Capacity Scaling in MIMO Wireless Systems Under Correlated Fading

Chen-Nee Chuah, *Student Member, IEEE*, David N. C. Tse, *Member, IEEE*, Joseph M. Kahn, *Fellow, IEEE*, and Reinaldo A. Valenzuela, *Fellow, IEEE*

***Abstract*—Previous studies have shown that single-user systems employing n-element antenna arrays at both the transmitter and the receiver can achieve a capacity proportional to n, assuming *independent* Rayleigh fading between antenna pairs. In this paper, we explore the capacity of dual-antenna-array systems under *correlated fading* via theoretical analysis and ray-tracing simulations. We derive and compare expressions for the asymptotic growth rate of capacity with n antennas for both independent and correlated fading cases; the latter is derived under some assumptions about the scaling of the fading correlation structure. In both cases, the theoretic capacity growth is linear in n but the growth rate is 10–20% smaller in the presence of correlated fading. We analyze our assumption of separable transmit/receive correlations via simulations based on a ray-tracing propagation model. Results show that empirical capacities converge to the limit capacity predicted from our asymptotic theory even at moderate $n = 16$. We present results for both the cases when the transmitter does and does not know the channel realization.**

***Index Terms*—Asymptotic capacity growth, correlated fading, multiantenna arrays, multiple-input–multiple-output (MIMO) systems, ray tracing.**

I. Introduction

In response to the demand for higher bit rates in wireless local-area networks (LANs), researchers have explored the use of multiple-element arrays (MEAs) at both the transmitter and the receiver. Signals propagating through the wireless channel experience path loss and distortion due to multipath fading and additive noise. These impairments, along with the constraints of power and bandwidth, limit the system capacity. In the past, multiple antennas have been used at the receiver to combat multipath fading, e.g., using maximal-ratio combining [1], or to suppress interfering signals, e.g., using optimal combining [2]. Recent studies report that in single-user, point-to-point links, using MEAs at both transmitter and receiver increases the capacity significantly over single-antenna systems [3], [4].

Manuscript received June 13, 2000; revised June 3, 2001. The material in this paper was presented in part at IEEE GLOBECOM'98, Sydney, Australia, November 8–12, 1998.

C.-N. Chuah is with the Department of Electrical and Computer Engineering, University of California, Davis, Davis, CA 95616 USA (e-mail: chuah@ece.ucdavis.edu).

D. N. C. Tse and J. M. Kahn are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720-1770 USA (e-mail: dtse@eecs.berkeley.edu; jmk@eecs.berkeley.edu).

R. A. Valenzuela is with Bell Laboratories, HOH R-255, Holmdel, NJ 07733 USA (e-mail: rav@dnrc.bell-labs.com).

Communicated by M. L. Honig, Associate Editor for Communications.

Publisher Item Identifier S 0018-9448(02)00635-1.

Foschini and Gans have analyzed the information-theoretic capacity of MEA systems in a narrow-band Rayleigh-fading environment [3]. They consider independent and identically distributed (i.i.d.) fading at different antenna elements, and assume that the transmitter does not know the channel. With n transmitting and n receiving antennas, the MEA mutual information with equal-power allocation I_n is reported to grow linearly with n for a given fixed average transmitter power. An MEA system achieves almost n more bits per hertz for every 3-dB increase in signal-to-noise ratio (SNR) at high SNR, compared to the single-antenna case, which only achieves one additional bit per hertz for every 3-dB increase in SNR.

In practice, correlation exists between the signals transmitted by or received at different antenna elements. Correlation can arise if the elements are not spaced sufficiently far apart. For example, Lee pointed out in [5] that in order to obtain a correlation coefficient at adjacent elements less than 0.7, the elements must be spaced by about 15–20 wavelengths in the broadside case and 70 wavelengths in the inline case. The presence of a dominant line-of-sight component can also affect the MEA capacities. It is important to understand the impact of these factors on MEA system capacity.

The goal of this paper is to explore the capacities of single-user MEA systems in a more realistic propagation environment, where the fading is correlated. We consider the performance in two scenarios: 1) the transmitter knows the channel, so that optimal transmit power allocation (also known as water filling) can be used; 2) the transmitter does not know the channel, so that equal power is allocated to each of the transmit antenna elements. In both cases, it is assumed that the receiver knows the channel perfectly. We study the behavior of MEA capacities through analysis and simulation.

The multiple-input–multiple-output (MIMO) fading channel is modeled as a random matrix H. The water-filling capacity C_n and the mutual information under equal power allocation I_n of a n by n system are random variables, being functions of the singular values of the random H. We construct a fading correlation structure assuming separable transmit/receive correlations, and derive the large system limiting distribution of the singular values of H in two cases: a) when the fades between different antenna pairs are independent and b) when these fades are correlated. Using these results, we show almost-sure convergence of the asymptotic growth rate C_n/n and I_n/n. In both the independent and correlated fading cases, the capacity and mutual information grow linearly with n but the growth rate is different when the fades are correlated. In particular, we show that under correlated fading, the growth rate of I_n is smaller at all SNRs

compared to the independent fading case, while the growth rate of C_n is smaller at high SNR but larger at low SNR.

Our hypothesized fading correlation structure is studied carefully via simulation based on a ray-tracing propagation model. We use the WiSE (Wireless System Engineering) [6] software tool to model explicitly the channel response between a transmitter and a receiver placed inside an office building. Comparing the empirical capacity distribution with the asymptotic theory, reasonable agreement is found even for moderate $n \leq 16$. Initial results can be found in [7] and [8]. We also quantify the capacity improvements achieved by water filling over the equal power strategy empirically at different SNR levels.

An alternative approach to ray-tracing simulations is to use scattering models [9], [12] to characterize the spatial fading correlations. In independent work, Shiu *et al.* quantify the effect of fading correlations on MEA capacity in [13] by employing such an abstract scattering model.

The remainder of this paper is organized as follows. In Section II, we model the channel as a MIMO system with flat frequency response. Using this mathematical model, we define information-theoretic capacity and mutual information of MEA systems in Section III, and analyze their asymptotic behavior as $n \to \infty$ in Section IV. In Section V, we present capacity estimates for the simulated channels and discuss the discrepancies between these results and the asymptotic capacities predicted by theory. We briefly describe how WiSE is used to model the indoor propagation environment that our numerical analysis is based on. We also include details about placements of transmitting and receiving MEAs, arrangement of antennas in an array, and basic assumptions about the antenna elements. Conclusions are presented in Section VI.

To simplify notations, we will focus exclusively on the case when the number of transmit antenna is equal to the number of receive antenna (n by n systems). The extension of the analytical results to the case with unequal number of transmit and receive antennas is straightforward.

II. Channel Model

The following notation is used throughout the paper: $'$ for vector transpose, $\dagger$ for transpose conjugate, $I_{n\times n}$ for the identity matrix, $E[\cdot]$ for expectation, and underline for vectors. All logarithms are with respect to base 2.

We consider a single-user,[1] point-to-point communication channel with n transmitting and n receiving antenna elements, denoted as an (n, n)-MEA system. We assume that the transmitted signal occupies a bandwidth W, over which the channel frequency response is essentially constant. For this assumption to be valid, W must be much smaller than the channel coherence bandwidth, which is approximately the reciprocal of the channel delay spread.[2] Since the maximum delay spread of our channels is about 25 ns, we require that W be much less than 40 MHz. Assuming zero excess bandwidth, this requires a symbol rate much less than 40 Mbaud.

[1] n transmitting antennas are colocated, and so are the receiving antennas.

[2] Here, delay spread refers to the difference in arrival times of the earliest and latest strong rays.

For the remaining analysis and discussions, we assume that the channel is linear and time-invariant and use the following discrete-time equivalent model:

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

$\underline{X} = [x_1, x_2, \ldots, x_T]'$ 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 $n \times 1$ vectors, $\underline{Y}$ and $\underline{Z}$, respectively, where y_i and z_i represent the signal and noise received at the ith antenna. The complex path gain between transmitter j and receiver i is represented by $\{H_{ij}\colon i, j = 1, 2, \ldots, n\}$.

We further assume the following.

- The total average power (sum over all transmitting antennas) is P_{tot}, regardless of n.
- The noise vector $\underline{Z}$ is an additive white complex Gaussian random vector, whose entries $\{Z_i, i = 1, 2, \ldots, n\}$are i.i.d. circularly symmetric complex Gaussian random variables with variance

$$E[|Z|^2] = N_{\text{o}}W$$

  where W is the signal bandwidth.

We consider the following two cases.

1) H is known only to the receiver but not the transmitter. Power is distributed equally over all transmitting antennas in this case.
2) H is known at the transmitter and receiver, so that power allocation can be optimized to maximize the achievable rate over the channel.

In this work, we treat H as quasi-static. H is considered fixed for the whole duration of communication, so that the capacity is computed for each realization of H without time averaging. On the other hand, H changes if the receiver is moved from one place to the other, and we assume this will happen over a time scale much longer than the duration of communication. The associated capacity and mutual information C_n and I_n for each specific realization of H can be viewed as random variables. We are interested in studying the statistics of these random variables, in particular, the averages, $\overline{C}_n$ and $\overline{I}_n$ and the values at 5% channel outage, $C_n^{0.05}$ and $I_n^{0.05}$.

III. MEA Capacity and Mutual Information

Channel capacity is defined as the highest rate at which information can be sent with arbitrarily low probability of error. Since the channel H is considered quasi-static, it is reasonable to associate the capacity to a specific realization of H, given a fixed average total power P_{tot} and noise variance $N_{\text{o}}W$ (see Section II for channel model and assumptions). Throughout our analysis, we assume that $\{H_{ij}\colon i, j = 1, 2, \ldots, n\}$ are identically distributed with the variance normalized to be 1. Therefore, the average received SNR is defined as

$$\rho = \frac{P_{\text{tot}}}{N_{\text{o}}W}.$$

When n antennas are used, we denote the MEA capacity with water filling and mutual information with equal-power allocation as C_n and I_n, respectively. For the case with $n = 1$, the Shannon capacity is

$$C_1 = I_1 = \log\left(1 + \rho|H_{11}|^2\right) \text{ b/s/Hz.} \tag{2}$$

In the high-SNR regime, each 3-dB increase of ρ yields a capacity increase of 1 b/s/Hz.

