

# Environmental Issues for MIMO Capacity

Daniel W. Bliss, Keith W. Forsythe, Alfred O. Hero, III, *Fellow, IEEE*, and Ali F. Yegulalp

**Abstract**—Wireless communication using multiple-input multiple-output (MIMO) systems enables increased spectral efficiency for a given total transmit power. Increased capacity is achieved by introducing additional spatial channels that are exploited using space–time coding. In this paper, the environmental factors that affect MIMO capacity are surveyed. These factors include channel complexity, external interference, and channel estimation error. The maximum spectral efficiency of MIMO systems in which both transmitter and receiver know the channel (using channel estimate feedback) is compared with MIMO systems in which only the receiver knows the channel. Channel complexity is studied using both simple stochastic physical scattering and asymptotic large random matrix models. Both uncooperative (worst-case) and cooperative (amenable to multiuser detection) interference are considered. An analysis for capacity loss associated with channel estimation error at the transmitter is introduced.

**Index Terms**—Channel capacity, channel phenomenology, information theory, interference cancellation, MIMO communication, multiuser detection, space–time coding.

## I. INTRODUCTION

MULTIPLE-INPUT multiple-output (MIMO) systems are a natural extension of developments in antenna array communication. While the advantages of multiple receive antennas, such as gain and spatial diversity, have been known and exploited for some time [1]–[3], the use of transmit diversity has more recently been investigated [4], [5]. Finally, the advantages of MIMO communication, exploiting the physical channel between many transmit and receive antennas, are currently receiving significant attention [6]–[9]. While it is possible for the channel to be so nonstationary that it cannot be estimated in any useful sense [10], in this paper, a quasistationary channel assumption is employed.

MIMO systems provide a number of advantages over single-antenna communication. Sensitivity to fading is reduced by the spatial diversity provided by multiple spatial paths. Under certain environmental conditions, the power requirements associated with high spectral efficiency communication can be significantly reduced by avoiding the compressive region of the information theoretic capacity bound. Here,

Manuscript received June 12, 2001; revised May 28, 2002. This work was supported by the Defense Advanced Research Projects Agency under Air Force Contract F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government. The associate editor coordinating the review of this paper and approving it for publication was Dr. Dennis R. Morgan.

D. W. Bliss, K. W. Forsythe, and A. F. Yegulalp are with Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02420-9185 USA (e-mail: bliss@ll.mit.edu; forsythe@ll.mit.edu; yegulalp@ll.mit.edu).

A. O. Hero, III is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122 USA (e-mail: hero@eecs.umich.edu).

Publisher Item Identifier 10.1109/TSP.2002.801914.

spectral efficiency is defined as the total number of bits per second per Hertz transmitted from one array to the other. Capacity increases linearly with signal-to-noise-ratio (SNR) at low SNR but increases logarithmically with SNR at high SNR. A given total transmit power can be divided among multiple spatial paths (or modes), driving the capacity closer to the linear regime for each mode, thus increasing the aggregate spectral efficiency. As seen in Fig. 1, which assumes an optimal high spectral efficiency MIMO channel [a channel matrix with a flat singular value distribution (SVD)], MIMO systems enable high spectral efficiency at much lower required energy per bit. Because MIMO systems use antenna arrays, interference can be mitigated naturally.

### A. Environment

The environmental factors that affect MIMO system capacity, namely channel complexity, external interference, and channel stationarity, are addressed in this paper in Sections III–V, respectively. The first category (channel complexity) is a function of the richness of scatterers. In general, capacity increases as the singular values of the channel matrix increase. The distribution of singular values is a measure of the usefulness of various spatial paths through the channel.

The second category (external interference) adversely affects the usefulness of paths through the channel. Given that the most useful portion of the channel lives in a subspace of the channel matrix, capacity loss is a function of the overlap of the interference with this subspace. Generally, interference is assumed to be uncooperative (worst-case). However, if the interference source is cooperative, that is, the various users share system parameters and control, the adverse effects of interference can be reduced significantly through the use of multiuser detectors.

The third category is channel stationarity. If the environment is stationary, then channel estimation error vanishes asymptotically. However, in practical systems, channel stationarity limits the useful period over which a channel can be estimated. Because the transmitter will generally have access to older channel estimates than the receiver, one would expect the channel estimation error to be greater at the transmitter.

### B. Channel Estimation Feedback

In implementing MIMO systems, one must decide whether channel estimation information will be fed back to the transmitter so that it can adapt. Most MIMO communication research has focused on systems without feedback. A MIMO system with an *uninformed transmitter* (without feedback) is simpler to implement, and at high SNR, its spectral efficiency bound approaches that of an *informed transmitter* (with feedback).

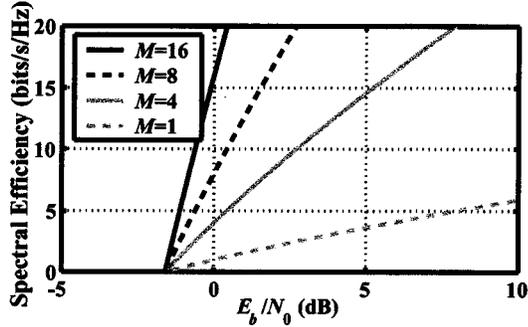


Fig. 1. Spectral efficiency bound as a function of noise density normalized energy per bit ( $E_b/N_0$ ) comparison of  $M \times M$  MIMO systems assuming channel matrices with flat SVD.

### C. Space–Time Coding

The focus of this paper is the environmental sensitivity of MIMO communication; however, for completeness, a few space–time coding references are discussed. In order to implement a MIMO communication system, a particular coding scheme must be selected. Most space–time coding schemes have a strong connection to well-known single-input single-output (SISO) coding approaches and assume an *uninformed transmitter*. Space–time coding can exploit the MIMO degrees of freedom to increase redundancy, spectral efficiency, or some combination of these characteristics [11]. Preliminary ideas are discussed in [6]. A simple and elegant solution that maximizes diversity and enables simple decoupled detection is proposed in [12]. More generally, orthogonal space–time block codes are discussed in [13] and [14]. A general discussion of distributing data across transmitters (linear dispersive codes) is given in [15]. High SNR design criteria and specific examples are given for space–time trellis codes in [16]. Unitary codes optimized for operation in Rayleigh fading are presented in [17]. More recently, MIMO extensions of turbo coding have been suggested [18], [19]. Finally, coding techniques for *informed transmitter* systems have received some interest [20], [21].

## II. INFORMATION THEORETIC CAPACITY

The information theoretic capacity of MIMO systems has been widely discussed, for example, in [7]. The development of the *informed transmitter* “water filling” and *uninformed transmitter* approaches is repeated here. This is useful as an introduction to MIMO capacity and to the notation used in this paper. In addition, the spectral efficiency bounds in the presence of interference are introduced.

### A. Informed Transmitter (IT)

For narrowband MIMO systems, the coupling between the transmitter and receiver for each sample in time can be modeled using

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (1)$$

where

$\mathbf{z}$  complex receive array output;

$\mathbf{H}$   $n_{Rx} \times n_{Tx}$  (number of receive by transmit antenna) channel matrix;

$\mathbf{x}$  transmit array vector;

$\mathbf{n}$  zero mean complex Gaussian noise.

The capacity is defined as the maximum of the mutual information [22]

$$\mathcal{I}(\mathbf{z}, \mathbf{x}|\mathbf{H}) = \left\langle \log_2 \left[ \frac{p(\mathbf{z}|\mathbf{x}, \mathbf{H})}{p(\mathbf{z}|\mathbf{H})} \right] \right\rangle \quad (2)$$

over the source probability density  $p(\mathbf{x}|\mathbf{H})$  subject to average transmit power constraints, where the expectation value is indicated using the notation  $\langle \cdot \cdot \cdot \rangle$ . Noting that the mutual information can be expressed as the difference between two conditional entropies

$$\mathcal{I}(\mathbf{z}, \mathbf{x}|\mathbf{H}) = h(\mathbf{z}|\mathbf{H}) - h(\mathbf{z}|\mathbf{x}, \mathbf{H}) \quad (3)$$

that  $h(\mathbf{z}|\mathbf{x}, \mathbf{H}) = h(\mathbf{n}) = n_{Rx} \log_2(\pi e \sigma_n^2)$ , and that  $h(\mathbf{z}|\mathbf{H})$  is maximized for a zero mean Gaussian source  $\mathbf{x}$ , the capacity is given by

$$C = \sup_{\langle \mathbf{x}\mathbf{x}^\dagger \rangle} \log_2 \frac{|\sigma_n^2 \mathbf{I}_{n_{Rx}} + \mathbf{H}\langle \mathbf{x}\mathbf{x}^\dagger \rangle \mathbf{H}^\dagger|}{|\sigma_n^2 \mathbf{I}_{n_{Rx}}|} \quad (4)$$

where

$|\cdot \cdot \cdot|$  determinant;

$\dagger$  Hermitian conjugate;

$\mathbf{I}_{n_{Rx}}$  identity matrix of size  $n_{Rx}$ .

There are a variety of possible constraints on  $\langle \mathbf{x}\mathbf{x}^\dagger \rangle$ , depending on the assumed transmitter limitations. Here, it is assumed that the fundamental limitation is the total power transmitted. The optimization of the  $n_{Tx} \times n_{Tx}$  noise-normalized transmit covariance matrix  $\mathbf{P} = \langle \mathbf{x}\mathbf{x}^\dagger \rangle / \sigma_n^2$  is constrained by the total noise-normalized transmit power  $P_o$ . Allowing different transmit powers at each antenna, this constraint can be enforced using the form  $\text{tr}\{\mathbf{P}\} \leq P_o$ . The channel capacity is achieved if the channel is known by both the transmitter and receiver, giving

$$C_{IT} = \sup_{\mathbf{P}; \text{tr}(\mathbf{P})=P_o} \log_2 |\mathbf{I}_{n_{Rx}} + \mathbf{H}\mathbf{P}\mathbf{H}^\dagger|. \quad (5)$$

To avoid radiating negative power, the additional constraint  $\mathbf{P} > \mathbf{0}$  is imposed by using only a subset of channel modes.

Substituting  $\mathbf{U}\mathbf{S}\mathbf{W}^\dagger$ , the magnitude-ordered singular value decomposition, for  $\mathbf{H}$ , (5) can be written as

$$C_{IT} = \sup_{\mathbf{Q}; \text{tr}\{\mathbf{Q}(\mathbf{S}^\dagger \mathbf{S})^{-1}\}=P_o} \log_2 |\mathbf{I}_{n_{\min}} + \mathbf{Q}| \quad (6)$$

$$\mathbf{Q} \equiv \mathbf{S}\mathbf{W}^\dagger \mathbf{P} \mathbf{W} \mathbf{S}^\dagger \quad (7)$$

where  $\mathbf{S}$  is a diagonal  $n_{\min} \times n_{\min}$  matrix,  $n_{\min} = \min(n_{Tx}, n_{Rx})$ , and  $\mathbf{U}$  and  $\mathbf{W}$  are  $n_{Rx} \times n_{\min}$  and  $n_{Tx} \times n_{\min}$  matrices containing the selected columns of unitary matrices. The maximum under the total power constraint can be found by differentiating with respect to  $q$  an arbitrary parameter of  $\mathbf{Q}$

$$\frac{\partial}{\partial q} \{ \log_2 |\mathbf{I}_{n_{\min}} + \mathbf{Q}| - \lambda \text{tr}\{\mathbf{Q}(\mathbf{S}^\dagger \mathbf{S})^{-1}\} \} = 0 \quad (8)$$

where  $\lambda$  is the undetermined parameter associated with the Lagrangian constraint. Evaluating the derivative

$$\text{tr} \left\{ (\mathbf{I}_{n_{\min}} + \mathbf{Q})^{-1} \frac{\partial \mathbf{Q}}{\partial q} \right\} - \text{tr} \left\{ (\mathbf{S}^\dagger \mathbf{S})^{-1} \frac{\partial \mathbf{Q}}{\partial q} \right\} \quad (9)$$

this relationship is satisfied for all  $\partial \mathbf{Q} / \partial q$  if  $\mathbf{Q}$  is a diagonal matrix given by

$$\mathbf{Q} = \frac{1}{\lambda} \mathbf{S}^\dagger \mathbf{S} - \mathbf{I}_{n_{\min}}. \quad (10)$$

