On the Asymptotic Capacity of MIMO Systems with Antenna Arrays of Fixed Length

Shuangqing Wei, Dennis Goeckel and Ramakrishna Janaswamy

Department of Electrical and Computer Engineering,

University of Massachusetts, Amherst, MA, 01003-9292

e-mail: {swei,goeckel,janaswamy}@ecs.umass.edu

Abstract

Previous authors have shown that the asymptotic capacity of a multiple element antenna (MEA) system with N transmit and N receive antennas (termed an (N, N) MEA) grows linearly with N if, for all l, the correlation of the fading for two antenna elements whose indices differ by l remains fixed as antennas are added to the array. However, in practice, the total size of the array is often fixed, and thus the correlation of the fading for two elements separated in index by some value l will change as the number of antenna elements is increased. In this paper, under the condition that the size of an array of antennas is fixed, and assuming that the transmitter does not have access to the channel state information (CSI) while the receiver has perfect CSI, the asymptotic properties of the instantaneous mutual information $I_{N,N}$ of an (N, N) MEA wireless system in a quasi-static fading channel are derived analytically and tested for accuracy for finite N through simulations. For many channel correlation structures, it is demonstrated that the asymptotic performance converges almost surely, implying that such MEA systems have a certain strong robustness to the instantiation of the channel fading values.

[1]This paper is based upon work supported in part by the National Science Foundation under Grants NCR-9714597, ECS-9979282, EIA-0080119, and CAREER Award No. CCR-9875482.

[2]This paper has appeared in part at the Conference on Information Sciences and Systems, March 2002, and at the International Conference on Communications, May 2003.

1 Introduction

Multiple element antenna (MEA) wireless systems have demonstrated the theoretical [1, 2] and practical [3] potential to increase system bandwidth efficiencies well beyond those previously imagined. In this paper, the instantaneous mutual information between the transmitter and receiver in an (N_T, N_R) MEA system, with N_T transmit antennas and N_R receive antennas, is considered. Early work in this area, which motivated much of the MEA work to follow, assumed that the fading between different element pairs was independently and identically distributed (i.i.d). Under this assumption, it has been shown [1, 2] that, even if the transmitter has no knowledge of the channel fading values, the capacity divided by $\hat{N} = \min(N_T, N_R)$ approaches a non-zero constant for a fixed average transmit power, as $\hat{N} \to \infty$.

The aforementioned assumption of an i.i.d distribution of channel path gains can often be violated due to the insufficient spacing of antennas and/or the absence of a rich scattering environment around the transmitter and/or receiver. For example, for a given angular spreading of the incoming waves, the spatial correlation of the signals received at two points will generally increase with decreasing distance between the points [4]. Electromagnetic mutual coupling between the elements will also change the correlation between the signals received from adjacent points. Recent work investigating the impact of correlated fading on the capacity of MEA systems can be found in [5] and [6]. In [6], a (N, N) MEA wireless system is assumed, and antennas are arranged in a regular grid, the total size of which scales upward with the number of antennas, thus preserving the relative position of adjacent antennas. Under this assumption, [6] employed random matrix theory to show that, as N approaches infinity, the instantaneous mutual information $I_{N,N}$ of such MEA systems still increases linearly, albeit with a smaller rate than in the i.i.d fading case.

In practice, the maximum physical size of the antenna array is fixed due to physical constraints imposed by the application (e.g. on a mobile unit). In this work, the asymptotic characteristics of $I_{N,N}$ are investigated in a scenario where the length of a linear array is fixed, and no channel state information (CSI) is available to the transmitter while the receiver has the perfect CSI. Two different sets of assumptions for the signal-to-noise ratio will be considered. In one case, it is assumed that the total average transmit power from the N_T antennas will be fixed, which implies that the average receiver power grows linearly with N_R. This set of assumptions would apply, for example, if a systems engineer were considering adding small dipoles to a sparse linear array at the receiver, in which case the total effective area of the receive array would scale with N_R. In contrast, [7] independently has considered asymptotic *mean* capacity for antenna arrays of fixed size has been considered assuming that the total average *received* power is fixed, thereby implying that the total effective area of the receive antennas does not grow with N_R. In this case, the results will indicate the gain as RF chains are added at the receiver.

Traditional information theory for fading channels has often been devoted to finding the mean capacity for a given channel, as is investigated in [7]. However, recent work in information theory for fading channels, particularly for multiple antenna element systems, has been devoted to outage capacity (e.g. [1]); that is, it is often of interest to know how often the capacity of a system will be above some required rate. By demonstrating almost sure convergence of the asymptotic capacity of fixed-size linear arrays, the mean and outage capacities are both addressed. In particular, almost sure convergence to a limiting expression implies: (1) the mean capacity converges to the same expression, and (2) the system will almost always be in outage for desired rates above the derived capacity and almost never be in outage for rates below the derived capacity.

To understand the practical significance of almost sure convergence of the mutual information of (N, N) MIMO systems, contrast such results with those for asymptotic mean capacity (e.g. [7]). For a given type of propagation environment characterized by the correlation properties of the channel fading matrix, the mean capacity characterizes how well a system does on average but says little about how the system would operate in any single environment, thus greatly complicating wireless network planning. The demonstration of almost sure convergence for the same system for a given channel correlation structure implies that the system will operate reliably - regardless (almost surely) of the actual instantiation of the channel fading values that are encountered. Figure 1 shows a representative example of the results derived here. For the case with a fixed size array at the receiver end only (e.g. a base-to-mobile communication), the upper set of curves demonstrates that the mutual information under each of the instantiations of the fading converges rapidly to the mean capacity, thus demonstrating the robustness of the performance of the MEA system to the specific instantiation of the fading values. The lower set of curves of Figure 1 shows results when there are fixed size arrays at both the transmitter and receiver (e.g. a mobile-to-mobile communication), where mean convergence occurs but almost sure convergence does not occur.

The main results of this work are shown in Table 1. In Section 2, the system model is presented, which follows that in [6]. Note that the approaches taken in [6] cannot be employed here. Instead, the asymptotic characteristics of eigenvalues of Hermitian matrices and the statistical characteristics of eigenvalues of large sample covariance matrices are investigated in Section 3.1 and Section 3.2, respectively. Section 4 applies the results to the analysis of the mutual information of MEA systems. Simulation results are presented in Section 5, and conclusions are drawn in Section 6.

2 System Model

Throughout the paper, the following notations will be used: I_N for the $N \times N$ identity matrix, $A^\dagger$ for transpose conjugate of the matrix A, A^* for conjugate of the matrix A, $\det(A)$ for determinant of the square matrix A, A' for transpose of the matrix A, and $\underline{X}$ for column vector.

A single-user, point-to-point, narrowband wireless communication system with N transmit antennas and N receive antennas is assumed. The case where the number of transmit antennas and the number of receive antennas differ can be considered in an analogous manner. Let H be the $N \times N$ channel fading matrix, whose (i, j)th entry $H_{i,j}$ is the complex path gain between transmitter j and receiver i. Then, the discrete-time equivalent system model is given by:

$$\underline{Y} = H\underline{X} + \underline{Z} \tag{1}$$

where $\underline{X}$ is an $N \times 1$ vector whose jth component represents the signal transmitted by the jth antenna. Similarly, the received signal and received noise are represented by $N \times 1$ complex vectors, $\underline{Y}$ and $\underline{Z}$, respectively. The noise vector $\underline{Z}$ is an additive white Gaussian random vector, whose entries $\{Z_i, i = 1, \ldots, N\}$ are i.i.d circularly symmetric complex Gaussian random variables with mean zero, where Z_i is the additive noise in the ith receiver. Let σ_0^2 be the variance of Z_i, which will be normalized to one. Thus, $Z_i \sim \tilde{N}(0, 1)$, where $\sim \tilde{N}(\mu, \sigma^2)$ indicates a random variable possesses a circularly symmetric complex Gaussian distribution with mean μ and variance σ^2.

As noted in the Section 1, two different assumptions regarding the total average transmit and receiver power will be considered in this work: (1) the total average receive power is fixed, and (2) the total average transmit power is fixed. For the reminder of this section, quantities will be written only for the second case to make the exposition smoother, but they are easily modified for the first case as shown in Section 4.1. Let the total average power transmitted across the N transmit antennas be $E\left[\sum_{k=1}^{N}|X_k|^2\right]=\rho$, regardless of N. Entries of the channel fading matrix H are assumed to be circularly symmetric complex Gaussian random variables with zero mean and $E[|H_{i,j}|^2]=1$, and thus a Rayleigh fading channel is being assumed. Therefore, the average signal-to-noise ratio (SNR) at a single receive antenna is ρ for the second case as stated above. In this work, H will be treated as quasi-static, which means entries of H are constant during a data frame and vary from frame to frame. It is assumed that the *transmitter* has neither knowledge of the entries of H nor knowledge of the correlation statistics of the entries, but that the *receiver* has perfect knowledge of the entries of H (i.e. no transmitter channel state information (CSI) is assumed, but perfect receiver CSI is assumed). Hence, as in [2], if the input vector $\underline{X}$ is a proper complex Gaussian random vector, whose covariance matrix is $E[\underline{X}\cdot\underline{X}^\dagger]=Q$, the mutual information $I_{N,N}$ of this MEA system (conditioned on H) is

$$I_{N,N}=\log_2\det\left(I_N+H\cdot Q\cdot H^\dagger\right)\text{ bps/Hz} \tag{2}$$

Since there is no CSI nor knowledge of the correlation of the entries of H available at the transmitter, a reasonable Q is $\frac{\rho}{N}I_N$ [2], which implies transmitting data independently with the same average power ρ/N across each of the N antennas. Then, (2) simplifies to

$$I_{N,N}=\log_2\det\left(I_N+\frac{\rho}{N}H\cdot H^\dagger\right)\text{ bps/Hz.} \tag{3}$$

It is assumed that the covariance matrix of the random variables $H_{i,j}$ has the following general covariance structure, as described in [6]:

$$E[H_{i,k}H_{j,l}^*]=\Psi_{k,l}^T\Psi_{i,j}^R, \tag{4}$$

where Ψ^T and Ψ^R are $N\times N$ covariance matrices generated by the transmit and receive antennas, respectively. In [6], it was assumed that as N was increased, the relative position of adjacent antennas is fixed for some regular arrays, such as square or linear grids, implying that the total size of the array grows with N. In contrast to [6], assume that the total length of the linear array at the receiver (mobile unit) side is fixed. The length of the linear array at the transmitter side (base station) will be assumed to be either: (1) fixed, or (2) large enough to make $\Psi^T = I_N$.

As in [6], matrix H can be factorized in the form $H \stackrel{\mathcal{D}}{=} \left(\Psi^R\right)^{\frac{1}{2}} W \left(\Psi^T\right)'^{\frac{1}{2}}$, where the entries of W are i.i.d with $\tilde{N}(0,1)$, and $x \stackrel{\mathcal{D}}{=} y$ means random variables x and y have the same distribution. In order to analyze the asymptotic performance of (3), as $N \to \infty$, the unitary transformation of matrices yields

$$I_{N,N} \stackrel{\mathcal{D}}{=} \log_2 \det \left(I_N + \frac{\rho}{N} \Psi^R W \left(\Psi^T\right)' W^\dagger \right) \stackrel{\mathcal{D}}{=} \log_2 \det \left(I_N + \frac{\rho}{N} D_N^R W D_N^T W^\dagger \right), \tag{5}$$

where D_N^R and D_N^T are diagonal matrices, whose diagonal entries are the eigenvalues of Ψ^R and Ψ^T, respectively, in descending order of their magnitudes: i.e., $D_N^R(1,1) \geq \cdots \geq D_N^R(N,N)$ and $D_N^R(1,1) \geq \cdots \geq D_N^T(N,N)$.

3 Asymptotic Analysis of Eigenvalues of Large Matrices

From (5), it is clear that the eigenvalues of the random matrix $\frac{1}{N} D_N^R W D_N^T W^\dagger$ determine the characteristics of $I_{N,N}$. First, in Section 3.1, the asymptotic behavior of the deterministic covariance matrix Ψ^R is considered. Next, in Section 3.2, the asymptotic behavior of eigenvalues of the random matrix $\frac{1}{N}\left(D_N^R\right)^{1/2} W W^\dagger \left(D_N^R\right)^{1/2}$ is studied.

3.1 Characteristics of Eigenvalues of Large Covariance Matrices

In this section, interest is in the number of nonzero eigenvalues of Ψ^R and the rate at which the eigenvalues converge to their limiting values. Without loss of generality, let $\tilde{\psi}^R(r)$ be the normalized ($\tilde{\psi}^R(0) = 1$) spatial correlation function at the receiver end for a linear array of fixed length, such that

$$\Psi_{i,j}^R = \tilde{\psi}^R \left(\frac{i-j}{N-1} L_R \right),\ i, j = 1, \ldots, N, \tag{6}$$

where L_R is the total length of the linear array. Therefore, Ψ^R is a non-negative definite Hermitian and Toeplitz matrix. As noted in [7], the eigenvalues of the matrix Ψ^R/N will be converging to the point spectrum (i.e. eigenvalues in this case) $\left\{\lambda_k^{(R,\infty)}\right\}$ of the non-negative definite Hermitian operator $\psi^R(x,y) = \tilde{\psi}^R\left[(x-y)L_R\right]$ on the Hilbert space $L_2[0,1]$ [15], where $x, y \in [0,1]$. The operator $\psi^R(x,y)$ is completely continuous and square summable over $[0,1]\times[0,1]$ [15],

$$\int_0^1 \int_0^1 \left|\psi^R(x,y)\right|^2 \, dx \, dy < \infty. \tag{7}$$

Eigenvalues $\{\lambda_k^{(R,\infty)}\}$ of $\psi^R(x,y)$ can be determined by

$$\int_0^1 \tilde{\psi}^R\left[(x-y)L_R\right]\phi_k^{(R)}(y)dy = \lambda_k^{(R,\infty)}\phi_k^{(R)}(x),\ k = 0, 1, \cdots, \infty, \tag{8}$$

where $x \in [0,1]$, and $\left\{\phi_k^{(R)}(x)\right\}$ are the eigen-functions of the operator $\psi^R(x,y)$. From [15, pp. 365], zero is the only limit point of the spectrum of $\psi^R(x,y)$. In addition, the nonzero eigenvalues of $\psi^R(x,y)$ have finite multiplicity and form a sequence tending to zero if they are denumerable infinite in number [15, pp. 233].

