Efficient Use of Side Information in Multiple-Antenna Data Transmission over Fading Channels

Aradhana Narula, Michael J. Lopez, Mitchell D. Trott, and Gregory W. Wornell, *Member, IEEE*

***Abstract*— We derive performance limits for two closely related communication scenarios involving a wireless system with multiple-element transmitter antenna arrays: a point-to-point system with partial side information at the transmitter, and a broadcast system with multiple receivers. In both cases, ideal beamforming is impossible, leading to an inherently lower achievable performance as the quality of the side information degrades or as the number of receivers increases. Expected signal-to-noise ratio (SNR) and mutual information are both considered as performance measures. In the point-to-point case, we determine when the transmission strategy should use some form of beamforming and when it should not. We also show that, when properly chosen, even a small amount of side information can be quite valuable. For the broadcast scenario with an SNR criterion, we find the efficient frontier of operating points and show that even when the number of receivers is larger than the number of antenna array elements, significant performance improvements can be obtained by tailoring the transmission strategy to the realized channel.**

***Index Terms*—Antenna arrays, fading channels, feedback communication, space–time codes, spatial diversity, wireless communication.**

I. Introduction

MULTIPLE-ELEMENT transmitter antenna arrays have an increasingly important role to play in emerging wireless communication networks, particularly at base stations in cellular systems. Indeed, when used in conjunction with appropriately designed signal processing algorithms, such arrays can dramatically enhance performance.

Transmitter arrays have long been used for beamforming in radio communications. The potential for using such arrays in switched diversity schemes similar to those used for receiver arrays has also been recognized for some time [8], [9]. Beamforming methods that rely on accurate channel knowledge at the transmitter remain an active area of study from both a communication-theoretic [5] and information-theoretic [19] perspective. At the other extreme, transmitter arrays may be used in point-to-point scenarios in which the transmitter has no knowledge of the channel parameters, or equivalently in broadcast scenarios where there are infinitely many recipients [7], [20], [26]. The characterization and development of antenna coding and precoding strategies that approach the fundamental performance limits of such systems are comparatively recent; see, e.g., [3], [14], [15], [24], [27], and [28], and references therein.

In this paper, we characterize the performance limits of transmitter arrays applied to two related scenarios: a point-to-point system with partial side information at the transmitter,[1] and a broadcast system communicating common information to multiple receivers. In our analysis, we focus on two measures of performance: signal-to-noise ratio (SNR) and mutual information. While the two metrics are closely related, they have important differences. SNR characterizes the performance of typical uncoded systems, while mutual information measures the maximum rate of reliable communication achievable with coded systems (in the absence of delay and processing constraints).

When the message is intended for a single recipient, a beamforming[2] strategy is optimal [8]. With beamforming, the transmissions from the M different antenna elements at the base are designed to add coherently at the intended receiver, yielding an average factor of M enhancement of SNR and a corresponding enhancement of mutual information over single-element antenna systems [14], [15]. However, this improvement requires that the transmitter have accurate knowledge of the parameters of the channel to the intended recipient, which is difficult to achieve when the parameters are time-varying. Gains obtained in practice with only partial information at the transmitter are more modest as a result. In addition, in broadcast scenarios this factor of M enhancement cannot be obtained at each receiver even when the parameters of all channels are perfectly known. This is because it is generally not possible to simultaneously beamform to multiple recipients. We explore the degree to which it is possible to approach the factor of M beamforming performance limit in both the point-to-point and broadcast problems.

An outline of the paper is as follows. Section II describes the basic modeling assumptions for the channel. Section III considers the point-to-point scenario with a multiple-element transmit antenna and a single receive antenna. Achievable performance is characterized when only partial information

Manuscript received September 1997; revised April 1998. This work was supported in part by NSF under Grants NCR-9314341 and MIP-9502885, by ONR under Grant N00014-96-1-0930, by ARL under Cooperative Agreement DAAL01-96-2-0002, and by the USAF under Contract F19628-95-C-0002.

A. Narula is with Motorola, Inc., Mansfield, MA 02048 USA.

M. J. Lopez and G. W. Wornell are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

M. D. Trott is with ArrayComm, Inc., San Jose, CA 95134 USA.

Publisher Item Identifier S 0733-8716(98)07893-7.

[1] While this problem has not been explored in any detail to the best of our knowledge, interestingly, the dual of this scenario—a system with a single-element transmit antenna and a multiple-element receiver array, with the channel partially known at the receiver—has. See, for example, [9].

[2] We assume a narrowband plane wave propagation model in which the delay spreads at the antenna elements are small compared to the symbol period.

about the channel parameters in the form of noisy or quantized measurements is available at the transmitter. We show that even limited side information can be quite valuable, and we suggest useful ways to choose its form. When mutual information is the performance measure, we show that as the quality of the side information degrades, the transmitter should transition from a beamforming strategy to a more general approach, with the switching point dependent on the channel SNR.

Section IV develops the complementary problem in which there are multiple receivers with the associated channel parameters known at the transmitter. We find the efficient frontier of operation that balances performance at the various receivers. Among other results, we develop and analyze practical strategies whose average and worst-case (outage) performance—even when the number of receivers is relatively large—is significantly better than that of approaches that do not exploit channel information.

II. System and Channel Model

We consider a narrowband system with an M-element transmitter antenna array and L receiver antennas. At one extreme, each of these L antennas may be associated with a distinct receiver. Or, at the other extreme, they may correspond to an L-element suitably spaced antenna array for a single receiver. Regardless of the number of receivers, we assume a broadcast scenario in which a common message is transmitted to all (using potentially different symbols at each of the M antenna elements).

As the associated channel model, the complex baseband received signal at the lth receive antenna is the noisy superposition of the M transmitted symbols $X_1, X_2, \cdots, X_M$, each attenuated and phase-shifted by a complex coefficient $\alpha_{i,l}$ representing the fading encountered between transmit antenna i and receive antenna l, i.e.,

$$Y_l = \sum_{i=1}^{M} \alpha_{i,l}^* X_i + V_l, \qquad l = 1, 2, \cdots, L. \tag{1}$$

This fading model can be derived from the effects of multiple copies of the signal arriving at the receiver at slightly different times due to reflections off of objects in the transmit path [8], under the assumption that the delay spread of these arrivals is less than the symbol duration. Moreover, we consider Rayleigh fading, whereby the $\alpha_{i,l}$ are modeled as zero-mean, identically distributed circularly symmetric (or "proper" [18]) Gaussian random variables, and assume that all antennas and the constituent elements are sufficiently spatially separated that the $\alpha_{i,l}$ are all mutually independent. We use σ_α^2 to denote the (common) variance of each of the $\alpha_{i,l}$'s. Finally, in (1), V_l at each receiver captures both receiver noise and cochannel interference, and is modeled as circularly symmetric zero-mean white Gaussian noise of variance N_0, independent for each l.

For future convenience, we collect the channel parameters for the lth receiver into an M-dimensional complex vector $\boldsymbol{\alpha}_l$, i.e.,

$$\boldsymbol{\alpha}_l = [\alpha_{1,l} \;\; \alpha_{2,l} \;\; \cdots \;\; \alpha_{M,l}]^T, \qquad l = 1, 2, \cdots, L. \tag{2}$$

For the special case $L = 1$, we adopt the simpler notation $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1$.

We assume throughout that all channel parameters $\alpha_{i,l}$ are perfectly known at the respective receivers. While unrealistic in practice (since such parameters generally must be estimated from the received waveform), this assumption allows us to isolate the impact on system performance of different degrees of knowledge about the channel at the transmitter, which is the focus of the paper.

The Rayleigh model described above is most appropriate when the transmitting array is cited near many local RF scatterers, as is typically the case indoors [25]. The model also applies, for example, to an outdoor array in an urban setting when communicating with street-level users. The minimum required spacing of the antenna elements depends on the RF environment; less than one wavelength may suffice indoors, while an outdoor array may require 20 wavelengths or more [11].

Because the channel parameters are random, system performance whether measured by SNR or mutual information is also a random variable. However, although the channel parameters can remain essentially constant over transmission intervals of reasonable duration in practice, ergodic fluctuations are generally experienced over sufficiently long time scales. In such cases, then, expected SNR or expected mutual information provide the corresponding measures of average performance of interest in such systems. And it is these measures upon which we will focus.[3]

As we will see in Section III-A, the expected SNR $E[|\boldsymbol{\alpha}_l^H \boldsymbol{X}|^2]/N_0$ obtained at receiver l in such systems is determined by the system designer's choice of the second-moment statistics of the complex channel input vector

$$\boldsymbol{X} = [X_1 \;\; X_2 \;\; \cdots \;\; X_M]. \tag{3}$$

To maximize expected SNR, the designer need only determine the complex correlation matrix

$$\Gamma_{\boldsymbol{X}} = E\big[\boldsymbol{X}\boldsymbol{X}^H\big] \tag{4}$$

appropriate for the available side information.

When expected mutual information is the performance metric, it is straightforward to verify that optimum performance is achieved by a Gaussian input that is zero-mean and circularly symmetric [16]. The input distribution is likewise characterized by (4) in this case.