A. Capacity With Water-Filling Power Allocation

In this subsection, we derive the MEA capacity C_n assuming the transmitter has perfect knowledge about the channel. With this knowledge of the channel, the total transmit power can be allocated in the most efficient way over the different transmitters to achieve the highest possible bit rate. Based on the model in Section II and definitions in [14], the MEA capacity with optimal power allocation is

$$C_n = \max_{Q} \log\det\left[I_{n\times n} + \rho HQH^{\dagger}\right] \text{ b/s/Hz} \tag{3}$$

where Q is the $n \times n$ covariance matrix of $\underline{X}$ and Q must satisfy the average power constraint

$$tr(Q) = \sum_{i=1}^{n} E\left[|x_i|^2\right] \le P_{\text{tot}}. \tag{4}$$

The optimal solution is

$$C_n = \sum_{i=1}^{n} \log(\lambda_i \mu)^{+} \tag{5}$$

where μ satisfies

$$\sum_{i} \left(\mu - \frac{1}{\lambda_i}\right)^{+} = \rho \tag{6}$$

and the λ_i are the eigenvalues of $HH^{\dagger}$.

The optimal solutions given in (5) and (6) are analogous to the optimal power allocation calculated through the water-filling algorithm for parallel Gaussian channels [14]. Intuitively, (5) and (6) suggest that the original MIMO channel can be decomposed into n parallel independent subchannels, and we allocate more power to the subchannels with higher SNR $\rho\lambda_i$. Here, μ is the "water level" that marks the height of the power that is poured into the "water vessel" formed by the function $\{1/\lambda_i, i = 1, 2, \ldots, n\}$. Each of these subchannels contributes to the total capacity through $\log 2(\lambda_i\mu)^{+}$. If $\lambda_i\mu \gg 1$, we say that this subchannel provides an effective mode of transmission and is called a *strong eigenmode*.

B. Mutual Information With Equal-Power Allocation

In this case, we assume that equal power is radiated from each transmitting antenna, which is a natural thing to do when the transmitter does not know the channel. The mutual information of (n, n)-MEAs with equal-power allocation is

$$I_n = \log\det\left[I_{n\times n} + \frac{\rho}{n} HH^{\dagger}\right] \text{ b/s/Hz.} \tag{7}$$

Applying singular-value decomposition to H, we can write (7) as

$$I_n = \sum_{i=1}^{n} \log\left(1 + \frac{\rho\lambda_i}{n}\right).$$

IV. Asymptotic Analysis

The capacity C_n and mutual information I_n depend on H, which is random in a fading environment. We analyze the asymptotic behavior of I_n and C_n as $n \to \infty$ for two cases: a) when the H_{ij} are independent, and b) when the H_{ij} are correlated. Our analysis is based on the channel model and properties described in Sections II and III. In all cases, we normalize $E[|H_{ij}|^2] = 1$ for all i, j. For clarity, let us use $I_n(H)$ and $C_n(H)$ to explicitly denote the dependency on H.

A. Independent Fading

We first assume that the path gains H_{ij} are i.i.d. for all i and j. We scale up the size of the MEA by letting n grow large. For each n, let F_n be the empirical distribution of the eigenvalues of $HH^{\dagger}$, i.e., for each λ

$$F_n(\lambda) := \frac{1}{n}\left|\left\{i\colon \lambda_t^{(i)} \le \lambda\right\}\right|$$

the fraction of squared singular values of H less than or equal to λ. Note that since the singular values are random, so is the empirical distribution. An important observation is that, from the expressions (5)–(7), both the capacity and the mutual information under equal-power allocation depends on H only through the empirical distribution of the eigenvalues. The asymptotic properties of the random variables $C_n(H)$ and $I_n(H)$ hinge on how the (random) empirical distribution of the singular values behaves as $n \to \infty$. We have the following theorem (see, e.g., [15]).

Theorem IV.1: Define $G_n(\lambda) := F_n(n\lambda)$. Then almost surely, G_n converges in distribution to a limit G^*, which has a density given by

$$g^*(\lambda) = \begin{cases} \frac{1}{\pi}\sqrt{\frac{1}{\lambda} - \frac{1}{4}}, & 0 \le \lambda \le 4 \\ 0, & \text{else.} \end{cases} \tag{8}$$

Moreover, if $\lambda_{\max}(HH^{\dagger})$ is the largest eigenvalue of $HH^{\dagger}$, then almost surely

$$\lim_{n\to\infty} \frac{\lambda_{\max}(HH^{\dagger})}{n} = 4.$$

This result says several interesting things. First, the scaling by n in the definition of G_n means that the eigenvalues are growing of the order of n. After rescaling, the random distribution converges to a *deterministic* limiting distribution, i.e., for large n, the empirical distribution of the eigenvalues looks similar for almost all realizations of H. Moreover, the limit does not depend on the distribution of the entries H_{ij}.

The asymptotic behavior of the mutual information $I_n(H)$ follows directly from this proposition

$$\frac{I_n(H)}{n} = \frac{1}{n}\sum_{i=1}^{n} \log\left(1+\frac{\rho}{n}\lambda_i\right) \to \int_0^4 \log(1+\rho\lambda) g^*(\lambda)\, d\lambda$$

where the convergence is almost surely. This observation was previously made by Foschini [3]. The integral can actually be computed in closed form, as was done in [16] in the context of a related capacity analysis problem for randomly spread code-division multiple-access (CDMA) systems.

$$\begin{aligned} I^*(\rho) &\equiv \int_0^4 \log(1+\rho\lambda) g^*(\lambda)\, d\lambda \\ &= 2\log\left(1+\sqrt{4\rho+1}\right) - \frac{\log e}{4\rho}\left(\sqrt{4\rho+1}-1\right)^2. \end{aligned} \tag{9}$$

We now turn to the water-filling capacity. By relabeling the parameter μ as μ_n/n we can rewrite (5) and (6) as

$$\frac{C_n(H)}{n} = \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{\lambda_i}{n}\mu_n\right)^+$$

where μ_n satisfies

$$\frac{1}{n}\sum_{i=1}^{n}\left(\mu_n - \frac{n}{\lambda_i}\right)^+ = \rho. \tag{10}$$

As $n \to \infty$, the empirical distribution of λ_i/n converges almost surely to a limit with density g^*. From (10), we see that μ_n converges to μ^* satisfying the equation

$$\int_0^4 \left(\mu^* - \frac{1}{\lambda}\right)^+ g^*(\lambda)\, d\lambda = \rho$$

and $C_n(H)/n$ converges almost surely to

$$C^*(\rho) \equiv \int_0^4 \log(\lambda\mu^*)^+ g^*(\lambda)\, d\lambda. \tag{11}$$

Thus, when both the transmitter and the receiver have perfect knowledge of the fading channel, the capacity scales like $nC^*(\rho)$, where $C^*(\rho)$ can be interpreted as the capacity of a fading channel with fading distribution g^* when water filling over the fading state is performed [17]. Similarly, when only the receiver has knowledge of the channel and the transmitter allocates an equal amount of power to each transmit antenna, the achievable mutual information scales like $nI^*(\rho)$, where $I^*(\rho)$ can be interpreted as the mutual information achieved by using constant transmit power in a fading channel with the gain distributed as g^*. We conclude that both $C_n(H)$ and $I_n(H)$ scale linearly with n but the rate of growth is larger for C_n than is for I_n. Moreover, if we let C_n^ϵ be the ϵ-outage capacity, i.e., such that

$$P(C_n(H) \le C_n^\epsilon) = \epsilon$$

then the above results implies that

$$\lim_{n\to\infty} \frac{C_n^\epsilon}{n} = C^*(\rho)$$

for all $\epsilon > 0$. This is because almost-sure convergence implies convergence in probability. This means that for large n, the capacity becomes insensitive to the realization of H. Similar comments apply to the scaling of $I_n(H)$.

We now compare $I^*(\rho)$ and $C^*(\rho)$ in both the low- and high-SNR regimes.

As a first-order approximation, at low SNR

$$\begin{aligned} I^*(\rho) &= \int_0^4 \log(1+\rho\lambda)) g^*(\lambda)\, d\lambda \\ &\approx \rho \int_0^4 \lambda g^*(\lambda)\, d\lambda \\ &= \rho. \end{aligned} \tag{12}$$

We observe that at low SNR, $I^*(\rho)$ depends only on the average SNR and not on the eigenvalue distribution g^*.

For the capacity, we calculate

$$\frac{dC^*(\rho)}{d\rho} = \frac{1}{\mu^*} \tag{13}$$

where μ^* is the water-filling level. As ρ approaches 0, μ^* approaches $\frac{1}{4}$. To first order, at low SNR

$$C^*(\rho) \approx 4\rho$$

and we conclude that

$$\lim_{\rho\to 0} \frac{C^*(\rho)}{I^*(\rho)} = 4.$$

Hence, the water-filling strategy affords a significant performance gain over the constant-power strategy at low SNR. The intuition is that when there is little transmit power, it is much more effective to expend it on the strongest eigenmode of the system (with gain 4) rather than spread the power evenly across all modes.

Next we consider the high-SNR regime. Using the explicit expression (9), we see that as $\rho \to \infty$

$$I^*(\rho) = \log(\rho/e) + o(1)$$

a result already noted in [3].

At high SNR, it is well known that the water-filling and the constant power strategies yield almost the same performance

$$\lim_{\rho\to\infty} [C^*(\rho) - I^*(\rho)] = 0$$

and hence $C^*(\rho)$ has the same high-SNR approximation of $\log(\rho/e)$.

Although water filling does not always give significant capacity improvements over the equal-power strategy, the perfect channel knowledge at the transmitter often leads to easier and more reliable implementations of the receiver, since the receiver can now be dealing with decoupled channels instead of having to perform cancellation and nulling.