This discussion assumes that  $\mathbf{Q}$  is full rank. The additional positive power constraint is satisfied by employing only a subset of channel modes. This intuitively satisfying but arbitrary enforcement of the positive power constraint is justified with greater precision in the Appendix. The total power is given by

$$\begin{aligned} \text{tr}\{\mathbf{Q}(\mathbf{S}^\dagger \mathbf{S})^{-1}\} &= P_o \\ &= \text{tr} \left\{ \frac{\mathbf{I}_{n_{\min}}}{\lambda} - (\mathbf{S}^\dagger \mathbf{S})^{-1} \right\} \end{aligned} \quad (11)$$

$$\frac{1}{\lambda} = \frac{P_o + \text{tr}\{(\mathbf{S}^\dagger \mathbf{S})^{-1}\}}{n_{\min}}. \quad (12)$$

The constraint  $\mathbf{P} \geq \mathbf{0}$  is enforced by employing only the top  $n_+$  modes of the  $n_{\min}$  channel modes. The optimum  $\mathbf{Q}_{IT}$  is given by

$$\mathbf{Q}_{IT} = \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad (13)$$

$$\Delta = \left( \frac{P_o + \text{tr}(\mathbf{D}^{-1})}{n_+} \right) \mathbf{D} - \mathbf{I}_{n_+} \quad (14)$$

where the entries  $d_m$  in the diagonal matrix  $\mathbf{D}$  contain the  $n_+$  top eigenvalues of  $\mathbf{S}\mathbf{S}^\dagger$  (or, equivalently, of  $\mathbf{H}\mathbf{H}^\dagger$ ). The values  $d_m$  must satisfy

$$\Delta \mathbf{D}^{-1} > \mathbf{0} \quad (15)$$

$$d_m > \frac{n_+}{P_o + \text{tr}\{\mathbf{D}^{-1}\}}. \quad (16)$$

If (16) is not satisfied for some  $d_m$ , it will not be satisfied for any smaller  $d_m$ . The resulting capacity is given by

$$C_{IT} = \log_2 \left| \frac{P_o + \text{tr}\{\mathbf{D}^{-1}\}}{n_+} \mathbf{D} \right|. \quad (17)$$

The receive and transmit beamforming pairs are given by the columns of  $\mathbf{U}$  and  $\mathbf{W}$  associated with the selected eigenvalues contained in  $\mathbf{D}$ .

In this discussion, it is assumed that the environment is stationary over a period long enough for the error associated with channel estimation to vanish asymptotically. In order to study typical performance of quasistationary channels sampled from a given probability distribution, capacity is averaged over an ensemble of quasistationary environments. Under the ergodic assumption (that is, the ensemble average is equal to the time average), the mean capacity  $\langle C_{IT} \rangle$  is the channel capacity.

### B. Uninformed Transmitter (UT)

If the channel is not known at the transmitter, then the optimal transmission strategy is to transmit equal power with each

antenna  $\mathbf{P} = P_o/n_{Tx} \mathbf{I}_{n_{Tx}}$ , [7]. Assuming that the receiver can accurately estimate the channel but the transmitter does not attempt to optimize its output to compensate for the channel, the maximum spectral efficiency is given by

$$C_{UT} = \log_2 \left| \mathbf{I}_{n_{Rx}} + \frac{P_o}{n_{Tx}} \mathbf{H}\mathbf{H}^\dagger \right|. \quad (18)$$

This a common transmit constraint as it may be difficult to provide the transmitter channel estimates.

### C. Capacity Ratio $C_{IT}/C_{UT}$

At high SNR,  $C_{IT}$  and  $C_{UT}$  converge. This can be observed in the large  $P_o$  limit of the ratio of (17) and (18)

$$\begin{aligned} \frac{C_{IT}}{C_{UT}} &\rightarrow \frac{\log_2 \left| \frac{P_o + \text{tr}\{(\mathbf{S}^\dagger \mathbf{S})^{-1}\}}{n_{\min}} \mathbf{S}^\dagger \mathbf{S} \right|}{\log_2 \left| \mathbf{I}_{n_{\min}} + \frac{P_o}{n_{Tx}} \mathbf{S}^\dagger \mathbf{S} \right|} \\ &\rightarrow \frac{\log_2(P_o) - \log_2(n_{\min}) + \frac{\log_2 |\mathbf{S}^\dagger \mathbf{S}|}{n_{\min}}}{\log_2(P_o) - \log_2(n_{Tx}) + \frac{\log_2 |\mathbf{S}^\dagger \mathbf{S}|}{n_{\min}}} \\ &\rightarrow 1. \end{aligned} \quad (19)$$

If  $n_{Tx} > n_{Rx}$ , then the convergence to one is logarithmically slow.

At low SNR, the ratio  $C_{IT}/C_{UT}$  is given by

$$\begin{aligned} \frac{C_{IT}}{C_{UT}} &\rightarrow \frac{\log_2[(P_o + 1/d_{\max})d_{\max}]}{\log_2 \left| \mathbf{I}_{n_{Rx}} + \frac{P_o}{n_{Tx}} \mathbf{H}\mathbf{H}^\dagger \right|} \\ &= \frac{\log(1 + P_o \text{maxeig}\{\mathbf{H}\mathbf{H}^\dagger\})}{\text{tr} \left\{ \log \left( \mathbf{I}_{n_{Rx}} + \frac{P_o}{n_{Tx}} \mathbf{H}\mathbf{H}^\dagger \right) \right\}} \\ &\approx \frac{\text{maxeig}\{\mathbf{H}\mathbf{H}^\dagger\}}{\frac{1}{n_{Tx}} \text{tr}\{\mathbf{H}\mathbf{H}^\dagger\}} \end{aligned} \quad (20)$$

using (17) with  $n_+ = 1$  and (18). Given this asymptotic result, a few observations can be made. The spectral efficiency ratio is given by the maximum to the average eigenvalue ratio of  $\mathbf{H}^\dagger \mathbf{H}$ . If the channel is rank one, such as in the case of a multiple-input single-output (MISO) system, the ratio is approximately equal to  $n_{Tx}$ . Finally, in the special case where  $\mathbf{H}^\dagger \mathbf{H}$  has a flat eigenvalue distribution, the optimal transmit covariance matrix is not unique. Nonetheless, the ratio  $C_{IT}/C_{UT}$  approaches one.

### D. Interference

Extending the previous discussion [8], [23], capacity is calculated in the presence of uncooperative (worst-case) external interference  $\boldsymbol{\eta}$  in addition to the spatially-white complex Gaussian noise  $\mathbf{n}$ , which was considered previously. The mutual information is again given by (2) and (3), where entropy  $h(\mathbf{z}|\mathbf{x}, \mathbf{H})$  in the presence of the external interference becomes  $h(\mathbf{n} + \boldsymbol{\eta})$

$$h(\mathbf{z}|\mathbf{x}, \mathbf{H}) \leq \log_2 \{ \pi e |\sigma_n^2 \mathbf{I} + \sigma_n^2 \mathbf{R}| \} \quad (21)$$

and  $\sigma_n^2 \mathbf{R}$  is the spatial interference covariance matrix. Equality is achieved if and only if the interference amplitudes have a

Gaussian distribution. Thus, the worst-case *informed* capacity (the maximum–minimum mutual information)

$$C_{int} = \sup_{p(\mathbf{z}|\mathbf{H})} \inf_{p(\boldsymbol{\eta})} \mathcal{I}(\mathbf{z}, \mathbf{x}|H) \quad (22)$$

becomes

$$C_{IT,int} = \sup_{\tilde{\mathbf{P}}; \text{tr}(\tilde{\mathbf{P}})=P_o} \log_2 \left| \mathbf{I} + \tilde{\mathbf{H}}\tilde{\mathbf{P}}\tilde{\mathbf{H}}^\dagger \right| \quad (23)$$

using

$$\tilde{\mathbf{H}} \equiv (\mathbf{I} + \mathbf{R})^{-1/2} \mathbf{H}. \quad (24)$$

Gaussian interference corresponds to a saddle point of the mutual information at which the maximum–minimum capacity is achieved. The capacity in the presence of Gaussian interference has a form identical to (17) under the transformation  $\mathbf{D} \rightarrow \tilde{\mathbf{D}}$ , where  $\tilde{\mathbf{D}}$  contains the eigenvalues of  $\tilde{\mathbf{H}}\tilde{\mathbf{H}}^\dagger$ . The transmitted noise-normalized power covariance matrix  $\tilde{\mathbf{P}}$  is calculated using  $\tilde{\mathbf{H}}$ . Similarly, the *uninformed transmitter* spectral efficiency bound in the presence of noise is given by the same transformation of  $\mathbf{H} \rightarrow \tilde{\mathbf{H}}$ .

### E. Other Performance Metrics

The information-theoretic capacity is not the only possible metric of performance. As an example, another useful performance metric is the “outage capacity” [16]. “Outage capacity” is the achievable spectral efficiency bound, assuming a given probability of error-free decoding of a frame. In many practical situations, this metric may be the best measure of performance, for example, in the case where the system can resend frames of data. However, this metric is dependent on particular system choices (allowable probability of outage and frame size). For this paper, the information theoretic capacity is employed.

## III. CHANNEL COMPLEXITY

A variety of techniques are used to simulate the channel matrix [24]. The simplest approach is to assume that all the entries in the channel matrix are sampled from identical independent complex Gaussians  $\mathbf{H} \sim \mathbf{G}$ . While this approach is convenient from the perspective of performing analytic calculations, it may provide a channel eigenvalue distribution that is too flat. At the other extreme, channels can be characterized by a diversity order [25], which is used to indicate an effective cutoff in the eigenvalue distribution induced by spatial correlation. A number of approaches that introduce spatial correlations have been suggested. One approach uses the form

$$\mathbf{H} = \mathbf{R}_{left}^{1/2} \mathbf{G} \mathbf{R}_{right}^{1/2}. \quad (25)$$

The above model results in a covariance matrix of the Kronecker product form  $\mathbf{R}_{left} \otimes \mathbf{R}_{right}^*$  for the entries in the channel matrix  $\mathbf{H}$ . This product structure can arise from a spherical Green’s function model of propagation such as that used in Section III-C, provided several additional conditions are met. First, scatterers are concentrated around (but not too close to) the transmitter and receiver. Second, multiple scattering of a particular kind (from

transmitter element to transmitter scatterer to receiver scatterer to receiver element) dominates propagation. Third, scatterers are sufficiently separated in angle when viewed by their associated array. Finally, all transmitter scatterers couple with all receiver scatterers. Ray-tracing models of urban propagation indicate that the latter assumption, in particular, is often violated.

For this discussion, three approaches will be explored:

- 1) line-of-sight toy physical model;
- 2) large dimension random matrix model;
- 3) stochastic physical single scattering model.

### A. Toy $2 \times 2$ Channel Model

Because the distribution of channel matrix eigenvalues is essential to the effectiveness of MIMO communication, a toy example is employed for the purposes of introduction. The eigenvalue distribution of a  $2 \times 2$  narrowband MIMO system in the absence of environmental scatterers is discussed. To visualize the example, one can imagine two receive and two transmit antennas located at the corners of a rectangle. The ratio of channel matrix eigenvalues can be changed by varying the shape of the rectangle. The columns of the channel matrix  $\mathbf{H}$  can be viewed as the receiver array response vectors (one vector for each transmit antenna)

$$\mathbf{H} = \sqrt{2}(a_1 \mathbf{v}_1 \quad a_2 \mathbf{v}_2) \quad (26)$$

where  $a_1$  and  $a_2$  are constants of proportionality (equal to the root mean squared transmit-to-receive attenuation for transmit antennas 1 and 2 respectively) that take into account geometric attenuation and antenna gain effects, and  $\mathbf{v}_1$  and  $\mathbf{v}_2$  are unit norm array response vectors. For the purpose of this discussion, it is assumed that  $a = a_1 = a_2$ , which is valid if the rectangle deformation does not significantly affect overall transmitter-to-receiver distances.