We therefore focus our attention throughout on the correlation matrix $\Gamma_{\boldsymbol{X}}$ and its properties in different scenarios. Among other properties, the rank of $\Gamma_{\boldsymbol{X}}$ conveys important information: observe that a system can be viewed as implementing beamforming (to some location) whenever $\Gamma_{\boldsymbol{X}}$ has rank 1. Finally, we constrain the total transmitted power according

[3] Exploiting the small z approximation $\log(1+z) \approx z$, we obtain at low SNR that

$$E[\log(1 + \mathrm{SNR}_l)] \approx E[\mathrm{SNR}_l].$$

As a result, we can also view maximizing expected SNR as effectively maximizing the rate of reliable communication in coded systems in this SNR regime.

to $\sum_{i=1}^{M} E[|X_i|^2] \le \mathcal{E}_s$, which is conveniently expressed in terms of $\Gamma_{\boldsymbol{X}}$ as

$$\mathrm{tr}(\Gamma_{\boldsymbol{X}}) \le \mathcal{E}_s. \tag{5}$$

III. Point-to-Point Transmission with Partial Side Information

For an M-element transmit antenna and single-element receive antenna ($L = 1$), we consider two different and fairly general models for the partial side information S about the channel available to the transmitter. Collectively, these two models capture the salient features of practical systems that employ side information obtained via feedback or other means.

In one case, this side information takes the form of a random vector

$$\boldsymbol{S} = [S_1 \;\; S_2 \;\; \cdots \;\; S_M]^T \tag{6}$$

where each S_m represents a noisy measurement or estimate of the corresponding α_m. We consider the case in which the pairs (α_m, S_m) for $m = 1, 2, \cdots, M$ are independent, identically distributed (i.i.d.) and each jointly circularly symmetric Gaussian. With σ_S^2 denoting the (common) variance of each of the S_i's, the complex correlation coefficient $\rho = E[\alpha_i^* S_i]/(\sigma_\alpha \sigma_S)$ then provides a complete description of the dependence between α_i and S_i.

Among other properties, these conditions imply that the posterior mean and correlation[4] of $\boldsymbol{\alpha}$ take the form

$$E[\boldsymbol{\alpha}|\boldsymbol{S}] = \frac{\sigma_\alpha}{\sigma_S}\rho \boldsymbol{S} \tag{7}$$

$$\Gamma_{\boldsymbol{\alpha}|\boldsymbol{S}} = E[\boldsymbol{\alpha}\boldsymbol{\alpha}^H|\boldsymbol{S}] = \sigma_\alpha^2(1-|\rho|^2)I_M + |\rho|^2\frac{\sigma_\alpha^2}{\sigma_S^2}\boldsymbol{S}\boldsymbol{S}^H \tag{8}$$

where I_M denotes the $M \times M$ identity matrix. As will be apparent from our development, our results apply more generally to vectors $\boldsymbol{S}$ with any distribution such that (7) and (8) hold.

As an alternative model for partial side information that is matched to other classes of practical systems, we also consider the case in which S is an N-bit description of the vector $\boldsymbol{\alpha}$ obtained via some quantization process. This scenario arises naturally when a dedicated digital feedback channel with limited bandwidth exists from receiver to transmitter. When N is large and the quantization is appropriately chosen, the two models are often effectively equivalent.

A. Optimization of Expected SNR

Since the SNR of the received signal is $\boldsymbol{\alpha}^H \Gamma_{\boldsymbol{X}} \boldsymbol{\alpha}/N_0$, the signal design problem given side information $S = s$ can be expressed as one of choosing the input correlation matrix $\Gamma_{\boldsymbol{X}}$ to maximize

$$E[\mathrm{SNR}] = \frac{E[\boldsymbol{\alpha}^H \Gamma_{\boldsymbol{X}} \boldsymbol{\alpha} | S = s]}{N_0} \tag{9}$$

subject to the power constraints (5).

[4]Note that we distinguish correlation from covariance and use the former in what follows.

Substituting (4) into the numerator of (9) and interchanging the order of expectations to obtain

$$E[\mathrm{SNR}] = \frac{E[\boldsymbol{X}^H \Gamma_{\boldsymbol{\alpha}|S} \boldsymbol{X}]}{N_0} \tag{10}$$

in terms of the posterior channel correlation matrix $\Gamma_{\boldsymbol{\alpha}|S}$, we see that the optimum input correlation matrix is the one that maximizes a quadratic form. As is well known, the maximum occurs when $\boldsymbol{X} = \tilde{X}\hat{\boldsymbol{e}}$, where $\hat{\boldsymbol{e}}$ is the (normalized) principal eigenvector associated with the principal (largest) eigenvalue $\hat{\lambda}$ of $\Gamma_{\boldsymbol{\alpha}|S}$, and where $\tilde{X}$ has second moment $E[|\tilde{X}|^2] = \mathcal{E}_s$. The rank-one correlation matrix $\Gamma_{\boldsymbol{X}} = \hat{\boldsymbol{e}}\hat{\boldsymbol{e}}^H \mathcal{E}_s$ therefore attains the maximum achievable performance, $E[\mathrm{SNR}] = \hat{\lambda}\mathcal{E}_s/N_0$. Equivalently, the optimum SNR is achieved by beamforming in a direction determined by the eigenstructure of the posterior channel correlation matrix.

It is worth stressing that beamforming is not always necessary, however. When the largest eigenvalue is not unique, correlation matrices of higher rank can also achieve the optimum SNR. For example, when S is independent of $\boldsymbol{\alpha}$, i.e., when S provides no information about the channel parameters, then all eigenvalues equal σ_α^2 and a scaled identity matrix can be used for $\Gamma_{\boldsymbol{X}}$.

In what follows, a convenient parameter that allows schemes that exploit side information to be evaluated and related is the enhancement in expected SNR provided by the channel side information, which we denote by γ

$$\gamma = \frac{\text{expected SNR given side information}}{\text{expected SNR without side information}}. \tag{11}$$

1) Noisy Side Information: Having shown that expected SNR is optimized by beamforming, we now solve explicitly for the beamforming weights and the corresponding performance when the side information has the form (6). From (8), it is straightforward to verify that the principal eigenvalue and eigenvector are given by

$$\hat{\lambda} = \sigma_\alpha^2(1-|\rho|^2) + |\rho|^2\frac{\sigma_\alpha^2}{\sigma_S^2}\|\boldsymbol{S}\|^2, \qquad \hat{\boldsymbol{e}} = \frac{\boldsymbol{S}}{\|\boldsymbol{S}\|}. \tag{12}$$

Thus, regardless of how distorted a representation the side information is of the actual channel parameters, beamforming to the location implied by the noisy side information is optimum. The transmission paths combine coherently at the true receiver location if and only if the side information is perfect ($|\rho| = 1$). When $|\rho| < 1$, imperfect combining takes place, as reflected in the resulting maximum expected SNR

$$E[\mathrm{SNR}] = \sigma_\alpha^2\left[1 + |\rho|^2(M-1)\right]\frac{\mathcal{E}_s}{N_0}. \tag{13}$$

As (13) reveals, the SNR enhancement factor γ increases monotonically from one to M as the transmitter ranges from having no knowledge of the channel parameters ($\rho = 0$) to perfect knowledge ($\rho = 1$). Consistent with the discussion in Section III-A, when $\rho = 0$, any correlation matrix with trace $\mathcal{E}_s$ achieves the same performance.

2) Quantized Side Information: We now consider the case in which the side information consists of an N-bit description of $\boldsymbol{\alpha}$ and explore how these N bits should be chosen as a function of $\boldsymbol{\alpha}$ to optimize expected SNR. The key problem is equivalent to that of vector quantization. The quantizer divides the space of channel vectors $\boldsymbol{\alpha}$ into 2^N regions $\{R_1, R_2, \cdots, R_{2^N}\}$, and, for each region, the transmitter selects the transmission strategy that maximizes expected SNR.

Since the optimal strategy involves beamforming in the direction $\boldsymbol{\beta}_i$ implied by the principal component of $E[\boldsymbol{\alpha}\boldsymbol{\alpha}^H | \boldsymbol{\alpha} \in R_i]$, the Lloyd algorithm [6] can be used to find a locally optimal set of quantization regions. In particular, given an initial partition $\{R_1, R_2, \cdots, R_{2^N}\}$, the algorithm determines the corresponding antenna weights $\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_{2^N}\}$ that maximize expected SNR. Then, given these weights, a new partition is formed by associating each possible $\boldsymbol{\alpha}$ with the direction that produces the largest SNR (breaking ties arbitrarily)

$$R_i = \{\boldsymbol{\alpha}: |\boldsymbol{\alpha}^H \boldsymbol{\beta}_i| > |\boldsymbol{\alpha}^H \boldsymbol{\beta}_j| \text{ for all } j \neq i\}. \tag{14}$$

Each iteration either increases or leaves unchanged the expected SNR; the algorithm repeats until the expected SNR converges.

The power constraint at the transmitter requires the beam directions $\{\boldsymbol{\beta}_i\}$ to have equal norm. Thus the partitioning rule (14) depends only on the direction of $\boldsymbol{\alpha}$, not on its norm. Furthermore, when the components of $\boldsymbol{\alpha}$ are i.i.d. Gaussian, the norm of $\boldsymbol{\alpha}$ remains independent of its direction even after conditioning on an event $\boldsymbol{\alpha} \in R_i$. In effect, we are quantizing the surface of the complex unit hypersphere using the special metric $d(\boldsymbol{\alpha}, \boldsymbol{\beta}) = |\boldsymbol{\alpha}^H \boldsymbol{\beta}|$, which we emphasize has characteristics very different from the more familiar Euclidean distance.

The power constraints ultimately mean that only $2M - 2$ of the $2M$ real channel parameters need be quantized, e.g., the relative magnitudes and phases of any $M - 1$ of the M complex coefficients.

i) Example—$M = 2$ transmit antenna elements: This algorithm can be used with any number of antenna elements; for the purposes of illustration we describe the associated results in the case $M = 2$, for which the maximum possible SNR enhancement factor is $\gamma = M = 2$. The $2M - 2 = 2$ relevant degrees of freedom can be expressed in terms of two angles: the relative magnitude $\phi = \tan^{-1}(|\alpha_2|/|\alpha_1|)$ and the relative phase $\theta = \measuredangle\alpha_2 - \measuredangle\alpha_1$.