B. Correlated Fading

1) Correlation Model: In the previous subsection, we assumed that the fades between different antenna pairs are independent of each other. We will now consider the situation when

the fading between antenna pairs is correlated. While the results we obtained for the independent fading case holds for *any distribution* of the individual H_{ij}, the results we present here for the correlated case are only for the case of a Rayleigh-fading model. Each of the H_{ij} are assumed to be complex, zero-mean, circular symmetric Gaussian random variables with variance $E[|H_{ij}|^2] = 1$. The H_{ij} are jointly Gaussian with the following covariance structure:

$$E[H_{pj}H_{qk}^*] = \Psi_{jk}^T \Psi_{pq}^R$$

where Ψ^T and Ψ^R are n by n covariance matrices. This fading model embodies three assumptions.

- The correlation between the fading from transmit antennas p and q to the same receive antenna is Ψ_{pq}^T and does not depend on the receive antenna; Ψ^T describes the *transmit correlation*.
- The correlation between the fading from a transmit antenna to receive antenna j and to receive antenna k is Ψ_{jk}^R and does not depend on the transmit antenna; Ψ^R describes the *receive correlation.*
- The correlation between the fading of two distinct antenna pairs is the product of the corresponding transmit correlation and receive correlation.

The first two assumptions are usually quite accurate when antenna elements are colocated in the same physical unit at the transmitter and also at the receiver. The product-form assumption is made for analytical tractability and can be thought of as a first-order approximation of the correlation structure when the fading from two transmit antennas to the same receive antenna and the fading from two receive antennas to the same transmit antenna is much more highly correlated than that between two distinct antenna pairs. This product form assumption is studied through simulations in Section V.

To consider the scaling of capacity and mutual information with the number of antennas, we need to make further assumptions on the covariance matrices Ψ^R and Ψ^T as the system scales. In particular, we assume that the empirical eigenvalue distributions of Ψ^R and Ψ^T converge in distribution to some limiting distributions F_R and F_T, respectively. This will be true if

1) the correlation between the fading at two antennas depends only on the relative and not absolute positions of the antennas;
2) the antennas are arranged in some regular arrays, such as square or linear grids, and as we scale up the number of antennas, the relative position of adjacent antennas are fixed; and
3) the correlation decays sufficiently fast over space.

For example, if the antennas are arranged in a linear array, Ψ^R and Ψ^T are Toeplitz. If the power spectral densities of the stationary processes

$$\{H_{j1}: j = 1, 2, \ldots\} \quad \text{and} \quad \{H_{1k}: k = 1, 2, \ldots\}$$

exist at all frequencies, then the limiting eigenvalue distributions F_T and F_R of Ψ^T and Ψ^R exist. For a given $x \geq 0$, $F_T(x)$ is the fraction of frequencies in the power spectral density of $\{H_{1k}: k = 1, 2, \ldots\}$ with power less than or equal to x. Motivated by this example, we will in general define

$$S_R(\omega) := F_R^{-1}(\omega), \quad S_T(\omega) := F_T^{-1}(\omega), \qquad \omega \in [0, 1].$$

S_R and S_T defined in this way is always nondecreasing from 0 to 1. One can think of S_R and S_T as power spectral densities except that the frequencies are reordered such that they are always nondecreasing functions of ω. In the results to be presented, the ordering is immaterial and only the distribution of powers is relevant. We also observe that

$$\int_0^1 S_R(\omega)\, d\omega = \int_0^1 S_R(\omega)\, d\omega = 1$$

because $E[|H_{ij}|^2] = 1$ for all i, j.

It should be noted that the power spectral densities of some fading correlation models may not exist at all frequencies. An example is Jakes' model [1], with the "U-shaped" power spectral density which is bounded over only a finite interval. The reason is that the autocorrelation function decays slowly as a function of distance r, like $\frac{1}{r}$. The results below do not apply to such models.

2) Analysis: The starting point of the analysis is that H can be factorized in the form $H = (\Psi^R)^{\frac{1}{2}} W (\Psi^T)^{\frac{1}{2}}$, where the entries of W are i.i.d. complex circular symmetric Gaussian with mean 0 and variance 1. Hence,

$$HH^\dagger = (\Psi^R)^{\frac{1}{2}} W \Psi^T W^\dagger (\Psi^R)^{\frac{1}{2}}.$$

For capacity analysis, we are interested in the eigenvalue distribution of $HH^\dagger$, or equivalently $\Psi^R W \Psi^T W^\dagger$. Now W is isotropic, i.e., UW and WU have the same distribution as W for any deterministic unitary matrix U. We can factorize $\Psi^T = UD_TU^\dagger$ and $\Psi^R = VD_RV^\dagger$, where U and V are unitary and D_T, D_R are diagonal. The fact that W is isotropic allows us to conclude that the matrix $\Psi^R W \Psi^T W^\dagger$ has the same eigenvalue distribution as $D_R W D_T W^\dagger$. It should be noted that as $n \to \infty$, the eigenvalue distributions of D_R and D_T converge to F_R and F_T, respectively.

Theorem IV.1 tells us that the distribution of the eigenvalues of $WW^\dagger$, scaled by $1/n$, converges for large n. It turns out that the eigenvalue distribution of $D_R W D_T W^\dagger$ converges as well under the same scaling. However, in this case, no explicit expression for the limiting distribution is available. Instead, it is given in terms of its *Steltjes' transform* [18]. The Steltjes' transform of a distribution G is defined by

$$m_G(z) := \int \frac{1}{\lambda - z}\, dG(\lambda)$$

for $z \in \mathcal{C}$ with $\Im(z) > 0$. It can be shown by an inversion theorem that the Steltjes' transform uniquely specifies a distribution. The following result yields a characterization of the limiting eigenvalue distribution of $D_T W D_R W^\dagger$.

Theorem IV.2: Let $\tilde{F}_n$ be the empirical eigenvalue distribution of $D_T W D_R W^\dagger$. Define $\tilde{G}_n(\lambda) := \tilde{F}_n(n\lambda)$. Then almost

surely, $\tilde{G}_n$ converges in distribution to a limit G^o, whose Steltjes' transform is given by

$$m_{G^o}(z) = \int_0^1 u(\theta, z)\, d\theta$$

where $u(\theta, z)$ is the unique solution to the functional fixed-point equation

$$u(\theta, z) = \frac{1}{-z + S_T(\theta)\int_0^1 \frac{S_R(\omega)}{1+S_R(\omega)\int_0^1 u(\phi, z)S_T(\phi)\, d\phi}\, d\omega}.$$

The proof of this result, which is based on random matrix results in Girko [18], is given in Appendix II.

Using this result, it can now be shown, exactly as in the independent fading case, that almost surely as $n \to \infty$

$$\frac{I_n(H)}{n} \to I^o(S_R, S_T, \rho)$$
$$\frac{C_n(H)}{n} \to C^o(S_R, S_T, \rho)$$

where

$$I^o(S_R, S_T, \rho) = \int_0^\infty \log(1+\rho\lambda)\, dG^o(\lambda) \qquad (14)$$

and

$$C^o(S_R, S_T, \rho) = \int_0^\infty \log(\lambda\mu^*)^+\, dG^o(\lambda) \qquad (15)$$

with μ^* satisfying

$$\int_0^\infty \left(\mu^* - \frac{1}{\lambda}\right)^+ dG^o(\lambda) = \rho.$$

The important conclusion is that even with correlation, the capacity and mutual information still scale linearly with n. However, the rate of growth is different from the independent fading case. It should be emphasized that this conclusion is valid only under the specific scaling assumptions we made.

The constants C^o and I^o depend on the limiting distribution G^o, which is only indirectly characterized via its Steltjes' transform in Theorem IV.2. The following result gives a more direct characterization of the constant I^o without involving the Steltjes' transform of G^o.

Theorem IV.3:

$$I^o(S_R, S_T, \rho) = \int_0^1 \log[1 + S_T(x)\beta(x)]\, d = x$$

and for each $x \in [0, 1]$, $\beta(x)$ is the unique solution β to the fixed-point equation

$$\beta = \int_0^1 \frac{S_R(\omega)}{\frac{1}{\rho} + S_R(\omega)\int_{1-x}^1 \frac{S_T(\phi)}{1+S_T(\phi)\beta}\, d\phi}\, d\omega \equiv f_c(\beta). \qquad (16)$$

For the special case of no correlation at the receiver (i.e., $S_R(\omega) = 1$ for all $\omega \in [0, 1]$), the fixed-point equation that β must satisfy is simplified to

$$\beta = \frac{1}{\frac{1}{\rho} + \int_{1-x}^1 \frac{S_T(\phi)}{1+S_T(\phi)\beta}\, d\phi} \equiv f_{\text{ind}}(\beta). \qquad (17)$$

The proof of this result exploits the fact that the mutual information I_n can be achieved by a combination of successive decoding and linear minimum mean-square error (MMSE) demodulation. A sketch of this proof can be found in Appendix III.

Does correlation always reduce capacity? Let us fix the transmit correlation S_T and compare the performance when there is correlation at the receiver and when there is none. Since the function $h(y) = y/(a + by)$ is concave for $a, b > 0$, it follows from Jensen's inequality that in Theorem IV.3

$$f_c(\beta) \le f_{\text{ind}}(\beta)$$

for all $\beta > 0$. Hence, for a given x, if $\beta_c = f_c(\beta_c)$ and $\beta_{\text{ind}} = f_{\text{ind}}(\beta_{\text{ind}})$, then $\beta_c \le f_{\text{ind}}(\beta_c)$. By monotonicity of the fixed-point equation $\beta = f_{\text{ind}}(\beta)$, this implies $\beta_c < \beta_{\text{ind}}$. Hence, correlation at the receiver always decreases I^o. By the reciprocity property (see Appendix I), it can be seen that correlation at the transmit antenna always reduces I^o for a fixed receive correlation S_R.

A more general statement can be made to compare the performance under two different power spectra. A nondecreasing spectrum S_1 is defined to be *more spread out* than a nondecreasing S_2 if

$$\int_0^1 S_1(\omega)\, d\omega = \int_0^1 S_2(\omega)\, d\omega = 1$$

and for every $\theta \in [0, 1]$

$$\int_\theta^1 S_1(\omega)\, d\omega \ge \int_\theta^1 S_2(\omega)\, d\omega.$$

Note that the flat spectrum corresponding to independent fading is the least spread out according to this definition: there is the same amount of power at all frequencies. This notion of "spreading out" (also called *majorization* [19]) can be taken as a measure of the strength of correlation: the more spread out the spectrum, the stronger the correlation.