The capacity of the  $2 \times 2$  MIMO system is a function of the channel singular values and the total transmit power. Eigenvalues of  $\mathbf{H}\mathbf{H}^\dagger$  are given by

$$\mu_{1,2} = 2a^2(1 \pm \|\mathbf{v}_1^\dagger \mathbf{v}_2\|) \quad (27)$$

where the absolute value norm is denoted by  $\|\dots\|$ . The separation between receive array responses can be described in a convenient form in terms of generalized beamwidths [26]

$$b_{1,2} = \frac{2}{\pi} \arccos\{\|\mathbf{v}_1^\dagger \mathbf{v}_2\|\}. \quad (28)$$

For small angular separations, this definition of beamwidths closely approximates many *ad hoc* definitions for physical arrays. The eigenvalues  $\mu_1$  and  $\mu_2$  are displayed in Fig. 2 as a function of generalized beamwidth separation. When the transmit and receive arrays are small, as indicated by small separation in beamwidths, one eigenvalue is dominant. As the array apertures become larger, which is indicated by larger separation, one array’s individual elements can be resolved by the other array. Consequently, the smaller eigenvalue increases. Conversely, the larger eigenvalue decreases slightly.

Equations (16) and (17) are employed to determine the capacity for the  $2 \times 2$  system. The water-filling technique first

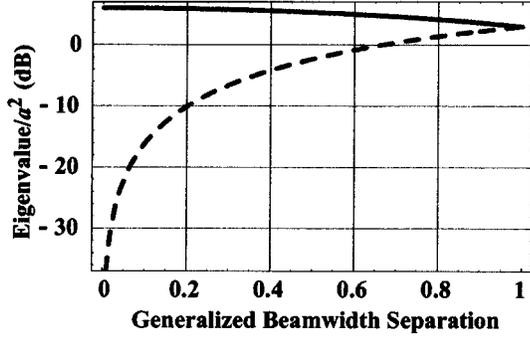


Fig. 2. Eigenvalues of  $\mathbf{H}\mathbf{H}^\dagger$  for a  $2 \times 2$  line-of-sight channel as a function of antenna separation.

must determine if both modes in the channel are employed. Both modes are used if the following condition is satisfied:

$$\begin{aligned} \mu_2 &> \frac{2}{P_o + \frac{1}{\mu_1} + \frac{1}{\mu_2}} \\ P_o &> \frac{1}{\mu_2} - \frac{1}{\mu_1} \\ &> \frac{\|\mathbf{v}_1^\dagger \mathbf{v}_2\|}{a^2(1 - \|\mathbf{v}_1^\dagger \mathbf{v}_2\|^2)} \end{aligned} \quad (29)$$

assuming  $\mu_1 > \mu_2$ .

If the condition is not satisfied, then only the stronger channel mode is employed, and the capacity, from (17), is given by

$$\begin{aligned} C_{IT} &= \log_2(1 + \mu_1 P_o) \\ &= \log_2(1 + 2a^2[1 + \|\mathbf{v}_1^\dagger \mathbf{v}_2\|]P_o). \end{aligned} \quad (30)$$

Otherwise, both modes are used, and the capacity is given by

$$\begin{aligned} C_{IT} &= \log_2 \left| \frac{P_o + \frac{1}{\mu_1} + \frac{1}{\mu_2}}{2} \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} \right| \\ &= \log_2 \left\{ \left( \frac{\mu_1 \mu_2 P_o + \mu_1 + \mu_2}{2} \right)^2 \frac{1}{\mu_1 \mu_2} \right\} \\ &= 2 \log_2 \{ a^2(1 - \|\mathbf{v}_1^\dagger \mathbf{v}_2\|^2)P_o + 1 \} \\ &\quad - \log_2 \{ 1 - \|\mathbf{v}_1^\dagger \mathbf{v}_2\|^2 \}. \end{aligned} \quad (31)$$

The resulting capacity as a function of  $a^2 P_o$  for two beamwidth separations 0.1 and 0.9, is displayed in Fig. 3. At low  $a^2 P_o$ , the capacity associated with small beamwidth separation performs best. In this regime, capacity is linear with receive power, and small beamwidth separation increases the coherent gain. At high  $a^2 P_o$ , large beamwidth separation produces a higher capacity as the optimal MIMO system distributes the energy between modes.

In complicated multipath environments, small arrays employ scatterers to create virtual arrays of a much larger effective aperture. The effect of the scatterers on capacity depends on their number and distribution in the environment. The individual antenna elements can be resolved by the larger effective aperture produced by the scatterers. As was demonstrated in Fig. 2, the ability to resolve antenna elements is related to the number of large singular values of the channel matrix and, thus, the capacity.

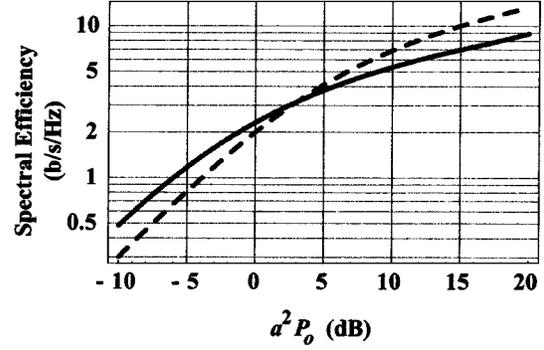


Fig. 3. Informed transmitter capacity of a  $2 \times 2$  line-of-sight channel, assuming antenna beamwidth separations of 0.1 (solid) and 0.9 (dashed).

1) *Received Power*: The choice of  $a^2 P_o$  for the horizontal axis of Fig. 3 is convenient because it can be employed to easily compare performance using different constraints and environments. This choice corresponds to the typical noise-normalized received power for a single receive and single transmit antenna radiating power  $\sigma_n^2 P_o$ . However, this choice can be mildly misleading because the total received power will, in general, be much larger than  $a^2 P_o$ . In general,  $a^2$  is defined by the Frobenius norm squared of the channel matrix normalized by the number of transmitters and receivers

$$a^2 = \frac{\text{tr}\{\mathbf{H}\mathbf{H}^\dagger\}}{n_{Tx} n_{Rx}}. \quad (32)$$

The total received noise-normalized power produced by a set of orthogonal receive beamformers is given by  $\text{tr}\{\mathbf{H}\mathbf{P}\mathbf{H}^\dagger\}$ . The *uninformed transmitter* rate is maximized by sending equal power to all transmit antennas so that  $\text{tr}\{\mathbf{H}\mathbf{P}\mathbf{H}^\dagger\}$  becomes  $P_o/n_{Tx} \text{tr}\{\mathbf{H}\mathbf{H}^\dagger\} = n_{Rx} a^2 P_o$ . It is worth noting that  $\mathbf{P}$  is not, in general, optimized by the *informed transmitter* to maximize received power but to maximize capacity. For the  $2 \times 2$  toy example, the total received power is given by  $2(1 + \|\mathbf{v}_1^\dagger \mathbf{v}_2\|)a^2 P_o$  and  $2a^2 P_o + 2\|\mathbf{v}_1^\dagger \mathbf{v}_2\|^2/(1 - \|\mathbf{v}_1^\dagger \mathbf{v}_2\|^2)$  when using one or two modes, respectively. In both cases, the total received power is much larger than  $a^2 P_o$ .

The total received power for the capacity-optimized *informed transmitter*, given an arbitrary channel matrix, is

$$\begin{aligned} \text{tr}\{\mathbf{Q}_{IT}\} &= \text{tr} \left\{ \left( \frac{P_o + \text{tr}\{\mathbf{D}^{-1}\}}{n_+} \right) \mathbf{D} - \mathbf{I}_{n_+} \right\} \\ &= P_o \frac{\text{tr}\{\mathbf{D}\}}{n_+} + \frac{\text{tr}\{\mathbf{D}^{-1}\} \text{tr}\{\mathbf{D}\} - n_+^2}{n_+} \end{aligned} \quad (33)$$

using (14). The first term in (33) is bounded from below by

$$\begin{aligned} P_o \frac{\text{tr}\{\mathbf{D}\}}{n_+} &\geq P_o \frac{\text{tr}\{\mathbf{H}\mathbf{H}^\dagger\}}{\min\{n_{Tx}, n_{Rx}\}} \\ &\geq \max\{n_{Tx}, n_{Rx}\} a^2 P_o. \end{aligned} \quad (34)$$

The second term in (33) is bounded from below by zero. Consequently, the total received power is greater than or equal to  $\max\{n_{Tx}, n_{Rx}\} a^2 P_o$ .

For very small  $a^2 P_o$ , far from the nonlinear regime of the Shannon limit, the optimal solution is to maximize received

power. This is done by transmitting the best mode only, setting  $n_+ = 1$ . In this regime, the total received power is given by

$$\text{tr}\{\mathbf{Q}_{IT}\} = P_o \text{maxeig}\{\mathbf{H}\mathbf{H}^\dagger\}. \quad (35)$$

This result is bounded from above by  $n_{Tx}n_{Rx}a^2P_o$ , which is achieved if there is only a single nontrivial mode in the channel.

### B. Large Dimension Gaussian Channel

A common channel modeling approach is to construct a matrix  $\mathbf{G}$  by independently drawing matrix elements from a unit-variance complex Gaussian distribution, mimicking independent Rayleigh fading

$$\mathbf{H} \rightarrow a\mathbf{G}. \quad (36)$$

This matrix is characterized by a relatively flat distribution of singular values and is an appropriate model for very rich multiple scattering environments.

In the limit of a large channel matrix, the eigenvalue probability density function for  $(1/n_{Tx})\mathbf{G}\mathbf{G}^\dagger$  asymptotically approaches a variant of the Wigner distribution [27]–[31]. Of course, implemented systems will have a finite number of antenna elements; however, because the shape of the typical eigenvalue distributions quickly converges to that of the asymptotic distribution, insight can be gained by considering the infinite-dimensional case. The probability that a randomly chosen eigenvalue of the  $n_{Rx} \times n_{Rx}$  matrix  $(1/n_{Tx})\mathbf{G}\mathbf{G}^\dagger$  is less than  $\mu$  is given by  $F_r(\mu)$ . Here,  $\mathbf{G}$  is an  $n_{Rx} \times n_{Tx}$  matrix, and the ratio of  $n_{Rx}$  to  $n_{Tx}$  is given by  $r = n_{Rx}/n_{Tx}$ . In the limit of  $n_{Rx} \rightarrow \infty$ , the probability measure is

$$f_r(\mu) + c_r\delta(\mu) \quad (37)$$

where the constant associated with the “delta function” at 0 is given by

$$c_r = \max\left(0, 1 - \frac{1}{r}\right). \quad (38)$$

The first term of the probability measure  $f_r(\mu)$  is given by

$$f_r(\mu) = \begin{cases} \frac{\sqrt{(\mu - a_r)(b_r - \mu)}}{2\pi\mu r}; & a_r \leq \mu \leq b_r \\ 0; & \text{otherwise} \end{cases} \quad (39)$$

where

$$a_r = (\sqrt{r} - 1)^2, \quad b_r = (\sqrt{r} + 1)^2. \quad (40)$$

The eigenvalue probability density function for this matrix expressed using a decibel scale is displayed in Fig. 4. Using the probability density function, the large matrix eigenvalue spectrum can be constructed and is depicted in Fig. 5.

1) *Uninformed Transmitter Spectral Efficiency Bound*: In the large matrix limit, the *uninformed transmitter* spectral efficiency bound, which is defined in (18) and discussed in [9] and [31], can be expressed in terms of a continuous eigenvalue distribution

$$C_{UT} = \log_2 \left| \mathbf{I}_{n_{Rx}} + \frac{P_o}{n_{Tx}} \mathbf{H}\mathbf{H}^\dagger \right|$$

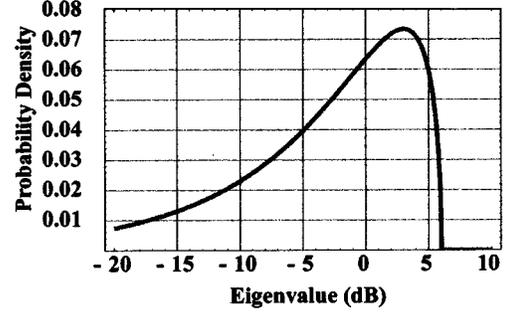


Fig. 4. Eigenvalue probability density function for the complex Gaussian channel  $((1/n_{Tx})\mathbf{G}\mathbf{G}^\dagger)$ , assuming an equal number of transmitters and receivers ( $r = 1$ ) in the infinite dimension limit.