First, consider the rate at which the eigenvalues of Ψ^R/N converge to their limiting values. Let $\left\{\lambda_k^{(R,N)}\right\}$ be the eigenvalues of the $N \times N$ matrix $\Psi^R/N = A_N$ listed in decreasing order. Then,

$$\sum_{k=0}^{N-1} \lambda_k^{(R,N)} = \mathrm{tr}\left(\frac{\Psi^R}{N}\right) = 1, \ \text{ for all } N, \tag{9}$$

where tr(A) is the trace of matrix A. From (6) and (8), observe that the eigenvalues $\left\{\lambda_k^{(R,N)}\right\}$ are obtained from the quadrature method using the rectangle rule [11, pp. 107] to approximate the eigenvalues of the homogeneous Fredholm's integral equation of the second kind in (8). Thus, as $N \to \infty$, the nonzero eigenvalue $\lambda_k^{(R,N)}$ converges to the corresponding eigenvalue $\lambda_k^{(R,\infty)}$ of the linear operator in (8) [11, pp. 248]. Regarding the rate of such, there exists a uniform error bound such that $\left|\lambda_k^{(R,N)} - \lambda_k^{(R,\infty)}\right| \leq C/N$, for any $k \leq N$, where C is a positive constant [11, pp. 270].

Next, it will be shown that the number of nonzero eigenvalues of Ψ^R/N is in the order of $o(N)$, for large N. For any matrix A with real eigenvalues, let F^A denote the empirical distribution function (i.e. e.d.f) of the eigenvalues of A; if A is $n \times n$, then

$$F^A(x) = \frac{1}{n}\left(\text{ number of eigenvalues of } A \leq x \right). \tag{10}$$

Therefore, there exists a sequence of cumulative distribution functions (CDF) $\left\{F^{A_N}\right\}$ defined accordingly to (10) for each A_N. Due to the non-negative definiteness of A_N and (9), it can be easily observed that $\lambda_k^{(R,N)} \in [0,1]$. Hence, F^{A_N} is concentrated on $[0,1]$. Let the kth moment of F^{A_N} be defined as

$$E_N\left[x^k\right] = \frac{\sum_{k=0}^{N-1}\left[\lambda_k^{(R,N)}\right]^k}{N} = \int_0^1 x^k F^{A_N}\{dx\}. \tag{11}$$

If $k \neq 0$, then $E_N\left[x^k\right] \to 0$ as $N \to \infty$; if $k = 0$, $\int_0^1 F^{A_N}\{dx\} = 1$, for any N. This implies that the sequence of the kth moment of F^{A_N} converges (in N) to a number u_k (in this case, 0 or 1) for each k. Such a convergence of the moments implies that the sequence $\left\{F^{A_N}\right\}$ converges to a PDF F_0 in distribution [16, pp. 251] (i.e. $F^{A_N} \xrightarrow{\mathcal{D}} F_0$) with $\int_0^1 xF_0\{dx\} = 0$, and this implies that $F_0(\{0\}) = 1$ [19, pp. 51]. Let the support of a e.d.f defined like that in (10) be the set of all eigenvalues. Therefore, the support of F^{A_N} is the set of $\left\{\lambda_k^{(R,N)}\right\}$, and the support of F_0 over positive values corresponds to the non-zero eigenvalues of the operator $\psi^R(x,y)$. Thus, an eigenvalue $\lambda_k^{(R,\infty)}$ is zero almost surely when measured by F_0. In another word, $F^{A_N}(\{\lambda : \lambda > 0\}) \to 0$, as $N \to \infty$. It can be concluded that $f_R(N)$, the number of nonzero values in $\left\{\lambda_k^{(R,N)}, k = 0, \cdots N-1\right\}$, satisfies $f_R(N)/N \to 0$ as N approaches infinity. The speed of $f_R(N)/N \to 0$ depends on smoothness of the kernel $\tilde{\psi}^R(r)$ in (8) in the sense of continuous differentiability of various orders [17], with smoother $\tilde{\psi}^R(r)$ corresponding to a faster convergence of $f_R(N)/N \to 0$, as will be shown in the simulation results.

3.2 Eigenvalues of Large Dimensional Sample Covariance Matrices

In this section, convergence issues regarding the eigenvalues of (random) sample covariance matrices of the form

$$B_N = \frac{1}{N}\left(\tilde{D}_N^R\right)^{1/2} X_N X_N^{\dagger} \left(\tilde{D}_N^R\right)^{1/2}, \tag{12}$$

are addressed, where $\tilde{D}_N^R = D_N^R/N$, and X_N is a $N \times N$ random matrix with entries that are i.i.d complex Gaussian random variables distributed as $\tilde{N}(0,1)$. In particular, it is established that the eigenvalues of the random matrix B_N are related to those of the deterministic matrix $\tilde{D}_N^R$.

Let the diagonal entries of $\tilde{D}_N^R$ be the set $\left\{\lambda_k^{(R,N)}, k = 0, \cdots, N-1\right\}$, whose asymptotic properties were studied in Section 3.1. Then, the desired result is established by showing that $\forall k \geq 0$, such that $\lambda_k^{(R,\infty)} > 0$, $\lambda_k^{(B_N)}$ converges almost surely to $\lambda_k^{(R,\infty)}$. In (12), the number of nonzero entries in the diagonal matrix $\tilde{D}_N^R$ is $f_R(N)$; therefore, the number of nonzero eigenvalues of B_N in (12) is not larger than $f_R(N)$. Let $\hat{B}_N$ denote the $f_R(N) \times f_R(N)$ random matrix

$$\hat{B}_N = \frac{1}{N} \left(\hat{D}_N^R\right)^{1/2} \hat{X}_N \hat{X}_N^\dagger \left(\hat{D}_N^R\right)^{1/2}, \tag{13}$$

where $\hat{D}_N^R$ is a $f_R(N) \times f_R(N)$ diagonal matrix with diagonal entries equal to the nonzero diagonal entries of $\tilde{D}_N^R$ in the same order, and $\hat{X}_N$ is the $f_R(N) \times N$ random matrix with i.i.d elements distributed as $\tilde{N}(0,1)$. Hence, the $f_R(N)$ eigenvalues (in decreasing order) of $\hat{B}_N$ in (13) are the same as those first $f_R(N)$ eigenvalues of B_N in (12), i.e. $\lambda_k^{B_N} = \lambda_k^{\hat{B}_N}$, $k = 0, \cdots, f_R(N) - 1$. In [12], it has been shown that for a matrix of the form of $\hat{B}_N$ in (13), the following inequality holds:

$$\lambda_k^{(R,N)} \lambda_{f_R(N)-1}^{\frac{1}{N}\hat{X}_N \hat{X}_N^\dagger} \leq \lambda_k^{B_N} \leq \lambda_k^{(R,N)} \lambda_0^{\frac{1}{N}\hat{X}_N \hat{X}_N^\dagger}, \tag{14}$$

where $k = 0, \cdots, f_R(N) - 1$, and $\lambda_{f_R(N)-1}^{\frac{1}{N}\hat{X}_N \hat{X}_N^\dagger}$ and $\lambda_0^{\frac{1}{N}\hat{X}_N \hat{X}_N^\dagger}$ are the smallest and the largest eigenvalues of the random matrix $\frac{1}{N}\hat{X}_N \hat{X}_N^\dagger$, respectively. As stated in Section 3.1, $\lambda_k^{(R,N)} \to \lambda_k^{(R,\infty)}$, for any k such that $\lambda_k^{(R,\infty)} > 0$. Since $f_R(N)/N \to c = 0$, it can be shown that [13],

$$\lambda_0^{\frac{1}{N}\hat{X}_N \hat{X}_N^\dagger} \xrightarrow{\text{a.s.}} \left(1 + \sqrt{c}\right)^2 = 1, \tag{15}$$

and

$$\lambda_{f_R(N)-1}^{\frac{1}{N}\hat{X}_N \hat{X}_N^\dagger} \xrightarrow{\text{a.s.}} \left(1 - \sqrt{c}\right)^2 = 1. \tag{16}$$

Therefore, it can be concluded that the the number of nonzero eigenvalues of B_N is $f_R(N)$ as well for large N, and they converge to the nonzero eigenvalues of $\psi^R(x,y)$ pointwisely, with probability one, as $N \to \infty$.

4 Asymptotic Analysis of Mutual Information $I_{N,N}$

4.1 Asymptotic Analysis of $I_{N,N}$ with Fixed Total Received Signal Power

In Section 2, the motivating equations were written for the case when the total transmitted power is ρ after normalizing the variance of the additive noise. For the case of a fixed total received power, the modification is obtained mathematically by further scaling the transmitted power by $1/N$, thus making the correlation matrix Q in (2) equal to $\frac{\rho}{N^2} I_N$ [7].

4.1.1 Fixed Length Linear Array at the Receiver Side Only

In this section, it will be assumed that as $N \to \infty$, Ψ^T can be maintained as I_N. However, at the receiver side, the antennas will need to be fit into a fixed-length linear array. This scenario is the case when the base station can afford a large array, while the mobile unit cannot due to its physical constraints. In such a scenario, the asymptotic characteristics of $I_{N,N}$, which is now

$$I_{N,N} = \log_2 \det \left(I_N + \frac{\rho}{N^2} \left(D_N^R \right)^{1/2} W W^\dagger \left(D_N^R \right)^{1/2} \right), \tag{17}$$

will be investigated, where D_N^R is defined in (5). Based on the assumptions stated above, the authors in [7] argued that if there exists only one-sided correlation caused by the receive antennas, $E\left[I_{N,N}\right]$ (i.e. the average of $I_{N,N}$ in (2)) converges to a constant. In this section, by employing the results of Section 3, a stronger result will be shown. In particular, with a fixed length linear array on the receiver side, as well as the total received signal power fixed, the instantaneous mutual information $I_{N,N}$ will be shown to converge almost surely to a deterministic constant as $N \to \infty$. That constant can be determined by $\left\{ \lambda_k^{(R,\infty)} \right\}$, the nonzero eigenvalues of $\psi^R(x, y)$.

By the definition of det$(\cdot)$, (17) becomes

$$I_{N,N} \;\; = \;\; \sum_{k=0}^{N-1} \log_2 \left(1 + \rho \lambda_k^{(B_N)} \right), \tag{18}$$

where B_N is defined in (12), and where $\lambda_k^{(B_N)}$ is the $(k+1)$st largest eigenvalue of the random matrix B_N. Recall from Section 3 that the number of non-zero eigenvalues $\lambda^{(B_N)}$ is almost surely $f_R(N)$, the number of non-zero eigenvalues of D_N^R/N. Thus, almost surely,

$$I_{N,N} = \sum_{k=0}^{f_R(N)-1} \log_2 \left(1 + \rho \lambda_k^{(\hat{B}_N)} \right) \leq \frac{\rho}{\ln 2} \sum_{k=0}^{f_R(N)-1} \lambda_k^{(R,N)} \lambda_0^{\frac{1}{N} \hat{X}_N \hat{X}_N^\dagger} = \frac{\rho}{\ln 2} \lambda_0^{\frac{1}{N} \hat{X}_N \hat{X}_N^\dagger}, \tag{19}$$

where $\hat{B}_N$ is defined in (13), the inequality is due to (14) and $\ln(1+x) \leq x$, and the second equality is because $\sum_{k=0}^{f_R(N)-1} \lambda_k^{(R,N)} = 1$. In (15), it has been shown that $\lambda_0^{\frac{1}{N} \hat{X}_N \hat{X}_N^\dagger}$ converges to 1 almost surely. Therefore, $I_{N,N}$ in (19) is upper-bounded by $\rho/\ln 2$ with probability one, as $N \to \infty$.

As shown in Section 3.2, $\left\{\lambda_k^{(\hat{B}_N)}\right\}$ converge pointwisely to the nonzero eigenvalues $\left\{\lambda_k^{(R,\infty)}\right\}$ of $\psi^R(x,y)$, with probability one, which results in the following theorem.

Theorem 1:

$$\sum_{k=0}^{f_R(N)-1} \log_2\left(1+\rho\lambda_k^{(B_N)}\right) \xrightarrow{\text{a.s.}} \sum_{k=0}^{f_R(N)-1} \log_2\left(1+\rho\lambda_k^{(R,\infty)}\right), \quad \text{as } N\to\infty \tag{20}$$

The proof of (20) is accomplished by Lemmas 1 and 2, which are stated and proven in Appendix A. Hence, if the fixed length linear array is located at the receiver side, as N approaches infinity, $I_{N,N}$ in (18) converges almost surely,

$$I_{N,N} \xrightarrow{\text{a.s.}} \sum_{k=0}^{f_R(N)-1} \log_2\left(1+\rho\lambda_k^{(R,\infty)}\right), \tag{21}$$

which is upper bounded by a finite number $\rho/\ln 2$. Theorem 1 is a more precise statement of Lemma 2 of the independent work [21], which employs a quite different proof technique.