In this scenario, typical random initial conditions lead to a quantization strategy with the behavior illustrated in Fig. 1. As this figure reflects, the SNR gap between perfect and no side information falls exponentially in the number of bits of side information. In particular, each additional bit of side information effectively halves the gap to the maximum possible SNR enhancement factor, i.e., $\gamma \approx 2 - 2^{-N}$.

Interestingly, the SNR enhancement factor of $\gamma = 1.5$ shown in Fig. 1 for the specific case $N = 1$ can also be obtained by either of two simple quantization strategies: the bit indicates whether $|\alpha_1| > |\alpha_2|$ or $|\alpha_1| < |\alpha_2|$ (i.e., the sign of $|\alpha_1| - |\alpha_2|$), or the bit indicates the sign of $\measuredangle\alpha_2 - \measuredangle\alpha_1$, where the relative angle ranges from $-\pi$ to π. For $N > 1$, however, the performance of schemes generated from a randomly initialized Lloyd algorithm cannot be matched with simple heuristically developed quantization strategies [16].

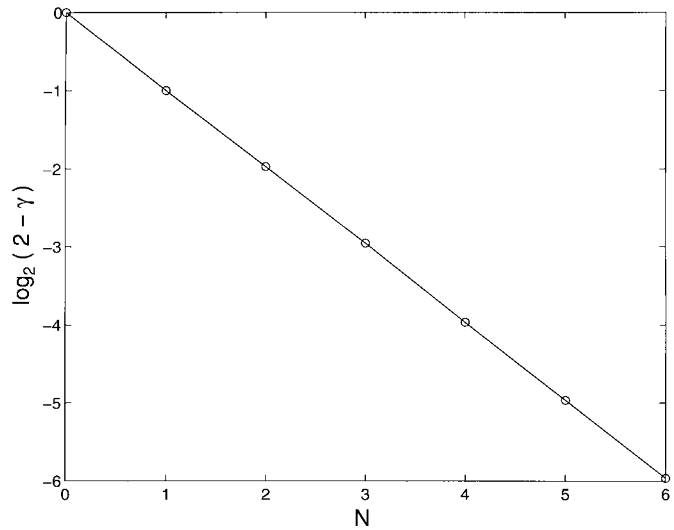

Fig. 1. Gap between perfect side information and N bits of side information using randomly generated initial partition regions.

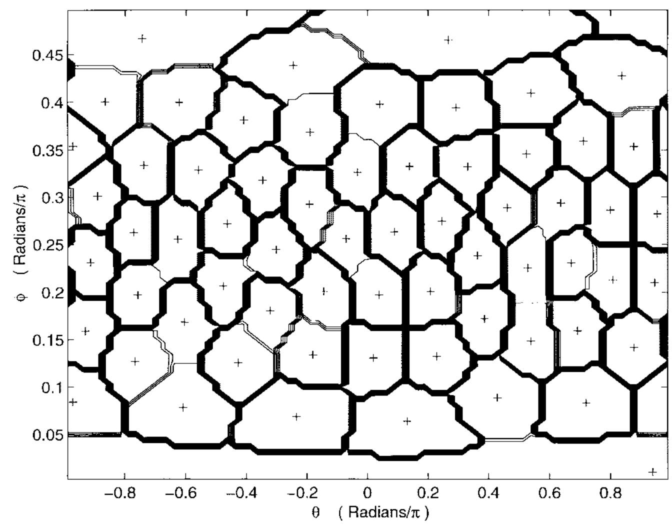

Fig. 2. Quantization regions and codebooks obtained via the Lloyd algorithm. The jagged region boundaries are an artifact of the approximations resulting from the discretization of the θ–ϕ sample space.

A typical resulting vector quantizer for $N = 6$ is shown in Fig. 2. The jagged region boundaries are an artifact of the numerical discretizations in the implementation of the algorithm. The fact that the quantization regions are smaller in the middle than near the edges reflects, in part, that points in the middle of the θ–ϕ space as drawn in Fig. 2 are more probable than those near the edges.

ii) Suboptimal strategies for $M > 2$ antenna elements: Heuristic quantization strategies can often be employed effectively in some scenarios as an alternative to strategies generated by the Lloyd algorithm. These are particularly appealing for large M since the computational requirements of the Lloyd algorithm increase rapidly with M.

One is a transmitter-based selection diversity strategy, whereby $N = \log_2 M$ bits of side information specify which of the M antenna elements is associated with the largest gain; the transmitter then uses this element exclusively. The performance of transmitter selection diversity is determined by the probability density of the gain of the element with highest gain, which is the maximum of M independent Rayleigh gains, i.e., $\max\{|\alpha_1|, |\alpha_2|, \cdots, |\alpha_M|\}$. The resulting SNR enhancement factor is [8]

$$\gamma = \sum_{k=1}^{M} \frac{1}{k} \tag{15}$$

which is well-approximated by $\log M$ when M is large. Hence, with this form of side information, SNR enhancement is effectively proportional to the number of bits of side information.

An alternative is a transmitter-based strategy realizing equal-gain combining, whereby $M-1$ bits of side information specify one bit of information about each of the relative phases of the channel coefficients. In particular, the side information specifies the sign of the principal value of the relative angles $\measuredangle\alpha_i - \measuredangle\alpha_1$, for $i = 2, \cdots, M$.

To compute the SNR enhancement achieved by this scheme, we first note that the posterior density of each $\psi_{ij} = \measuredangle\alpha_i - \measuredangle\alpha_j$ is given by, for $-\pi \le \psi \le \pi$

$$f_{\Psi|S}(\psi|s) = \begin{cases} \dfrac{1}{\pi} - \dfrac{1}{\pi^2}|\psi|, & \text{sgn}\,\measuredangle\alpha_i = \text{sgn}\,\measuredangle\alpha_j \\ f_\Psi(\psi) = \dfrac{1}{\pi^2}|\psi|, & \text{otherwise} \end{cases} \tag{16}$$

from which we see $E[|\alpha_i|^2|S] = \sigma_\alpha^2$, $E[\alpha_i\alpha_1^*|S] = \sigma_\alpha^2\sqrt{-1}/2$, and $E[\alpha_i\alpha_j^*|S] = \sigma_\alpha^2/\pi$. Thus, the associated SNR enhancement factor follows as

$$\gamma = 1 + \frac{M-1}{2\pi} + \frac{\sqrt{(M-1)^2\pi^2 + (M-2)^2}}{2\pi}. \tag{17}$$

As in the case of transmitter-based selection diversity, the SNR enhancement is also roughly linear in the number of bits of side information, although the number of bits required per antenna element and the resulting performance are significantly larger.

3) Asymptotic Results—Large N and M: When the number of bits N and the number of antennas M is large, we can use rate distortion theory to heuristically relate the quantization results in Section III-A-2 to the results on correlated side information in Section III-A-1, thereby gaining insight on the achievable performance of the vector quantizer.

Rather than directly pursuing a vector quantizer that maximizes expected SNR, assume instead that N bits are used to describe the complex M-vector $\boldsymbol{\alpha}$ using a vector quantization codebook that minimizes mean squared error

$$\sigma_\epsilon^2 = \frac{1}{M}\sum_{i=1}^{M} E\left[|\alpha_i - \beta_i(\boldsymbol{\alpha})|^2\right] \tag{18}$$

where $\boldsymbol{\beta}(\boldsymbol{\alpha})$ is the codeword used to represent $\boldsymbol{\alpha}$.

Unlike the absolute inner product metric $|\boldsymbol{\alpha}^H\boldsymbol{\beta}|$, squared Euclidean distance is a per-letter distortion measure. Rate distortion results therefore apply. Since the components of $\boldsymbol{\alpha}$ are i.i.d. complex Gaussian with variance σ_α^2, the mean squared error is lower-bounded by the corresponding distortion-rate function, i.e.,

$$\sigma_\epsilon^2 \ge D(R) = \sigma_\alpha^2 2^{-R} \tag{19}$$

where $R = N/M$ is the number of descriptive bits per antenna element. Holding R constant, this lower bound can be approached arbitrarily closely as $M \to \infty$.

When N is large, the quantization error $\boldsymbol{\epsilon} = \boldsymbol{\alpha} - \boldsymbol{\beta}(\boldsymbol{\alpha})$ may be modeled as a zero-mean Gaussian random vector, independent of $\boldsymbol{\beta}$, with i.i.d. components each having variance σ_ϵ^2. Under this assumption, $\boldsymbol{\beta}(\boldsymbol{\alpha})$ has the same form as the side information $\boldsymbol{S}$ considered in Section III-A1, where the correlation between $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ is

$$|\rho|^2 = \left|\frac{E[\alpha_i^*\epsilon_i]}{\sigma_\alpha\sigma_\epsilon}\right|^2 = 1 - \frac{\sigma_\epsilon^2}{\sigma_\alpha^2} \le 1 - 2^{-N/M}. \tag{20}$$

Therefore, for large M, from (13) the increase in expected SNR achieved by using N bits to describe the channel coefficients to minimize mean-square error is roughly

$$\gamma \approx 1 + \left(1 - 2^{-N/M}\right)(M-1). \tag{21}$$

The factor of M gap between perfect and zero side information decreases exponentially with N.

Recall from Section III-A2i that vector quantization for $M = 2$ also reduced the gap between perfect and zero side information exponentially with N. In particular, the reduction was proportional to 2^{-N}. The estimate (21) for $M = 2$ instead gives a reduction proportional to $2^{-N/2}$. This result is overly pessimistic for two reasons. First, minimizing mean-squared error rather than maximizing the absolute inner product needlessly quantizes the M absolute phases $\{\measuredangle\alpha_1, \measuredangle\alpha_2, \cdots, \measuredangle\alpha_M\}$ instead of the $M-1$ relative phases $\{\measuredangle\alpha_2 - \measuredangle\alpha_1, \measuredangle\alpha_3 - \measuredangle\alpha_1, \cdots, \measuredangle\alpha_M - \measuredangle\alpha_1\}$. Second, from our previous results, there is no need to quantize the norm of $\boldsymbol{\alpha}$. The combination of these two effects is to quantize $2M$ real parameters instead of $2M-2$. We expect the difference between the performance of the empirically trained vector quantizer and the rate-distortion approximation to be small when M is large.