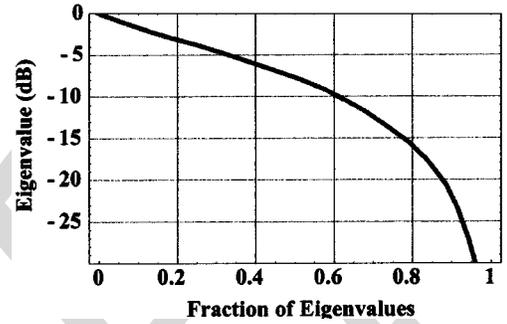


Fig. 5. Peak-normalized eigenvalue spectrum for the complex Gaussian channel  $((1/n_{Tx})\mathbf{G}\mathbf{G}^\dagger)$  assuming an equal number of transmitters and receivers ( $r = 1$ ) in the infinite dimension limit.

$$\begin{aligned} &= \log_2 \left| \mathbf{I}_{n_{Rx}} + a^2 P_o \frac{1}{n_{Tx}} \mathbf{G}\mathbf{G}^\dagger \right| \\ &\approx n_{Rx} \int_0^\infty d\mu f_r(\mu) f_r(\mu) \log_2(1 + \mu a^2 P_o) \quad (41) \end{aligned}$$

where the continuous form is asymptotically exact. This integral is discussed in [31].<sup>1</sup> The normalized asymptotic capacity as a function of  $a^2 P_o$  and  $r$ ,  $C_{UT}/n_{Rx} \approx \Phi(a^2 P_o; r)$  is given by

$$\begin{aligned} \Phi(x; r) &= \nu \left\{ \log_2 \left( \frac{x}{\nu} w_+ \right) + \frac{1-\rho}{\rho} \log_2 \left( \frac{1}{1-w_-} \right) \right. \\ &\quad \left. - \frac{w_-}{\rho \ln(2)} \right\} \\ w_\pm &= \frac{1}{2} + \frac{\rho}{2} + \frac{\nu}{2x} \pm \frac{1}{2} \sqrt{\left(1 + \rho + \frac{\nu}{x}\right)^2 - 4\rho} \\ \rho &= \min\left(r, \frac{1}{r}\right), \quad \nu = \frac{1}{\max(1, r)}. \quad (42) \end{aligned}$$

In the special case of  $M = n_{Tx} = n_{Rx}$ , the capacity is given by

$$\frac{C_{UT}}{M} \approx \frac{a^2 P_o}{\ln(2)} {}_3F_2([1, 1, 3/2], [2, 3], -4a^2 P_o) \quad (43)$$

where  ${}_pF_q$  is the generalized hypergeometric function [32].

2) *Informed Transmitter Capacity*: Similarly, in the large matrix limit, the *informed transmitter* capacity, which is defined in (17), can be expressed in terms of a continuous eigenvalue distribution [9]. To make connection with the continuous eigenvalue probability density defined in (37),  $\mathbf{D}$  from (17) is re-

<sup>1</sup>Equation (42) is expressed in terms of bits rather than nats as it is in [31].

placed with  $\mathbf{D} = a^2 n_{Tx} \mathbf{\Lambda}$ , where diagonal entries of  $\mathbf{\Lambda}$  contain the selected eigenvalues of  $(1/n_{Tx}) \mathbf{G} \mathbf{G}^\dagger$ .

$$C_{IT} = \log_2 \left| \frac{P_o + \frac{1}{a^2 n_{Tx}} \text{tr} \mathbf{\Lambda}^{-1}}{n_+} a^2 n_{Tx} \mathbf{\Lambda} \right|$$

$$\approx g n_{Rx} \log_2 \left[ \frac{a^2 n_{Tx} P_o + n_{Rx} \int_{\mu_{cut}}^{\infty} d\mu' f_r(\mu') \frac{1}{\mu'}}{g n_{Rx}} \right]$$

$$+ n_{Rx} \int_{\mu_{cut}}^{\infty} d\mu f_r(\mu) \log_2(\mu) \quad (44)$$

where  $g$  is the fraction of channel modes used by the transmitter

$$g = \frac{n_+}{n_{Rx}} \approx \int_{\mu_{cut}}^{\infty} d\mu f_r(\mu) \quad (45)$$

and  $\mu_{cut}$  is the minimum eigenvalue used by the transmitter, given by the continuous version of (16)

$$d_m = n_{Tx} a^2 \mu > \frac{n_+}{P_o + n_{Rx} \frac{1}{a^2 n_{Tx}} \int_{\mu_{cut}}^{\infty} d\mu' f_r(\mu') \frac{1}{\mu}}$$

$$\mu_{cut} = \frac{r \int_{\mu_{cut}}^{\infty} d\mu' f_r(\mu')}{a^2 P_o + r \int_{\mu_{cut}}^{\infty} d\mu' f_r(\mu') \frac{1}{\mu}}. \quad (46)$$

The approximations are asymptotically exact in the limit of large  $n_{Rx}$ .

For a finite transmit power, the capacity continues to increase as the number of antennas increases. Each additional antenna increases the effective area of the receive system. Eventually, this model breaks down as the number of antennas becomes so large that any additional antenna is electromagnetically shielded by existing antennas. However, finite random channel matrices quickly approach the shape of the infinite model. Consequently, it is useful to consider the antenna-number normalized capacity as a function of  $a^2 P_o$  and  $r$ ,  $C_{IT}/n_{Rx}$ , which is given by

$$\frac{C_{IT}}{n_{Rx}} \approx g \log_2 \left[ \frac{a^2 P_o + r \int_{\mu_{cut}}^{\infty} d\mu' f_r(\mu') \frac{1}{\mu'}}{r g} \right]$$

$$+ \int_{\mu_{cut}}^{\infty} d\mu f_r(\mu) \log_2(\mu). \quad (47)$$

Using the asymptotic eigenvalue probability density function given in (39), the integrals in (46) and (47) can be evaluated. The relatively concise results for  $r = 1$  are displayed here as

$$\int_{\mu_{cut}}^{\infty} d\mu f_1(\mu)$$

$$= 1 - \frac{\sqrt{(4 - \mu_{cut}) \mu_{cut}} + 4 \arcsin\left(\frac{\sqrt{\mu_{cut}}}{2}\right)}{2\pi} \quad (48)$$

and

$$\int_{\mu_{cut}}^{\infty} d\mu f_1(\mu) \frac{1}{\mu}$$

$$= -\frac{1}{2} + \frac{1}{\pi} \sqrt{\frac{4 - \mu_{cut}}{\mu_{cut}}} + \frac{1}{\pi} \arcsin\left(\frac{\sqrt{\mu_{cut}}}{2}\right). \quad (49)$$

To calculate the capacity, the following integral must also be evaluated:

$$\int_{\mu_{cut}}^{\infty} d\mu f_1(\mu) \log_2(\mu)$$

$$= \left\{ 4 {}_3F_2 \left( \left[ \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right], \left[ \frac{3}{2}, \frac{3}{2} \right], \frac{\mu_{cut}}{4} \right) \right.$$

$$+ \left( \sqrt{4 - \mu_{cut}} - \frac{4}{\sqrt{\mu_{cut}}} \arccsc \left[ \frac{2}{\sqrt{\mu_{cut}}} \right] \right)$$

$$\left. \times (1 - \ln[\mu_{cut}]) - \frac{2\pi}{\sqrt{\mu_{cut}}} \ln(\mu_{cut}) \right\} \frac{\sqrt{\mu_{cut}}}{\pi \ln[4]}. \quad (50)$$

Implicitly solving for  $\mu_{cut}$ , capacity as a function of  $a^2 P_o$  is displayed in Fig. 6. The *uninformed transmitter* spectral efficiency bound is plotted for comparison. For small  $a^2 P_o$ ,  $\mu_{cut}$  approaches the maximum eigenvalue supported by  $f_r(\mu)$ . In this regime, the ratio of  $C_{IT}/C_{UT}$  approaches 4. Conversely, at large  $a^2 P_o$ , the normalized *informed transmitter* and *uninformed transmitter* spectral efficiency bounds converge.

### C. Stochastic Physical Scattering Model

For many physical environments, the random channel matrix assumption may be inappropriate because it produces an eigenvalue spectrum that is overly optimistic in terms of the number of large eigenvalues. To investigate more realistic channel matrices, a simple scattering model is employed. This model was relatively successful in matching the spatial decorrelation of antenna elements measured at cellular phone frequencies and bandwidths [33]. Assuming a particular density, a field of point scatterers is generated randomly, and the channel matrix is calculated explicitly using

$$H_{m,l} \sim \sum_n \frac{e^{-2\pi i[d_{Rx,m}(n) + d_{Tx,l}(n)]}}{d_{Rx,m}(n) d_{Tx,l}(n)} \quad (51)$$

where distances  $d_{Rx,m}(n)$  and  $d_{Tx,l}(n)$  between antennas and scatterers are expressed in terms of wavelengths, and  $m, l$ , and  $n$  index the receive antenna, transmit antenna, and scatterer, respectively. The model does not include multiple scattering.

Given an ensemble of matrices constructed using this technique, the distribution of channel matrices is primarily a function of the number of transmit and receive antennas and the density of scatterers in units of  $1/L^2$ , where  $L$  is the distance between the arrays. If the field of scatterers is large compared with  $L$ , the size of the field does not overwhelm the contribution to an element in the scattering matrix. At some large distance  $R \sim d_{Rx,m} \sim d_{Tx,l}$ , the contribution of a scatterer to an entry in the channel matrix is attenuated by the inverse of the distance squared  $1/R^2$ . The number of scatterers in a differential annulus increases linearly with distance, but the effects of the scatterers combine incoherently so that the contribution grows more slowly than  $R$ , and the integrated contribution from radius  $R$  to  $\infty$  is finite.

The local distribution of antenna elements has a subtle effect on the channel SVD. As was discussed in Section III-A, the eigenvalue distribution depends on the ability of one array to resolve the individual elements of the opposing array. In the presence of scatterers, the issue is whether or not the virtual

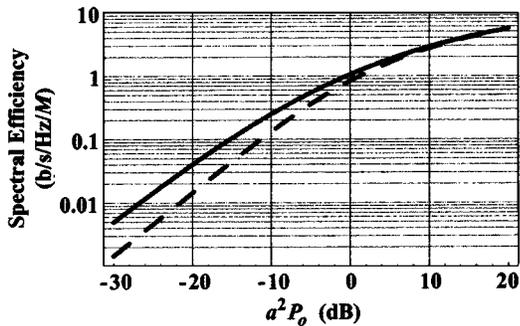


Fig. 6. Asymptotic large dimension Gaussian channel antenna-number-normalized spectral efficiency bounds  $C_{IT}/M$  (solid) and  $C_{UT}/M$  (dashed) (b/s/Hz/M) as a function of attenuated noise-normalized power ( $a^2 P_o$ ), assuming an equal number of transmitters and receivers ( $r = 1$ ,  $M = n_{Tx} = n_{Rx}$ ).

array (consisting of scatterers) can resolve the antennas in the opposing array. However, the effect is dominated by the density of scatterers. Assuming that the array is not oversampled spatially, the dependence on intra-array spacing is weak.

1) *Eigenvalue Spectrum Examples*: The sensitivity of eigenvalue spectra and capacity to variations in the dominant parameters (number of antennas and scatterer density) of the model are analyzed here. The median eigenvalues of an ensemble of eigenvalue spectra are displayed with the largest eigenvalue normalized to 0 dB in Figs. 7(a) and 8(a). The  $n$ th point in the median eigenvalue distribution indicates the median of the  $n$ th eigenvalue for each spectrum in the ensemble. The median eigenvalues are a helpful diagnostic tool but cannot be used as an input to other calculations because of correlations between eigenvalues. In Figs. 7(b) and 8(b), the corresponding capacities are displayed.

The median eigenvalue distribution as a function of the number of MIMO antenna elements is displayed in Fig. 7(a) for the same total aperture (16 wavelengths). As the number of antennas increases, in a fixed environment, the value of the smallest eigenvalue in each spectrum decreases. There are two reasons for this. First, the typical ratio of the maximum to minimum of a set of random numbers grows as the number in the set grows. Second, as the number of antennas increases, more scatterers are required to take advantage of the new degrees of freedom. The *informed transmitter* capacity ratio for each array size to the  $16 \times 16$  random matrix is displayed in Fig. 7(b). Over a wide range of SNR, the performance is a simple function of the number of antennas.