4.1.2 Fixed Length Linear Array at Both Sides

In this section, the case with fixed length linear arrays located at both the transmitter and receiver will be investigated. In this case, let $\psi^R(r)$ and $\psi^T(r)$ be the spatial correlation functions caused by the receive and transmit antennas, and let L_R and L_T be the length of the linear arrays. By following the same approach taken above, it can be concluded that the number of nonzero eigenvalues of matrices Ψ^R/N and Ψ^T/N will be in the order of $o(N)$, which can be represented by $f_R(N)$ and $f_T(N)$, respectively, with $f_T(N)/N \to 0$ and $f_R(N)/N \to 0$, as $N\to\infty$. Let these eigenvalues (in decreasing order) be the diagonal entries of the diagonal matrix $\hat{D}_N^R = \text{diag}\left\{\lambda_i^{(R,N)}, i=0,\ldots,f_R(N)-1\right\}$ and $\hat{D}_N^T = \text{diag}\left\{\lambda_i^{(T,N)}, i=0,\ldots,f_T(N)-1\right\}$, respectively. Therefore, with $Q=\frac{\rho}{N^2}I_N$, (5) becomes:

$$I_{N,N} \stackrel{\mathcal{D}}{=} \log_2\det\left(I_N+\frac{\rho}{N^2}D_N^R W D_N^T W^\dagger\right) \stackrel{\mathcal{D}}{=} \log_2\det\left(I_{f_R(N)}+\rho\hat{D}_N^R Y_N \hat{D}_N^T Y_N^\dagger\right) \tag{22}$$

where Y_N is a $f_R(N)\times f_T(N)$ random matrix, whose entries are i.i.d and distributed as $\tilde{N}(0,1)$. As $N\to\infty$, $\lambda_k^{(R,N)}\to\lambda_k^{(R,\infty)}$, $\lambda_k^{(T,N)}\to\lambda_k^{(T,\infty)}$, where $\lambda_k^{(R,\infty)}$ and $\lambda_k^{(T,\infty)}$ are the kth largest nonzero eigenvalues of $\psi^R(x,y)$, and $\psi^T(x,y)$, respectively, which are determined by the receive and transmit spatial correlation function in the same way as that in (8).

As one might expect, it can be shown, using Hadamard's inequality, Theorem 1, and Lemma 2 (see [22]) that the expected value of $I_{N,N}$ in (22) can be upper bounded by $I_{N,N}$ with the fixed length linear array at the receiver end only. As more antennas are put in this MEA system, the ergodic capacity will be non-decreasing under the conditions assumed throughout this work; thus, since $E\left[I_{N,N}\right]$ has a finite upper bound, it can be concluded that $E\left[I_{N,N}\right] \rightarrow C$, which agrees with what is claimed in [7] for the mean capacity, and C is a finite constant that will depend on the spectrum of the Hermitian operators, $\psi^R(x,y)$ and $\psi^T(x,y)$, respectively.

4.2 Asymptotic Analysis of $I_{N,N}$ with Fixed Total Transmit Signal Power

4.2.1 Linear Array of Fixed Length at Receiver Side Only

Following Section 3.2, $I_{N,N}$ in (5) will have the same asymptotic characteristics as

$$\hat{I}_{N,N} = \log_2 \det\left(I_{f_R(N)} + \rho\left(\hat{D}_N^R\right)^{1/2} \hat{X}_N \hat{X}_N^{\dagger} \left(\hat{D}_N^R\right)^{1/2}\right) = \sum_{k=0}^{f_R(N)-1} \log_2\left(1 + N\rho\lambda_k^{(\hat{B}_N)}\right), \quad (23)$$

as $N \rightarrow \infty$, where $\hat{B}_N$ and $\hat{D}_N^R$ are defined in (13).

In this case of fixed total transmit power, we are unable to get a result similar to almost sure convergence to a deterministic constant as that in Theorem 1 for the case with the fixed total received power. This is directly attributed to the extra factor N in (23), and can be explained intuitively by noticing that as N increases, the total received power will be increased accordingly, which will make the mutual information $I_{N,N}$ grow as well. Thus, in this section, two results that are much weaker than the precise almost sure convergence in Theorem 1 are established. In particular, it will be shown that, for the case of fixed total transmit power: (1) the normalized mutual information converges almost surely to zero, and (2) the *mean* capacity is upper bounded by an expression that is analogous to that on the right side of Theorem 1. Much stronger results (analogous to Theorem 1) can be justified by exploiting properties of specific spatial correlation functions, as will be discussed in Section 4.3.

First, it will be shown that the normalized mutual information $\hat{I}_{N,N}/N$ converges to zero almost surely.

$$
\begin{aligned}
\frac{\hat{I}_{N,N}}{N} &= \frac{1}{N}\sum_{k=0}^{f_R(N)-1}\log_2\left(1+N\rho\lambda_k^{(\hat{B}_N)}\right)\\
&\leq \frac{f_R(N)}{N}\log_2\left(1+\frac{N}{f_R(N)}\sum_{k=0}^{f_R(N)-1}\rho\lambda_k^{(\hat{B}_N)}\right)\\
&\leq \frac{f_R(N)}{N}\log_2\left(1+\frac{N}{f_R(N)}\sum_{k=0}^{f_R(N)-1}\rho\lambda_k^{(R,N)}\lambda_0^{\frac{1}{N}\hat{X}_N\hat{X}_N^\dagger}\right)\\
&= \frac{f_R(N)}{N}\log_2\left(1+\frac{N}{f_R(N)}\rho\lambda_0^{\frac{1}{N}\hat{X}_N\hat{X}_N^\dagger}\right)\xrightarrow{\text{a.s.}} 0, \qquad (24)
\end{aligned}
$$

where the first inequality is because of the concavity of the function $\log_2(x)$, and the second inequality is because of that in (14). The last equality is the consequence of $\sum_{k=0}^{f_R(N)-1}\lambda_k^{(R,N)} = 1$. Finally, the almost sure convergence is based on the a.s. convergence in (15), $f_R(N)/N \to 0$, and $x\log_2(1+ax) \to 0$, as $x \to 0$, where $a > 0$.

Next, the upper bound of the ergodic capacity $E[I_{N,N}]$ can be shown, similarly to that of Section 4.1.2 that [22]

$$
E[I_{N,N}] \leq \sum_{k=0}^{f_R(N)-1}\log_2\left(1+N\rho\lambda_k^{(R,N)}\right), \text{ for all } N. \qquad (25)
$$

4.2.2 Fixed Size Linear Array at Transmitter and Receiver

If the transmitting antennas are also spatially dense (i.e. with fixed length linear arrays located at both the base station and mobile), and $Q = \frac{\rho}{N}I_N$, then the analog to (22) with fixed total transmit power becomes

$$
I_{N,N} = \log_2\det\left(I_{f_R(N)} + N\rho\hat{D}_N^R Y_N \hat{D}_N^T Y_N^\dagger\right). \qquad (26)
$$

By taking the same approach as that in Section 4.1.2, it can be shown as well that [22]

$$
E[I_{N,N}] \leq \sum_{k=0}^{f_R(N)-1}\log_2\left(1+N\rho\lambda_k^{(R,N)}\right), \text{ for all } N. \qquad (27)
$$

where the right hand side of (27) can be approximated as $\sum_{k=0}^{N_0-1}\log_2\left(1+N\rho\lambda_k^{(R,\infty)}\right)$, for large N, which is the tight upper-bound (25) of the ergodic capacity of the $N\times N$ MIMO system with a fixed length linear array at the receiver side only, when the spatial correlation functions satisfy certain conditions discussed in Section 4.3.

4.3 Exploiting Properties of Common Correlation Functions

In Section 4.1 and 4.2, attention has been restricted to a general spatial correlation function $\psi^R(x, y)$ and to results that can be proven very formally. In this section, properties of the eigenvalues resulting from common spatial correlation functions are studied to assist in the evaluation of the quantities derived in Section 4.1 and 4.2 and to develop some well-justified (although less formal) approximations for the asymptotic behavior of the mutual information. In particular, a result similar to Theorem 1 for the case of fixed total transmit power is developed, and consideration is given to how the expressions for the asymptotic mutual information in all cases can be approximated very simply from properties of the power spectral density corresponding to the spatial correlation function.

Significant simplifications are obtained by approximating the number of non-zero eigenvalues in $\left\{\lambda_k^{(R,N)}\right\}$, for large N, as a finite constant N_0 for a broad class of spatial correlation functions. First, consider $f_R(N)$ for the case of a bandlimited spatial correlation function, which is true in many applications [5] [4, p. 134]. Let $F_R(\Omega)$ be the power spectral density corresponding to $\psi^R(r)$, and assume the support of $F_R(\Omega)$ is on the interval $[-\Omega_0^R, \Omega_0^R]$. Let $\{\lambda_k^{(R,N)}, k = 0, \ldots, N-1\}$ be the eigenvalues of Ψ^R/N. Based on Toeplitz matrix theory [9], by taking the same approach as that in our work on power control [10], it can be shown that as $N \to \infty$, $\{\lambda_k^{(R,N)}\}$ is asymptotically equally distributed with $\left\{f^{(\infty)}\left(\omega_k = \frac{k\pi}{N}\right),\ k = 0, 1, \ldots, N-1\right\}$ [22]. Let $\omega_k = \Omega_k L_R/N$, and thus $\Omega_k = k\pi/L_R$. Then, $f^{(\infty)}(\omega_k) = f^{(\infty)}(\Omega_k)$ can be determined as follows,

$$f^{(\infty)}\left(\Omega_k\right) \quad = \quad \lim_{N\to\infty} \frac{1}{N} \sum_{l=-N+1}^{N-1} \psi_R\left(\frac{lL_R}{N-1}\right) e^{-jl\omega}|_{\omega_k=\Omega_k L_R/N} \tag{28}$$

where $f^{(\infty)}(\Omega) = \frac{1}{L_R}\int_{-L_R}^{L_R} \psi_R(r) e^{-j\Omega r} d\,r$. Therefore, as $N \to \infty$, eigenvalues of the large matrix Ψ^R/N behave the same as the sampling points of the power spectral density determined from (28) in an average sense. What this indicates is that for a given linear array of fixed length, if more and more antennas are allocated within this array, the same segment of fading correlation function of $\psi^R(r)$ over $[-L_R, L_R]$ is sampled with higher and higher spatial frequency. Since $\psi^R(r)$ is bandlimited, the number of sampling points lying in the nonzero part of $f^{(\infty)}(\Omega)$ will be approximated as N_0, which results in the dominant N_0 nonzero eigenvalues of the Toeplitz

Hermitian matrix Ψ^R/N, if N is sufficiently large, and all other eigenvalues of it will fall outside the nonzero region of $f^{(\infty)}(\Omega)$.

Since $f^{(\infty)}(\Omega)$ is the Fourier transform of the spatial correlation function *truncated* to $[-L_R, L_R]$, the strictly bandlimited nature of $F_R(\Omega)$ implies that $f^{(\infty)}(\Omega)$ is not bandlimited. However, as will be shown below, the numerical results suggest that a bandlimited approximation to such is quite useful for numbers of antenna elements of interest. In particular, the power spectral density decays rapidly outside of a transition region. Thus, while the $\left\{\lambda_k^{(R,N)}\right\}$ are converging point-wisely to the point spectrum $\left\{\lambda_k^{(R,\infty)}\right\}$ of the Hermitian and Toeplitz operator specified in (8), the number of nonzero eigenvalues $f_R(N)$ can be approximated as a finite number N_0. It should be noted that *any* non-zero eigenvalue will eventually (N large enough) have a significant absolute impact on the capacity in the case when the total average transmit power is fixed (i.e. in (30) below), but this does not happen until N approaches the inverse of that eigenvalue, and thus the threshold below which a sample of the power spectral density is ignored can be chosen small enough to place those N beyond the values of interest. Numerical results will firmly support this approach.

Even if the spatial correlation function is not bandlimited, the approximation of $f_R(N) \approx N_0$ can still be obtained if $\tilde{\psi}(r)$ satisfies certain analytic properties. As noted in [17], for a kernel $\psi(x, y)$ like that in (8), if $\psi(x, y)$ is an analytic function in y on the whole segment $[0, 1]$ including the endpoints, uniformly in x, then $\left|\lambda_k^{(R,\infty)}\right| < \exp\left(-\alpha k - \beta\right)$, where α and β are positive constants, and it can be inferred that $\lambda_k^{(R,\infty)}$ is decreasing to zero very fast. Since the nonzero $\lambda_k^{(R,N)}$ converges to the nonzero eigenvalue $\lambda_k^{(R,\infty)}$, as $N \to \infty$, for $k > N_0$, $\lambda_k^{(R,N)}$ can be approximated as zero. Hence, in this case the $f_R(N)$ can be approximated as N_0 as well, which will be illustrated in the simulation results.

Therefore, if the spatial correlation function $\tilde{\psi}(r)$ is bandlimited or $\psi(x-y)$ satisfies the analytic properties stated above, as $N \to \infty$, the diagonal matrix $\hat{D}_N^R$ in (23) can be approximated as a diagonal matrix whose N_0 upper left diagonal entries are positive, and all other entries will be vanishing. Hence, for large N, the mutual information for the case of fixed transmit power (23) becomes:

$$I_{N,N} \approx \hat{I}_{N_0,N} = \log_2 \det\left(I_{N_0} + \rho \left(\bar{D}_N^R\right)^{1/2} \hat{W}\hat{W}^\dagger \left(\bar{D}_N^R\right)^{1/2}\right), \tag{29}$$

where $\bar{D}_N^R$ is a $N_0 \times N_0$ diagonal matrix with diagonal entries as $\{\lambda_k^{(R,N)}\}$, and $\hat{W}$ is a $N_0 \times N$ random matrix, whose entries are circularly symmetric complex Gaussian random variables: $\sim$ i.i.d $\tilde{N}(0,1)$.