B. Optimization of Expected Mutual Information

Since the mutual information of the channel is $\log(1 + \boldsymbol{\alpha}^H\Gamma_{\boldsymbol{X}}\boldsymbol{\alpha}/N_0)$, the signal design problem given side information S can be expressed as one of choosing the input correlation matrix $\Gamma_{\boldsymbol{X}}(S)$ to maximize

$$E_{\boldsymbol{\alpha}|S}\left[\log\left(1 + \frac{\boldsymbol{\alpha}^H\Gamma_{\boldsymbol{X}}(S)\boldsymbol{\alpha}}{N_0}\right)\right] \tag{22}$$

subject to the power constraints (5).

In general, this optimization is less tractable than that corresponding to the SNR criterion. However, it is straightforward to show that for the mutual information criterion—unlike for the SNR criterion—beamforming is not always optimum.

In particular, as developed in [14] and [15], when the side information provides no information about the channel parameters, the expected mutual information satisfies

$$E_{\boldsymbol{\alpha},S}\left[\log\left(1+\frac{\boldsymbol{\alpha}^H\Gamma_{\boldsymbol{X}}(S)\boldsymbol{\alpha}}{N_0}\right)\right] \le E_{\boldsymbol{\alpha}}\left[\log\left(1+\|\boldsymbol{\alpha}\|^2\frac{\mathcal{E}_s}{MN_0}\right)\right] \tag{23}$$

with equality holding if and only if $\Gamma_{\boldsymbol{X}}(S) = I_M\ \mathcal{E}_s/M$. Since this matrix obviously has full rank, beamforming cannot be used to achieve channel capacity.

There are scenarios, however, in which capacity is achieved via beamforming. For example, when the side information provides the transmitter with perfect knowledge of channel parameters, it is straightforward to show that the mutual information satisfies

$$E_{\boldsymbol{\alpha},\,S}\left[\log\left(1+\frac{\boldsymbol{\alpha}^H\Gamma_{\boldsymbol{X}}(S)\boldsymbol{\alpha}}{N_0}\right)\right] \le E_{\boldsymbol{\alpha}}\left[\log\left(1+\|\boldsymbol{\alpha}\|^2\frac{\mathcal{E}_s}{N_0}\right)\right] \tag{24}$$

with equality holding if and only if $\Gamma_{\boldsymbol{X}}(S) = \boldsymbol{\alpha}\boldsymbol{\alpha}^H\mathcal{E}_s/\|\boldsymbol{\alpha}\|^2$, i.e., if and only if the transmission paths combine coherently at the receiver.

Novel coding schemes that generate vector-valued channel symbols, such as those developed in [24], are needed to achieve the performance gains afforded by input correlation matrices of rank greater than one. As shown in [15], the cost of using a rank 1 correlation matrix when a larger rank is optimal can be as high as 0.833 bits per channel use. However, results in Section III-B1 confirm the conjecture in [15] that this gap decreases to zero as the quality of the side information increases, so that SNR-based design and mutual information-based design become equivalent.

It is difficult to develop fully general necessary and sufficient conditions for when beamforming is a capacity-achieving strategy. We develop partial answers again in the cases of our two models for side information.

1) Noisy Side Information: One approach to analyzing the scenario in which the side information takes the form of the vector (6) involves exploiting the diagonalization $\Gamma_{\boldsymbol{X}} = UDU^H$, where U is unitary and D is diagonal, and expressing $\boldsymbol{\alpha}$ in innovations form, i.e., $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}} + \boldsymbol{\epsilon}$, where $\hat{\boldsymbol{\alpha}} = \boldsymbol{S}\rho\sigma_\alpha/\sigma_S$ and where

$$\boldsymbol{\epsilon} = [\epsilon_1\ \ \epsilon_2\ \ \cdots\ \ \epsilon_M]^T \tag{25}$$

is independent of $\boldsymbol{S}$ and has i.i.d. Gaussian components each with variance $\sigma_\epsilon^2 = \sigma_\alpha^2(1-|\rho|^2)$. This allows the optimization to be expressed in the form

$$\max_{U,\,D:\,\mathrm{tr}(D)\le\mathcal{E}_s} E_{\boldsymbol{\epsilon}|\hat{\boldsymbol{\alpha}}}\left[\log\left(1+\frac{(\hat{\boldsymbol{\alpha}}+\boldsymbol{\epsilon})^H UDU^H(\hat{\boldsymbol{\alpha}}+\boldsymbol{\epsilon})}{N_0}\right)\right]. \tag{26}$$

In the remainder of this section we use this formulation to develop a condition under which beamforming is optimal for a transmitter antenna array with $M = 2$ elements. For $M = 2$, expanding the quadratic in (26) and exploiting that $U\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}$ are identically distributed we can write the optimization in the form as shown in (27) at the bottom of the page.

Both U and D in (27) can each be described via a single parameter. We can use the parameterization

$$\hat{\boldsymbol{\alpha}}^H U = \left[\sqrt{\beta}\|\hat{\boldsymbol{\alpha}}\|\ \ \sqrt{1-\beta}\|\hat{\boldsymbol{\alpha}}\|\right]^H \tag{28}$$

in terms of a scalar $\beta \in [0,\,1]$ because all the numerator terms in (27) have distributions that are independent of the phases of the components of $\hat{\boldsymbol{\alpha}}^H U$. The parameter β describes the principal eigenvector of $\Gamma_{\boldsymbol{X}}$; the larger β, the more closely this eigenvector is to being aligned with the beamforming direction associated with the side information. Likewise, we can write

$$D = \mathcal{E}_s\begin{bmatrix}\lambda & 0\\ 0 & 1-\lambda\end{bmatrix} \tag{29}$$

in terms of a single parameter $\lambda \in [0,\,1]$. The parameter λ describes how much power is transmitted in the direction of the principal eigenvector; the closer λ is to one, the better beamforming approximates the optimum solution. Moreover, $\beta = \lambda = 1$ corresponds to an exact beamforming solution.

The optimization problem (27) now reduces to

$$\begin{aligned}\max_{0\le\lambda,\,\beta\le1}\ & E_{\{\epsilon_1,\,\epsilon_2\}}\\ &\cdot\left[\log\left(1+\frac{\mathcal{E}_s}{N_0}\left\{[\lambda\beta+(1-\lambda)(1-\beta)]\|\hat{\boldsymbol{\alpha}}\|^2\right.\right.\right.\\ &\quad+\lambda|\epsilon_1|^2+(1-\lambda)|\epsilon_2|^2+2\sqrt{\beta}\lambda\|\hat{\boldsymbol{\alpha}}\|\,\mathrm{Re}\{\epsilon_1\}\\ &\quad\left.\left.\left.+2\sqrt{1-\beta}(1-\lambda)\|\hat{\boldsymbol{\alpha}}\|\,\mathrm{Re}\{\epsilon_2\}\right\}\right)\right]\end{aligned} \tag{30}$$

where $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the real and imaginary parts, respectively, of their arguments.

Note that the optimum parameters are generally not unique: if $\lambda_\circ$, $\beta_\circ$ achieve (30), so do $1-\lambda_\circ$, $1-\beta_\circ$. Hence, it suffices to restrict our attention to the range $1/2 \le \lambda,\,\beta \le 1$.

In general, the optimization (30) can be performed numerically, and simulations of this type confirm that the maximum is achieved when $\beta = 1$, as intuition suggests. Assuming $\beta = 1$ is optimum more generally, we can derive a sufficient condition on the side information quality metric $|\rho|$ for beamforming to be optimal.

We begin by introducing some simplifying notation. Let $R = \|\hat{\boldsymbol{\alpha}}\|^2$ and $\nu = \mathcal{E}_s/N_0$. With $\beta = 1$, the maximum expected mutual information is

$$\max_{0\le\lambda\le1} E[\log(1+\lambda A+(1-\lambda)B)] \tag{31}$$

where $A = \nu(\mathrm{Re}\{\epsilon_1\}+\sqrt{R})^2+\nu\mathrm{Im}\{\epsilon_1\}^2$ and $B = \nu(\mathrm{Re}\{\epsilon_2\}^2+\mathrm{Im}\{\epsilon_2\}^2)$.