In the theory of majorization, a real-valued function H is said to be *Schur-concave* (resp., Schur-convex) if S_1 is more (resp., less) spread out than S_2 implies that $H(S_1(\cdot)) < H(S_2(\cdot))$. A basic result says that the function:

$$H(S_1(\cdot)) = \int_0^1 h[S_1(\omega)] d\omega$$

is Schur-concave (resp., Schur-convex) if the function h is concave (resp., convex). Applying this result to our problem, it follows that the right-hand side of (16) is a Schur-concave function of S_T. It then follows that I^o is, in fact, a Schur-concave function of S_T, i.e., stronger correlation always decreases I^o.

The above discussion focuses on the effect of fading correlation on the mutual information I^o. But, in fact, something more basic is going on. It is shown in Appendix IV that the more spread out are the power spectra S_T and S_R, the more spread out is G^o, the limiting spectrum of $D_R W D_T W^\dagger$. The mutual information I^o (14), being a Schur-concave function of G^o, therefore, decreases with stronger correlation. However, the water-filling capacity is *not* a Schur-concave function of G^o, and hence similar conclusions cannot be drawn.

We now focus on the low-SNR and high-SNR regimes. At low SNR, it follows from (12) that I^o depends only on the average received SNR and does not depend on the eigenvalue distribution G^o. Hence, at low SNR, fading correlation has no effect on the mutual information achieved by the equal transmit power strategy. On the other hand, the water-filling capacity at low SNR is approximately the average received SNR amplified by the upper limit of the eigenvalue distribution G^o. Since the stronger the correlation, the more G^o is spread out, this upper limit increases and hence the water-filling capacity at low SNR actually increases as the correlation becomes stronger.

At high SNR, the difference $C^o - I^o$ approaches 0 and hence both are reduced by fading. To calculate I^o, let us make the substitution $\beta(x) = \eta_R(x)\rho + \epsilon(x, \rho)$. Taking the limit as $\rho \to \infty$ in (16), we see that $\epsilon(x, \rho) \to 0$ and $\eta_R(x)$ satisfies the fixed-point equation

$$1 = \int_0^1 \frac{S_R(\omega)}{\eta_R(x) + xS_R(\omega)}\, d\omega \tag{18}$$

for each $x \in [0, 1]$. The high-SNR approximation of I^o is therefore,

$$\begin{aligned} I^o(S_R, S_T, \rho) &= \int_0^1 \log[1 + \rho S_T(x)\eta(x)]\, dx + o(1) \\ &= \log\rho + \int_0^1 \log S_T(\omega)\, d\omega \\ &\quad + \int_0^1 \log \eta_R(x)\, dx + o(1). \end{aligned} \tag{19}$$

This can be simplified further. By the reciprocity property (see Appendix I), we know that $I_n(H) = I_n(H^\dagger)$. From this we can conclude that

$$I^o(S_T, S_R, \rho) = I^o(S_R, S_T, \rho).$$

Let us set $S_T(\omega) = 1$ for all $\omega \in [0, 1]$. From (19), we get

$$I(S_T, S_R, \rho) = \log\rho + \int_0^1 \log \eta_R(x)\, dx + o(1).$$

On the other hand,

$$\begin{aligned} I(S_R, S_T, \rho) &= \log\rho + \int_0^1 \log S_R(\omega)\, d\omega \\ &\quad + \int_0^1 \log(1 - x)\, dx + o(1). \end{aligned}$$

Equating these two expressions, we get

$$\int_0^1 \log \eta_R(x)\, dx = \int_0^1 \log S_R(\omega)\, d\omega + \int_0^1 \log(1 - x)\, dx.$$

Now

$$\int_0^1 \log(1 - x)\, dx = \log e.$$

Substituting these equations into (19), we get

$$\begin{aligned} I(S_T, S_R, \rho) &= \log(\rho/e) + \int_0^1 \log S_T(\omega)\, d\omega \\ &\quad + \int_0^1 \log S_R(\omega)\, d\omega + o(1). \end{aligned} \tag{20}$$

We observe that the first term is the high–SNR capacity for independent fading. Hence, the second and third terms represent the capacity penalty due to correlation at the transmit and receive sides, respectively. The fact that they are not positive follows directly from Jensen's inequality.

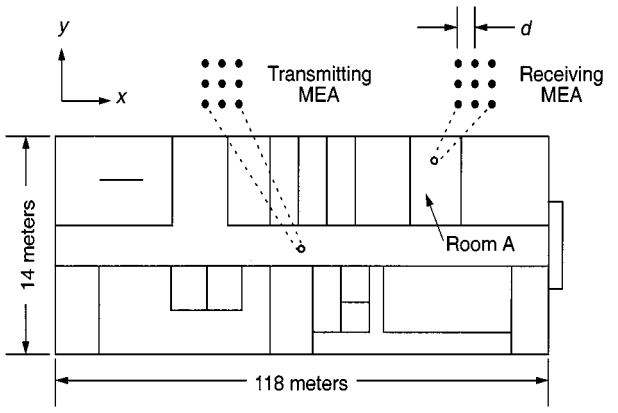

Fig. 1. Floor plan for the first floor of an office building at Crawford Hill, NJ. Receivers with antennas positioned in linear or square grids are placed randomly at 1000 locations in Room A. The transmitting MEA is placed with its adjacent sides parallel to x-axis and y-axis, respectively. The receiving MEA is placed in a random orientation at each of the sample locations.

V. Simulation Experiments

A. Methodology and Assumptions

We use the WiSE ray-tracing simulator [6] to construct random instances of channel matrix H for indoor wireless environment. WiSE allows us to specify the floor plan of a building (e.g., location of vertical walls, ceilings, corridors, etc.) and generate the corresponding propagation models inside the building. As described in [6], the reflection/refraction coefficients and scattering effect for different building materials are derived from a multilayer dielectric model. For our numerical study, we consider the indoor wireless environment of a two-floor office building at Crawford Hill, NJ (see Fig. 1). We place the transmitting MEA on the first floor ceiling near the middle of the office building throughout our study. Receiving MEAs are placed with random rotations at 1000 randomly chosen positions in Room A, which is at intermediate distance from the transmitter. We consider a carrier frequency of 5.2 GHz, i.e., wavelength $\lambda_o = 5.8$ cm. The MEAs consist of multiple omnidirectional antennas, arranged either in square grids or linear arrays within horizontal planes. The separation between antenna elements d is the same at both the transmitting and receiving MEAs. We consider $d = 0.5\lambda_o$ and $d = 5\lambda_o$, unless specified otherwise.

The power of the rays impinging on the receiving antennas is recorded when the carrier is launched from the transmitting MEA with power $10 \log P_{\max}$ dBm. The impulse response between a specific transmitting–receiving antenna pair is modeled as the vector sum of all the rays arriving at the receiving antenna as

$$g_{ij}(t) = \sum_{k=0}^{M} \sqrt{P_k} \cdot e^{i\theta_k} \cdot \delta_k(t - \tau_k) \tag{21}$$

where P_k, θ_k, and τ_k are the received power, phase angle, and time delay of the kth ray, respectively. M is the total number of rays and $\delta_k(t)$ is the delta impulse function. With narrow-band

assumption, we compute the frequency response at infinitesimally small bandwidth centered at the carrier frequency as

$$h_{ij} = \sum_{k=0}^{M} \sqrt{P_k} \cdot e^{i\theta_k} \cdot e^{i2\pi f_o \tau_k}. \tag{22}$$

H is computed using (22) and P_k, θ_k, and τ_k are obtained from the WiSE simulation. All the n^2 elements h_{ij} are complex numbers in this case.

Since H varies for different receiver locations, we estimate the channel variance σ^2 by averaging over 1000 realizations of H, and over all possible antenna pairs, j to i. We assume that the average received SNR ρ, as defined in Section III, should be high enough for low-error-rate communication. If SNR is too low, we need very complex codes to provide enough redundancy to combat the noise so that we can recover the desired signal with low error probability at the receiver. The practical constraints on analog-to-digital (A/D) converters that are available with current technology limit the maximum SNR that can be exploited effectively. Thus, we consider SNRs in the 18–22 dB range. For all our simulations, we assume W to be 10 MHz, and N_o to be -170 dBm/Hz, giving a total noise variance N_oW of -100.8 dBm. When we take expectation with respect to different realizations of H, we mean taking the ensemble average over the 1000 sample receiver locations. The capacities with and without water filling, C_n and I_n, are computed for different n. The results are presented in terms of complementary cumulative distribution functions (CCDFs), the averages $\overline{C}_n$ and $\overline{I}_n$, and capacities at 5% channel outage, $C_n^{0.05}$ and $I_n^{0.05}$.

B. Fading Correlation

As mentioned before, H_{ij} are correlated for finite separation between antenna elements. For an illustration, we consider the case of a two-antenna MEA system. 1000 realizations of the channel matrix H are generated using WiSE for different antenna spacing d. Using the notations in Section IV, Ψ_{12}^T and Ψ_{12}^R are determined. The magnitude of Ψ_{12}^T and Ψ_{12}^R resemble zero-ordered Bessel functions that decay very slowly, as shown in Fig. 2(a). At $d = 0.5\lambda_o$, a strong correlation of 0.65 exists between path gains originating from different transmitters. The correlation between path gains arriving at different receivers is 0.34. The asymmetry is due to the different local scattering environments around the transmitter and receiver.

In Section IV, we modeled the two-dimensional correlation function in product form. To verify the appropriateness of this approach, we plot the product of Ψ_{12}^T and Ψ_{12}^R in Fig. 2(b), together with the correlation $E[H_{11}H_{22}^*]$ inferred from WiSE simulation results. Close agreement is found consistently between these two curves over the range of d we consider. This implies that our assumption of separable transmit/receive correlations in Section IV-B1 is a reasonable first approximation.