Let $\bar{B}_N = \frac{1}{N}\left(\bar{D}_N^R\right)^{1/2} \hat{W}\hat{W}^\dagger \left(\bar{D}_N^R\right)^{1/2}$, which is a Wishart matrix of dimension $N_0 \times N_0$. It is trivial to show that $\lambda_k^{(\bar{B}_N)}$ converges to $\lambda_k^{(R,\infty)}$, for $k = 0, \cdots, N_0 - 1$, almost surely, where $\lambda_k^{(\hat{B}_N)}$ is the $(k+1)$th largest eigenvalue of the random matrix $\hat{B}_N$. Therefore, the following convergence result concerning the mutual information $\hat{I}_{N_0,N}$ in (29) will hold:

$$\lim_{N\to\infty}\left(\hat{I}_{N_0,N} - \sum_{i=0}^{N_0-1}\log_2\left(1 + N\rho\lambda_i^{(R,\infty)}\right)\right) = 0, \text{ a.s.}, \tag{30}$$

which can be proven by employing the results that $\lambda_k^{(\bar{B}_N)} \to \lambda_k^{(R,\infty)}$, a.s., $k = 0, \cdots, N_0 - 1$, and $N_0 < \infty$ [13]. Thus, the desired analog to Theorem 1 for the case of fixed total transmit power is obtained. Further interpretation of (30) can be found in [8].

5 Simulation Results

The key to the applicability of the results of Section 4 is how well they hold for large but finite N. The parameters employed are as follows: linear array lengths at the mobile unit of length $L_R = 5\lambda$ or λ, where λ is the carrier wavelength, and signal-to-noise ratios (SNR) of $\rho = -5$ dB, or $\rho = 22$ dB. Consider the correlation function $\psi_1^R(r) = \mathrm{sinc}\,(2r/\lambda)$, which has bandlimited power spectral density; it is obtained for uniform angles of arrival in both the azimuth and elevation planes. A spatial correlation function that is not bandlimited, $\psi_2^R(r) = e^{-1/2(2\pi r\sigma_\theta/\lambda)^2}$ [20], where $\sigma_\theta = 0.25$, is also taken as one of the examples. Finally, the artificial example $\psi_3^R(r) = e^{-|r|/d_\lambda}$, where $d_\lambda = \lambda$, is considered to demonstrate analytic requirements for $\psi^R(r)$. For the case with a fixed length linear array at transmitter and receiver, let $\psi^R(r) = \psi^T(r)$.

Figure 2 shows the discrete Fourier transform

$$f^{(N)}(\omega) = \frac{1}{N}\sum_{l=-N+1}^{N-1}\psi_R\left(\frac{lL_R}{N-1}\right)e^{-jl\omega}, \tag{31}$$

for $N = 200$ and $L_R = 5\lambda$, with the corresponding set of dominant eigenvalues $\left\{\lambda_k^{(R,N)}\right\}$ of Ψ^R/N

in increasing order. If $\psi^R(r) = \text{sinc}\,(2r/\lambda)$, then the number of dominant eigenvalues is chosen as $N_R = 15$ for $L_R = 5\lambda$, $N_R = 6$ for $L_R = \lambda$.

Almost sure convergence is demonstrated through a combination of two approaches. First, a number of randomly generated realizations of $I_{N,N}$ are shown to behave as expected. Second, histograms are generated to demonstrate convergence in distribution of $I_{N,N}$, which is, of course, implied by almost sure convergence. Figures 3, 4, 5 and 6 show the characteristics of instantaneous mutual information $I_{N,N}$ of a (N, N) MEA system versus N, with the normalized total average received power, for various spatial correlation functions. In the case with a fixed length array at the receiver side only, $I_{N,N}$ is converging to the analytical results as N grows large as claimed in Theorem 1 (20). In the case of the fixed length arrays put at both sides, more randomness is observed, as expected. In addition, the mean value of $I_{N,N}$ is demonstrated to converge to a constant that is smaller than the limit if fixed length linear array exists at only one side.

For spatial correlation function $\psi_3^R(r)$, since $e^{-|r|/\lambda}$ does not have derivative at $r = 0$, and thus does not satisfy the analytical conditions stated in Section 4.2.1, the decreasing rate of $\lambda_k^{(R,\infty)}$ is only of the order of k^{-2} for large k [18]. However, the conditions required by Theorem 1 are still satisfied for $\psi_3^R(r)$. Therefore, when the total receive power is fixed, the similar convergence conclusions can still be drawn as claimed by Theorem 1 and verified in Figure 6, but with slower convergence as predicted in proofs of Lemma 1 and 2 than that in Figures 3 and 4.

In Figures 7 and 8, histograms of $I_{N,N}$ agree with the results obtained in Figure 5, because the PDFs of $I_{N,N}$ are becoming more concentrated around $\sum_{k=0}^{f_R(N)-1} \log_2\left(1 + \rho\lambda_k^{(R,N)}\right) = 0.4419$, as N increases from 40 to 200, when the fixed length array is put at only one side; the PDF does not show any convergence as N increases when such arrays are put at both sides, but the mean of $I_{N,N}$ shows little variation with the increasing of N.

Figures 1, 9, 10 and 11 demonstrate the performance of the instantaneous mutual information $I_{N,N}$ of a (N, N) MEA system versus N with the fixed total average transmit power. It can be observed that the sum term in (30) is a very accurate approximation for $I_{N,N}$ if there exists a fixed length linear array at one side with spatial correlation functions being bandlimited or having nice analytical properties as that of $\psi_2^R(r)$. Therefore, the upper-bound of the ergodic capacity

$E\left[I_{N,N}\right]$ in (25) is tight in this scenario. If there are fixed length linear arrays at both sides, $I_{N,N}$, indicated by the dash-dot lines in those figures, does not show convergence, which is as expected from analysis, and the average value is upper bounded by the asymptotic value of $I_{N,N}$ when the fixed length array is used at one side only, as given in (27).

6 Conclusion

In this paper, the convergence of the instantaneous mutual information $I_{N,N}$ of a (N,N) MEA system is investigated analytically and tested through simulations, for the case when spatial correlations are caused by the restriction that the elements of the array must occupy a fixed length at either the mobile unit or at both sides (see Table 1). The main contribution of this paper is that the almost sure convergence of the mutual information $I_{N,N}$ under certain conditions has been shown in Theorem 1 by exploiting the relationships between the eigenvalues of the random matrix B_N and the eigenvalues of the linear operator $\psi^R(x,y)$. In addition, for those common spatial correlation functions described in Section 4.3, some simple approximations of $I_{N,N}$ can be achieved for the case with total fixed transmit power. This implies that, when the fixed length array is put at the receiver side only, and N is large, $I_{N,N}$ can be approximated well by a deterministic figure, which only relies on a finite number of non-zero eigenvalues determined by the spatial correlation function. Similar results can be obtained in a straightforward way in the case when an arbitrary 2-D antenna array is used at either the receiver side or both sides by exploiting the Hermitian characteristic of the covariance matrix Ψ^R and Ψ^T [22].

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.* vol. 6, pp. 311-335, March 1998.

[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecommunications,*" vol. 10, no. 6, pp. 586-595, 1999.

[3] G. J. Foschini, “Layered space-Time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Technical Journal*, pp. 41-59, Autumn 1996.

[4] R. Janaswamy, *Radiowave Propagation and Smart Antennas for Wireless Communications*, Kluwer: 2000.

[5] D. S. Shiu, G. J. Foschini, M. J. Gans and J. M. Kahn, “Fading correlation and its effect on the capacity of multielement antenna systems,” *IEEE Trans. Comm.*, vol. 48, pp. 502-513, March 2000.

[6] C. N. Chuah, D. Tse, J. M. Kahn and R. A. Valenzuela, “Capacity scaling in MIMO wireless systems under correlated fading,” *IEEE Trans. Inform. Theory,* vol. 48, pp. 637-651, March 2002.

[7] N. Chiurtu, B. Rimodi and E. Telatar, “Dense multiple antenna systems,” *Information Theory Workshop,* 2001.

[8] S. Wei, D. Goeckel and R. Janaswamy, “On the capacity of fixed length linear arrays under bandlimited correlated fading,” *CISS Proc.*, April 2002.

[9] U. Grenander and G. Szego, *Toeplitz Forms and Their Applications,* Chelsea Publishing Company: 1984.

[10] S. Wei and D. L. Goeckel, “Error statistics for average power measurements in wireless communication systems,” *IEEE Trans. Comm.*, pp. 1535-1546, September 2002.

[11] C. T. H. Baker, *The Numerical Treatment of Integral Equations,* Clarendon, 1977.

[12] Z. D. Bai and J. W. Silverstein, Exact separation of eigenvalues of large dimensional sample covariance matrices,” *Ann. Probab.*, vol. 27, 1536-1555, 1999.

[13] Z. D. Bai and J. W. Silverstein, "CLT for linear spectral statistics of large dimensional covariance matrices," submitted.

[14] R. Durrett, *Probability: Theory and Examples, second edition,* Duxbury, 1996.

[15] F. Riesz and B. Nagy, *Functional Analysis,* Frederick Ungar, 1955.

[16] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume II, Second Edition*, Wiley, 1971.

[17] M .H. Zahar-Itkin, "Growth of the eigenvalues of a linear integral equation," *Soviet Math.,* vol. 5, pp. 1348-1351, 1964.

[18] J. Capon, "Asymptotic eigenfunctions and eigenvalues of a homogeneous integral equation," *IRE Transactions on Information Theory,* pp. 2-4, January, 1962.

[19] G. B. Folland, *Real Analysis,* Wiley, 1999.

[20] H. Bölcskei, D. Gesbert and A. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.,* pp. 225-234, Feb. 2002.

[21] D. Gesbert, T. Ekman and N. Christophersen, "Capacity limits of dense palm-sized MIMO arrays," *Proc. of IEEE Globecomm*, Taiwan, 2002.

[22] S. Wei, *Convergence Results on Broad-band Wireless Communication Systems and Their Implications,* Ph.D. dissertation, University of Massachusetts, Amherst, 2003.

A Proofs of Results from Section 4.1.1

Lemma 1:

$$\sum_{k=0}^{f_R(N)-1} \left[\log_2\left(1+\rho\lambda_k^{(B_N)}\right) - \log_2\left(1+\rho\lambda_k^{(R,N)}\right)\right] \xrightarrow{\text{a.s.}} 0. \tag{32}$$

Proof: Let Ω be the set of outcomes of the underlying probability space [14], and let ω refer to an outcome in Ω. Since $\lambda_0^{\frac{1}{N}\hat{X}_N\hat{X}_N^\dagger}$ and $\lambda_{f_R(N)-1}^{\frac{1}{N}\hat{X}_N\hat{X}_N^\dagger}$ in (14) converge to 1 almost surely, $\exists$ a set A, $P(A)=1$, such that for $\omega \in A$, $\lambda_0^{\frac{1}{N}\hat{X}_N(\omega)\hat{X}_N^\dagger(\omega)} \to 1$, and $\lambda_{f_R(N)-1}^{\frac{1}{N}\hat{X}_N(\omega)\hat{X}_N^\dagger(\omega)} \to 1$. Therefore, for any $1>\delta>0$, $\omega \in A$, there exists $N_0(\omega,\delta)$, such that when $N > N_0(\omega,\delta)$, $\left|\lambda_0^{\frac{1}{N}X(\hat{\omega})_N\hat{X}_N^\dagger(\omega)} - 1\right| < \delta$, $\left|\lambda_{f_R(N)-1}^{\frac{1}{N}X(\hat{\omega})_N\hat{X}_N^\dagger(\omega)} - 1\right| < \delta$, and for any $0 \le k \le f_R(N)$, by inequality in (14), the following bound of $\lambda_k^{(B_N)}$ can be obtained,

$$(1-\delta)\lambda_k^{(R,N)} \le \lambda_k^{(B_N(\omega))} \le (1+\delta)\lambda_k^{(R,N)}. \tag{33}$$

By substituting the bounds of $\lambda_k^{(B_N(\omega))}$ in (33) into $I_{N,N}$, and employing the two inequalities,

$$\ln(1+x+\epsilon) \le \ln(1+x)+\epsilon \qquad \ln(1+x-\epsilon) \ge \ln(1+x)-\epsilon$$

for any $x>0$, $\epsilon>0$ and $x>\epsilon$, we have the following bounds concerning $\log_2\left(1+\rho\lambda_k^{(B_N(\omega))}\right)$,

$$\log_2\left(1+\rho\lambda_k^{(R,N)}\right) - \frac{\rho}{\ln 2}\delta\lambda_k^{(R,N)} \le \log_2\left(1+\rho\lambda_k^{(B_N(\omega))}\right) \le \log_2\left(1+\rho\lambda_k^{(R,N)}\right) + \frac{\rho}{\ln 2}\delta\lambda_k^{(R,N)}. \tag{34}$$

By summing up the inequality in (34) from 0 to $f_R(N)$, and taking advantage of $\sum_{k=0}^{f_R(N)-1}\lambda_k^{(R,N)} = 1$, it yields

$$\left|\sum_{k=0}^{f_R(N)-1}\left[\log_2\left(1+\rho\lambda_k^{(B_N(\omega))}\right) - \log_2\left(1+\rho\lambda_k^{(R,N)}\right)\right]\right| \le \frac{\rho}{\ln 2}\delta, \tag{35}$$

which is true for any $\omega \in A$, and $0<\delta<1$. Recalling $P(A)=1$, Lemma 1 is shown.

Lemma 2:

$$\sum_{k=0}^{f_R(N)-1}\left[\log_2\left(1+\rho\lambda_k^{(R,N)}\right) - \log_2\left(1+\rho\lambda_k^{(R,\infty)}\right)\right] \longrightarrow 0. \tag{36}$$

Proof: Using the uniform upperbound $\left|\lambda_k^{(R,N)} - \lambda^{(R,\infty)}\right| \le C/N$, for all k, as stated in Section 3.1 yields

$$\log_2\left(1-\frac{\rho}{N}\right)^{f_R(N)} \le \sum_{k=0}^{f_R(N)-1}\left[\log_2\left(1+\rho\lambda_k^{(R,N)}\right) - \log_2\left(1+\rho\lambda_k^{(R,\infty)}\right)\right] \le \log_2\left(1+\frac{\rho}{N}\right)^{f_R(N)}. \tag{37}$$

Note $\left(1-\frac{\rho}{N}\right)^{f_R(N)} \to e^0 = 1$, and $\left(1+\frac{\rho}{N}\right)^{f_R(N)} \to e^0 = 1$, since $f_R(N)/N \to 0$ [14, pp. 80].