The following general result, proved in the Appendix, characterizes the behavior of (31) using easily computed second moment properties of A and B.

$$\max_{U,\,D:\,\mathrm{tr}(D)\le\mathcal{E}_s} E_{\boldsymbol{\epsilon}|\hat{\boldsymbol{\alpha}}}\left[\log\left(1+\frac{\hat{\boldsymbol{\alpha}}^H UDU^H\hat{\boldsymbol{\alpha}}+\boldsymbol{\epsilon}^H D\boldsymbol{\epsilon}+\hat{\boldsymbol{\alpha}}^H UD\boldsymbol{\epsilon}+\boldsymbol{\epsilon}^H DU^H\hat{\boldsymbol{\alpha}}}{N_0}\right)\right] \tag{27}$$

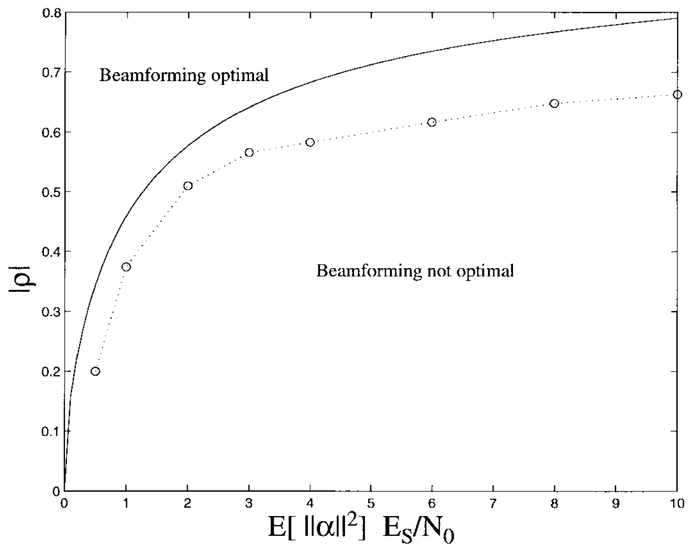

Fig. 3. Regions where a rank 1 (beamforming) or rank 2 (not beamforming) input correlation matrix maximize expected mutual information, as a function of the SNR $\tilde{\nu} = E\|\boldsymbol{\alpha}\|^2\mathcal{E}_s/N_0$ and the side information quality $|\rho|$.

Lemma 1: Let A and B be independent nonnegative random variables with nonzero variance. Then a sufficient condition for the maximum in (31) to be achieved at $\lambda = 1$ is

$$\frac{E[A]}{1+\dfrac{\sigma_A^2}{E[A]}} \geq E[B]. \qquad (32)$$

In other words, if the mean of A is greater than the mean of B (i.e., if $R > 0$) and if the variance of A is not too large, a rank 1 correlation matrix maximizes expected mutual information. To convert this into a condition on $|\rho|$ is a matter of some algebra.

The random variable A is a noncentral χ^2 distribution with two degrees of freedom. Its mean and variance [10] are

$$E[A] = \nu\left(R+\sigma_\epsilon^2\right) \qquad \sigma_A^2 = \nu^2\sigma_\epsilon^2\left(\sigma_\epsilon^2+2R\right). \qquad (33)$$

Similarly, the expected value of B is

$$E[B] = \nu\sigma_\epsilon^2. \qquad (34)$$

Substituting (33) and (34) into (32), defining $\tilde{\nu} = 2\sigma_\alpha^2\nu = E[||\boldsymbol{\alpha}||^2]\mathcal{E}_s/N_0$, and using both $R = 2(\sigma_\alpha^2 - \sigma_\epsilon^2)$ and $1-|\rho|^2 = \sigma_\epsilon^2/\sigma_\alpha^2$ yields the following sufficient condition for ensuring that the maximum expected mutual information occurs at $\lambda = 1$:

$$-3\tilde{\nu}|\rho|^6 + (5\tilde{\nu}+4)|\rho|^4 + (-\tilde{\nu}+4)|\rho|^2 - \tilde{\nu} \geq 0. \qquad (35)$$

For $\tilde{\nu} > 0$, this cubic polynomial in $|\rho|^2$ is positive between one and its greatest root less than one, which we denote by $|\hat{\rho}|^2$.

The solid line in Fig. 3 indicates $|\hat{\rho}|$ as a function of $\tilde{\nu}$. When $|\rho|$ is above this line, beamforming maximizes expected mutual information. For comparison, the dashed line indicates the result of a numerical search for the critical values of $|\rho|$ below which a rank 2 correlation matrix $\Gamma_{\boldsymbol{X}}$ yields greater expected mutual information than a rank 1 matrix. The bound (35), though not tight, has the same characteristic shape as the numerical results.

2) Quantized Side Information: The approach used in Section III-A2 to design an N-bit vector quantizer to maximize expected SNR can also be used to maximize expected mutual information. A difficulty with this approach is that there appears to be no simple expression for the correlation matrix $\Gamma_{\boldsymbol{X}}$ (of rank greater than 1 in general) that maximizes expected mutual information as a function of the region shape R_i. A numerical search for $\Gamma_{\boldsymbol{X}}$ makes the Lloyd algorithm computationally burdensome. Some suggestive results about the behavior of the Lloyd algorithm can, however, be obtained for the special case of $M = 2$ antennas and $N = 1$ bits.

The correlation matrix can be parameterized as

$$\Gamma_{\boldsymbol{X}} = \begin{bmatrix} \lambda & \sqrt{\lambda(1-\lambda)}re^{-j\hat{\theta}} \\ \sqrt{\lambda(1-\lambda)}re^{j\hat{\theta}} & (1-\lambda) \end{bmatrix}\mathcal{E}_s \qquad (36)$$

where $0 \leq \lambda,\ r \leq 1$, and $-\pi \leq \hat{\theta} \leq \pi$. Two natural possibilities for the side information are i) designating the antenna with the larger gain and ii) specifying whether the relative phase is between $-\pi$ and zero or between zero and π. In the remainder of this section, we show that rank 1 correlation matrices are optimal in both cases.

Expanding $\boldsymbol{\alpha}^H\Gamma_{\boldsymbol{X}}\boldsymbol{\alpha}$ yields maximum expected mutual information

$$\begin{aligned}\max_{\Gamma_{\boldsymbol{X}}:\,\mathrm{tr}(\Gamma_{\boldsymbol{X}})\leq\mathcal{E}_s} & E\left[\log\left(1+\boldsymbol{\alpha}^H\Gamma_{\boldsymbol{X}}\boldsymbol{\alpha}\,\frac{\mathcal{E}_s}{N_0}\right)\right] \\ &= \max_{0\leq\lambda,\,r\leq 1;\,\hat{\theta}} E_{\boldsymbol{\alpha}|S}\left[\log\left(1+\left\{\lambda|\alpha_1|^2+(1-\lambda)|\alpha_2|^2\right.\right.\right. \\ &\qquad \left.\left.\left. +2|\alpha_1||\alpha_2|r\sqrt{\lambda(1-\lambda)}\,\cos(\Theta)\right\}\frac{\mathcal{E}_s}{N_0}\right)\right] \qquad (37)\end{aligned}$$

where $\Theta = \measuredangle\alpha_1 - \measuredangle\alpha_2 + \hat{\theta}$. If S is chosen to specify the larger gain antenna, Θ is uniformly distributed between $-\pi$ and π and is independent of $|\alpha_1|$, $|\alpha_2|$. By Jensen's inequality, the expectation over Θ in (37) is maximized by choosing $r = 0$. The problem thus reduces to

$$\max_{0\leq\lambda\leq 1} E_{\{|\alpha_1|,\,|\alpha_2|\}|S}\log\left(1+\left\{\lambda|\alpha_1|^2+(1-\lambda)|\alpha_2|^2\right\}\frac{\mathcal{E}_s}{N_0}\right). \qquad (38)$$

If S specifies $|\alpha_1| > |\alpha_2|$, the expected mutual information is monotonically increasing with λ and is thus maximized when all energy is radiated from the first antenna. For $|\alpha_1| < |\alpha_2|$, the maximizing λ is zero, and all energy is radiated from the second antenna. This solution, which is equivalent to selection diversity, is thus a stationary point of the Lloyd algorithm.

If, instead, S is used to provide one bit of information about the relative phase $\measuredangle\alpha_1 - \measuredangle\alpha_2$, $\hat{\theta}$ can be chosen to keep Θ between $-\pi/2$ and $\pi/2$, so that the cosine term in (37) is always positive. Therefore, to maximize expected mutual information, r should be as large as possible: $r = 1$. Assume without loss of generality that $E[|\alpha_i|^2] = 1$. Averaging (37) over $a_1 = |\alpha_1|$ and $a_2 = |\alpha_2|$, which are Rayleigh distributed, yields

$$\max_{0\le\lambda\le 1} E_{\Theta|S} \int_{a_2=0}^{\infty}\int_{a_1=0}^{\infty} \log\Bigg(1+\Big\{\lambda a_1^2+(1-\lambda)a_2^2 + 2a_1a_2\sqrt{\lambda(1-\lambda)}\,\cos(\Theta)\Big\}\frac{\mathcal{E}_s}{N_0}\Bigg) \cdot 4a_1a_2e^{-a_1^2}e^{-a_2^2}\,da_1\,da_2. \tag{39}$$

The maximum is found by setting the derivative of (39) with respect to λ equal to zero as shown in (40) at the bottom of the page. By symmetry, (40) is satisfied at $\lambda = \frac{1}{2}$. Evaluating the second derivative confirms that this is indeed a maximum. The resulting correlation matrix is again rank 1.

Thus, for $M = 2$ and $N = 1$, the same partitioning schemes and rank 1 correlation matrices yield local maxima of the Lloyd algorithm for both expected mutual information and expected SNR. However, as we would expect from the results in Section III-B1, for other cases rank 2 correlation matrices may be optimal.

IV. Broadcast Transmission with Perfect Side Information

The problem of point-to-point transmission with partial side information is closely related to the problem of broadcast transmission, as we now develop. We explore the problem of transmitting a single message to L receivers simultaneously using an M-element transmit antenna array given complete channel knowledge at the transmitter. Moreover, we focus on expected SNR as a performance metric for the purposes of illustration, though we remark in advance that, as in earlier parts of the paper, the ideas extend naturally to the mutual information metric, though the analysis is less tractable.

For point-to-point transmission, we established the optimality of beamforming given side information for optimizing received SNR. In broadcast scenarios, it is generally not possible to simultaneously beamform to multiple receivers, so not every receiver can experience performance equivalent to that of point-to-point systems. Phrased differently, the need to share a message with multiple receivers in general reduces the average received SNR, where the average is over all receivers. Neverthless, as we will see, beamforming in an appropriately chosen direction that depends on the collective channel parameters can significantly enhance the performance of a broadcast system relative to one whose transmitter ignores the channel information.

