k,l} = [\bar{\mathbf{b}}_k^H \bar{\mathbf{b}}_l] \circ (\mathbf{A}^H \mathbf{A})^T, \forall k, l = 1, \dots, M$. To characterize the distribution, an additional assumption is required over the matrices that form the quadratic terms in $\mathbf{H}\mathbf{H}^H$.

Assumption A-2. Under the setting of Assumption A-1,

$$\frac{\|[\bar{\mathbf{b}}_k^H \bar{\mathbf{b}}_l] \circ (\mathbf{A}^H \mathbf{A})^T\|_{sp}}{\|[\bar{\mathbf{b}}_k^H \bar{\mathbf{b}}_l] \circ (\mathbf{A}^H \mathbf{A})^T\|_F} \rightarrow 0, \quad \forall k, l = 1, \dots, M, \quad (14)$$

where $\|\cdot\|_{sp}$ denotes the spectral norm and $\|\cdot\|_F$ denotes the Forbenius norm of a matrix. This is a technical assumption required to observe the convergence behavior.

Denoting the $(k, l)^{th}$ quadratic term by $T^{k,l} = \frac{1}{N} \alpha^H \mathbf{C}^{k,l} \alpha$, every quadratic term can be decomposed into its real and imaginary parts, which are also quadratic forms in α as,

$$T^{k,l} = \frac{1}{N} \Re(\alpha^H \mathbf{C}^{k,l} \alpha) + \frac{j}{N} \Im(\alpha^H \mathbf{C}^{k,l} \alpha), \quad (15)$$

where $\Re(\alpha^H \mathbf{C}^{k,l} \alpha)$ and $\Im(\alpha^H \mathbf{C}^{k,l} \alpha)$ are given by (16) and (17) respectively. Equivalently $\mathbf{H}\mathbf{H}^H$ can be written as a sum of two matrices with the real and imaginary parts.

With this decomposition at hand, we now state in the following theorem the asymptotic behavior of the joint distribution of the entries of $\Re(\mathbf{H}\mathbf{H}^H)$ and $\Im(\mathbf{H}\mathbf{H}^H)$ stacked in vector \mathbf{x} ,

$$\mathbf{x} = [\text{vec}(\Re(\mathbf{H}\mathbf{H}^H))^T \text{vec}(\Im(\mathbf{H}\mathbf{H}^H))^T]^T, \quad (18)$$

where the operator $\text{vec}(\cdot)$ maps an $M \times M$ matrix to an $M^2 \times 1$ vector by stacking the rows of the matrix.

Theorem 1: Let \mathbf{H} be a $N_{MS} \times N_{BS}$ MIMO channel matrix from (8), then under **A-1** and **A-2**, \mathbf{x} behaves as multivariate Gaussian such that the MGF of \mathbf{x} converges as,

$$\mathbb{E} \left[\exp \left(\frac{\sqrt{N} \mathbf{s}^T (\mathbf{x} - \mathbf{m})}{\sqrt{\mathbf{s}^T \Theta \mathbf{s}}} \right) \right] - \exp \left(\frac{1}{2} \right) \rightarrow 0, \quad (19)$$

where $\mathbf{m} = \frac{1}{2N} \times$

$$[Tr(\mathbf{C}_{\Re}^{1,1}) \dots Tr(\mathbf{C}_{\Re}^{M,M}), Tr(\mathbf{C}_{\Im}^{1,1}) \dots Tr(\mathbf{C}_{\Im}^{M,M})]^T \quad (20)$$

and Θ is given by (21).