	Fixed length linear array at the RX only	Fixed Length linear array at TX and RX
Fixed total RX Power	Section 4.1.1: $I_{N,N}$ converges almost surely to a deterministic constant C_1, as $N \to \infty$, $C_1 = \sum_{k=0}^{f_R(N)-1} \log_2 \left(1 + \rho\mu_k^{(R,\infty)}\right)$, $f_R(N)/N \to 0$, which is upper bounded by $\rho/\ln 2$. See (20).	Section 4.1.2: $E\left[I_{N,N}\right]$ converges to a constant C_0, which is upper bounded by the limit C_1.
Fixed total TX Power	Section 4.2.1: $I_{N,N} \approx \hat{I}_{N_0,N}$, see (29). The difference between $\hat{I}_{N_0,N}$ and $\sum_{i=0}^{N_0-1} \log_2 \left(1 + N\rho\lambda_i^{(R,\infty)}\right)$ converges almost surely to zero. See (30).	Section 4.2.2: $E\left[I_{N,N}\right]$ is smaller than that with fixed length array at RX only. See (27).

Table 1: The main results of the paper, where $I_{N,N}$ is the instantaneous mutual information of a (N, N) MEA system, and $E\left[I_{N,N}\right]$ is its expected value for various transmit (TX) and receive (RX) assumptions. The variable ρ represents the total received power in the first row (fixed total RX power), and the received SNR at each antenna in the second row (fixed total TX power), respectively. $\left\{\lambda^{(R,\infty)}\right\}$ is the set of nonzero eigenvalues of the Hermitian operator $\psi^R(x, y)$, as determined in (8), and C_0 and C_1 are deterministic constants.

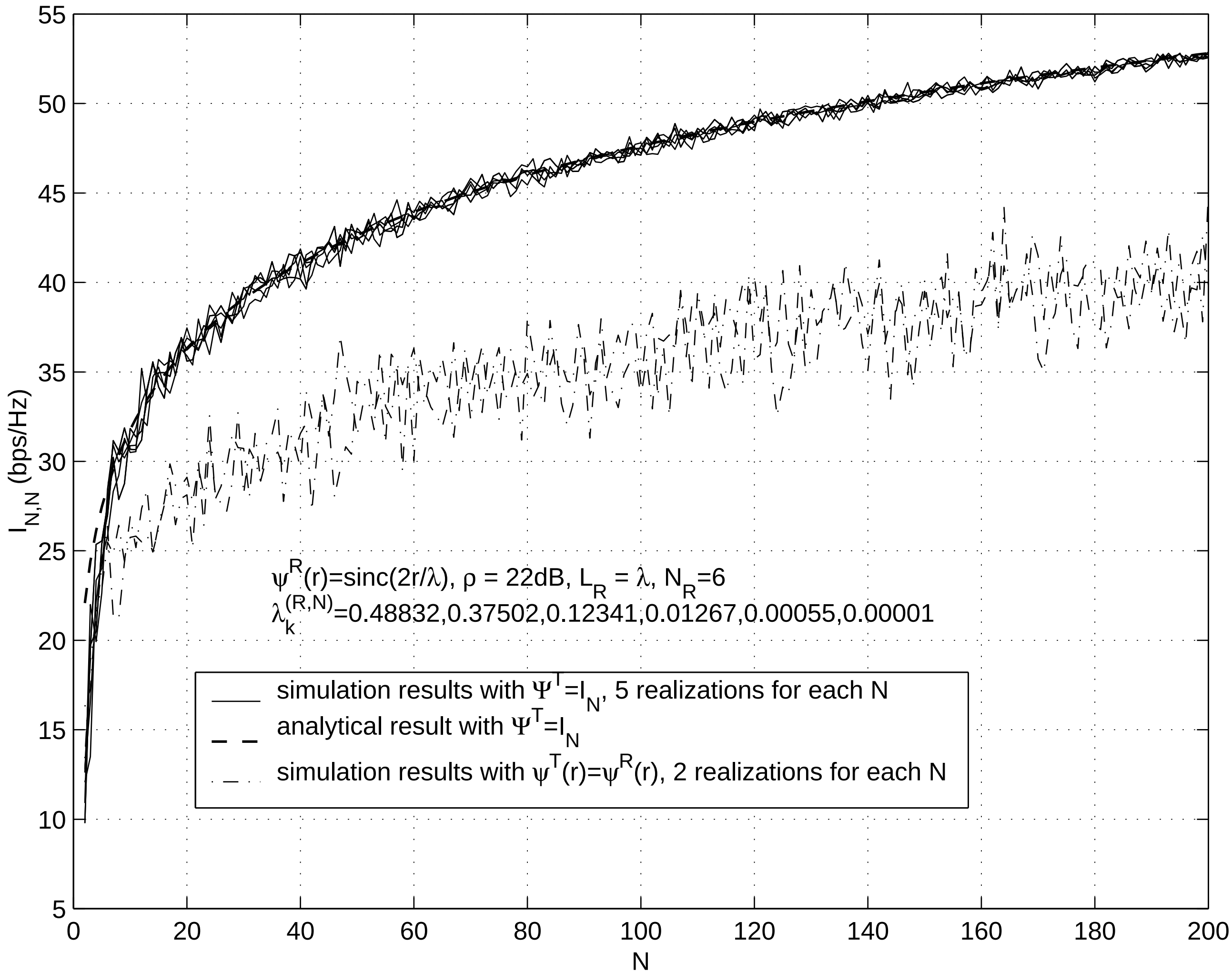

Figure 1: Instantaneous mutual information $I_{N,N}$ of a (N,N) MEA system versus N, with fixed total average transmit power. Simulation results are obtained through generating one realization of random variable $I_{N,N}$ for each N in terms of the first equation in (5), where $L_R = \lambda$, $\rho = 22\,\text{dB}$. For the case when the fixed length linear array is put at the receiver side only, 5 realizations are generated for each N, and the Toeplitz matrix Ψ^R is determined by $\psi^R(r) = \text{sinc}\,(2r/\lambda)$, $\Psi^T = I_N$. For the case when the fixed length linear arrays are put at both the transmitter and receiver side, 2 realizations are generated for each N, and $\Psi^R = \Psi^T$ are determined by $\psi^T(r) = \psi^R(r) = \text{sinc}\,(2r/\lambda)$. Analytical results are obtained by the sum term in (30), where $f_R(N)$ is approximated as $N_R = 6$ and eigenvalues $\left\{\lambda_k^{(R,N)}\right\}$ are listed in the figure. It can be observed that for each realization in the case with a fixed length array at the receiver side only, $I_{N,N}$ is converging rapidly for large N, and the upper-bound in (25) is tight. The case of a fixed length array located at both sides is indicated by the dash-dot lines, and more randomness is observed as expected.

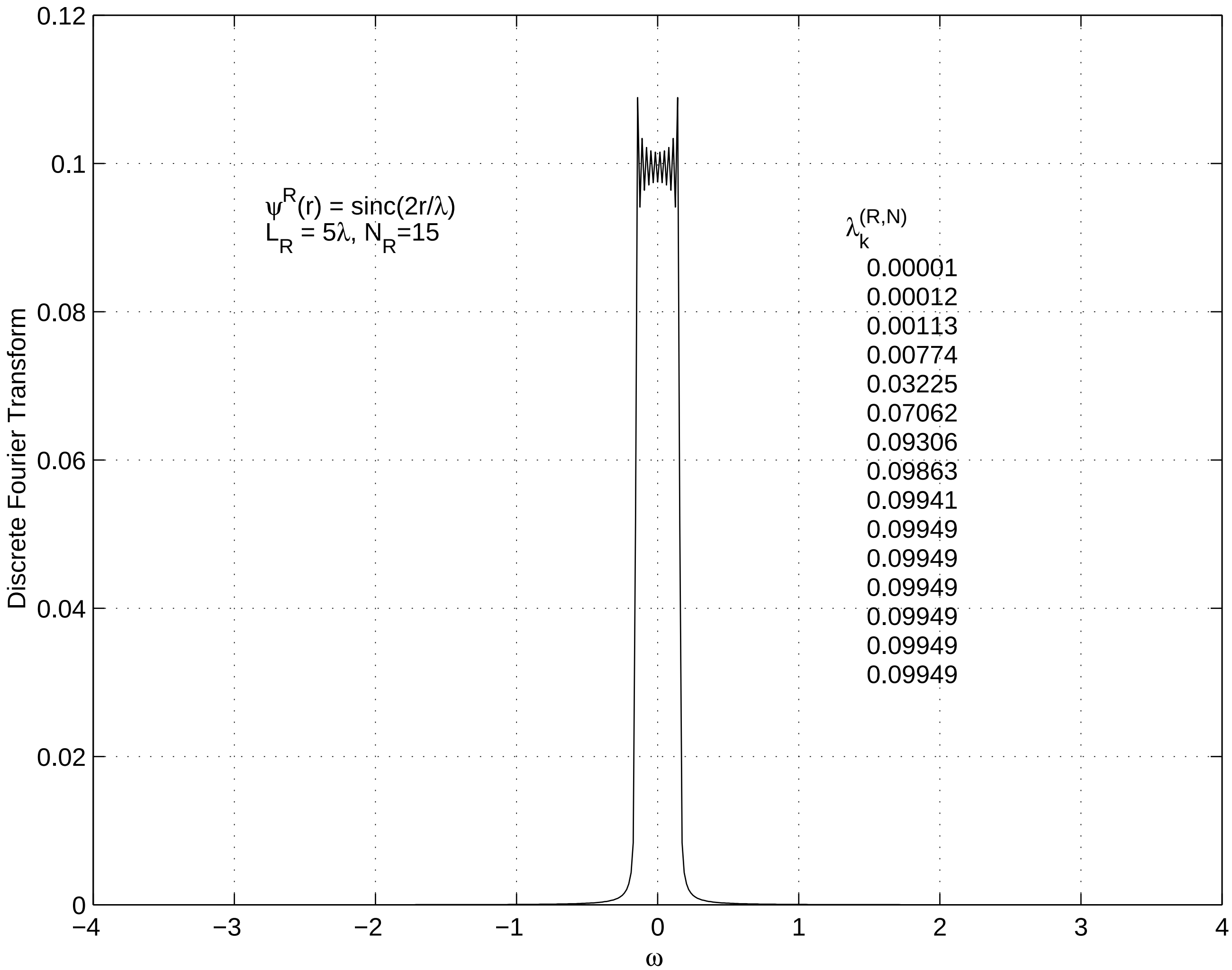

Figure 2: Discrete Fourier transform of the first column and the first row of the covariance matrix Ψ^R/N as determined in (31), where Ψ^R is determined by (6), and $N = 200$, $\psi^R(r) = \text{sinc}\,(2r/\lambda)$, $L_R = 5\lambda$. The number of dominant eigenvalues of Ψ^R/N is $N_R = 15$, and they are listed in increasing order.