C. Capacity of MEA Systems

In this subsection, we consider square MEAs, which are more compact than linear arrays for a given n. The receiver MEA is placed in Room A. Fig. 3 shows the CCDFs of C_n and I_n for $n = 1, 4, 9, 16, 25$, and 36, assuming $d = 0.5\lambda_o$ and $\rho =$ 18 dB. Recall that C_n and I_n are defined as the capacity with optimal water-filling power allocation and with equal-power allocation, respectively. Examining Fig. 3, we see that as n increases, the CCDFs of both C_n and I_n shift to the right, indicating that MEAs yield a capacity gain that increases steadily with n. We see that as n increases, the horizontal gap between

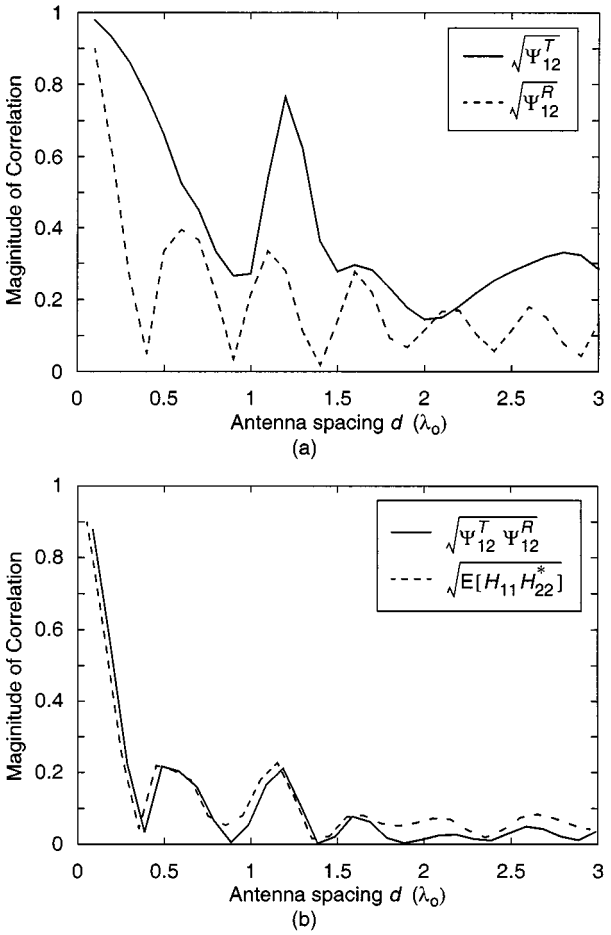

Fig. 2. (a) Magnitude of correlations Ψ_{12}^T and Ψ_{12}^R (as defined in Section IV) for antenna spacings ranging from 0 to $3\lambda_o$. (b) Magnitude of the normalized correlation $E[H_{11}H_{22}^*]$ compared to the magnitude of the product $\Psi_{12}^T\Psi_{12}^R$.

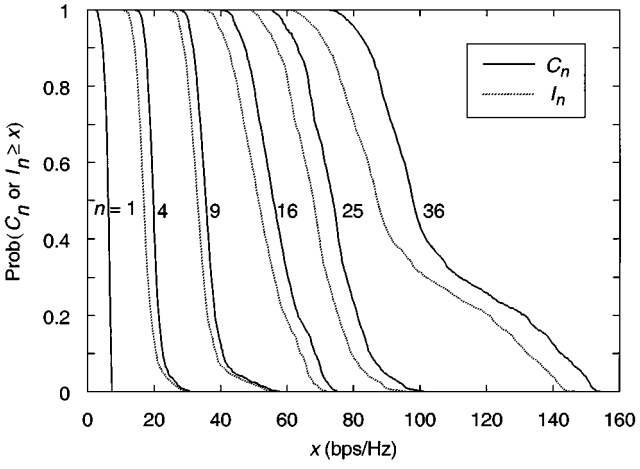

Fig. 3. The CCDFs of C_n (achieved via water filling) and I_n (with equal power allocation) for $n = 1, 4, 9, 16, 25$, and 36 at $\rho = 18$ dB. MEAs are arranged in square grids with $d = 0.5\lambda_o$.

TABLE I
THE PERCENTAGE DIFFERENCE: $(C_n^{0.05} - I_n^{0.05})/I_n^{0.05}$ FOR MEAS PLACED IN ROOM A. HERE $\rho = 18$ dB AND $d = 0.5\lambda_o$

Number of antennas, n	1	4	9	16	25	36
$C_n^{0.05}$ (bps/Hz)	5.9	20	36	57	75	106
$I_n^{0.05}$ (bps/Hz)	5.9	19	34	52	68	95
% difference	0	5.7	7.2	8.8	10	11

C_n and I_n increases, i.e., water filling yields a larger gain over equal-power allocation.

A reasonable performance indicator is the capacity that can be supported with 5% outage. Table I presents values of $C_n^{0.05}$ and $I_n^{0.05}$ extracted from the CCDFs shown in Fig. 3. When $n = 1$, $C_n^{0.05} = I_n^{0.05} = 5.9$ b/s/Hz. Increasing n can yield dramatic increases in $C_n^{0.05}$ and $I_n^{0.05}$. When $n = 4$, $C_n^{0.05} = 20$ b/s/Hz and $I_n^{0.05} = 19$ b/s/Hz, which are nearly three-and-a-half times higher than for $n = 1$. Increasing to $n = 36$, we obtain $C_n^{0.05} = 106$ b/s/Hz and $I_n^{0.05} = 95$ b/s/Hz, which are, respectively about 18 and 16 times higher than for $n = 1$.

Table I also presents $(C_n^{0.05} - I_n^{0.05})/I_n^{0.05}$, the fractional gain yielded by water filling over equal-power allocation. This fractional gain increases from 0 $(n = 1)$ to 11.3% $(n = 36)$.

The capacity improvement of water filling over equal-power allocation depends not only on n, but on the SNR ρ as well. Fig. 4(a) and (b) shows the ratios $\overline{C}_n/\overline{I}_n$ and $C_n^{0.05}/I_n^{0.05}$ versus ρ for $n = 4, 9, 16$. The figure assumes MEAs on square grids with $d = 0.5\lambda_o$. The receiving MEA is placed in Room A. The ratios $\overline{C}_n/\overline{I}_n$ and $C_n^{0.05}/I_n^{0.05}$ are substantial at low SNR ρ, and decrease asymptotically toward unity as ρ increases. When ρ is low, it is important to allocate the available power to the strongest subchannels, while as ρ increases, there is sufficient power to be distributed over all subchannels.

D. Asymptotic Behavior of MEA Capacities

In this section, we study the asymptotic behavior of the capacity as n grows large. We focus on the high-SNR regime, considering $\rho = 22$ dB. Since $C_n \approx I_n$ for high SNR, we consider only the water-filling capacity C_n here. We consider linear MEAs for two different values of antenna spacing: $d = 0.5\lambda_o$ and $d = 5\lambda_o$. The transmitting MEA is placed either parallel to the long dimension of the hallway (inline case) or perpendicular to it (broadside case). In all cases, the receiving MEA is placed in a random angular orientation in Room A. In this section, we consider the average capacities obtained on simulated channels $\overline{C}_n$ as opposed to the 5% outage capacity $C_n^{0.05}$ considered in the previous section.

In order to compute the asymptotic growth rate of capacity assuming independent fading, as derived in Section IV-A, we use simulated channel matrices H' whose entries H'_{ij} are generated by placing individual transmitter and receiver antenna elements at i.i.d. random locations in Room A, instead of placing them in a linear array separated by a fixed distance (e.g., $d = 0.5\lambda_o$) as in the regular case. With such an arrangement, the fades between antenna pairs are almost mutually independent. We use these matrices H' to estimate the variance σ^2 and the equivalent SNR ρ. We then compute the asymptotic growth rate C^*

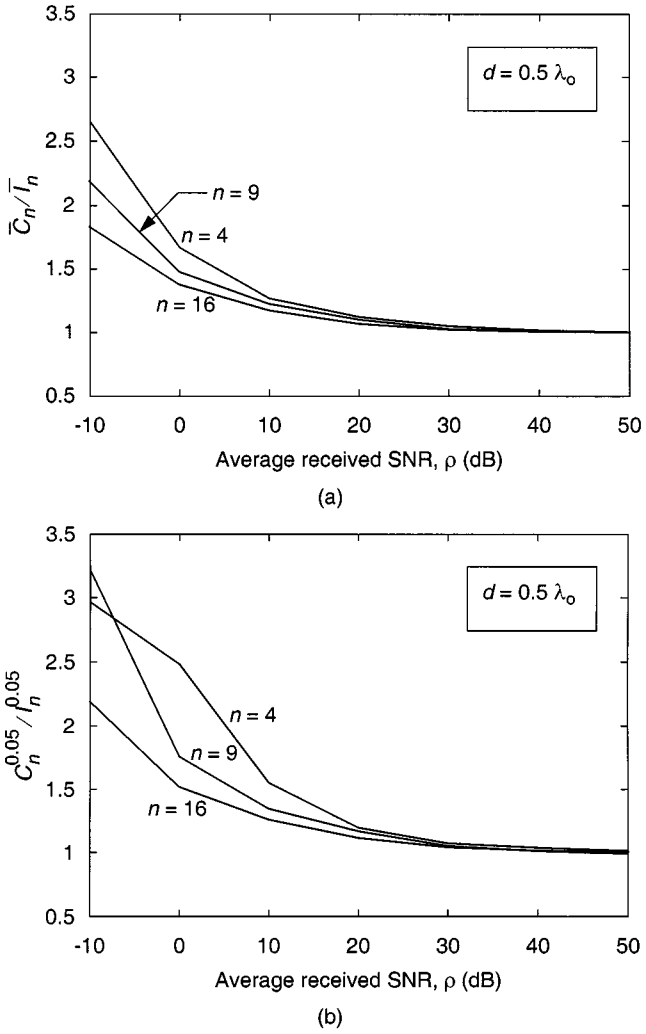

Fig. 4. (a) The ratio of average capacity with water filling to that with equal-power allocation $\overline{C}_n/\overline{I}_n$ at varying average received SNR ρ, for $n = 4, 9$, and 16. (b) Ratio of 5% outage capacity with water filling to that with equal-power allocation $C_n^{0.05}/I_n^{0.05}$ over different ρ, for $n = 4, 9$, and 16 . In both (a) and (b) MEAs at both the transmitter and the receiver are arranged in square grids.