The proof follows from the fact that the mean and covariance of the quadratic terms can be computed using (3.20) in [15]. The Central Limit Theorem for the $(k, l)^{th}$ entry of $\Re(\mathbf{H}\mathbf{H}^H)$ and $\Im(\mathbf{H}\mathbf{H}^H)$ can be established from the result in [16] under **A-1** and **A-2**, implying that the MGF behaves as,

$$\mathbb{E} \left[\exp \left(\frac{s \sqrt{N} [\Re(T^{k,l}) - \frac{1}{2N} Tr(\mathbf{C}_{\Re}^{k,l})]}{\sqrt{\frac{[Tr(\mathbf{C}_{\Re}^{k,l} \mathbf{C}_{\Re}^{k,l}) + Tr(\mathbf{C}_{\Re}^{k,l} (\mathbf{C}_{\Re}^{k,l})^T)]}{4N}}} \right) \right] - \exp \left(\frac{s^2}{2} \right) \rightarrow 0.$$

To characterize the joint distribution of $\Re(\mathbf{H}\mathbf{H}^H)$, $\Im(\mathbf{H}\mathbf{H}^H)$, the behaviour of their linear combination $\mathbf{g}^T [\mathbf{x} - \mathbf{m}]$, where \mathbf{g} is any arbitrary vector, is studied. Theorem 1 follows from it.

B. Distribution of Mutual Information

The analysis starts by expressing (10) as,

$$I(\sigma^2) = \log \det((\mathbf{R} + \sigma^2 \mathbf{I}_M) + \mathbf{H}\mathbf{H}^H) - \log \det(\mathbf{R} + \sigma^2 \mathbf{I}_M). \quad (22)$$

Decomposing $((\mathbf{R} + \sigma^2 \mathbf{I}_M) + \mathbf{H}\mathbf{H}^H)$ into real and imaginary parts as $\mathbf{Y} + j \mathbf{Z}$, and denoting the entries of $\mathbf{R} + \sigma^2 \mathbf{I}_M$ by ζ_{ij} , we can extend the result of Theorem 1 to $\Re((\mathbf{R} + \sigma^2 \mathbf{I}_M) + \mathbf{H}\mathbf{H}^H)$ and $\Im((\mathbf{R} + \sigma^2 \mathbf{I}_M) + \mathbf{H}\mathbf{H}^H)$, which will also show the convergence in (19) under assumptions **A-1** and **A-2** with mean vector \mathbf{m} given by (23) and the same covariance. The mean matrix \mathbf{M} for $((\mathbf{R} + \sigma^2 \mathbf{I}_M) + \mathbf{H}\mathbf{H}^H)$ is expressed as,

$$\mathbf{M} = \mathbf{M}_1 + j \mathbf{M}_2, \quad (24)$$

where \mathbf{M}_1 and \mathbf{M}_2 are $M \times M$ matrices, containing the real and imaginary parts of the means of $((\mathbf{R} + \sigma^2 \mathbf{I}_M) + \mathbf{H}\mathbf{H}^H)$.

$$\Re(\boldsymbol{\alpha}^H \mathbf{C}^{k,l} \boldsymbol{\alpha}) = [\Re(\boldsymbol{\alpha})^T \quad \Im(\boldsymbol{\alpha})^T] \begin{bmatrix} \Re(\mathbf{C}^{k,l}) & -\Im(\mathbf{C}^{k,l}) \\ \Im(\mathbf{C}^{k,l}) & \Re(\mathbf{C}^{k,l}) \end{bmatrix} \begin{bmatrix} \Re(\boldsymbol{\alpha}) \\ \Im(\boldsymbol{\alpha}) \end{bmatrix} = [\Re(\boldsymbol{\alpha})^T \quad \Im(\boldsymbol{\alpha})^T] (\mathbf{C}_{\Re}^{k,l}) \begin{bmatrix} \Re(\boldsymbol{\alpha}) \\ \Im(\boldsymbol{\alpha}) \end{bmatrix}. \quad (16)$$