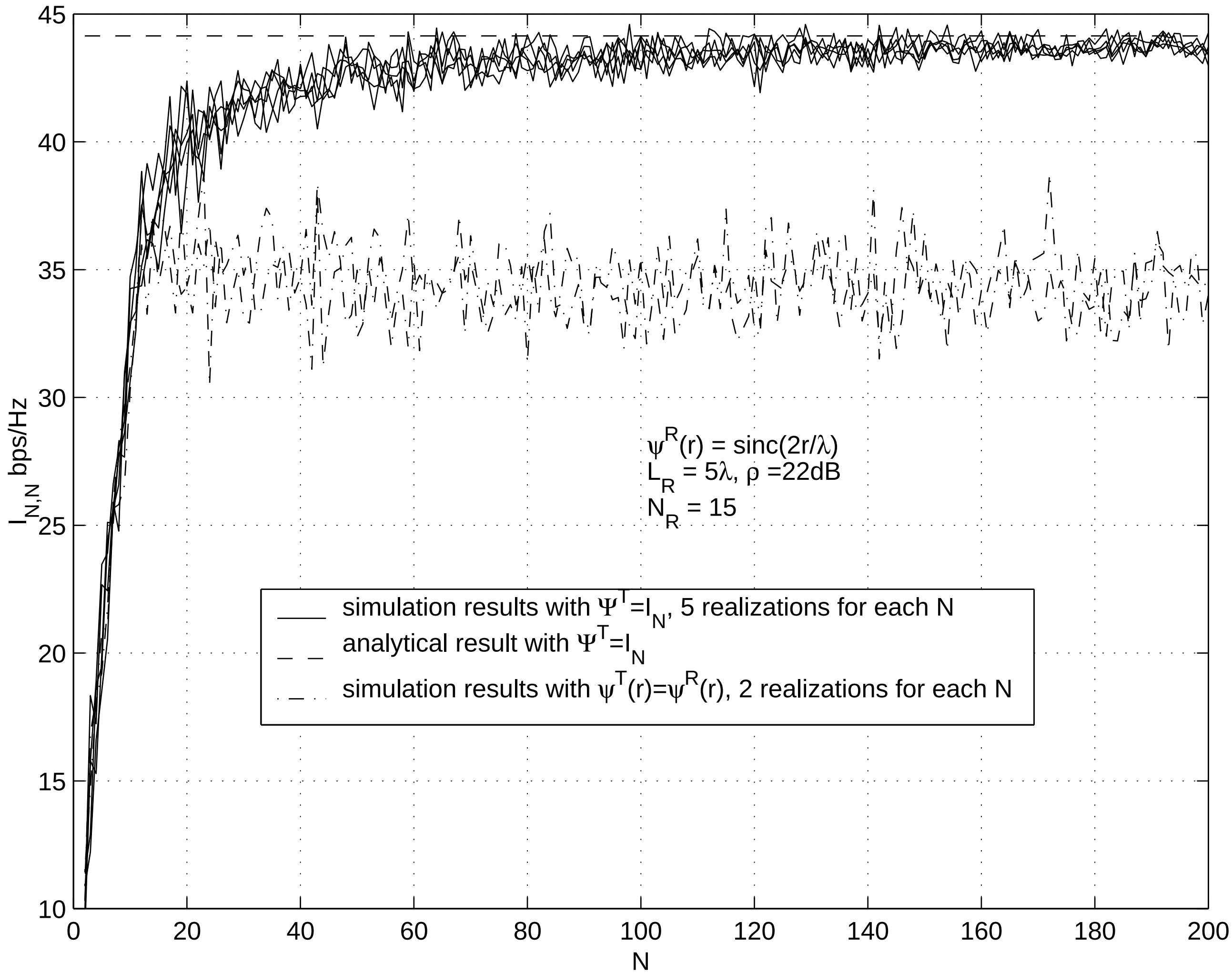

Figure 3: Instantaneous mutual information $I_{N,N}$ of a (N, N) MEA system versus N, with fixed total average received power, such that $Q = \frac{\rho}{N^2} I_N$ in (2). Simulation results are obtained through generating one realization of random variable $I_{N,N}$ for each N, where $L_R = 5\lambda$, $\rho = 22\,\text{dB}$. For the case when the fixed length linear array is put at the receiver side only, 5 realizations are generated for each N, and the Toeplitz matrix Ψ^R is determined by $\psi^R(r) = \text{sinc}\,(2r/\lambda)$, $\Psi^T = I_N$. For the case when the fixed length linear arrays are put at both the transmitter and receiver side, 2 realizations are generated for each N, and $\Psi^R = \Psi^T$ are determined by $\psi^T(r) = \psi^R(r) = \text{sinc}\,(2r/\lambda)$. Analytical results are obtained by the sum term in (21), where $f_R(N)$ is approximated as $N_R = 15$ and eigenvalues $\left\{\lambda_k^{(R,N)}\right\}$ are obtained in Figure 2, and $\sum_{k=0}^{N_R-1} \log_2\left(1 + \rho\lambda_k^{(R,N)}\right) = 44.1497$. It can be observed that for each realization in the case with a fixed length array at the receiver side only, $I_{N,N}$ is converging to the analytical result as N grows large. The case of a fixed length array located at both sides is indicated by the dash-dot lines, and more randomness is observed as expected.

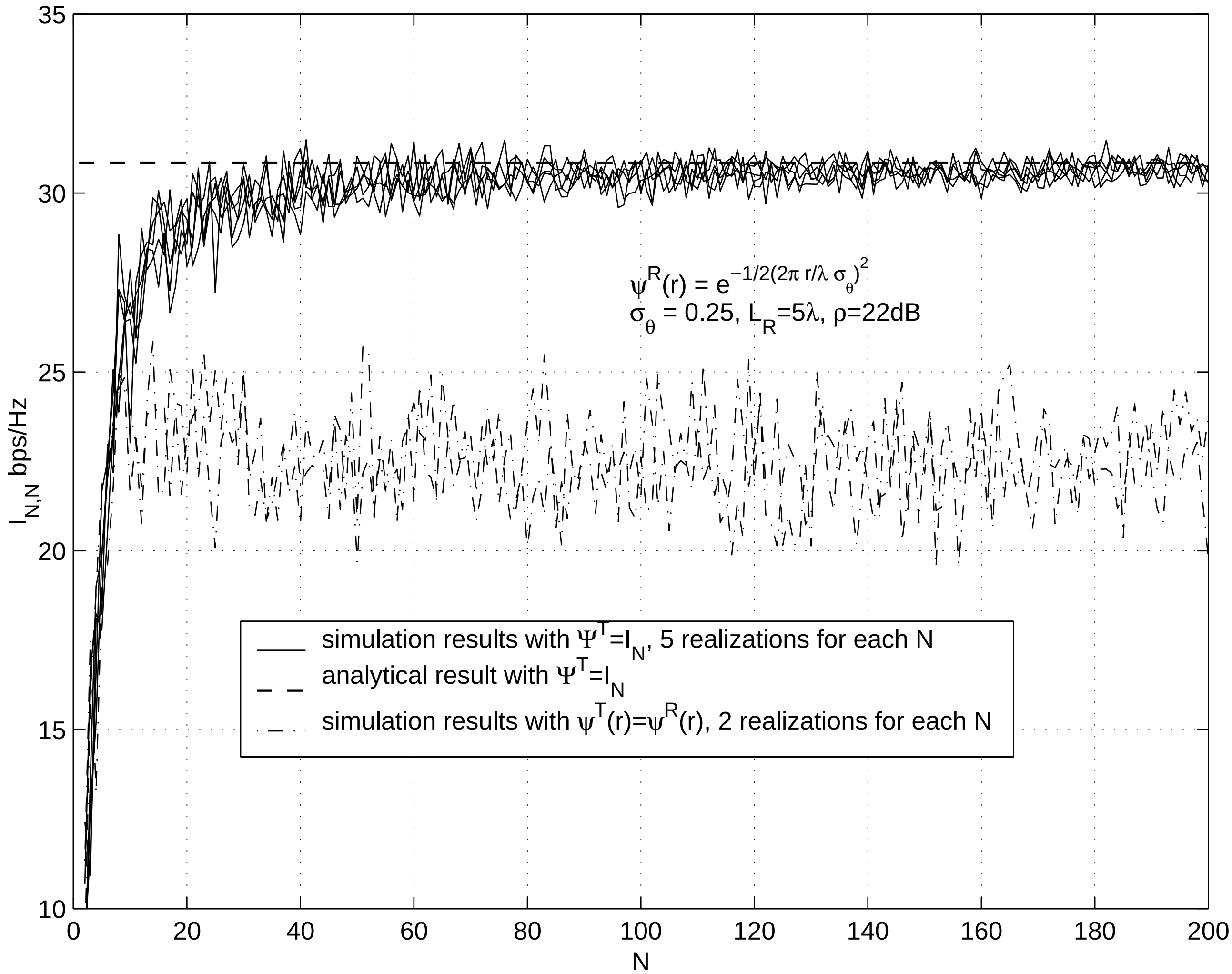

Figure 4: Instantaneous mutual information $I_{N,N}$ of a (N, N) MEA system versus N, with fixed total average received power, such that $Q = \frac{\rho}{N^2} I_N$ in (2). Simulation results are obtained through generating one realization of random variable $I_{N,N}$ for each N, where $L_R = 5\lambda$, $\rho = 22\,\text{dB}$. For the case when the fixed length linear array is put at the receiver side only, 5 realizations are generated for each N, and the Toeplitz matrix Ψ^R is determined by $\psi^R(r) = e^{-1/2(2\pi r \sigma_\theta/\lambda)^2}$, $\Psi^T = I_N$, where $\sigma_\theta = 0.25$. For the case when the fixed length linear arrays are put at both the transmitter and receiver side, 2 realizations are generated for each N, and $\Psi^R = \Psi^T$ are determined by $\psi^T(r) = \psi^R(r) = e^{-1/2(2\pi r \sigma_\theta/\lambda)^2}$. Analytical results are obtained by the first sum term in (21), where the eigenvalues $\left\{\lambda_k^{(R,N)}\right\}$ are obtained through numerical computations, and $\sum_{k=0}^{f_R(200)-1} \log_2\left(1 + \rho\lambda_k^{(R,N)}\right) = 30.8435$. It can be observed that for each realization in the case with a fixed length array at the receiver side only, $I_{N,N}$ is converging to the analytical result as N grows large. The case of a fixed length array located at both sides is indicated by the dash-dot lines, and more randomness is observed as expected.

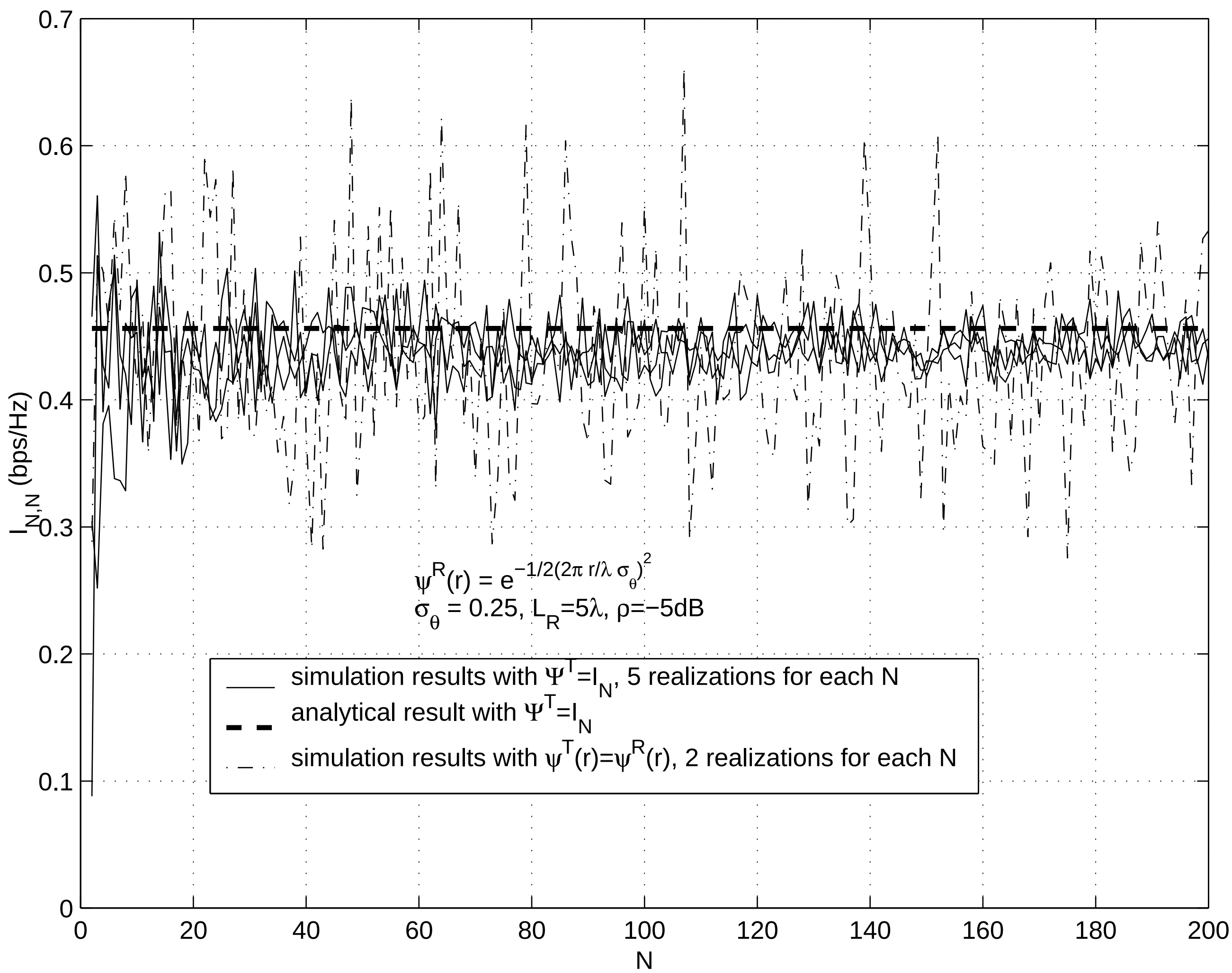

Figure 5: Instantaneous mutual information $I_{N,N}$ of a (N, N) MEA system versus N, with fixed total average received power, such that $Q = \frac{\rho}{N^2} I_N$ in (2). Simulation results are obtained through generating one realization of random variable $I_{N,N}$ for each N, where $L_R = 5\lambda$, $\rho = -5\,\mathrm{dB}$. For the case when the fixed length linear array is put at the receiver side only, 5 realizations are generated for each N, and the Toeplitz matrix Ψ^R is determined by $\psi^R(r) = e^{-1/2(2\pi r \sigma_\theta/\lambda)^2}$, $\Psi^T = I_N$, where $\sigma_\theta = 0.25$. For the case when the fixed length linear arrays are put at both the transmitter and receiver side, 2 realizations are generated for each N, and $\Psi^R = \Psi^T$ are determined by $\psi^T(r) = \psi^R(r) = e^{-1/2(2\pi r \sigma_\theta/\lambda)^2}$. Analytical results are obtained by the sum term in (21), where eigenvalues $\left\{\lambda_k^{(R,N)}\right\}$ are obtained through numerical computations, and $\sum_{k=0}^{f_R(200)-1} \log_2\left(1 + \rho\lambda_k^{(R,N)}\right) = 0.4419$. It can be observed that for each realization in the case with a fixed length array at the receiver side only, $I_{N,N}$ is converging to the analytical result as N grows large. The case of a fixed length array located at both sides is indicated by the dash-dot lines, more randomness is observed as expected.