using (11). To compute the growth rate of capacity including fading correlation, as derived in Section IV-B, we use simulated channel matrices H to estimate the variance of H_{ij} (i.e., σ^2), λ_i^R and λ_j^T. In this case, both the transmitting and the receiving MEAs are linear arrays, and individual elements are placed at a fixed spacing of either $0.5\lambda_o$ or $5\lambda_o$ apart. Recall that λ_i^R and λ_j^T are eigenvalues of Ψ^R and Ψ^T, respectively. For each case, we generate 1000 random channel matrices H and estimate the covariance matrices Ψ^R and Ψ^T as $E[H_{i\cdot}^{\dagger} H_{i\cdot}]$ and $E[H_{\cdot j} H_{\cdot j}^{\dagger}]$, respectively. Since the asymptotic growth rate C^o given by (15) is difficult to compute, we can approximate C^o by computing I^o (since $C^o \approx I^o$ at high SNR) as given by (20). We can approximate $S_T(w)$ and $S_R(w)$ in (20) with piecewise-linear curves, and replace the integrals with summations

$$C_n^o \approx I_n^o \approx \frac{1}{n}\left(\log(\rho) + \sum_{i=1}^{n} \log \lambda_i^R + \sum_{j=1}^{n} \log \lambda_j^T\right) \quad (23)$$

where λ_i^R's and λ_j^T's are eigenvalues of Ψ^R and Ψ^T, respectively.

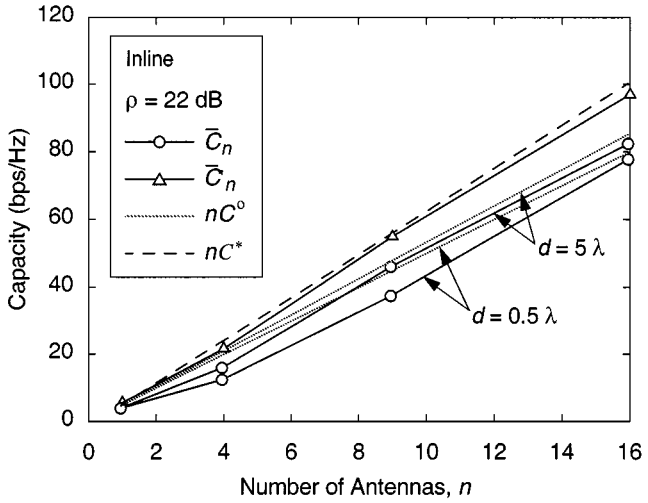

Fig. 5. Average capacity $\overline{C}_n$ versus n for inline case. We consider linear arrays with two antenna spacings: $d = 0.5\lambda_o$ and $d = 5\lambda_o$. The transmitting MEA is placed parallel to the x-axis. The receiving MEA is placed in a random orientation at 1000 random locations in Room A. $\overline{C}'_n$ is the average capacity obtained when the transmit and receive antenna elements lie at i.i.d. locations within their respective workspaces, i.e., they are not constrained to regular linear arrays. nC^* and nC^o are asymptotic results for correlated and independent H_{ij}, respectively.

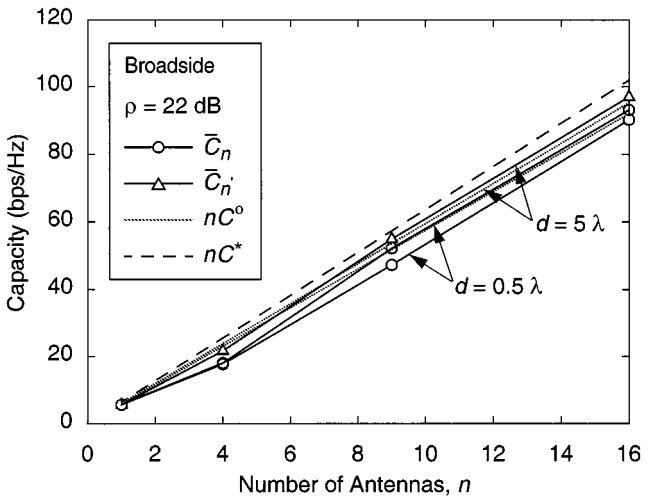

Fig. 6. Average capacity $\overline{C}_n$ versus n for the broadside case. We consider linear arrays with two antenna spacings: $d = 0.5\lambda_o$ and $d = 5\lambda_o$. The transmitting MEA is placed parallel to the y-axis. The receiving MEA is placed in a random orientation at 1000 random locations in Room A. $\overline{C}'_n$ is the average capacity obtained when the transmit and receive antenna elements lie at i.i.d. locations within their respective workspaces, i.e., they are not constrained to regular linear arrays. nC^* and nC^o are asymptotic results for correlated and independent H_{ij}, respectively.

We first consider the inline case. Fig. 5 shows the average capacity $\overline{C}_n$ versus n for $d = 0.5\lambda_o$ and $d = 5\lambda_o$. For each d, $\overline{C}_n$ grows roughly linearly with n, from about 5.57 b/s/Hz at $n = 1$ to about 79.3 b/s/Hz for $d = 0.5\lambda_o$ and 84.0 b/s/Hz for $d = 5\lambda_o$ at $n = 16$. At each n, $\overline{C}_n$ is larger for $d = 5\lambda_o$ than for $d = 0.5\lambda_o$, because the larger d reduces fading correlation, which was shown in Section IV-B to reduce capacity. In Fig. 5, we see that the asymptotic growth rate of capacity, assuming independent fading nC^*, significantly exceeds the observed average capacities $\overline{C}_n$ even for $d = 5\lambda_o$. The discrepancy between nC^* and $\overline{C}_n$ grows with increasing n. On the other hand, the asymptotic growth rates of capacity including correlation nC^o form better upper bounds for $\overline{C}_n$ than nC^* for both values of d. nC^o are in better agreement with $\overline{C}_n$ for $d = 5\lambda$ than $d = 0.5\lambda$, for all values of n.

To further explore the effects of fading correlation, we have computed channel matrices H'_{ij} in which the transmitter and receiver antenna elements lie in i.i.d. random locations within their respective workspaces, rather than lying in a regular linear or square array. Thus, the elements of H'_{ij} should be more nearly independent than those of H_{ij}. We have used these H'_{ij} to compute the average capacity $\overline{C}'_n$, which is also shown in Fig. 5. We observe that at each n, $\overline{C}'_n$ is larger than $\overline{C}_n$, and that $\overline{C}'_n$ is nearly as large as nC^*, the asymptotic growth rate assuming independent fading.

Similar results are obtained for the broadside case, as shown in Fig. 6. The average capacity $\overline{C}_n$ is generally higher than for the inline case for both $d = 0.5\lambda_o$ and $5\lambda_o$. Indeed, $\overline{C}_n$ in this case lies closer to the asymptotic value nC^* than $\overline{C}_n$ for the inline case. At $n = 16$ and $d = 5\lambda_o$, $\overline{C}_n/nC^* = 91.3\%$ for the broadside case, but only 82.4% for the inline case. This suggests that there is less correlation between path gains for the broadside configuration than for the inline configuration.

Our results indicate that fading correlation can significantly reduce MEA system capacity, even for antenna element spacing as large as $d = 5\lambda_o$. Moreover, our results show that the asymptotic growth rate nC^o, which considers correlation, provides a good estimate of the observed average capacity $\overline{C}_n$. This tends to validate the assumptions under which the formula for C^o was derived in Section IV-B, including the correlation model for the fades between different antenna pairs. If the assumptions in Section IV-B hold, C_n/n should converge almost surely to C^o (see (23)) in the limit of large n and high SNR. In Figs. 7 and 8, we illustrate this asymptotic behavior of C_n/n by plotting the empirical probability density functions (pdfs) of C_n/n for $n = 4$, 9, and 16, considering $\rho = 22$ dB. Fig. 7 considers $d = 0.5\lambda_o$ (where there is strong correlation between elements of H_{ij}), while Fig. 8 considers $d = 5\lambda_o$ (where there is less correlation between elements of H_{ij}). As n increases, the pdf becomes narrower and has a higher peak value, i.e., C_n/n becomes less random. In the limit of large n, we expect the pdf of C_n/n to converge to an impulse function centered at the value C^o. The narrowing pdfs in Figs. 7 and 8 illustrate the almost-sure convergence of C_n/n to C^o. Note that when $d = 5\lambda_o$, the pdfs are narrower and taller than when $d = 0.5\lambda_o$. This indicates that the rate of convergence is higher when d is larger, i.e., when the correlation between elements of H_{ij} is lower.

VI. Conclusion

MEA systems offers potentially large capacity gains over single-antenna systems. With perfect channel knowledge at the transmitter, power can be allocated optimally over different transmitting antennas (water filling) to achieve capacity C_n. The water-filling gain C_n/I_n is most significant when there are fewer strong eigenmodes, i.e., when the average received

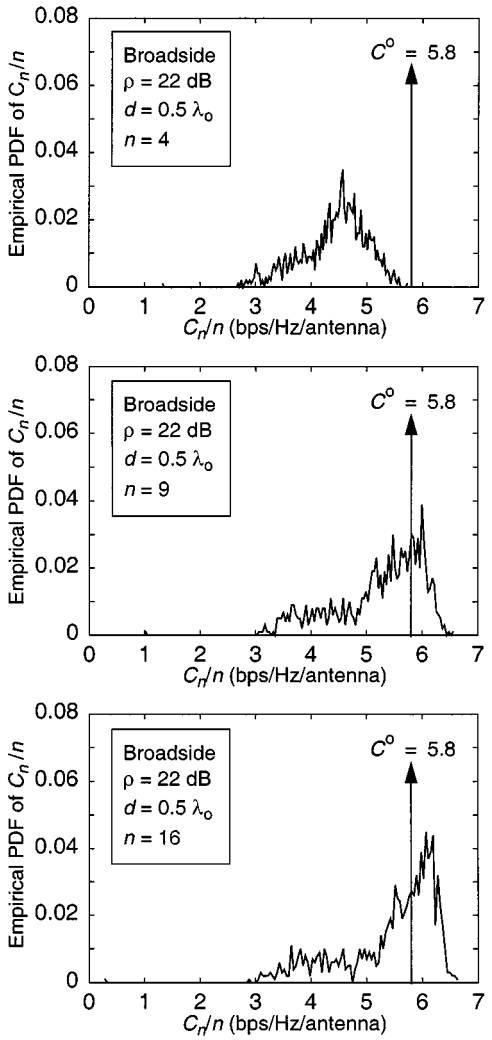

Fig. 7. Empirical pdf of the normalized capacity for $n = 4, 9,$ and 16. We consider linear arrays with antenna elements separated by $0.5\lambda_o$. The reference value is C^o as predicted by the asymptotic theory considering correlated H_{ij}.