$$\Im(\boldsymbol{\alpha}^H \mathbf{C}^{k,l} \boldsymbol{\alpha}) = [\Re(\boldsymbol{\alpha})^T \quad \Im(\boldsymbol{\alpha})^T] \begin{bmatrix} \Im(\mathbf{C}^{k,l}) & \Re(\mathbf{C}^{k,l}) \\ -\Re(\mathbf{C}^{k,l}) & \Im(\mathbf{C}^{k,l}) \end{bmatrix} \begin{bmatrix} \Re(\boldsymbol{\alpha}) \\ \Im(\boldsymbol{\alpha}) \end{bmatrix} = [\Re(\boldsymbol{\alpha})^T \quad \Im(\boldsymbol{\alpha})^T] (\mathbf{C}_{\Im}^{k,l}) \begin{bmatrix} \Re(\boldsymbol{\alpha}) \\ \Im(\boldsymbol{\alpha}) \end{bmatrix}. \quad (17)$$

$$\Theta = \begin{bmatrix} \frac{1}{4N} [Tr(\mathbf{C}_{\Re}^{1,1}(\mathbf{C}_{\Re}^{1,1})) + Tr(\mathbf{C}_{\Re}^{1,1}(\mathbf{C}_{\Re}^{1,1})^T)] & \dots & \frac{1}{4N} [Tr(\mathbf{C}_{\Re}^{1,1}(\mathbf{C}_{\Im}^{M,M})) + Tr(\mathbf{C}_{\Re}^{1,1}(\mathbf{C}_{\Im}^{M,M})^T)] \\ \vdots & \ddots & \vdots \\ \frac{1}{4N} [Tr(\mathbf{C}_{\Re}^{M,M}(\mathbf{C}_{\Re}^{1,1})) + Tr(\mathbf{C}_{\Re}^{M,M}(\mathbf{C}_{\Re}^{1,1})^T)] & \dots & \frac{1}{4N} [Tr(\mathbf{C}_{\Re}^{M,M}(\mathbf{C}_{\Im}^{M,M})) + Tr(\mathbf{C}_{\Re}^{M,M}(\mathbf{C}_{\Im}^{M,M})^T)] \\ \frac{1}{4N} [Tr(\mathbf{C}_{\Im}^{1,1}(\mathbf{C}_{\Re}^{1,1})) + Tr(\mathbf{C}_{\Im}^{1,1}(\mathbf{C}_{\Re}^{1,1})^T)] & \dots & \frac{1}{4N} [Tr(\mathbf{C}_{\Im}^{1,1}(\mathbf{C}_{\Im}^{M,M})) + Tr(\mathbf{C}_{\Im}^{1,1}(\mathbf{C}_{\Im}^{M,M})^T)] \\ \vdots & \ddots & \vdots \\ \frac{1}{4N} [Tr(\mathbf{C}_{\Im}^{M,M}(\mathbf{C}_{\Re}^{1,1})) + Tr(\mathbf{C}_{\Im}^{M,M}(\mathbf{C}_{\Re}^{1,1})^T)] & \dots & \frac{1}{4N} [Tr(\mathbf{C}_{\Im}^{M,M}(\mathbf{C}_{\Im}^{M,M})) + Tr(\mathbf{C}_{\Im}^{M,M}(\mathbf{C}_{\Im}^{M,M})^T)] \end{bmatrix}. \quad (21)$$

$$\mathbf{m} = [\Re(\zeta_{11}) + \frac{Tr(\mathbf{C}_{\Re}^{1,1})}{2N}, \dots, \Re(\zeta_{MM}) + \frac{Tr(\mathbf{C}_{\Re}^{M,M})}{2N}, \Im(\zeta_{11}) + \frac{Tr(\mathbf{C}_{\Im}^{1,1})}{2N}, \dots, \Im(\zeta_{MM}) + \frac{Tr(\mathbf{C}_{\Im}^{M,M})}{2N}]^T \quad (23)$$

Before presenting the Gaussian approximation to the statistical distribution of the MI in the asymptotic limit, we define a matrix $\tilde{\mathbf{M}}$ and a $2M^2 \times 1$ vector $\mathbf{f}(\tilde{\mathbf{M}})$ as,

$$\tilde{\mathbf{M}} = \begin{bmatrix} \mathbf{M}_1 & -\mathbf{M}_2 \\ \mathbf{M}_2 & \mathbf{M}_1 \end{bmatrix}, \quad (25)$$