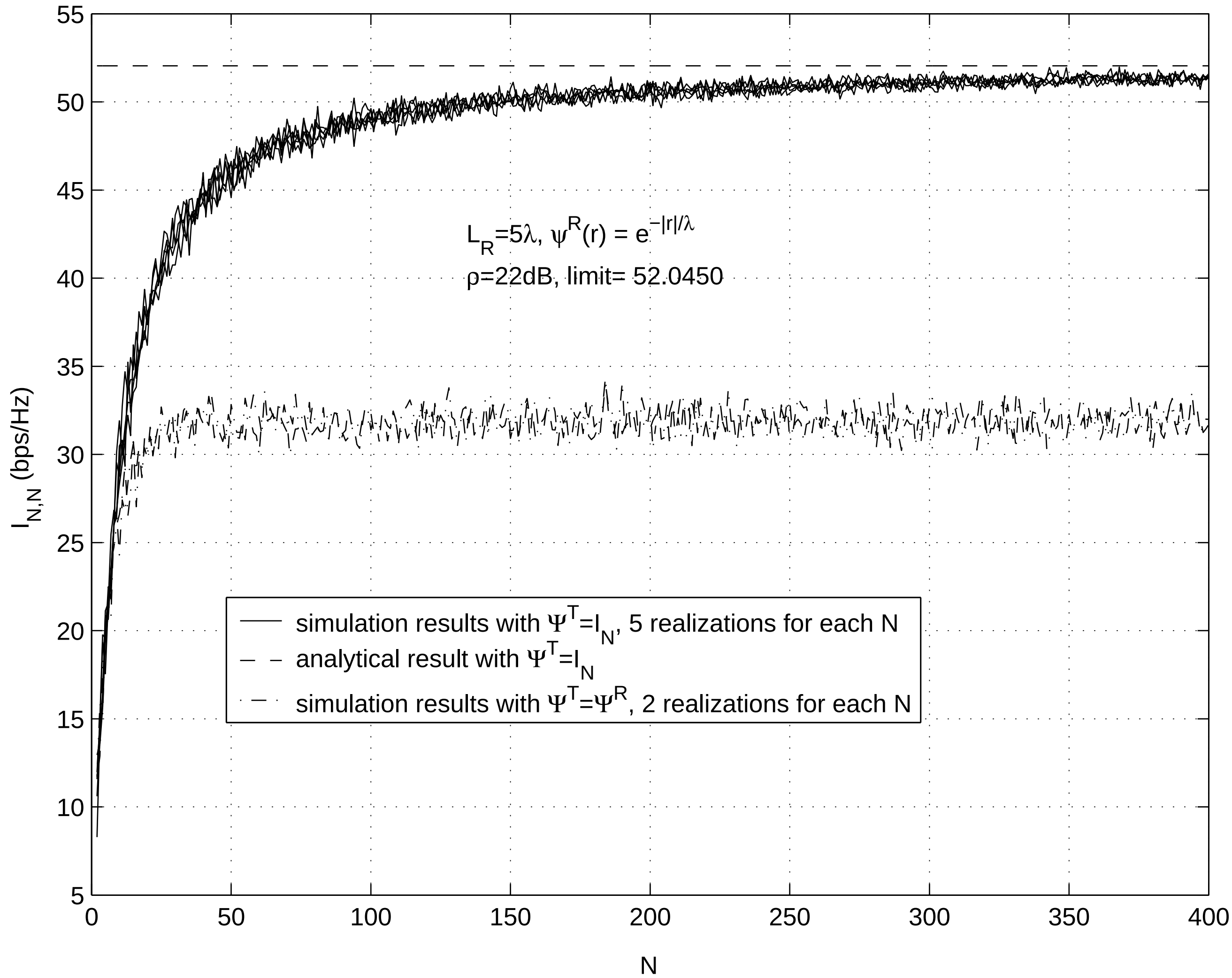

Figure 6: Instantaneous mutual information $I_{N,N}$ of a (N,N) MEA system versus N, with fixed total average receive power, such that $Q = \frac{\rho}{N^2} I_N$ in (2). Simulation results are obtained through generating one realization of random variable $I_{N,N}$ for each N, where $L_R = 5\lambda$, $\rho = 22\,\text{dB}$. For the case when the fixed length linear array is put at the receiver side only, 5 realizations are generated for each N, and the Toeplitz matrix Ψ^R is determined by $\psi^R(r) = e^{-|r|/\lambda}$, $\Psi^T = I_N$. For the case when the fixed length linear arrays are put at both the transmitter and receiver side, 2 realizations are generated for each N, and $\Psi^R = \Psi^T$ are determined by $\psi^T(r) = \psi^R(r) = e^{-|r|/\lambda}$. Analytical results are obtained by the sum term in (21), with $\sum_{k=0}^{N-1} \log_2 \left(1 + \rho\lambda_k^{(R,N)}\right) = 52.0460$. It can be observed that for each realization in the case with a fixed length array at the receiver side only, $I_{N,N}$ is converging to the analytical result as N grows large. The case of a fixed length array located at both sides is indicated by the dash-dot lines, and more randomness is observed as expected.

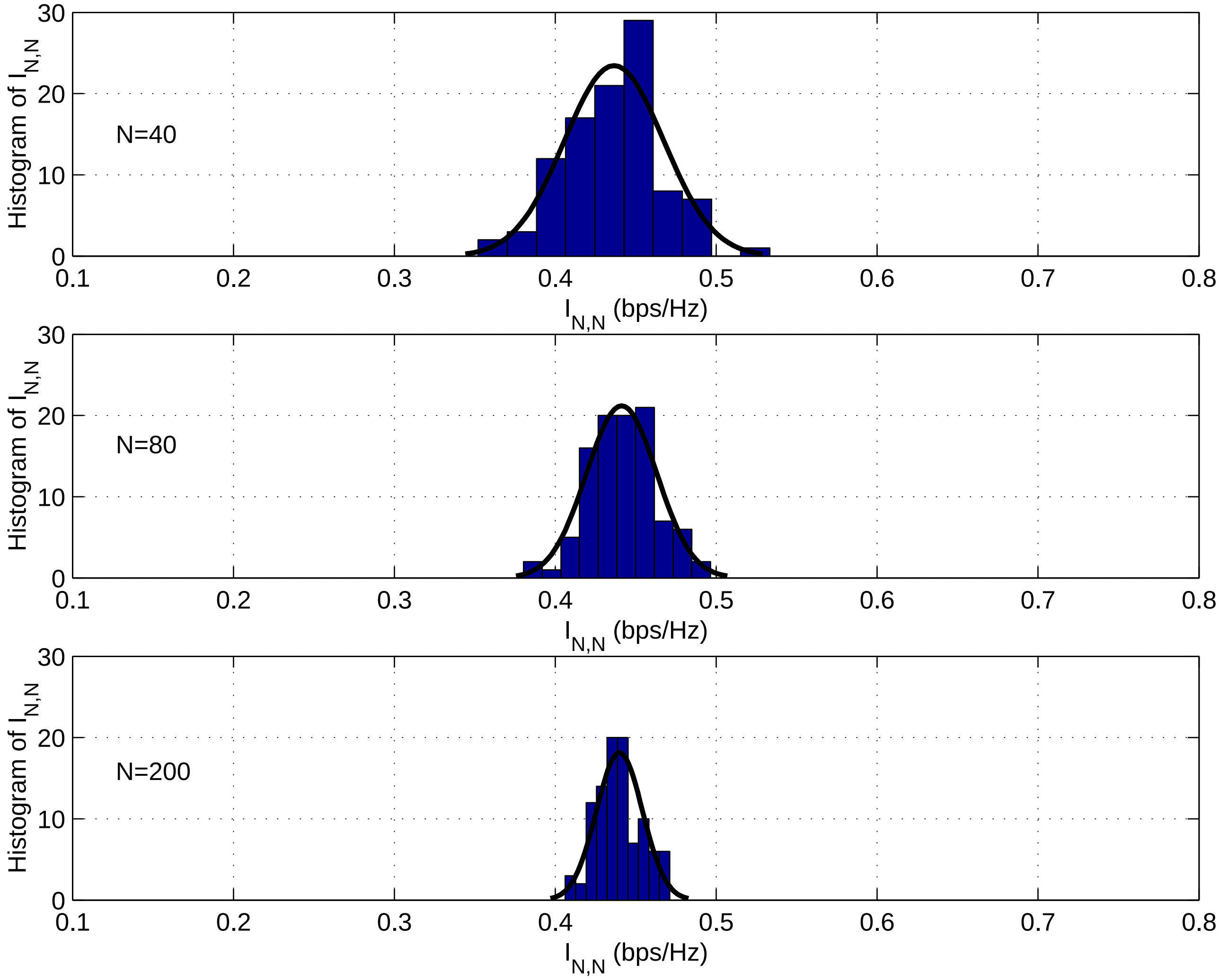

Figure 7: With the same parameters as those in Fig. 5, the three figures shown above are the histograms of $I_{N,N}$, when the total received signal power is fixed, and the fixed length array is put at the receiver side only. The histograms are obtained through generating 100 independent $I_{N,N}$ samples for $N = 40, 80, 200$, respectively. It can be observed that as N is becoming larger, the probability density function (PDF) of $I_{N,N}$ is becoming more concentrated around the constant 0.4419.

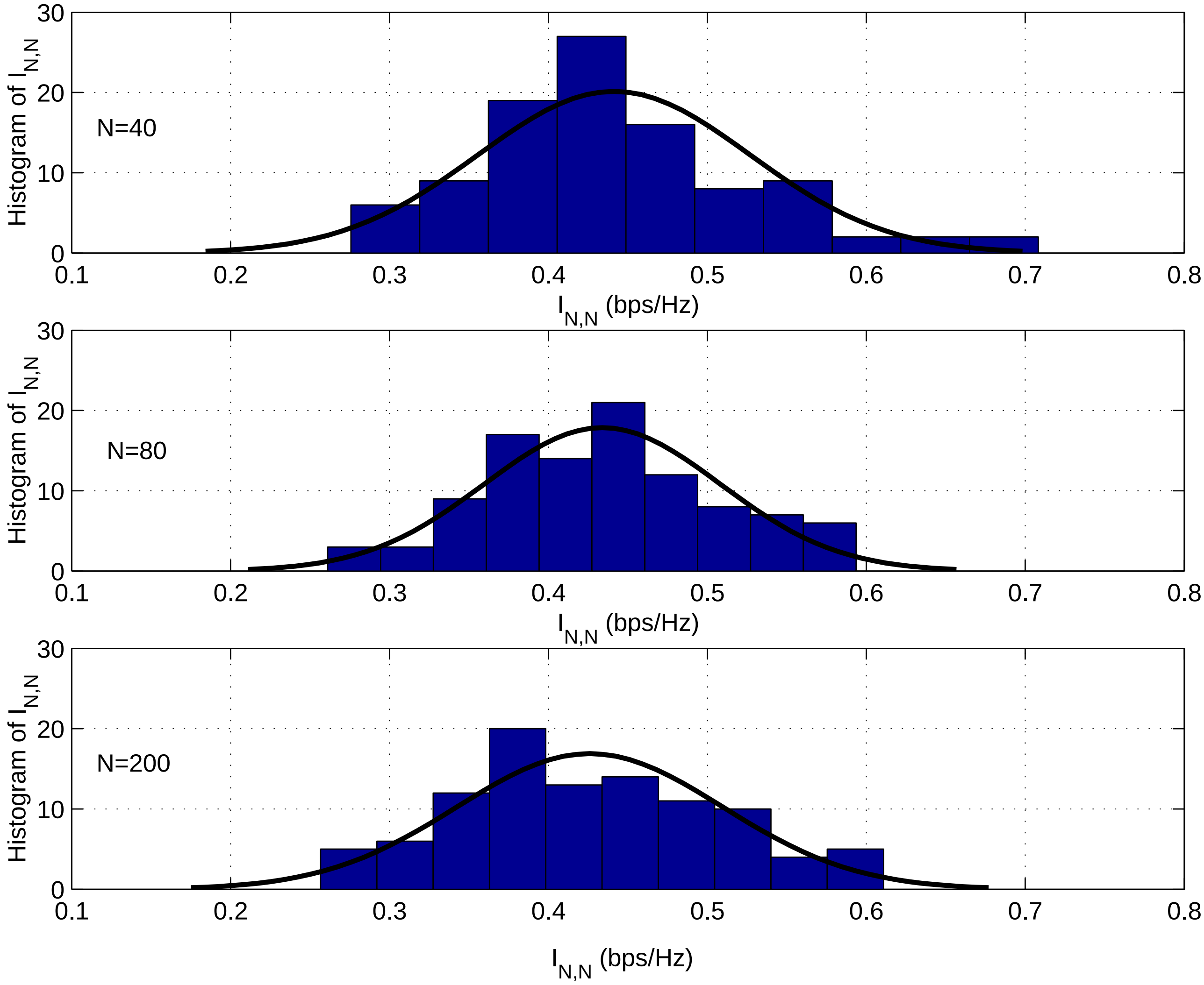

Figure 8: With the same parameters as those in Fig. 5, the three figures shown above are the histograms of $I_{N,N}$, when the total received signal power is fixed, and the fixed length array is put at both the transmitter and the receiver side. The histograms are obtained through generating 100 independent $I_{N,N}$ samples for $N = 40, 80, 200$, respectively. It can be observed that as N is becoming larger, there does not exist the tendency of the more concentration of the PDF of $I_{N,N}$, as expected.