In this section we explore the design of such systems and examine the manner in which achievable performance degrades as the number of receivers increases (and therefore knowledge of the channel parameters at the transmitter becomes less useful).

In the sequel, it is convenient to describe the set of received signals in vector form

$$\boldsymbol{Y} = [Y_1\ \ Y_2\ \ \cdots\ \ Y_L]^T = A\boldsymbol{X} + \boldsymbol{V} \tag{41}$$

where

$$\boldsymbol{V} = [V_1\ \ V_2\ \ \cdots\ \ V_L]^T \tag{42}$$

is the channel noise vector whose components are i.i.d., and where A is the $L \times M$ matrix of channel parameters

$$A = [\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \cdots\ \ \boldsymbol{\alpha}_L]^H \tag{43}$$

with $\boldsymbol{\alpha}_l$ as given in (2). The $\boldsymbol{\alpha}_l$ vectors for different receivers are independent. We restrict all the $\alpha_{i,j}$ channel coefficients to have a common variance σ_α^2, which corresponds to all receivers being at similar distances from the transmitter; extensions to more general scenarios are possible but are beyond the scope of the present paper.

A. Broadcast Performance Frontier: Transmitter Operating Characteristic

When there are multiple receivers corresponding to distinct users, the transmitter has conflicting objectives, since maximizing one receiver's SNR is typically at the expense of the SNR's of the remaining receivers. Nevertheless, as we now develop, there is a convenient way to describe the tradeoffs available to the transmitter and to assess whether the transmitter is operating efficiently. In particular, when the transmitter input correlation matrix $\Gamma_{\boldsymbol{X}}$ is selected, there is an associated point $(\mathrm{SNR}_1, \mathrm{SNR}_2, \cdots, \mathrm{SNR}_L)$ in L-dimensional "SNR-space" that describes the associated SNR's experienced at the receivers in the system. Moreover, given the transmitter power constraint, there is a well-defined surface that defines the boundary of those points that are attainable in SNR-space. We refer to this frontier of achievable points as the "transmitter operating characteristic" (TOC) for the realized channel and power constraint. As will become apparent, a transmitter operates efficiently if and only if it results in an SNR vector lying on the TOC.

To develop and illustrate these ideas further, we restrict our attention to the case of two receivers $L = 2$. Then the TOC is defined as the set of points $(\mathrm{SNR}_1, \mathrm{SNR}_2)$ for which SNR_1 is maximized subject to a constraint of the form $\mathrm{SNR}_2 \ge \eta$, for various prescribed thresholds η.

It is straightforward to verify that the curve in SNR-space formed by this set of points has some equivalent characterizations that emphasize its universality. As one example, the same curve is obtained by maximizing SNR_2 subject to an analogous constraint on SNR_1. As another example, the same curve is obtained as the locus of points obtained when the input

$$\int_{a_2=0}^{\infty}\int_{a_1=0}^{\infty} 4\,\frac{(a_1^2-a_2^2)\,\dfrac{\mathcal{E}_s}{N_0}\,a_1a_2e^{-a_1^2}e^{-a_2^2}}{1+\Big\{\lambda a_1^2+(1-\lambda)a_2^2+2a_1a_2\sqrt{\lambda(1-\lambda)}\,\cos(\Theta)\Big\}\dfrac{\mathcal{E}_s}{N_0}}\,da_1\,da_2 = 0 \tag{40}$$

correlation is chosen so as to maximize the weighted sum of the constituent SNR's, i.e., an objective function of the form

$$w_1\,\mathrm{SNR}_1 + w_2\,\mathrm{SNR}_2 \tag{44}$$

with various nonnegative weights w_1 and w_2, again subject to the system power constraint.

These various characterizations are collectively useful both in developing key properties of this curve, as well as in its numerical evaluation. For example, the former characterizations imply that the two pairs

$$\left(\|\boldsymbol{\alpha_1}\|^2\,\frac{\mathcal{E}_s}{N_0},\,\frac{|\boldsymbol{\alpha_1}^H\boldsymbol{\alpha_2}|^2}{\|\boldsymbol{\alpha_1}\|^2}\,\frac{\mathcal{E}_s}{N_0}\right)$$
$$\left(\frac{|\boldsymbol{\alpha_1}^H\boldsymbol{\alpha_2}|^2}{\|\boldsymbol{\alpha_2}\|^2}\,\frac{\mathcal{E}_s}{N_0},\,\|\boldsymbol{\alpha_2}\|^2\,\frac{\mathcal{E}_s}{N_0}\right) \tag{45}$$

which correspond to beamforming directly to each of the first and second receivers, respectively, must lie on the TOC. This follows from the fact that one of the receivers experiences the maximum possible SNR in each case, i.e., all achievable SNR pairs must lie somewhere within the rectangle $0 \le \mathrm{SNR}_1 \le \|\boldsymbol{\alpha}_1\|^2\mathcal{E}_s/N_0$, $0 \le \mathrm{SNR}_2 \le \|\boldsymbol{\alpha}_2\|^2\mathcal{E}_s/N_0$.

The weighted SNR characterization can be used to establish that all points on the TOC are achieved by beamforming toward some location. To see this, note that via an approach analogous to that used in Section III-A, the optimum $\Gamma_{\boldsymbol{X}}$ is that maximizing

$$\frac{E[\boldsymbol{X}^H A^H W A \boldsymbol{X}]}{N_0} \tag{46}$$

subject to the power constraint (5), where $W = \mathrm{diag}(w_1,\, w_2)$. Thus, the weighted sum of SNR's is maximized by beamforming, where the appropriate antenna weight vector is the principal eigenvector $\hat{\boldsymbol{e}}$ of $A^H W A$.

The TOC curve for a particular channel realization and power constraint is depicted in Fig. 4. The two circles "∘" correspond to the points (45), and the two dashed line segments that connect them to the axes are also achieved by the same respective beamforming solutions. As we would expect, the TOC must be convex: points on a line segment connecting any two points on the TOC can be achieved by a timesharing strategy, which cannot be better than the optimum beamforming strategy.

The operating points corresponding to other performance criteria of interest can also be developed from the TOC. For example, it is often desirable to maximize the *minimum* performance among receivers—thereby ensuring that both receivers achieve a sufficient quality of service—which is achieved by operating at the intersection of the TOC with the line $\mathrm{SNR}_1 = \mathrm{SNR}_2$; in Fig. 4 this point is indicated via the symbol "▽." In scenarios where the line does not intersect the curve, we operate at the nearest of the points (45). In other cases, maximizing the *average* (or, equivalently, total) SNR over all receivers is more appropriate. This is achieved by operating at the point where the TOC has slope -1; in Fig. 4 this point is indicated via the symbol "□," and corresponds to weights in (44) satisfying $w_1 = w_2$. For both of these criteria, there are natural extensions for $L > 2$.

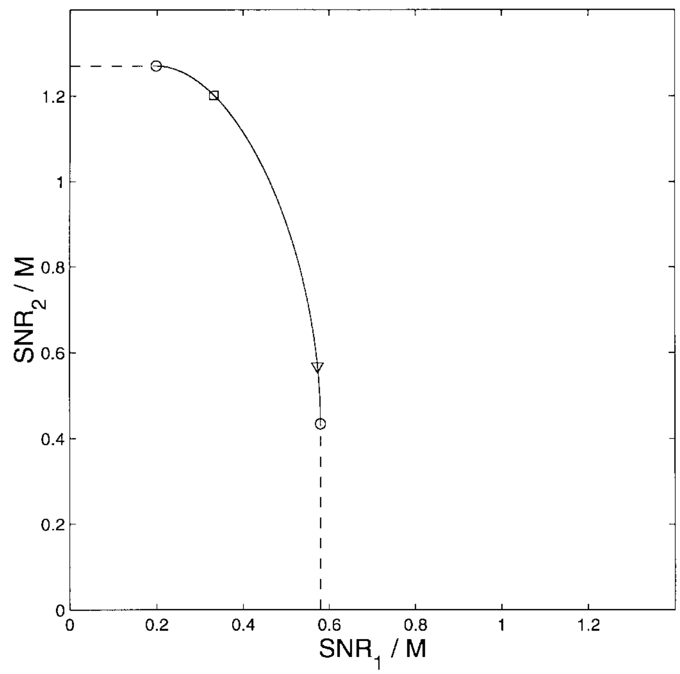

Fig. 4. A typical transmitter operating characteristic (TOC) for a ten-element transmitter antenna array broadcasting to two receivers and $\sigma_\alpha^2\mathcal{E}_s/N_0 = 1$. SNR pairs are achievable if and only if they lie on or inside the TOC. The symbols "∘" denote solutions corresponding to beamforming to one of the two receivers; the symbol "▽" denotes the point of maximum worst-case SNR among receivers; the symbol "□" denotes the point of maximum average of the two received SNR's.

While the solution that optimizes the worst-case SNR is often of most interest in practice, its performance is generally difficult to analyze. In contrast, optimizing average SNR is highly tractable, but can lead to solutions in which performance among receivers is uneven, with some receivers experiencing excellent SNR at the expense of others. However, this solution does provide an upper bound on achievable average per-receiver performance and can be interpreted as maximizing worst-case time-averaged SNR among receivers when the receivers are in motion. If such moving receivers undergo ergodic variations in the channel parameters when such a strategy is used, then each of the L receivers will achieve an equal share of the maximum total time-averaged SNR across all receivers. Additionally, in a scenario where the multiple receive antennas correspond to a single receiver in which maximal ratio combining [8] is used, this solution leads to maximum received SNR. In the sequel, we therefore focus on the maximization of average SNR.