The median eigenvalue distribution as a function of scatterer density is displayed in Fig. 8(a). At low density, the relatively low number of scatterers dominate the channel matrix with strong spatial correlation at the transmit and receive arrays. This causes the eigenvalue distribution to decrease quickly. As the density of scatterers increases, the environment becomes more random, and the eigenvalue distribution of  $\mathbf{H}\mathbf{H}^\dagger$  moves closer to the random matrix distribution. However, the distribution does not converge to the random matrix distribution. In the figure, it can be observed that once the density of scatterers (in units of  $1/L^2$ ) has exceeded the number of antennas, there is little effect on the distribution. The channel matrix in the high scatterer density limit is affected by two fields of scatterers:

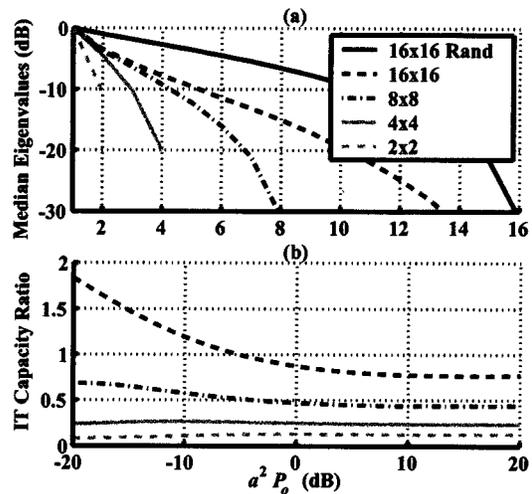


Fig. 7. (a) Median eigenvalue distribution of  $\mathbf{H}\mathbf{H}^\dagger$  for  $2 \times 2$ ,  $4 \times 4$ ,  $8 \times 8$ , and  $16 \times 16$  channel MIMO systems, assuming a dense field of scatterers ( $10/L^2$ ) and an antenna array separation  $L$ . The median eigenvalue distribution for a  $16 \times 16$  random matrix MIMO system is provided for comparison. (b) IT capacity ratio with respect to the  $16 \times 16$  random Gaussian channel.

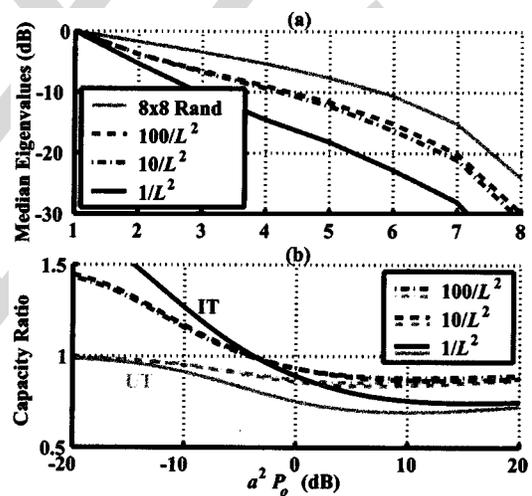


Fig. 8. (a) Median eigenvalue distributions of  $\mathbf{H}\mathbf{H}^\dagger$  for an  $8 \times 8$  channel MIMO systems, assuming scatterer densities of  $1/L^2$ ,  $10/L^2$ , and  $100/L^2$  for antenna array separations of  $L$ . The median eigenvalue distribution for an  $8 \times 8$  random matrix MIMO system is provided for comparison. (b) Capacity ratio with respect to  $8 \times 8$  random Gaussian channel.

one near the transmit array and one near the receive array. This is because at high density, there are a large number of scatterers near both the transmit and receive arrays, and the contribution increases inversely with distance. A scatterer near one of the arrays is necessarily far from the other. The field of scatterers near the transmit array is spatially uncorrelated from the transmit array's perspective, but this field of scatterers subtends a small angle from the receiver's perspective and is consequently highly correlated. Similarly, there is a dense field surrounding the receive array. These scattering fields contribute low rank components to the channel matrix. This effect competes with the much larger number of scatterers far from both arrays. The corresponding *informed transmitter* physical scatterer to the *informed transmitter* random matrix spectral efficiency bound ratios and the *uninformed transmitter*

physical scatterer to the *uninformed transmitter* random matrix capacity spectral efficiency bound ratios are displayed in Fig. 8(b). At low SNR, the relative performance of the *informed transmitter* is better in simpler environments, taking advantage of the dominant mode. At higher SNR, channels with higher complexity perform better.

#### IV. INTERFERENCE EFFECTS

##### A. Interference Model

A given MIMO communication system may be required to operate in the presence of other MIMO or wireless communication systems. This is certainly true in the case of wireless local area networks operating in the uncontrolled industrial, scientific, and medical (ISM) bands [34]. The effects of interference will be addressed using random infinite dimension and stochastic physical scattering models.

While one can certainly imagine a nearly limitless number of interference scenarios, three interference regimes are of particular interest:

- 1) a small number of strong interferers;
- 2) an uncooperative competing MIMO system;
- 3) a cooperative interfering MIMO system.

1) *Strong Interference*: In an environment populated by a relatively small number of strong interferers, the spatial whitening performed in (24) can be replaced with a projection operator, removing the spatial subspace associated with the interferers. Noting that the Hermitian interference matrix  $\mathbf{R}$  can be expressed as some power scaling multiplied by the outer product of two matrices  $\alpha\mathbf{V}\mathbf{V}^\dagger$  so that  $(\mathbf{I} + \mathbf{R})^{-1} = (\mathbf{I} + \alpha\mathbf{V}\mathbf{V}^\dagger)^{-1}$ , in the limit of high power,  $(\mathbf{I} + \mathbf{R})^{-1}$  becomes

$$\begin{aligned} \lim_{\alpha \rightarrow \infty} (\mathbf{I} + \alpha\mathbf{V}\mathbf{V}^\dagger)^{-1} &= \lim_{\alpha \rightarrow \infty} \mathbf{I} - \alpha\mathbf{V}(\mathbf{I} + \alpha\mathbf{V}^\dagger\mathbf{V})^{-1}\mathbf{V}^\dagger \\ &= \mathbf{I} - \mathbf{V}(\mathbf{V}^\dagger\mathbf{V})^{-1}\mathbf{V}^\dagger \equiv \mathcal{P}^\perp \end{aligned} \quad (52)$$

where  $\mathcal{P}^\perp$  is a projection matrix, which projects onto the complement of the column space of  $\mathbf{V}$ . Because projection matrices are idempotent, this is also the solution for  $(\mathbf{I} + \mathbf{R})^{-1/2}$ .

The strong interference-mitigated spectral efficiency bound can be written as

$$\tilde{C} = \log_2 \left| \mathbf{I}_{n_{Rx}} + a^2 \mathcal{P}^\perp \mathbf{G} \mathbf{P} \mathbf{G}^\dagger \mathcal{P}^\perp \right|. \quad (53)$$

The effect of strong interference on capacity is calculated, exploiting the fact that unitary transformation of independent identically distributed (i.i.d.) Gaussian matrices produces matrices with the same Gaussian statistics and that there exists a unitary matrix  $\mathbf{K}$  that transforms the projection matrix to a diagonal matrix with the form

$$\mathbf{K} \mathcal{P}^\perp \mathbf{K}^\dagger = \begin{pmatrix} \mathbf{I}_{n_{Rx}\gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad (54)$$

where the projection removes  $(1 - \gamma)n_{Rx}$  degrees of freedom. Using the large dimension limit discussed in Section III-B, this

bound is explicitly calculated. For the *uninformed transmitter*, the spectral efficiency bound is given by

$$\begin{aligned} \tilde{C}_{UT} &= \log_2 \left| \mathbf{I}_{\gamma n_{Rx}} + \frac{a^2 P_o}{n_{Tx}} \tilde{\mathbf{G}} \tilde{\mathbf{G}}^\dagger \right| \\ &\approx \gamma n_{Rx} \Phi(a^2 P_o; \gamma r) \end{aligned} \quad (55)$$

where  $\tilde{\mathbf{G}}$  is a  $(\gamma n_{Rx}) \times n_{Tx}$  matrix with entries sampled from a unit-norm complex Gaussian distribution,  $\Phi(x; r)$  is defined in (42), and  $r$  is defined in Section III-B. For the *informed transmitter*, the spectral efficiency bound is given by modifying (44) as

$$\begin{aligned} C_{IT} &= \log_2 \left| \frac{P_o + \frac{1}{a^2 n_{Tx}} \text{tr} \tilde{\mathbf{A}}^{-1}}{n_+} a^2 n_{Tx} \tilde{\mathbf{A}} \right| \\ &\approx \tilde{g} n_{Rx} \gamma \log_2 \left[ \frac{a^2 P_o + r \gamma \int_{\tilde{\mu}_{cut}}^{\infty} d\mu' f_{r\gamma}(\mu') \frac{1}{\mu'}}{\tilde{g} r \gamma} \right] \\ &\quad + n_{Rx} \gamma \int_{\tilde{\mu}_{cut}}^{\infty} d\mu f_{r\gamma}(\mu) \log_2(\mu) \end{aligned} \quad (56)$$

where the  $n_{Rx}\gamma$  diagonal elements of  $\tilde{\mathbf{A}}$  contain the selected eigenvalues of  $(1/n_{Tx})\tilde{\mathbf{G}}\tilde{\mathbf{G}}^\dagger$

$$\tilde{g} = \frac{n_+}{\gamma n_{Rx}} \approx \int_{\tilde{\mu}_{cut}}^{\infty} d\mu f_{r\gamma}(\mu) \quad (57)$$

and  $\tilde{\mu}_{cut}$  is given by

$$\tilde{\mu}_{cut} = \frac{r \gamma \int_{\tilde{\mu}_{cut}}^{\infty} d\mu' f_{r\gamma}(\mu')}{a^2 P_o + r \gamma \int_{\tilde{\mu}_{cut}}^{\infty} d\mu f_{r\gamma}(\mu) \frac{1}{\mu}}. \quad (58)$$

The spectral efficiency loss ratio is depicted in Fig. 9 for  $\gamma$  of 0.9 and 0.5. In the limit of large  $a^2 P_o$ , the ratio  $\tilde{C}/C$  converges to  $\gamma \max(1, r) / \max(1, r\gamma)$ .

2) *Competing MIMO Systems*: A reasonable model for the interference is to assume that it is associated with a channel matrix that is statistically independent of, but otherwise has characteristics similar to, the channel matrix associated with the intended transmitters. Using the statistical scattering model, an ensemble of channel matrix pairs is constructed. The first of the pair is associated with the intended transmitter, and the second is associated with the interfering MIMO system. Depending on the nature of the interference, the received signal can be much stronger or weaker than the intended signal. In Fig. 10, the median eigenvalue distribution is displayed for environments that contain competing MIMO systems with total interference-to-noise ratios  $\text{tr} \mathbf{R}$  of 20 and 40 dB. The eigenvalue distributions are peak-normalized in the absence of interference. The effect of interference changes the shape of the distribution and causes an overall downward shift.

In the case of the interfering MIMO system displayed in Fig. 10, the story is somewhat complicated. As one would expect, the adverse effects of the interference on the eigenvalue spectra become worse for stronger interference. Because the interfering MIMO system uses multiple transmit antennas, the interference affects all of the modes of the channel matrix. Interestingly, the loss of large eigenvalues for the sparse field matrix is less severe than that for the random channel matrices because the dominant portions of both the signal of interest

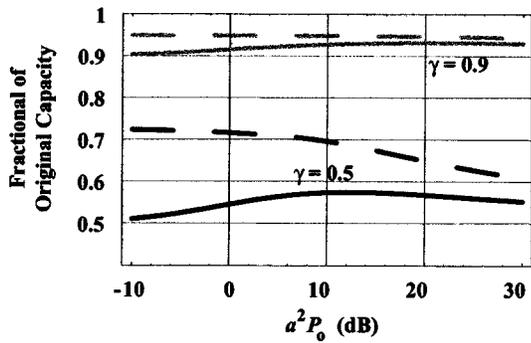


Fig. 9. Asymptotic interference loss capacity ratio  $\bar{C}_{IT}/C_{IT}$  (dashed) and  $\bar{C}_{UT}/C_{UT}$  (solid) assuming an equal number of transmitters and receivers ( $r = 1$ ) for the surviving degree of freedom fraction  $\gamma = 0.9$  (gray) and 0.5 (black).