$\mathbf{f}(\tilde{\mathbf{M}})|_{l+(k-1)M} = [\det(\mathbf{D}_1) + \dots + \det(\mathbf{D}_{2M})]|_{l+(k-1)M}$, $k = 1, \dots, 2M, l = 1, \dots, M$, where \mathbf{D}_i is identical to $\tilde{\mathbf{M}}$ except that the entries in the i^{th} row are replaced by their derivatives with respect to $(k, l)^{th}$ entry of $\begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \end{bmatrix}$. Every entry of the $2M^2 \times 1$ vector, $\mathbf{f}(\tilde{\mathbf{M}})$ will involve the sum of only two non-zero determinants as every $(k, l)^{th}$ entry occurs in only two rows. Additionally we make the following assumption, **Assumption A-3.** In the large N_{BS}, N regime under **A-1**,

$$\liminf \mathbf{f}(\tilde{\mathbf{M}})^T \Theta \mathbf{f}(\tilde{\mathbf{M}}) > 0.$$

We now present the Gaussian approximation to the statistical distribution of the MI in the asymptotic limit.

Theorem 2: Let \mathbf{H} be a $N_{MS} \times N_{BS}$ MIMO channel matrix from (8), such that assumptions **A-1**, **A-2** and **A-3** hold, then the statistical distribution of $I(\sigma^2)$ can be approximated as,

$$\mathbb{P}[\sqrt{N}I(\sigma^2) \leq x] - \frac{1}{2} \left(1 + \text{erf} \left(\frac{x - \sqrt{N}\mu}{\sqrt{2\sigma_a^2}} \right) \right) \rightarrow 0, \quad (26)$$

where erf is the error function, $\mu = 0.5 \log \det \tilde{\mathbf{M}} - \log \det(\mathbf{R} + \sigma^2 \mathbf{I}_M)$ and σ_a^2 is given by,

$$\left(\frac{.5}{\det \tilde{\mathbf{M}}} \right)^2 \times \mathbf{f}(\tilde{\mathbf{M}})^T \Theta \mathbf{f}(\tilde{\mathbf{M}}). \quad (27)$$

The proof follows from succumbing to the Taylor series expansion of a real-valued function $f(\mathbf{x}_n)$,

$$f(\mathbf{x}_n) = f(\mathbf{m}) + \nabla^T f(\mathbf{m})(\mathbf{x}_n - \mathbf{m}) + R_n, \quad (28)$$

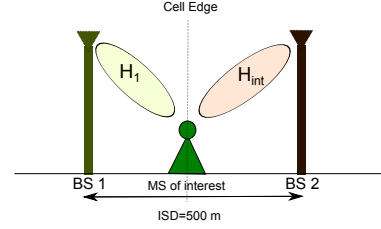


Fig. 2. Example scenario.

where \mathbf{m} is the mean vector in (23) and $R_n = o(1)$.

$$\mathbb{E}[\exp(f(\mathbf{x}_n) - f(\mathbf{m}))] = \mathbb{E} \left[\exp \left(\nabla^T f(\mathbf{m})(\mathbf{x}_n - \mathbf{m}) \right) \right] + o(1).$$

Note that $\mathbb{E}[\exp(\nabla^T f(\mathbf{m})(\mathbf{x}_n - \mathbf{m}))]$ is analogous to $\mathbb{E}[\exp(\mathbf{s}^T[\mathbf{x}_n - \mathbf{m}])]$ with $\mathbf{s} = \nabla f(\mathbf{m})$. Now using the convergence result from last section and invoking Slutsky's Theorem, under **A-3** results in,

$$\mathbb{E} \left[\exp \left(\frac{\sqrt{N}[f(\mathbf{x}_n) - f(\mathbf{m})]}{\sqrt{\nabla^T f(\mathbf{m}) \Theta \nabla f(\mathbf{m})}} \right) \right] - \exp \left(\frac{1}{2} \right) \rightarrow 0.$$

The expression for ∇f is worked out using $(\det(\mathbf{M}_1 + j\mathbf{M}_2))^2 = \det \tilde{\mathbf{M}}$ and the expression for derivative of the determinant from [Ch 6, [17]]. Theorem 2 follows from it.