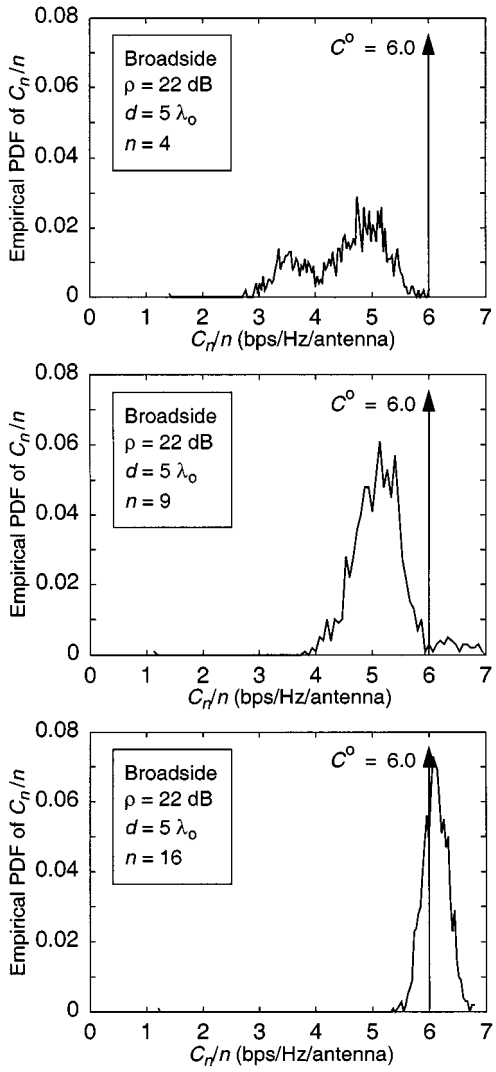

Fig. 8. Empirical pdf of the normalized capacity for $n = 4, 9,$ and 16. We consider linear arrays with antenna elements separated by $5\lambda_o$. The reference value is C^o as predicted by the asymptotic theory considering correlated H_{ij}.

SNR ρ is small. For example, $C_n^{0.05}/I_n^{0.05} = 3.5$ when $\rho = -10$ dB, but at $\rho = 50$ dB, water-filling gain is negligible, $C_n^{0.05}/I_n^{0.05} \approx 1$.

Assuming i.i.d. path gains between different antenna pairs, theoretical analysis shows that the capacity grows linearly with the number of antennas n in the limit of large n. In a more realistic propagation environment, correlation does exist between antenna pairs and affects the rate of growth of C_n and I_n, although it was shown that they still grow linearly with n. The rate of growth of I_n is reduced by correlation over the entire range of SNRs, while that for C_n is reduced by correlation at high SNR but is increased at low SNR. Our simulation results show that for $0.5\lambda_o$ antenna spacing, the simulated average capacity $\overline{C}_n$ is only 88.5% of the predicted value based on independent fading assumptions, nC^* for $n = 16$ in the case of broadside with $\rho =$ 22 dB. When the antenna spacing is increased to $d = 5\lambda_o$, the effect of correlations on total capacity is smaller: $\overline{C}_n/nC^* =$ 91.3%. The approximation based on our asymptotic results for correlated fading nC^o forms a close upper bound for the average capacity observed on simulated channels $\overline{C}_n$ at high SNR.

APPENDIX I
RECIPROCITY PROPERTY

We note here two reciprocity properties discussed in [20] which will be useful in some of the analysis in Section IV-B. These properties also make it very easy to extend the analytical results to the case with unequal number of transmit and receive antennas.

First, the (water-filling) capacity of the multiantenna channel with t transmit and r receive antennas and channel matrix H is the same as that of a system with r transmit and t receive antennas and channel matrix $H^\dagger$. This is because the nonzero singular values of $HH^\dagger$ and $H^\dagger H$ are identical. Second, the mutual information achieved with equal transmit power Q at each antenna in a system with r transmit and t receiver antennas and channel matrix H is the same as that achieved using transmit power Q in a system with r transmit and t receive antennas and channel matrix $H^\dagger$. This is because

$$I_n = \log\det(I_{r\times r} + QHH^\dagger) = \log\det(I_{t\times t} + QH^\dagger H).$$

It should be noted that for the water-filling capacity, the reciprocity is with respect to two systems with the same *total* transmit power, whereas for the equal-power mutual information, the reciprocity is with respect to two systems with the same power *per transmit antenna*.

APPENDIX II
PROOF OF THEOREM IV.2

The following theorem captures the essence of [18, Corollary 10.1.2], which is the key random matrix result we need.

Theorem AII.1: Let A_n be an $n \times n$ random matrix with independent entries which are zero-mean and satisfy the condition

$$n \operatorname{Var}(A_{ij}) < B$$

for some uniform bound $B < \infty$. Moreover, suppose we define for each n a function v_n: $[0, 1] \times [0, 1] \to \Re$ by

$$v_n(x, y) = n \operatorname{Var}(\mathbf{A}_{ij}),$$

i, j, satisfying

$$\frac{i}{n} \le x \le \frac{(i+1)}{n}, \qquad \frac{j}{n} \le y \le \frac{(j+1)}{n}$$

and that v_n converges uniformly to a limiting bounded function v. Then, the limiting eigenvalue distribution H^* of $A_n A_n^\dagger$ exists and its Steltjes' transform $m(z)$ is given by

$$m(z) = \int_0^1 u(x, z)\, dx \tag{24}$$

and $u(x, z)$ satisfies the equation,

$$u(x, z) = \frac{1}{-z + \int_0^1 \frac{v(x,y)}{1+\int_0^1 u(w,z)v(w,y)\,dw}\, dy}. \tag{25}$$

The solution of (25) exists and is unique in the class of functions $u(x, t) \ge 0$, analytical in z and continuous on $x \in [0, 1]$.

To prove Theorem IV.1, we can apply this random matrix result to

$$A_n := \frac{1}{\sqrt{n}} D_R^{1/2} W D_T^{1/2}.$$

The desired result follows by noting that, in this case, $v(x, y) = S_T(x) S_R(y)$.

APPENDIX III
PROOF OF THEOREM IV.3

Let us define

$$\tilde{W} := \frac{1}{\sqrt{n}} D_R^{1/2} W \tag{26}$$

and let $\tilde{w}_k$ be the kth column of $\tilde{W}$. The mutual information I_n has the same distribution as

$$\log\det\left(I + \rho \tilde{W} D_T \tilde{W}^\dagger\right).$$

This can be interpreted as the sum capacity of a vector multiple-access channel with n users and n degrees of freedom at the receiver, treating each of the n transmit antennas as a separate user. The transmit power of the kth user is $\rho(D_T)_{kk}$, and its vector of channel gain at the receiver is $\tilde{w}_k$.

The sum capacity can be achieved by a combination of successive cancellation and linear MMSE demodulation [21]: first one user is demodulated by a linear MMSE receiver and decoded, treating all the other users as interference; then the signal from that user is subtracted off and the process is repeated for the remaining users. Moreover, the sum capacity is achieved regardless of the decoding order among the users. Let us decode then in increasing order of the transmit powers of the users, and without loss of generality assume that the diagonal elements of D_T are in increasing order. Thus, we have

$$\log\det\left(I + \rho \tilde{W} D_T \tilde{W}^\dagger\right) = \sum_{k=1}^{n} \log(1 + \mathrm{SIR}_k), \tag{27}$$

where SIR_k is the signal-to-interference ratio (SIR) achieved when demodulating the kth user. From MMSE estimation theory it can be computed that

$$\mathrm{SIR}_k = \rho(D_T)_{kk} \tilde{w}^\dagger \left(\rho \tilde{W} D_T^{(k)} \tilde{W}^\dagger + I\right)^{-1} \tilde{w}$$

where

$$(D_T^{(k)})_{ii} = 0, \qquad \text{if } i \le k$$

and

$$(D_T^{(k)})_{ii} = (D_T)_{ii}, \qquad \text{if } i > k.$$

Substituting (26), we get

$$\mathrm{SIR}_k = \frac{(D_T)_{kk}}{n} w_k^\dagger \left(\frac{1}{n} W D_T^{(k)} W^\dagger + \frac{1}{\rho} D_R^{-1}\right)^{-1} w_k$$

where w_k is the kth column of W.

We now let $n \to \infty$ but keep $k = xn$ for x fixed. We wish to compute the asymptotic limit of SIR_k.

The entries of w_k are i.i.d. We can apply [22, Lemma 3.2] to conclude that

$$\mathrm{SIR}_k - (D_T)_{kk} \frac{1}{n} \operatorname{Tr}\left(W D_T^{(k)} W^\dagger + \frac{1}{\rho} D_R^{-1}\right)^{-1} \to 0 \tag{28}$$

in probability. The trace term is the Steltjes' transform of the empirical eigenvalue distribution of $W D_T^{(k)} W^\dagger + \frac{1}{\rho} D_R^{-1}$ evaluated at 0. The following theorem, due to Marcenko and Pastur [23] and refined by Silverstein and Bai [15], computes the limit.