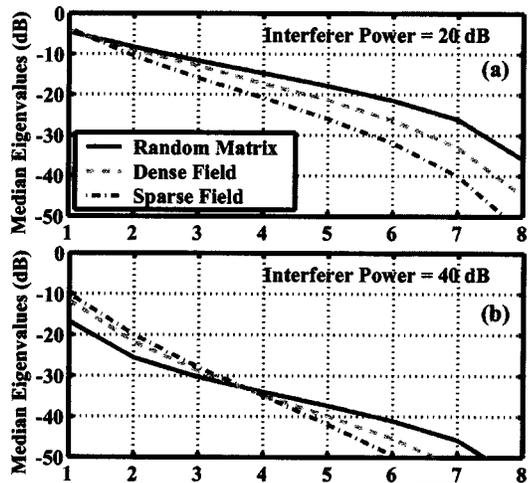


Fig. 10. Median eigenvalue distribution of  $\bar{\mathbf{H}}\bar{\mathbf{H}}^\dagger$  for an  $8 \times 8$  channel for random, dense, and sparse scattering fields, assuming random interference of (a) 20 dB and (b) 40 dB total noise-normalized power.

and the interference occupy smaller fractions of the total space in the sparse scatterer environment. This decreases the typical overlap between the associated subspaces and thus reduces the detrimental effects of mitigation. Of course, the channel matrix associated with the sparse scatterer environment had fewer useful modes to lose.

It is interesting to compare the capacity of an  $8 \times 8$  MIMO communication system with a  $1 \times 8$  SIMO system under the constraint that the total transmit power is equal. It is common to compare the capacity of MIMO systems to single-antenna transmit and receive systems. However, in the presence of strong interference, the capacity of single-to-single antenna systems is poor. The spectral efficiency bound ratio

$$\frac{\langle C(8 \times 8; \text{tr}\{\mathbf{P}\} = P_o) \rangle}{\langle C(1 \times 8; P_1 = P_o) \rangle} \quad (59)$$

is displayed in Fig. 11 for both *informed* and *uninformed transmitter* bounds, where the expectation is evaluated over an ensemble of scatterers and interferers for a given environment.

In Fig. 11, the sensitivity of MIMO capacity to environment is demonstrated. At very high SNR, the *uninformed* spectral efficiency bound and *informed transmitter* capacities converge. At low SNR, the *informed transmitter* avoids modes with small

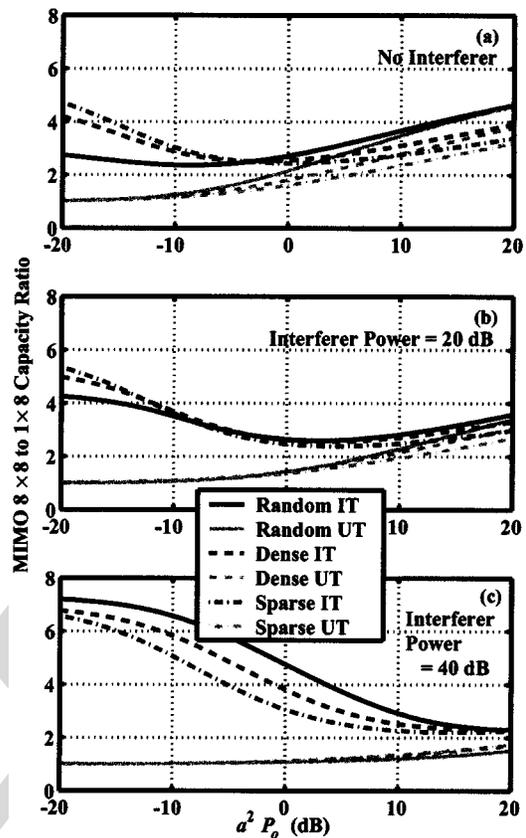


Fig. 11. Spectral efficiency bound ratio of  $8 \times 8$  MIMO to  $1 \times 8$  SIMO systems for random, dense, and sparse scattering fields, assuming (a) no interference, (b) interference of 20 dB, and (c) 40 dB total noise-normalized power for both *informed* and *uninformed transmitter*.

singular values, whereas the *uninformed transmitter* randomly spreads energy between modes. The loss is most significant for environments with relatively few large channel matrix singular values.

3) *Infinite-Dimension Competing MIMO System*: The maximum spectral efficiency for the *uninformed transmitter* in the presence of an uncooperative (worst-case) interfering MIMO system [9] is given by

$$\begin{aligned} \tilde{C}_{UT} &= \log_2 \left| \mathbf{I} + \frac{P_o}{n_{Tx}} (\mathbf{I} + \mathbf{R})^{-1} \mathbf{H} \mathbf{H}^\dagger \right| \\ &= \log_2 \left| \mathbf{I} + \mathbf{R} + \frac{P_o}{n_{Tx}} \mathbf{H} \mathbf{H}^\dagger \right| - \log_2 |\mathbf{I} + \mathbf{R}| \\ &= \log_2 \left| \mathbf{I} + \frac{P_{int}}{n_{int}} \mathbf{J} \mathbf{J}^\dagger + \frac{P_o}{n_{Tx}} \mathbf{H} \mathbf{H}^\dagger \right| \\ &\quad - \log_2 \left| \mathbf{I} + \frac{P_{int}}{n_{int}} \mathbf{J} \mathbf{J}^\dagger \right| \end{aligned} \quad (60)$$

where the noise-normalized interference plus noise covariance matrix is given by

$$\mathbf{R} = \frac{P_{int}}{n_{int}} \mathbf{J} \mathbf{J}^\dagger. \quad (61)$$

The notation  $\tilde{C}$  indicates the spectral efficiency bound in the presence of interference. The interference transmitter-to-receiver channel matrix is  $\mathbf{J} = a_{int} \mathbf{\Gamma}$ , which is similar to the

channel matrix defined in (36). The  $n_{int}$  interfering transmitters have total power  $P_{int}$ .

A particularly interesting interference environment occurs when a MIMO system attempts to operate in the presence of an uncooperative competing MIMO, where the average receive power per transmit antenna is equal for the interferer and the intended transmitter  $a^2 P_o / n_{Tx} = a_{int}^2 P_{int} / n_{int}$ . In this case, the spectral efficiency bound can be written using

$$\begin{aligned} \mathbf{A} &= (\mathbf{\Gamma} \mathbf{G}) \\ \mathbf{A}\mathbf{A}^\dagger &= \mathbf{\Gamma}\mathbf{\Gamma}^\dagger + \mathbf{G}\mathbf{G}^\dagger \end{aligned} \quad (62)$$

where the shape of  $\mathbf{A}$  is  $n_{Rx} \times (n_{Tx} + n_{int})$ . Assuming that  $\mathbf{\Gamma}$  and  $\mathbf{G}$  are independent and that the complex entries of each are selected from a unit-variance complex Gaussian distribution, as was previously assumed, the spectral efficiency bound can be expressed as

$$\begin{aligned} \tilde{C}_{UT, Eq} &= \log_2 \left| \mathbf{I} + \frac{a^2 P_o}{n_{Tx}} \mathbf{A}\mathbf{A}^\dagger \right| \\ &\quad - \log_2 \left| \mathbf{I} + \frac{a^2 P_o}{n_{Tx}} \mathbf{\Gamma}\mathbf{\Gamma}^\dagger \right| \\ &= \log_2 \left| \mathbf{I} + a^2 P_o \frac{n_{Tx} + n_{int}}{n_{Tx}} \frac{\mathbf{A}\mathbf{A}^\dagger}{n_{Tx} + n_{int}} \right| \\ &\quad - \log_2 \left| \mathbf{I} + a^2 P_o \frac{n_{int}}{n_{Tx}} \frac{\mathbf{\Gamma}\mathbf{\Gamma}^\dagger}{n_{int}} \right|. \end{aligned} \quad (63)$$

The asymptotic form of (63) can be expressed as the difference between two terms using (42) with two different sets of parameters. The maximum spectral efficiency bound in the presence of this interference for the *uninformed transmitter* is given by

$$\begin{aligned} \frac{\tilde{C}_{UT, Eq}}{n_{Rx}} &\approx \Phi \left( a^2 P_o \frac{n_{Tx} + n_{int}}{n_{Tx}}; \frac{n_{Rx}}{n_{Tx} + n_{int}} \right) \\ &\quad - \Phi \left( a^2 P_o \frac{n_{int}}{n_{Tx}}; \frac{n_{Rx}}{n_{int}} \right) \end{aligned} \quad (64)$$

where  $\Phi(x; r)$  is defined in (42). The effects of the interference for an uncooperative interfering equivalent MIMO system are displayed in Fig. 12. The effect can be significant.

4) *Cooperative MIMO Interference*: Assuming knowledge of the interfering MIMO system parameters (for example, all channel matrices) and cooperative control of the interfering users, the interference treated above can be mitigated by employing a MIMO extension to the multiuser detector (MUD) [30], increasing the capacity of each MIMO user beyond that achievable with the spatial interference cancellation alone. A simple example is provided by a system of *uninformed transmitter* MIMO users, each utilizing  $n_{Tx}$  transmitters communicating with a single receiver array fielding  $n_{Rx}$  elements. It is assumed that  $a^2 P_o$  is the same for all users, which can be achieved using power control.

The MIMO extension to the MUD spectral efficiency bound is given by the convex hull of a set of inequalities. In particular, the rates of all  $m$  users must satisfy

$$R_1 + \dots + R_m \leq \log_2 \left| \mathbf{I}_{n_{Rx}} + \sum \frac{a^2 P_o}{n_{Tx}} \mathbf{G}_k \mathbf{G}_k^\dagger \right| \quad (65)$$

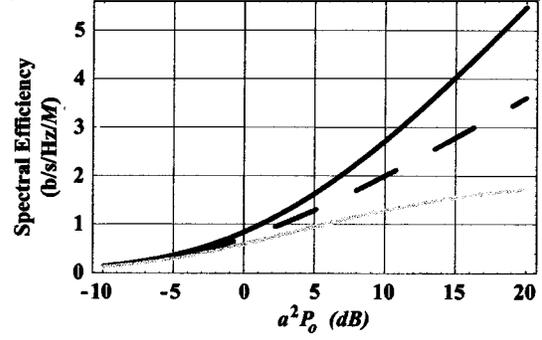


Fig. 12. Infinite-dimension antenna normalized capacity for an equal number of transmitters and receivers ( $r = 1$ ) given: no interference (black), cooperative interference (dashed), and an equivalent uncooperative interfering MIMO system (gray).

where  $R_k$  denotes the spectral efficiency of the  $k$ th user, and  $\mathbf{G}_k$  has dimensions  $n_{Rx} \times n_{Tx}$  and has i.i.d. complex Gaussian elements with zero mean and unit complex variance. Denoting  $\mathbf{A}_k = (\mathbf{G}_1 \dots \mathbf{G}_k)$ , the bound becomes

$$\log_2 \left| \mathbf{I}_{n_{Rx}} + m a^2 P_o \frac{\mathbf{A}_m \mathbf{A}_m^\dagger}{m n_{Tx}} \right| \approx n_{Rx} \Phi \left( m a^2 P_o; \frac{n_{Rx}}{m n_{Tx}} \right). \quad (66)$$

Using this relationship, the following asymptotic inequality is constructed:

$$R_1 + \dots + R_m \lesssim n_{Rx} \Phi(m a^2 P_o; r/m).$$

This asymptotic bound is achievable for a particular set of rates by a receiver employing successive interference cancellation (SIC). Recall that SIC detects signals (in this case, MIMO signals) in order, treating yet undetected signals as interference in the manner of Section IV-A3 and subtracting previously detected signals. More specifically, note that

$$\begin{aligned} &\log_2 \left| \mathbf{I}_{n_{Rx}} + m a^2 P_o \frac{\mathbf{A}_m \mathbf{A}_m^\dagger}{m n_{Tx}} \right| \\ &= \sum_{k=0}^{m-1} \left( \log_2 \left| \mathbf{I}_{n_{Rx}} + (m-k) a^2 P_o \frac{\mathbf{A}_{m-k} \mathbf{A}_{m-k}^\dagger}{(m-k) n_{Tx}} \right| \right. \\ &\quad \left. - \log_2 \left| \mathbf{I}_{n_{Rx}} + (m-k-1) \right. \right. \\ &\quad \left. \left. \cdot a^2 P_o \frac{\mathbf{A}_{m-k-1} \mathbf{A}_{m-k-1}^\dagger}{(m-k-1) n_{Tx}} \right| \right) \end{aligned}$$

where the  $k$ th term in the summand represents an achievable spectral efficiency bound  $R_k$  after the previously detected (lower  $k$ ) signals have been subtracted, and the remaining signals (higher  $k$ ) are treated as interference as in Section IV-A3. Thus, in the asymptotic limit, one can achieve  $\sum R_k \approx n_{Rx} \Phi(m a^2 P_o; r/m)$ . By averaging over all possible SIC orderings and controlling the corresponding user rates, the spectral efficiency bound is the same for all users:

$$\tilde{C}_{UT, MUD} \approx \frac{n_{Rx}}{m} \Phi(m a^2 P_o; r/m). \quad (67)$$

The spectral efficiency bound of a single user of the multiuser MIMO network described above (for two users and  $n_{Tx} = n_{Rx}$ ), given a MIMO multiuser detector as a receiver, is depicted in Fig. 12. Note that MUD receivers substantially increase capacity at higher SNRs over the capacity achieved using spatial interference cancellation alone.