IV. NUMERICAL RESULTS

We corroborate the results derived in this work with simulations. A simple but realistic multi-cell scenario shown in Fig. 2 is considered. The worst case performance when the MS is at the cell edge is studied. \mathbf{R} is computed as,

$$\mathbf{R} = \sum_{i=1}^{N_{int}} \mathbb{E}[\mathbf{H}_i \mathbf{H}_i^H], \quad \text{where} \quad (29)$$

$$\mathbb{E}[\mathbf{H}_i \mathbf{H}_i^H] = \frac{1}{N} \mathbf{B}_i \text{diag}(\boldsymbol{\alpha}) \mathbf{A}_i^H \mathbf{A}_i \text{diag}(\boldsymbol{\alpha}^*) \mathbf{B}_i^H, \quad i = 1, \dots, N_{int},$$

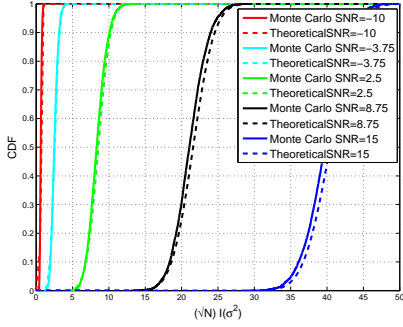


Fig. 3. Comparison of Monte Carlo simulated CDF and asymptotic theoretical CDF in (26).

where N_{int} is the number of interfering BSs (1 for Fig. 2).

In the standards, (θ_n, ϑ_n) are drawn from $\text{Laplace}(\theta_0, \sigma)$ density, while (ϕ_n, φ_n) are drawn from Wrapped Gaussian distribution, which can be approximated as Von Mises (μ, κ) .

The radio and channel parameters are set as $\theta_{3dB} = 15^\circ$, $\phi_{3dB} = 70^\circ$, $\sigma_{BS} = 7^\circ$, $\sigma_{MS} = 10^\circ$, $\kappa_{BS}, \kappa_{MS} = 5$ and $\mu = 0$. Moreover θ_0 is set equal to the elevation LoS angle between the BS and MS. Two thousand Monte Carlo realizations of the 3D channel in (1) are generated to obtain the simulated MI for comparison. SNR denotes $\frac{1}{\sigma^2}$ in dB. We now validate the result obtained for the CDF of MI for finite-sized systems. The first result deals with the case when BS 2 is operating in a different frequency band as BS 1. Fig. 3 illustrates a good fit between the asymptotic theoretical and simulated result for a finite-sized system with $N, N_{BS} = 60$ and $N_{MS} = 4$. Even at the highest value of SNR , the asymptotic mean and variance show only a 1.2 and 3.2 percent relative error respectively.

Finally the multi-cell scenario is considered in Fig. 4, where we plot the CDF of the MI at different values of antenna boresight angles of the serving BS. We again deal with the worst case scenario of the MS being in the direct boresight of the interfering cell. Comparison with the simulated CDF of MI obtained at $SNR = 5$ dB validates the accuracy of the derived theoretical asymptotic distribution and illustrates the impact 3D elevation beamforming can have on the system performance through the meticulous selection of downtilt angles. For the case on hand, the MI of the system is maximized when the serving BS also sets its antenna boresight angles equal to the elevation LoS angle with the MS, i.e. $\theta_{vilt} \approx 96^\circ$.

V. CONCLUSION

The prospect of enhancing system performance through elevation beamforming has stirred a growing interest among researchers in wireless communications. In this work, we used the principle of maximum entropy to determine the distribution of the MIMO channel matrix based on the prior angular information. An asymptotic analysis, in the number of paths and BS antennas, of the MI is provided and validated for finite-sized systems via simulations. We believe that the results presented will enable a fair evaluation of the 3D massive MIMO channels being outlined in the future generation of mobile communication standards.