Theorem AIII.1: Let A and B be two n by n diagonal matrices whose empirical distributions of the diagonal elements converge to F_A and F_B, respectively, as $n \to \infty$. Let m_A be the Steltjes' transform of F_A. Then, the empirical eigenvalue distribution of $\frac{1}{n} W B W^\dagger + A$ converges almost surely, and the Steltjes' transform $m(z)$ of the limiting distribution is the unique solution to the functional fixed-point equation

$$m(z) = m_A\left(z - \int \frac{\tau}{1 + \tau m(z)}\, dF_B(\tau)\right).$$

Applying this result to the above yields

$$\frac{1}{n} \operatorname{Tr}\left(W D_T^{(k)} W^\dagger + \frac{1}{\rho} D_R^{-1}\right)^{-1} \to m(0) \equiv \beta(x)$$

where $\beta(x)$ satisfies

$$\beta(x) = \int_0^1 \frac{S_R(\omega)}{\frac{1}{\rho} + S_R(\omega)\int_{1-x}^1 \frac{S_T(\phi)}{1+S_T(\phi)\beta(x)}\,d\phi}\,d\omega.$$

Hence, combining this with (27) and (28), we get

$$\frac{1}{n}\log\det\left(I+\rho\tilde{W}D_T\tilde{W}^\dagger\right) = \frac{1}{n}\sum_{k=1}^{n}\log(1+\mathrm{SIR}_k) \to \int_0^1 \log(1+\rho S_T(x)\beta(x))\,dx$$

thus establishing the desired result.

Appendix IV
A Majorization Result

In this appendix, we will explore the effect of correlation has on the limiting eigenvalue distribution of the key random matrix $D_R W D_T W^\dagger$. Just as for power spectrum, we can define an equivalent notion of a distribution being more spread out than another.

Theorem AIV.1: A distribution F is more spread out than G if they have the same expectation, and for every $\theta \in [0, 1]$

$$\int_\theta^1 F^{-1}(x)\,dx \geq \int_\theta^1 G^{-1}(x)\,dx.$$

The main result we want to prove is the following.

Theorem AIV.2: Let S_R, $\tilde{S}_R$ be two receiver correlation power spectra, and S_T, $\tilde{S}_T$ be two transmitter correlation power spectra. Let G and $\tilde{G}$ be the limiting eigenvalue distributions of $D_R W D_T W^\dagger$ and $\tilde{D}_R W \tilde{D}_T W^\dagger$, respectively, where the diagonal elements of D_R, $\tilde{D}_R$, D_T, $\tilde{D}_T$ approach the spectra S_R, $\tilde{S}_R$, S_T, $\tilde{S}_T$, respectively. If $\tilde{S}_R$ is more spread out than S_T and $\tilde{S}_T$ is more spread out than S_T, then $\tilde{G}$ is more spread out than G.

To prove the theorem, we need the following definition, which gives a slightly more general notion of Schur convexity than the one presented earlier in the main body of the paper.

Definition AIV.3: Consider a map $\mathcal{H}$: $\mathcal{L}^1[0, 1] \to \Re$, where $\mathcal{L}^1[0, 1]$ is the set of all integrable functions on $[0, 1]$. For a function $g \in \mathcal{L}^1[0, 1]$, let F_g be the empirical distribution of g, i.e., $F_g(x) = m\{u\colon g(u) \leq x\}$, where m is the Lebesque measure. The map $\mathcal{H}$ is said to be Schur-convex if for any two functions g_1 and g_2 such that F_{g_1} is more spread out than F_{g_2} implies that $\mathcal{H}(g_1) \geq \mathcal{H}(g_2)$.

The following is a key lemma in the proof of Theorem AIV.2.

Lemma AIV.4: Suppose the map $\mathcal{H}$: $\mathcal{L}^1[0, 1] \to \Re$ has the following properties.

1) For any g, $\mathcal{H}(g)$ depends on g only through the empirical distribution F_g.
2) $\mathcal{H}$ is convex, i.e.,

$$\mathcal{H}(\alpha g_1 + (1-\alpha)g_2) \leq \alpha\mathcal{H}(g_1) + (1-\alpha)\mathcal{H}(g_2)$$

or $\alpha \in [0, 1]$.

Then $\mathcal{H}$ is Schur-convex.

Proof: Fix n and consider the class of functions in $\mathcal{L}^1[0, 1]$ that are piecewise-constant functions on the intervals $[k/n, (k+1)/n)$, for $k = 0, \ldots, n-1$. The mapping $\mathcal{H}$ on this class can be viewed as a function of n variables, the n values that the functions in the class can take on. Condition 1) then implies that $\mathcal{H}$ is a symmetric function of these n variables. The Schur convexity of $\mathcal{H}$ on this class follows from [19, Proposition C.2, p. 67]. The Schur convexity of $\mathcal{H}$ on $\mathcal{L}^1[0, 1]$ follows from an approximating argument by taking n large. □

Proof of Theorem AIV.2: If A is an n by n Hermitian matrix and $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$ are its ordered eigenvalues, then it is known that for all k

$$\sum_{i=1}^{k}\lambda_i(A) = \max_{UU^\dagger = I_{k\times k}} UAU^\dagger$$

where the maximization is over $k \times n$ complex matrices U. This extremal representation shows that $\sum_{i=1}^k \lambda_i(A)$ is a convex function of the entries of A. Applying this observation to the matrix $A = D_R H D_T H^\dagger$, it follows that $\sum_{i=1}^k \lambda_i(D_R H D_T H^\dagger)$ is a convex function of the entries of D_R, and also a convex function of the entries of D_T. From Theorem IV.2, almost surely the empirical eigenvalue distribution of $D_R H D_T H^\dagger$ converges to a limiting distribution G. This implies that for each fixed $\theta \in [0, 1]$

$$\sum_{i=1}^{\lfloor(1-\theta)n\rfloor}\lambda_i(D_R H D_T H^\dagger) \to \int_\theta^1 (G)^{-1}(x)\,dx := \mathcal{C}(g_R, g_T)$$

where g_R, $g_T \in \mathcal{L}^1[0, 1]$ are the limiting functions of the diagonal elements of D_R, D_T, respectively. The convexity of $\mathcal{C}$ both as a function of g_R and as a function of g_T follows from a limiting argument. Moreover, the dependence of $\mathcal{C}$ on g_T and g_R is only through their empirical distributions. Hence, from Lemma AIV.4, it follows that $\mathcal{C}$ is a Schur-convex function of g_R for a fixed g_T, and also a Schur-convex function of g_T for a fixed g_R. Hence, for a fixed S_R, if $\tilde{S}_T$ is more spread out than S_T, then

$$\int_\theta^1 \left(\tilde{G}\right)^{-1}(x)\,dx \geq \int_\theta^1 (G)^{-1}(x)\,dx.$$

This holds for all θ, and hence $\tilde{G}$ is more spread out than G; similarly, for a fixed S_T, if $\tilde{S}_R$ is more spread out than S_R, then $\tilde{G}$ is more spread out than G. This proves the theorem.

Acknowledgment

The authors wish to thank J. Ling and D. Chizhik for their assistance with the WiSE simulation tools. Discussions with G. J. Foschini, J. Salz, and D. Shiu have been enlightening and are greatly appreciated.

References

[1] W. Jakes, Jr., *Microwave Mobile Communications*. New York: Wiley, 1974.
[2] J. Winters, "Optimum combining for indoor radio systems with multiple users," *IEEE Trans. Commun.*, vol. 35, pp. 1222–1230, Nov. 1987.

[3] G. J. Foschini and M. J. Gans, "Capacity when using diversity at transmit and receive sites and the Rayleigh-faded matrix channel is unknown at the transmitter," in *Proc. WINLAB Workshop on Wireless Information Network*, Mar. 1996.
[4] ——, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
[5] W. C.-Y. Lee, "Effects on correlation between two mobile radio base-station antennas," *IEEE Trans. Commun.*, vol. COM-21, pp. 1214–1224, Nov. 1974.
[6] S. J. Fortune, D. H. Gay, B. W. Kernighan, O. Landron, R. A. Valenzuela, and M. H. Wright, "WiSE design of indoor wireless systems: Practical computation and optimization," *IEEE Comput. Sci. Eng.*, vol. 2, pp. 58–68, Mar. 1995.
[7] G. J. Foschini and R. A. Valenzuela, "Initial estimation of communication efficiency of indoor wireless channels," *Wireless Networks*, vol. 3, no. 2, pp. 141–154, 1997.
[8] C. N. Chuah, J. M. Kahn, and D. Tse, "Capacity of multi-antenna array systems in indoor wireless environment," in *Proc. IEEE GLOBECOM*, vol. 4, Nov. 1998, pp. 1894–1899.
[9] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Commun.*, vol. 5, pp. 10–22, Feb. 1998.
[10] J. Salz and J. H. Winters, "Effect of fading correlation on adapting arrays in digital wireless communications," in *Proc. IEEE Vehicular Technology Conf.*, vol. 3, 1993, pp. 1758–1774.
[11] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified channel model for mobile radio systems with smart antennas," in *Proc. Inst. Elec. Eng.–Radar, Sonar Navig.*, vol. 145, Feb. 1998, pp. 32–41.
[12] G. G. Raleigh and J. M. Cioffi, "Spatio–temporal coding for wireless communications," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.
[13] D. 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. Commun.*, vol. 48, pp. 502–513, Mar. 2000.
[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[15] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *J. Multivariate Anal.*, vol. 54, no. 2, pp. 175–192, 1995.
[16] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.
[17] A. Goldsmith and P. Varaiya, "Capacity of fading channel with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.
[18] V. L. Girko, *Theory of Random Determinants*. Norwell, MA: Kluwer, 1990.
[19] A. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic, 1979.
[20] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 586–595, 1999.
[21] M. K. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization and successive decoding achieves the total capacity of the Gaussian multiple-access channel," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Nov. 1997, pp. 1405–1409.
[22] D. Tse and O. Zeitouni, "Performance of linear multiuser receivers in random environments," in *IEEE Communications Theory Mini-Conf.*, 1999, pp. 163–167.
[23] V. A. Marcenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR-Sbornik1*, pp. 457–483, 1967.