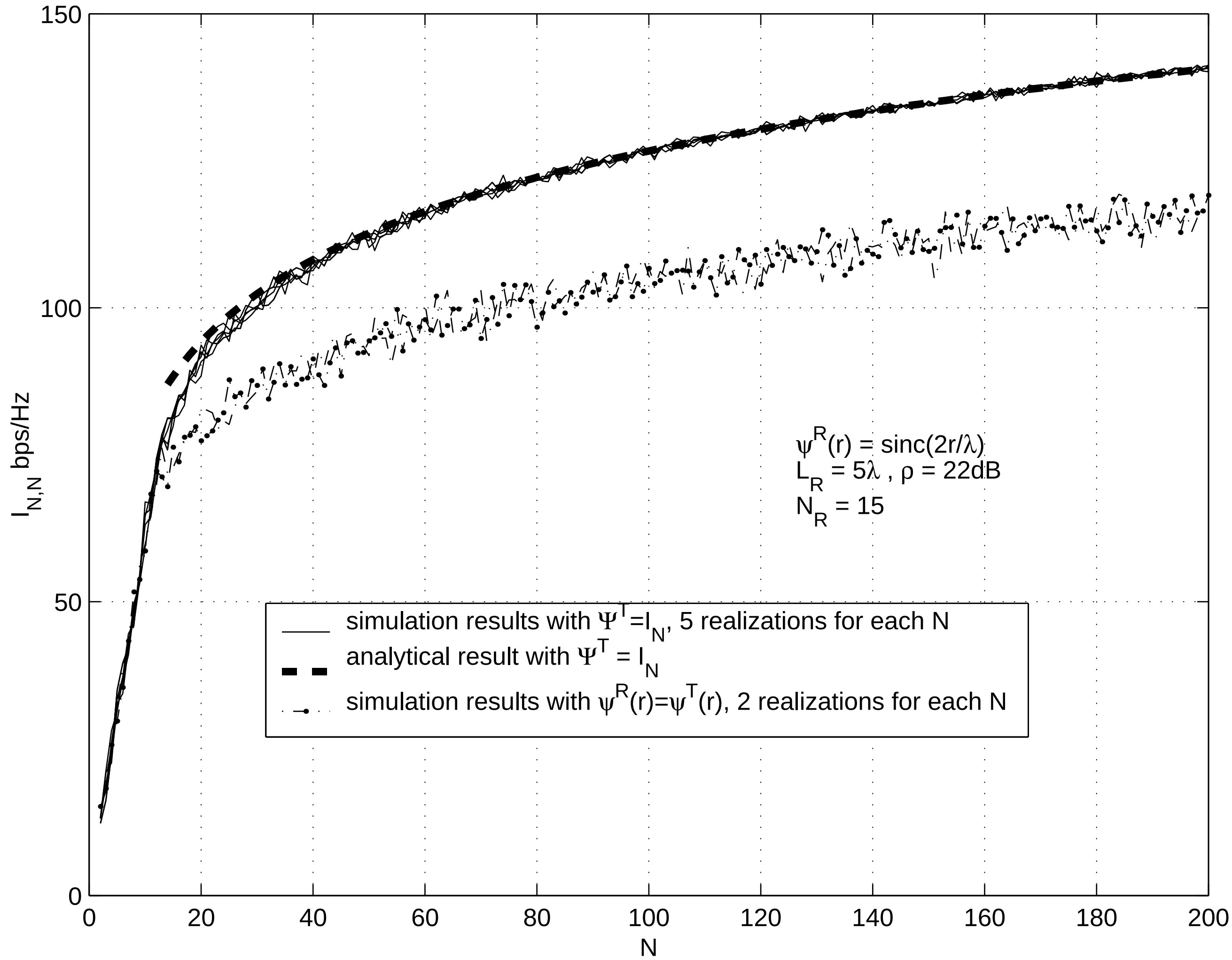

Figure 9: Instantaneous mutual information $I_{N,N}$ of a (N, N) MEA system versus N, with the fixed total average transmit power. Simulation results are obtained through generating one realization of random variable $I_{N,N}$ for each N in terms of the first equation in (5), where $L_R = 5\lambda$, $\rho = 22$ dB. For the case when the fixed length linear array is put at the receiver side only, 5 realizations are generated for each N, and the Toeplitz matrix Ψ^R is determined by $\psi^R(r) = \text{sinc}\,(2r/\lambda)$, $\Psi^T = I_N$. For the case when the fixed length linear arrays are put at both the transmitter and receiver side, 2 realizations are generated for each N, and $\Psi^R = \Psi^T$ are determined by $\psi^T(r) = \psi^R(r) = \text{sinc}\,(2r/\lambda)$. Analytical results are obtained by the sum term in (30), where $f_R(N)$ is approximated as $N_R = 15$ and eigenvalues $\left\{\lambda_k^{(R,N)}\right\}$ are obtained from Figure 2. It can be observed that for each realization in the case with a fixed length array at the receiver side only, $I_{N,N}$ is converging rapidly for large N, and the upper-bound in (25) is tight. The case of a fixed length array located at both sides is indicated by the dash-dot lines, and more randomness is observed as expected.

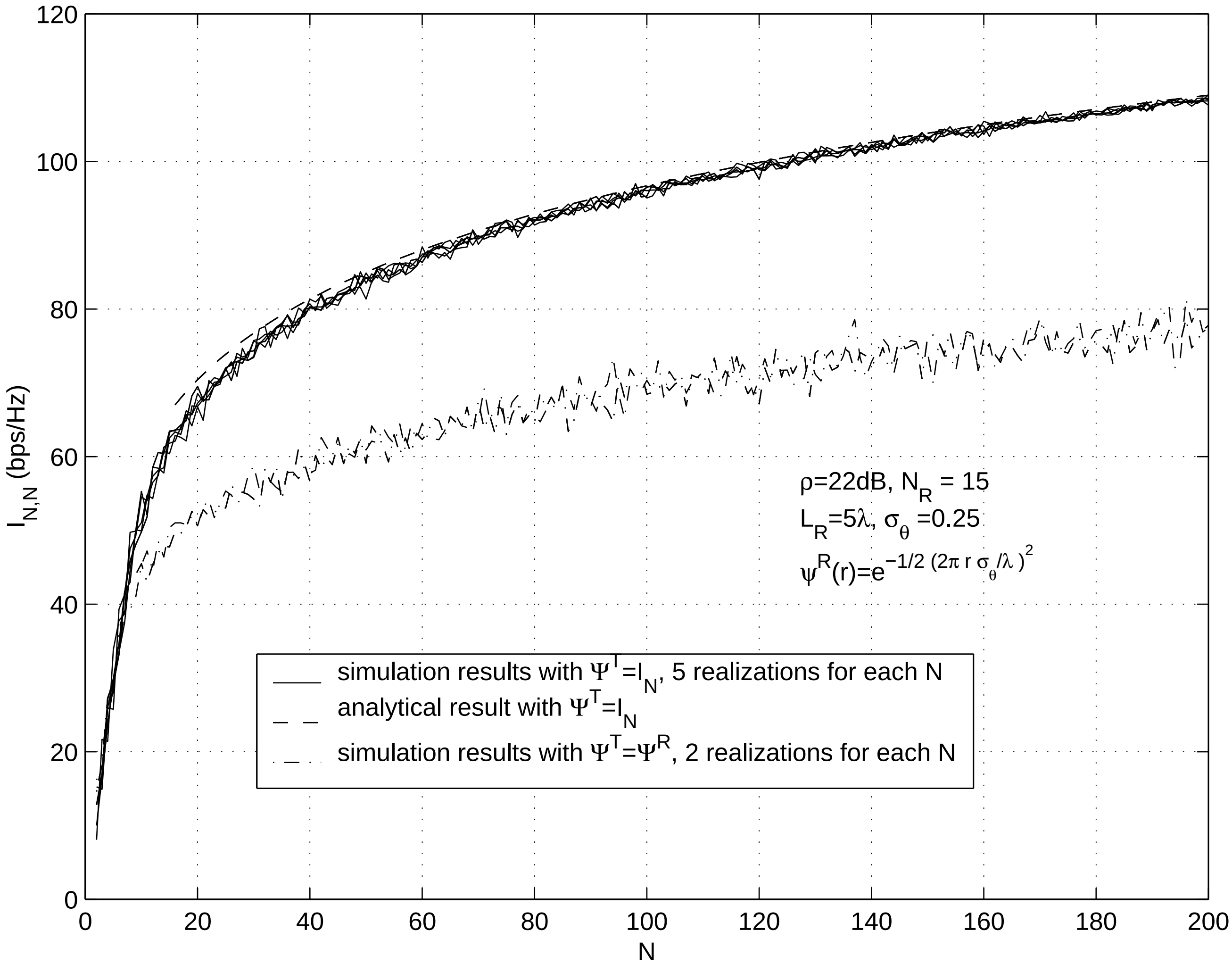

Figure 10: Instantaneous mutual information $I_{N,N}$ of a (N, N) MEA system versus N, with fixed total average transmit power, such that $Q = \frac{\rho}{N} I_N$ in (2). Simulation results are obtained through generating one realization of random variable $I_{N,N}$ for each N, where $L_R = 5\lambda$, $\rho = 22$ dB. For the case when the fixed length linear array is put at the receiver side only, 5 realizations are generated for each N, and the Toeplitz matrix Ψ^R is determined by $\psi^R(r) = e^{-1/2(2\pi r\sigma_\theta/\lambda)^2}$, $\Psi^T = I_N$, where $\sigma_\theta = 0.25$. For the case when the fixed length linear arrays are put at both the transmitter and receiver side, 2 realizations are generated for each N, and $\Psi^R = \Psi^T$ are determined by $\psi^T(r) = \psi^R(r) = e^{-1/2(2\pi r\sigma_\theta/\lambda)^2}$. Analytical results are obtained by the sum term in (30), where $f_R(N)$ is approximated as $N_R = 15$. It can be observed that for each realization in the case with a fixed length array at the receiver side only, $I_{N,N}$ is well approximated by the analytical result, and the upper-bound in (25) is tight, as N grows large. The case of a fixed length array located at both sides is indicated by the dash-dot lines, and more randomness is observed as expected.

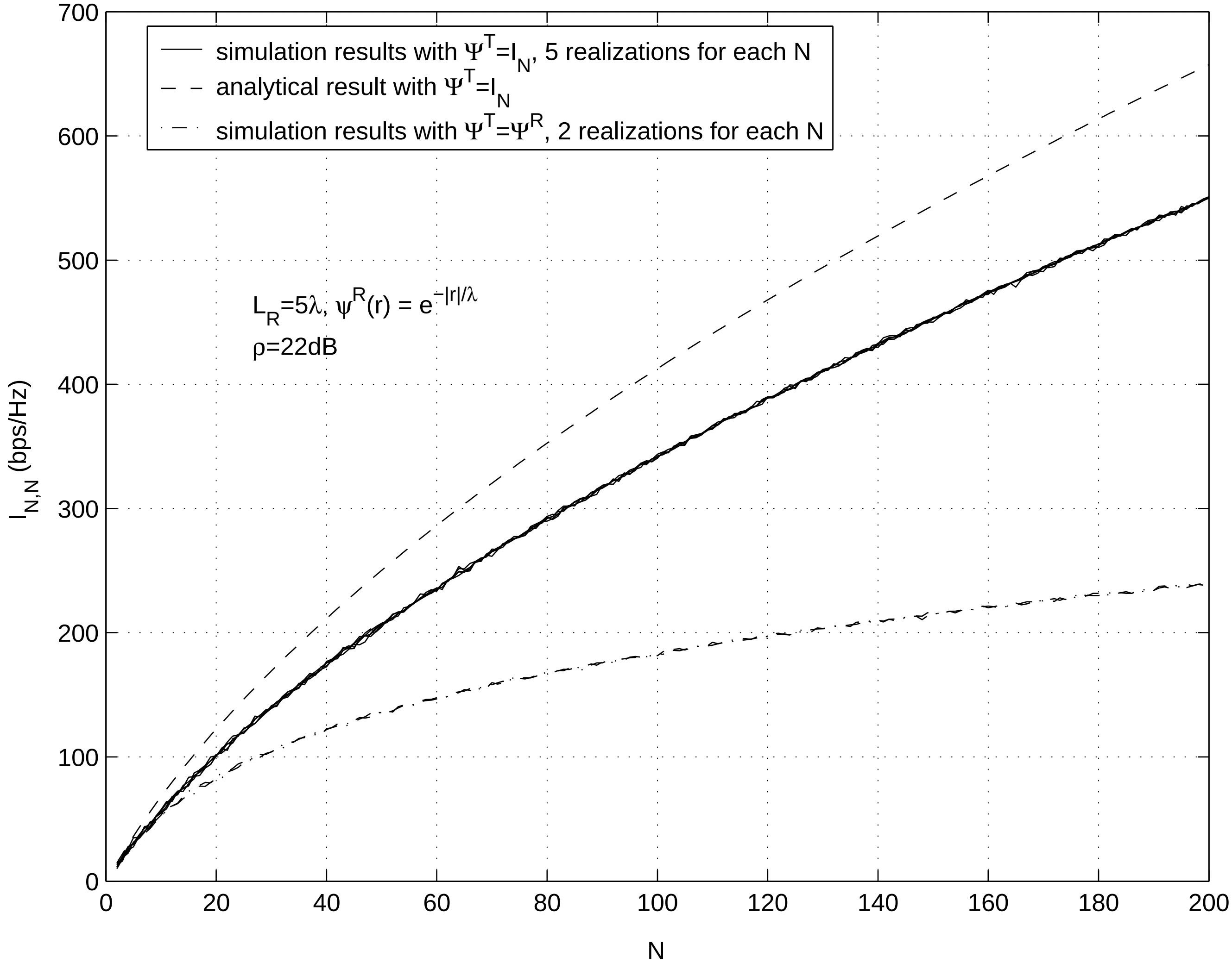

Figure 11: Instantaneous mutual information $I_{N,N}$ of a (N, N) MEA system versus N, with fixed total average transmit power, such that $Q = \frac{\rho}{N} I_N$ in (2). Simulation results are obtained through generating one realization of random variable $I_{N,N}$ for each N, where $L_R = 5\lambda$, $\rho = 22$ dB. For the case when the fixed length linear array is put at the receiver side only, 5 realizations are generated for each N, and the Toeplitz matrix Ψ^R is determined by $\psi^R(r) = e^{-|r|/\lambda}$, $\Psi^T = I_N$. For the case when the fixed length linear arrays are put at both the transmitter and receiver side, 2 realizations are generated for each N, and $\Psi^R = \Psi^T$ are determined by $\psi^T(r) = \psi^R(r) = e^{-|r|/\lambda}$. Analytical results are obtained by the sum term in (25), where $f_R(N)$ can no longer be approximated as N_0 because of the slow decreasing rate of $\lambda_k^{(R,N)}$. It can be observed that for each realization in the case with a fixed length array at the receiver side only, $I_{N,N}$ cannot be approximated by the analytical result, and the upper-bound in (25) is becoming looser, as N grows large. The case of a fixed length array located at both sides is indicated by the dash-dot lines, which agrees with what is expected based on (27).