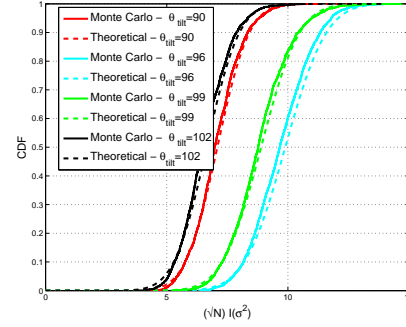


Fig. 4. Comparison of Monte Carlo simulated CDF and asymptotic theoretical CDF in a multi-cell environment.

REFERENCES

- [1] A. Kuchar, J.-P. Rossi, and E. Bonek, "Directional macro-cell channel characterization from urban measurements," *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 2, pp. 137–146, Feb. 2000.
- [2] J. Koppenborg, H. Halbauer, S. Saur, and C. Hoek, "3D beamforming trials with an active antenna array," in *Proc. ITG Workshop Smart Antennas*, pp. 110–114, 2012.
- [3] 3GPP TR 36.873 V12.0.0, "Study on 3D channel model for LTE," Sep. 2014.
- [4] F. Athley and M. Johansson, "Impact of electrical and mechanical antenna tilt on LTE downlink system performance," in *IEEE 71st Vehicular Technology Conference (VTC 2010-Spring)*, May 2010, pp. 1–5.
- [5] B. Partov, D. Leith, and R. Razavi, "Utility fair optimization of antenna tilt angles in LTE networks," *IEEE/ACM Transactions on Networking*, vol. PP, no. 99, pp. 1–1, Jan 2014.
- [6] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing - the large-scale array regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, Oct 2013.
- [7] J. Shore and R. Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Transactions on Information Theory*, vol. 26, no. 1, pp. 26–37, Jan 1980.
- [8] E. T. Jaynes, "Information theory and statistical mechanics, part 1," *Phys. Rev.*, vol. 106, pp. 620–630, 1957.
- [9] M. Debbah and R. R. Müller, "MIMO channel modeling and the principle of maximum entropy," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1667–1690, May 2005.
- [10] "Spatial channel model for multiple input multiple output (MIMO) simulations," [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/25996.htm>, Sep. 2003.
- [11] J. Meinilä, P. Kyösti *et al.*, "D5.3: WINNER+ final channel models V1.0," [Online]. Available: http://projects.celtic-initiative.org/winner+/WINNER+%20Deliverables/D5.3_v1.0.pdf, June. 2010.
- [12] Report ITU-R M.2135, "Guidelines for evaluation of radio interface technologies for IMT-advanced," [Online]. Available: <http://www.itu.int/pub/R-REP-M.2135-2008/en>, 2008.
- [13] IST-4-027756 WINNER II, "D1.1.2, WINNER II channel models," [Online]. Available: <https://www.ist-winner.org/WINNER2-Deliverables/D1.1.2v1.1.pdf>, Sept. 2007.
- [14] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli Esposti, H. Hofstetter, P. Kyösti, D. Laurenson, G. Matz, A. F. Molisch, C. Oestges, and H. Özcelik, "Survey of channel and radio propagation models for wireless MIMO systems," *EURASIP Journal on Wireless Communications and Networking Volume 2007*, url = <http://www.eurecom.fr/publication/2162>, 02 2007.
- [15] W. Hachem, M. Kharouf, J. Najim, and J. Silverstein, "A CLT for information-theoretic statistics of non-centered gram random matrices," *Random Matrices: Theory and Applications*, 01 (02), April 2012.
- [16] R. Bhansali, L. Giraitis, and P. Kokoszka, "Convergence of quadratic forms with nonvanishing diagonal," *Statistics & Probability Letters*, vol. 77, no. 7, pp. 726–734, April 2007. [Online]. Available: <http://ideas.repec.org/a/eee/stapro/v77y2007i7p726-734.html>
- [17] C. D. Meyer, *Matrix analysis and applied linear algebra*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2000.