B. Optimizing Average SNR Per Receiver

Analogous to the corresponding quantity in the point-to-point problem, it is convenient to define an SNR enhancement factor ζ in the broadcast problem. This enhancement factor is the additional average SNR per receiver obtained by tailoring the transmission strategy to the finite receiver population rather than assuming an infinite receiver population and ignoring the available side information

$$\zeta = \frac{\text{average SNR with finite receiver population}}{\text{average SNR for infinite receiver population}}. \tag{47}$$

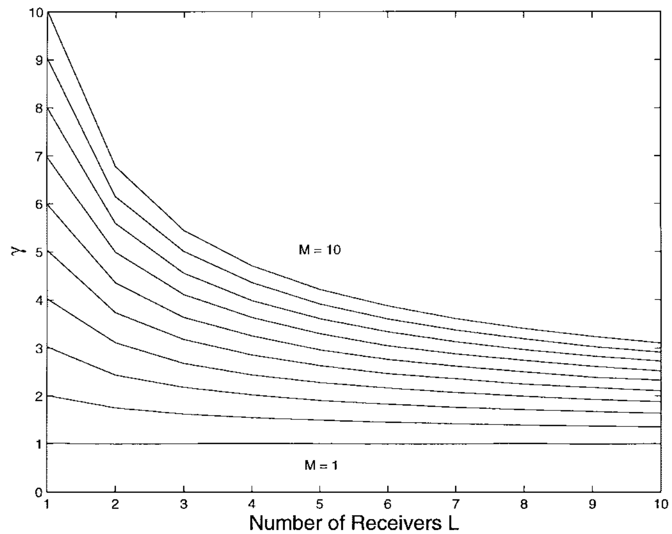

Fig. 5. Expected average SNR enhancement per receiver γ (over average SNR per receiver for an infinite receiver population) for an M-element transmitter antenna array broadcasting to L receivers.

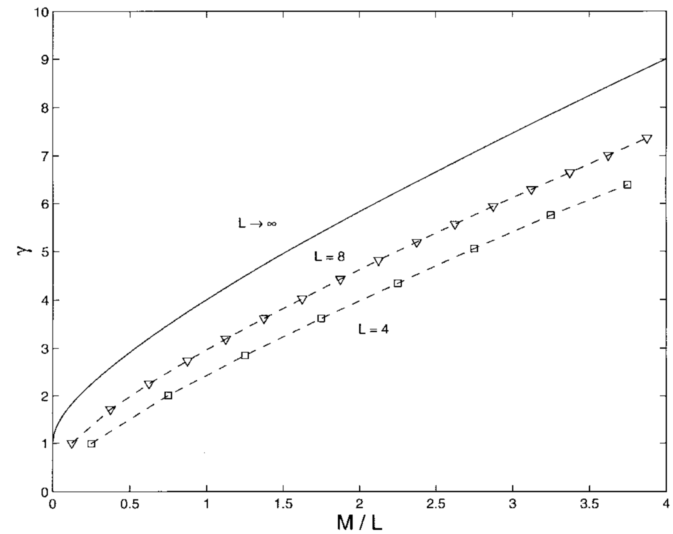

Fig. 6. Expected average SNR enhancement per receiver γ for various values of M/L. The solid line shows the deterministic ($\zeta = \gamma$) asymptotic values for a transmitter antenna array with M elements broadcasting to $L \rightarrow \infty$ receivers with the ratio M/L held fixed. The dashed curves denote representative points corresponding to finite M and L for $L = 4$ ("□") and $L = 8$ ("▽").

The expected receiver-average SNR enhancement is $\gamma = E[\zeta]$. As we will see, γ ranges from one to M as the number of receivers L decreases from ∞ to one.

From Section IV-A, we see that we can maximize average SNR per receiver by beamforming in the direction corresponding to the principal eigenvalue of $\Gamma_A = A^H A \mathcal{E}_s / L N_0$, and that the corresponding achieved average SNR per receiver is given by the largest eigenvalue $\hat{\lambda}$ of Γ_A (or, equivalently, $AA^H \mathcal{E}_s / L N_0$). Since A is a random matrix, the achievable average SNR is a random variable.

As $L \rightarrow \infty$, we have, via the strong law of large numbers

$$\Gamma_A = \frac{\mathcal{E}_s}{N_0} \frac{1}{L} A^H A \xrightarrow{\text{a.s.}} \frac{\mathcal{E}_s}{N_0} \sigma_\alpha^2 I_M \tag{48}$$

independent of the realized channel parameters. Hence, for the infinite receiver population beamforming in any direction maximizes average SNR per receiver, and this optimality is achieved without the transmitter having access to the channel parameters. Moreover, from (48) we see immediately that the resulting average SNR per receiver in this case, i.e., the denominator of (47), is $\mathcal{E}_s \sigma_\alpha^2 / N_0$, again independent of the realized channel.

For modest numbers of transmitter antenna elements and receivers, the expected value of the numerator of (47) can be determined from Monte Carlo simulations, from which the expected SNR enhancement γ follows. This enhancement is depicted as a function of the number of target receivers in Fig. 5 for several transmitter array sizes M. As we would expect, each of the SNR curves in Fig. 5 asymptotically approaches an SNR enhancement of unity as the number of receivers increases. However, as the figure reflects, this convergence is very slow; even in an $L = 10$ receiver scenario, each additional antenna element provides a significant SNR enhancement.

To develop the SNR enhancement characteristics further, we exploit that the matrix AA^H then has a complex Wishart distribution [13] when $M \leq L$; when $M > L$, it is the matrix $A^H A$ that is Wishart distributed. Thus, the SNR enhancement ζ is completely specified by the largest eigenvalue of a Wishart distributed matrix.

When M and L approach infinity in such a way that the ratio M/L of transmitter antenna elements per receiver approaches a positive constant, then the largest eigenvalue of Γ_A (and hence average SNR per receiver) can be shown to converge almost surely to [2], [4], [21]

$$\hat{\lambda} \xrightarrow{\text{a.s.}} \left(1 + \sqrt{\frac{M}{L}}\right)^2 \frac{\sigma_\alpha^2 \mathcal{E}_s}{N_0} \tag{49}$$

and thus

$$\zeta \xrightarrow{\text{a.s.}} \left(1 + \sqrt{\frac{M}{L}}\right)^2. \tag{50}$$

This asymptotic SNR enhancement behavior is shown by the solid curve in Fig. 6, from which we see that the SNR growth is effectively linear in the antenna/receiver ratio M/L for even modest ratios. Moreover, when the number of antenna elements M is significantly larger than the number of receivers L, there is a gain of approximately 3 dB in SNR for every doubling of M. We stress that the limit in (49) is no longer random: it does not depend on the realized channel parameters, so $\zeta = \gamma$.

It is also worth emphasizing that a ratio of $M/L = 0$ corresponds to a scenario in which M grows much more slowly than L, i.e., $M = o(L)$. A special case corresponds to using a fixed number of transmit antenna elements M while allowing the number of receivers L to increase to infinity. Consistent with our earlier analysis, this ratio leads to a deterministic SNR enhancement of unity.

Also shown in Fig. 6 is the expected SNR enhancement for representative scenarios involving antennas with finitely many elements and finite receiver populations (using Monte Carlo simulations). As the plot reflects, the asymptotic behavior (49) is approximated reasonably closely for even moderate values of M and L. Indeed, as Figs. 5 and 6 reflect, for a ratio of one transmitter antenna element per receiver ($M/L = 1$), for example, expected SNR enhancement γ increases from one to approximately three as $M = L$ increases from one to eight, achieving about 75% of the asymptotic performance of four shown in Fig. 6. Fig. 6 also demonstrates that the rates at which SNR grows with the number of antenna elements is closely approximated by the asymptotic results even for moderate values of M and L; for example, with $L = 8$ receivers, the additional SNR enhancement added by each of the eighth and ninth antennas is approximately 90% of the value suggested by the asymptotic curve at a ratio of $M/L = 1$.

For finite values of M and L, the SNR enhancement ζ is a random variable whose value depends on the realized channel. If more accurate performance statistics are desired than the asymptotic assumptions can provide, it is possible to calculate the probability distribution of the possible values the SNR may take on. In particular, the joint distribution of all the eigenvalues ζ_i of the normalized Wishart matrix $A^H A/\sigma_\alpha^2$ is [2]

$$f_{\zeta_1, \zeta_2, \cdots, \zeta_L}(\ell_1, \ell_2, \cdots, \ell_L) = \frac{\exp\left[-\sum_{i=1}^{M} \ell_i\right] \prod_{i=1}^{M} \ell_i^{L-M} \prod_{i<j} (\ell_i - \ell_j)^2}{\prod_{i=1}^{M} \Gamma(L-i+1)\Gamma(M-i+1)} \tag{51}$$

where

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt$$

denotes the usual Gamma function. Following Edelman [2], the density of the largest eigenvalue (the SNR enhancement ζ) can be computed by integrating over all but one of the ζ_i and dividing by $(M-1)!$ to remove the arbitrary ordering of the eigenvalues. When $M = 2$, the resulting probability density for the SNR enhancement ζ is shown in (52) (found at the bottom of the page) where

$$\gamma(z, a) = \int_0^a t^{z-1} e^{-t}\, dt \tag{53}$$

is the incomplete Gamma function. From these probability functions, it is possible to numerically calculate detailed statistics over the ensemble of possible channels. For example, Fig. 7 depicts the corresponding cumulative distribution function—the probability that average SNR per receiver falls below a particular threshold—for $M = 2$ antenna elements for various numbers of receivers L. As we would expect, the curves become less dispersed as the number of receivers increases, i.e., the average performance per receiver becomes less random and we approach the deterministic asymptotic behavior. Other useful statistics concerning average SNR per receiver for specific M and L such as the expectation and variance can be readily computed from probability densities as shown in (52).