## V. CHANNEL ESTIMATION ERROR

Channel estimation accuracy is limited by channel stationarity. For the purpose of this discussion, channel estimation error is modeled as a perturbing matrix  $\Sigma$  with i.i.d. elements. The estimated channel is then given by  $\hat{\mathbf{H}} \equiv \mathbf{H} + \|\mathbf{H}\|\Sigma$ . Here,  $\|\cdot\cdot\cdot\|$  indicates the Frobenius norm. The validity of this model depends on the details of the error source. It is assumed that there is no correlation between the source of error and the modes of channel matrix. While both *informed* and *uninformed transmitter* MIMO systems suffer loss as a result of channel estimation error, the *informed transmitter* suffers a loss due to using incorrect transmit spatial coding.

The losses peculiar to *informed transmitter* MIMO systems can be investigated by assuming that the receiver has an accurate estimate of the channel but that the transmitter has an inaccurate estimate. This model is reasonable for nonstationary channels. Assuming data is transmitted in blocks, the receiver can perform channel estimation using the current block of data; however, the transmitter must wait for that information to be fed back. Ignoring the possibility of channel prediction, the transmitter will employ channel estimates from a previous block. Using this estimated channel with error  $\Sigma$ , the “optimal” noise-normalized transmit covariance is constructed, solving for  $\hat{\mathbf{P}}$ , using (7) and (13), assuming the estimated channel is the true channel. As a result, the spectral efficiency bound with channel estimation error at the transmitter is given by

$$\langle C_{TxErr} \rangle = \left\langle \log_2 \left| \mathbf{I} + \mathbf{H}\hat{\mathbf{P}}\mathbf{H}^\dagger \right| \right\rangle \quad (68)$$

where the expectation is evaluated over an ensemble of scatterers and channel errors.

In Fig. 13, the fraction of the optimal capacity assuming transmit channel estimation error for  $\|\Sigma\|^2 = 0.01, 0.1, \text{ and } 1$  is displayed as a function of  $a^2 P_o$ . For this analysis, an ensemble of errors and realizations of the dense scatterer environment is used. For comparison, the spectral efficiency bound of the *uninformed transmitter* is presented. At high SNR, MIMO systems are very forgiving of transmit-channel estimation error for the same reason that the *uninformed transmitter* spectral efficiency bound approaches the optimal capacity at high SNR. At low SNR, the spectral efficiency remains remarkably insensitive to channel estimation error. Relatively few modes are used by the optimal transmitter. It is apparently difficult for random noise to significantly disturb the transmit beamformers even when the channel estimation error and the channel have the same Frobenius norm.

## VI. SUMMARY

The sensitivity of spectral efficiency bounds to environmental factors has been discussed. In Section II, the information theo-

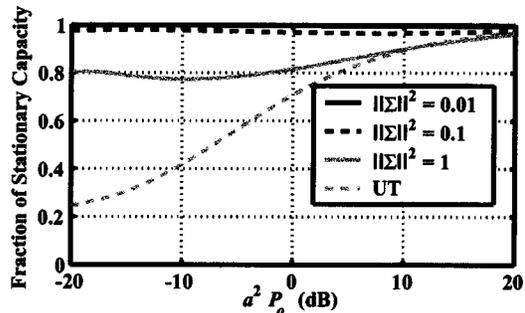


Fig. 13. Fraction of stationary capacity for an  $8 \times 8$  MIMO system with transmitter channel estimation error, assuming a dense scattering field and no interferers.

retic capacity for MIMO communication systems was reviewed for both the *informed* and *uninformed transmitter*. The spectral efficiency bounds in the presence of worst-case interference were discussed. In Section III, the complexity of channels expressed in terms of channel matrix SVDs was discussed. Line-of-sight and stochastic physical scattering models were introduced. Using the stochastic physical model, channel matrix SVDs and capacity sensitivity to the number of antennas and scatterer density were investigated. The asymptotic large Gaussian matrix channel SVD and corresponding *uninformed transmitter* spectral efficiency bound was reviewed. The corresponding *informed transmitter* capacity was introduced. In Section IV, three regimes of interference were investigated:

- 1) strong interference;
- 2) uncooperative competing MIMO system;
- 3) cooperative MIMO interference.

A strong interference asymptotic large Gaussian matrix capacity result was introduced for both the *informed* and *uninformed transmitter*. A competing MIMO interference, asymptotic large Gaussian matrix *uninformed transmitter* capacity result was introduced. Using the stochastic physical scattering model, the competing MIMO interference spectral efficiency bounds were investigated for both the *informed* and *uninformed transmitter*. Exploiting MUD, a competing cooperative MIMO interference asymptotic large Gaussian matrix *uninformed transmitter* capacity result was introduced. Finally, in Section V, the effects of channel estimation error on performance of the *informed transmitter* was investigated using the stochastic physical scattering model.

## APPENDIX

This appendix provides a more rigorous derivation of the *informed transmitter* capacity given in (17). The starting point is once again (5). Applying the generic matrix identity  $|\mathbf{I} + \mathbf{X}\mathbf{Y}| = |\mathbf{I} + \mathbf{Y}\mathbf{X}|$ , it can be rewritten as

$$C_{IT} = \sup_{\mathbf{A}; \text{tr}(\mathbf{A}\mathbf{A}^\dagger) = P_o} \log_2 |\mathbf{I}_{n_{Tx}} + \mathbf{B}\mathbf{A}\mathbf{A}^\dagger| \quad (69)$$

where  $\mathbf{B} \equiv \mathbf{H}^\dagger \mathbf{H}$ , and  $\mathbf{P} \equiv \mathbf{A}\mathbf{A}^\dagger$ . The matrices  $\mathbf{A}$  and  $\mathbf{B}$  both have dimensions  $n_{Tx} \times n_{Tx}$ . Note that positivity for  $\mathbf{P}$  is now automatic. The maximum in (69) is found by adding a Lagrange

multiplier to enforce the constraint  $\text{tr}(\mathbf{A}\mathbf{A}^\dagger) = P_o$  and differentiating with respect to the components of  $\mathbf{A}$ , yielding

$$(\mathbf{I} + \mathbf{B}\mathbf{A}\mathbf{A}^\dagger)^{-1}\mathbf{B}\mathbf{A} = \lambda\mathbf{A}. \quad (70)$$

Multiplying on the right by  $\mathbf{A}^\dagger$  produces

$$(\mathbf{I} + \mathbf{B}\mathbf{P})^{-1}\mathbf{B}\mathbf{P} = \lambda\mathbf{P} \quad (71)$$

where  $\lambda$  is the Lagrange multiplier constant that must be chosen to satisfy the constraint. Note that  $\mathbf{B}\mathbf{P}$  and  $(\mathbf{I} + \mathbf{B}\mathbf{P})^{-1}$  commute so that (71) can also be written as  $\mathbf{B}\mathbf{P}(\mathbf{I} + \mathbf{B}\mathbf{P})^{-1} = \lambda\mathbf{P}$ . Multiplying by a factor of  $(\mathbf{I} + \mathbf{B}\mathbf{P})$  on the left or right as appropriate produces

$$(\mathbf{B} - \lambda\mathbf{I} - \lambda\mathbf{B}\mathbf{P})\mathbf{P} = 0 \quad \text{and} \quad (\mathbf{B} - \lambda\mathbf{I} - \lambda\mathbf{P}\mathbf{B})\mathbf{P} = 0. \quad (72)$$

Subtracting (71) and (72) shows that

$$\lambda(\mathbf{B}\mathbf{P} - \mathbf{P}\mathbf{B})\mathbf{P} = 0. \quad (73)$$

Now, it can be shown that  $\mathbf{B}\mathbf{P} = \mathbf{P}\mathbf{B}$ . First, consider the case  $\lambda = 0$ . From (72), it follows that  $\mathbf{B}\mathbf{P} = 0$ , and thus,  $\mathbf{P}^\dagger\mathbf{B}^\dagger = \mathbf{P}\mathbf{B} = 0 = \mathbf{B}\mathbf{P}$ . In the case  $\lambda \neq 0$ , let  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n_{Tx}}$  be an eigenbasis for the Hermitian matrix  $\mathbf{P}$ . Computing the inner product of (73) between two arbitrary eigenvectors shows that

$$p_m \mathbf{v}_n^\dagger (\mathbf{B}\mathbf{P} - \mathbf{P}\mathbf{B}) \mathbf{v}_m = 0 \quad (74)$$

where  $p_m$  is the eigenvalue corresponding to  $\mathbf{v}_m$ . Taking the conjugate of the above equation and swapping  $n$  with  $m$  yields

$$p_n \mathbf{v}_n^\dagger (\mathbf{B}\mathbf{P} - \mathbf{P}\mathbf{B}) \mathbf{v}_m = 0. \quad (75)$$

If  $p_n \neq 0$  or  $p_m \neq 0$ , it follows from one of the above two equations that  $\mathbf{v}_n^\dagger (\mathbf{B}\mathbf{P} - \mathbf{P}\mathbf{B}) \mathbf{v}_m = 0$ . If  $p_n = p_m = 0$ ,  $\mathbf{v}_n^\dagger (\mathbf{B}\mathbf{P} - \mathbf{P}\mathbf{B}) \mathbf{v}_m = 0$  follows directly since  $\mathbf{P}$  annihilates both eigenvectors.

The above arguments show that the optimum value for  $\mathbf{P}$  must commute with  $\mathbf{B}$ , which means that they can be jointly diagonalized. Equation (69) for the capacity can be rewritten as

$$C_{IT} = \sup_{p_k \geq 0; P_o = \sum_k p_k} \sum_{k=1}^{n_{Tx}} \log_2(1 + b_k p_k) \quad (76)$$

where  $b_k$  and  $p_k$  are the eigenvalues of  $\mathbf{B}$  and  $\mathbf{P}$ . The optimization need only be performed with respect to the scalar values  $p_k$ , rather than the full matrix  $\mathbf{P}$ . Applying the method of Lagrange multipliers as before leads to the diagonalized analog of (72) as

$$(b_k - \lambda - \lambda b_k p_k) p_k = 0. \quad (77)$$

For each  $k$ , either  $p_k = 0$ , or  $b_k - \lambda - \lambda b_k p_k = 0$ . Define  $\mathcal{S}$  to be the set of  $k$  for which  $p_k \neq 0$  in the optimal solution. For  $k \in \mathcal{S}$ ,  $p_k = \lambda^{-1} - b_k^{-1}$ . Applying the total power constraint  $P_o = \sum_k p_k$  shows that the Lagrange multiplier must satisfy  $\lambda^{-1} = (1/|\mathcal{S}|)(P_o + \sum_{k \in \mathcal{S}} b_k^{-1})$ , where  $|\mathcal{S}|$  is the number of elements in the set  $\mathcal{S}$ .

The only remaining question follows: What elements are in  $\mathcal{S}$ ? First, suppose  $j \in \mathcal{S}$  and  $k \notin \mathcal{S}$ . It immediately follows that

$b_j \geq b_k$ ; otherwise,  $C_{IT}$  could be increased by swapping the values of  $p_j$  and  $p_k$ . Assuming that the eigenvalues are ordered so that  $b_1 \geq b_2 \geq \dots \geq b_{n_{Tx}}$ , the set  $\mathcal{S}$  must be given by  $\mathcal{S} = \{1, 2, \dots, n_+\}$  for some integer  $1 \leq n_+ \leq n_{Tx}$ . The value for  $n_+$  is determined by maximizing  $C_{IT}$  while maintaining the positivity condition  $p_k = \lambda^{-1} - b_k^{-1} \geq 0$  for  $1 \leq k \leq n_+$ . Noting that  $p_1 \geq p_2 \geq \dots \geq p_{n_+}$ , it suffices to require that  $p_{n_+} \geq 0$ .

To see which value of  $n_+$  to choose, it is useful to define the function  $C(m)$ :

$$C(m) = \sup_{P_o = \sum_k p_k(m)} \sum_{k=1}^m \log_2(1 + b_k p_k(m)). \quad (78)$$

Note that  $C(m)$  is similar to the capacity function, but there is no positivity constraint. It is clear that  $C(1) \leq C(2) \leq \dots \leq C(n_{Tx})$  since any set of values for  $p_k(m)$  can be extended to  $m+1$  by setting  $p_k(m+1) = p_k(m)$  and  $p_{m+1}(m+1) = 0$ . The optimization with respect to  $p_k(m)$  is performed using the method of Lagrange multipliers, leading to the solution  $p_k(m) = \lambda^{-1}(m) - b_k^{-1}$ . Applying the total power constraint shows that the Lagrange constant is given by  $\lambda^{-1}(m) = (1/m)(P_o + \sum_{k=1}^m b_k^{-1})$ . It follows that  $C(n_+) = C_{IT}$  since the values for  $p_k$ , where  $p_k \neq 0$ , are the same in both cases. Since the  $C(m)$  are monotonically increasing in  $m$ , we need to pick  $n_+$  to be the largest value for which  $p_{n_+}(n_+) \geq 0$ .

It is also easily shown that  $p_m(m) \geq 0$  for  $m \leq n_+$  and that  $p_m(m) < 0$  for  $m > n_+$ . First, note that  $p_1(1) = P_o > 0$ . Next, suppose  $p_m(m) < 0$  for some  $m$ . Plugging in the solution for  $p_m(m)$  gives the inequality

$$P_o + \sum_{k=1}^m b_k^{-1} < m b_m^{-1}. \quad (79)$$

Adding  $b_{m+1}^{-1}$  to both sides and factoring the right side gives

$$\begin{aligned} P_o + \sum_{k=1}^{m+1} b_k^{-1} &< (m+1)b_{m+1}^{-1} + m(b_m^{-1} - b_{m+1}^{-1}) \\ &\leq (m+1)b_{m+1}^{-1} \end{aligned}$$

which shows that  $p_{m+1}(m+1) < 0$  as well. Once  $p_m(m)$  is negative for some  $m$ , it must remain negative for all larger  $m$ .

To connect the results here to the main body of the paper, note that the eigenvalues  $b_k$  are the same as the entries  $d_k$  in the diagonal matrix  $\mathbf{D}$ . Plugging the solutions obtained for  $p_k$  into (76) leads directly to (17).

#### ACKNOWLEDGMENT

The authors would like to thank Prof. A. L. Swindlehurst of Brigham Young University, Dr. P. Hirschler-Marchand and G. Titi of MIT Lincoln Laboratory, Prof. V. Tarokh of Harvard University, and Dr. E. Young of Raytheon for their comments and contributions. The authors would also like to thank the referees for their thoughtful comments.

#### REFERENCES

- [1] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

- [2] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.
- [3] K. W. Forsythe, D. W. Bliss, and C. M. Keller, "Multichannel adaptive beamforming and interference mitigation in multiuser CDMA systems," in *Conf. Rec. Thirty-Third Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Oct. 1999, pp. 506–510.
- [4] A. Wittneben, "Base station modulation diversity for digital SIMULCAST," in *Proc. IEEE Veh. Technol. Conf.*, 1991, pp. 848–853.
- [5] V. Weerackody, "Diversity for direct-sequence spread spectrum using multiple transmit antennas," in *Proc. IEEE ICC*, vol. 3, Geneva, Switzerland, 1993, pp. 1775–1779.
- [6] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.
- [7] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.
- [8] D. W. Bliss, K. W. Forsythe, A. O. Hero, and A. L. Swindlehurst, "MIMO environmental capacity sensitivity," in *Conf. Rec. Thirty-Fourth Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Oct. 2000, pp. 764–768.
- [9] D. W. Bliss, K. W. Forsythe, and A. F. Yegulalp, "MIMO communication capacity using infinite dimension random matrix eigenvalue distributions," in *Conf. Rec. Thirty-Fifth Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2001.
- [10] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–158, Jan. 1999.
- [11] L. Zheng and D. Tse, "Optimal diversity-multiplexing tradeoff in multiple antenna fading channels," in *Conf. Rec. Thirty-Fifth Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2001.
- [12] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
- [13] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.
- [14] G. Ganesan and P. Stoica, "Space-time block codes: A maximum SNR approach," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1650–1656, May 2001.
- [15] B. Hassibi and B. Hochwald, "High-rate linear space-time codes," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 4, 2001, pp. 2461–2464.
- [16] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
- [17] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communication in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.
- [18] A. Stefanov and T. M. Duman, "Turbo coded modulation for wireless communications with antenna diversity," in *Proc. IEEE Veh. Technol. Conf.*, vol. 3, Amsterdam, The Netherlands, Sept. 1999, pp. 1565–1569.
- [19] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Full rate space-time turbo codes," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 969–980, Oct. 1998.
- [20] H. Sampath and A. J. Paulraj, "Joint transmit and receive optimization for high data rate wireless communication using multiple antennas," in *Conf. Rec. Thirty-Third Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, 1999, pp. 215–219.
- [21] N. Sharma and E. Geraniotis, "Analyzing the performance of the space-time block codes with partial channel state feedback," in *Proc. Wireless Commun. Networking Conf.*, vol. 3, 2000, pp. 1362–1366.
- [22] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [23] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, pp. 85–87, Mar. 2001.
- [24] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj, "Performance evaluation for scattering MIMO channel models," in *Conf. Rec. Thirty-Fourth Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Oct. 2000, pp. 748–752.
- [25] H. Bolcskei and A. J. Paulraj, "Performance of space-time codes in the presence of spatial fading correlation," in *Conf. Rec. Thirty-Fourth Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Oct. 2000, pp. 687–693.
- [26] K. W. Forsythe, "Utilizing waveform features for adaptive beamforming and direction finding with narrowband signals," *Mass. Inst. Technol. Lincoln Lab. J.*, vol. 10, no. 2, pp. 99–126, 1997.
- [27] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Ann. Math.*, vol. 62, no. 3, p. 548, Nov. 1955.
- [28] D. Jonsson, "Some limit theorems for the eigenvalues of a sample covariance matrix," *J. Multivariate Anal.*, vol. 12, pp. 1–38, 1982.
- [29] J. W. Silverstein, "Eigenvalues and eigenvectors of large dimensional sample covariance matrices," *Contemp. Math.*, vol. 50, pp. 153–159, 1986.
- [30] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.
- [31] P. B. Rapajic and D. Popescu, "Information capacity of a random signature multiple-input multiple-output channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1245–1248, Aug. 2000.
- [32] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. New York: Academic, 1994.
- [33] C. M. Keller and D. W. Bliss, "Cellular and PCS propagation measurements and statistical models for urban multipath on an antenna array," in *Proc. 2000 IEEE Sensor Array Multichannel Signal Process. Workshop*, Cambridge, MA, Mar. 2000, pp. 32–36.
- [34] T. S. Rappaport, *Wireless Communications: Principles & Practice*. Englewood Cliffs, NJ: Prentice Hall, 1996.



**Daniel W. Bliss** received the B.S.E.E. in electrical engineering from Arizona State University, Tempe, in 1989 and the M.S. and Ph.D. degrees in physics from the University of California at San Diego in 1995 and 1997, respectively.

He was with General Dynamics from 1989 to 1991, where he designed avionics for the Atlas-Centaur launch vehicle and performed research and development of fault-tolerant avionics. As a member of the superconducting magnet group at General Dynamics, from 1991 to 1993, he performed mag-

netic field calculations and optimization for high-energy particle accelerator superconducting magnets. His doctoral work was in the area of high-energy particle physics, searching for bound states of gluons, studying the two-photon production of hadronic final states, and investigating innovative techniques for lattice gauge theory calculations. Since 1997, he has been a staff member with the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, in the Sensor Technology Group, where he focuses on multi-antenna adaptive signal processing, primarily for communication systems, and on parameter estimation bounds, primarily for geolocation. His current research topics include algorithm development for multichannel multiuser detectors (MCMUDs) and information theoretic bounds and space-time coding for multiple-input multiple-output (MIMO) communication systems.



**Keith W. Forsythe** received the S.B. degree and the S.M. degree, both in mathematics, from the Massachusetts Institute of Technology (MIT), Cambridge.

He is a senior staff member at MIT Lincoln Laboratory, Lexington, in the Sensor Technology Group. In 1978 he joined Lincoln Laboratory, where he has worked in the areas of spread-spectrum communication, adaptive sensor array processing, and SAR imaging. His work on spread-spectrum systems includes electromagnetic modeling (geometric theory of diffraction) of antennas mounted on an airframe,

error-correction coding, jam-resistant synchronization techniques, and digital matched-filter design and performance. In the area of adaptive sensor array processing, he helped develop a number of signal processing algorithms that exploit waveform features to achieve levels of performance (beamforming, direction finding, geolocation, and other forms of parameter estimation) beyond those attainable by nonexploitive techniques. His work on SAR imaging involves techniques for resolution enhancement and interference rejection for foliage penetration systems.



**Alfred O. Hero, III** (F'98) was born in Boston, MA, in 1955. He received the B.S. degree in electrical engineering (*summa cum laude*) from Boston University in 1980 and the Ph.D degree from Princeton University, Princeton, NJ, in electrical engineering in 1984.

While at Princeton, he held the G.V.N. Lothrop Fellowship in Engineering. Since 1984, he has been a Professor with the University of Michigan, Ann Arbor, where he has appointments in the Department of Electrical Engineering and Computer Science, the

Department of Biomedical Engineering, and the Department of Statistics. He has held visiting positions at I3S at the University of Nice, Sophia-Antipolis, France, in 2001, Ecole Normale Supérieure de Lyon, Lyon, France, in 1999, Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1999, Scientific Research Labs of the Ford Motor Company, Dearborn, MI, in 1993, Ecole Nationale des Techniques Avancées (ENSTA), Ecole Supérieure d'Electricité, Paris in 1990, and the Lincoln Laboratory, Massachusetts Institute of Technology (MIT), Lexington, from 1987 to 1989. His research has been supported by NIH, NSF, AFOSR, NSA, ARO, ONR, and by private industry in the area of estimation and detection, statistical communications, signal processing, and image processing.

Dr. Hero has served as Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY. He was Chairman of the Statistical Signal and Array Processing (SSAP) Technical Committee and Treasurer of the Conference Board of the IEEE Signal Processing Society. He was Chairman for Publicity for the 1986 IEEE International Symposium on Information Theory (Ann Arbor, MI) and General Chairman of the 1995 IEEE International Conference on Acoustics, Speech, and Signal Processing (Detroit, MI). He was co-chair of the 1999 IEEE Information Theory Workshop on Detection, Estimation, Classification, and Filtering (Santa Fe, NM) and the 1999 IEEE Workshop on Higher Order Statistics (Caesaria, Israel). He is currently a member of the Signal Processing Theory and Methods (SPTM) Technical Committee and Vice President (Finance) of the IEEE Signal Processing Society. He is Chair of Commission C (Signals and Systems) of the United States National Commission of the International Union of Radio Science (URSI). He is a member of Tau Beta Pi, the American Statistical Association (ASA), and the Society for Industrial and Applied Mathematics (SIAM). He received the 1998 IEEE Signal Processing Society Meritorious Service Award, the 1998 IEEE Signal Processing Society Best Paper Award, and the IEEE Third Millennium Medal.



**Ali F. Yegulalp** received the B.A. degree in physics and mathematics from Columbia College, New York, NY, in 1990 and the Ph.D. degree in theoretical physics from Princeton University, Princeton, NJ, in 1996. His thesis work developed methods for computing exact solutions of boundary conformal field theories, which play a role in both string theory and dissipative quantum mechanics.

In 1996, he joined the Sensor Exploitation Group, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, where he has worked on

algorithm development and analysis in the areas of synthetic aperture radar (SAR), adaptive sensor array processing, and multiple-input multiple-output (MIMO) communications. His SAR work has included fast backprojection image formation, coherent change detection, autofocus, super-resolution imaging, automated target recognition, and interference/ambiguity cancellation. His work with adaptive array processing has focused mainly on new algorithms for moving target detection with multichannel SAR. Most recently, he has become involved with estimating performance bounds for MIMO systems.

IEEE  
PROOF