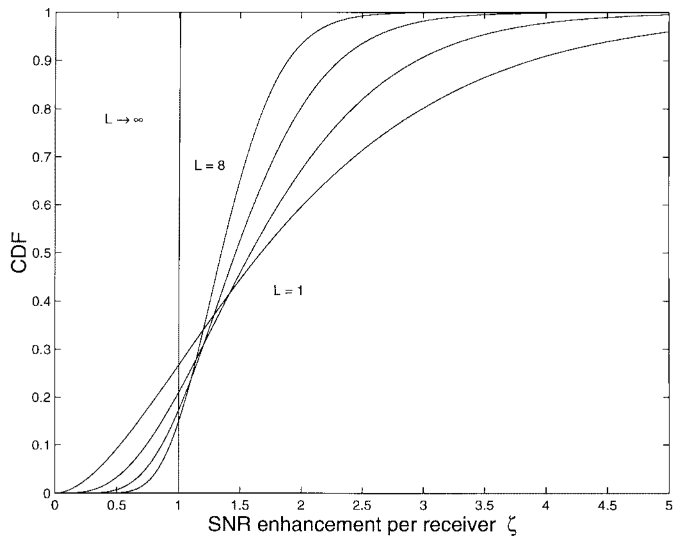

Fig. 7. Cumulative distribution for average SNR enhancement per receiver ζ for transmit antenna array with $M = 2$ elements broadcasting to L receivers, where $L = 1, 2, 4, 8$, and ∞.

C. Individual Receiver Performance

While the average SNR per receiver performance metric is a useful characterization of overall system performance, it does not reflect the SNR behavior experienced by any individual receiver in the system. In this section, we show that from an individual receiver's perspective, taking into account the channel parameters and beamforming as dictated by the average SNR criterion is preferable to ignoring the channel parameters—effectively assuming an infinite receiver population—even when the number of receivers is reasonably large.

To determine the performance of an individual receiver, we exploit the singular value decomposition

$$A = U \Sigma Q^H \tag{54}$$

and let $\hat{\boldsymbol{e}}$ denote the first column of Q, corresponding to the largest singular value. The SNR of the first receiver, which we consider without loss of generality, is then

$$\mathrm{SNR}_1 = \zeta |u_{1,1}|^2 \sigma_\alpha^2 \frac{\mathcal{E}_s}{N_0} \tag{55}$$

where $u_{1,1}$ is the upper left entry in the matrix U, and where, when $M = 2$ for example, ζ is as given by (52). Since U is a random circular unitary matrix [23], the probability density of $|u_{1,1}|^2$ is (see, e.g., [15])

$$f_{|u_{1,1}|^2}(\mu) = \begin{cases} (L-1)(1-\mu)^{L-2}, & 0 < \mu < 1 \\ 0, & \text{otherwise.} \end{cases} \tag{56}$$

$$f_\zeta(\ell) = \frac{e^{-\ell} \ell^{L-2} [\ell^L e^\ell - L\ell^{L-1} e^\ell + (\ell^2 + (L-1)(L-2\ell))\gamma(L-1, \ell)]}{(L-1)!(L-2)!} \tag{52}$$

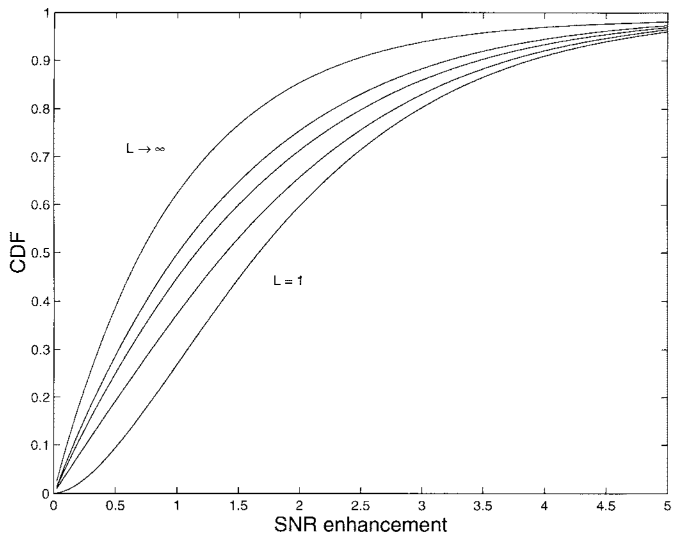

Fig. 8. Cumulative distribution for any particular receiver's SNR enhancement (over average SNR per receiver for an infinite receiver population) when using a transmitter antenna array with $M = 2$ elements to broadcast to L receivers, where $L = 1, 2, 4, 8$, and ∞.

The distribution for SNR_1 can be readily computed since random variables ζ and $|u_{1,1}|^2$ in (55) are independent—the principal eigenvector of $A^H A$ has no preferred direction. In the limiting case of $L \to \infty$ and M finite, it is straightforward to verify that SNR_1 has the same exponential distribution as if side information had been ignored when transmitting to an arbitrary number of receivers.

The resulting cumulative distribution function of a particular receiver's SNR for $M = 2$ and several values of L is shown in Fig. 8. As this figure reflects, outage probability—i.e., the probability that a receiver's SNR drops below some prescribed threshold—increases monotonically with the number of receivers, and that there is a significant outage difference between even an eight-receiver population and the infinite population case. Thus, in terms of both average and worst-case performance, transmission strategies that take into account the channel parameters according to the average per-receiver SNR criterion perform uniformly better than those that do not.

V. Conclusions

We have quantified the limits on system performance of a transmitter array that uses either partial side information to transmit to a single user or perfect side information to broadcast to a set of users.

For the point-to-point scenario with side information, beamforming in a direction determined by the eigenstructure of the posterior channel correlation matrix maximizes expected SNR. If the side information has the form of random variables S_i, representing noisy estimates of channel coefficients α_i, the improvement γ in expected SNR over a system without side information increases to M quadratically in the correlation coefficient $|\rho|$. Given N bits of side information, the transmitter can follow a vector-quantization-based approach to determine a locally optimal transmission strategy. For the case of $M = 2$ antennas, each additional bit of side information effectively halves the gap to the maximum possible SNR enhancement of M. Also, for large M and large N, the gap between perfect and zero side information decreases exponentially in the number of bits of side information N. Though a general analysis for other values of M is considerably more cumbersome, a scheme using $\log_2 M$ bits of side information provides an increase in an expected SNR logarithmic in M, and a scheme using $M - 1$ bits of side information produces an expected SNR improvement linear in M.

For sufficiently high quality side information, beamforming also maximizes mutual information. However, beamforming is not always optimal. For example, if the correlation between the side information and the true channel parameter is below 0.5, beamforming is suboptimal for channel SNR $(E|\alpha|^2 \mathcal{E}_s/N_0)$ greater than two. In such cases, more complex coding schemes may be used to achieve additional performance gains.

For the broadcast scenario, we considered the conflicting objectives inherent in transmitting to several receivers and introduced the "transmitter operating characteristic" as a useful tool in evaluating actual broadcasting schemes with respect to such objectives. From this characteristic, the frontier of operating points corresponding to SNR-based optimization criteria is easily identified.

We developed schemes for transmission to finite receiver populations and evaluated them relative to the performance of transmission schemes that assume an infinite receiver population and therefore ignore the channel parameters that constitute side information. In particular, we examined schemes that optimize average received SNR per receiver and used an expected SNR enhancement factor γ as a measure of the improvement in system performance obtained by explicitly taking into account the finite receiver population over the corresponding system designed for an infinite receiver population. As we saw, the expected SNR enhancement γ was significant even for relatively large receiver populations L. Specific degrees of SNR enhancement obtainable in systems of practical size could be efficiently predicted from asymptotic results we derived, which demonstrated approximately linear growth in SNR with the number of antenna elements.

Finally, while SNR enhancement per receiver is a measure of overall system performance, we showed that the schemes we developed for finite receiver populations are also significantly better than the corresponding systems for infinite receiver populations from an individual receiver's perspective. In particular, we showed that outage probabilities increase uniformly with the size L of the receiver population, so that explicitly taking into account side information in the form of channel parameters also enhances both the average and worst-case performance experienced by individual users. Systems with transmission strategies more closely related to minimizing individual receivers' outage probabilities than those explored in this paper will result in an even greater gain in worst-case performance and are an area of current research; for preliminary results, see, e.g., [12].

Many interesting and important problems remain. As one example, in scenarios where the antennas are too closely spaced or where the RF environment has few scatterers, then the amplitudes of the components of $\boldsymbol{\alpha}$ can become correlated; see, e.g., [1]. In this case, however, an uncorrelated model for the phases may still apply. We have not analyzed a randomized phase model in detail, but we anticipate qualitatively similar results to those that hold under a Rayleigh model. More complex models of correlation between array elements remain an area for future study.

APPENDIX
PROOF OF LEMMA 1

Define J as the expected mutual information

$$J = E[\log(1 + \lambda A + (1-\lambda)B)]. \tag{57}$$

We first show that J is strictly concave in λ by establishing that the second derivative is strictly negative

$$\frac{d^2 J}{d\lambda} = E\left[\frac{-(A-B)^2}{(1+\lambda A + (1-\lambda)B)^2}\right] \tag{58}$$

$$< 0. \tag{59}$$

The inequality is strict because A and B are independent random variables with nonzero variance. Consequently, the maximum of J occurs at $\lambda = 1$ if the first derivative $dJ/d\lambda \geq 0$ at $\lambda = 1$. Evaluating the first derivative at $\lambda = 1$ we find

$$\left.\frac{dJ}{d\lambda}\right|_{\lambda=1} = E\left[\frac{A-B}{1+A}\right] \tag{60}$$

$$= E\left[1 - \frac{1+B}{1+A}\right] \tag{61}$$

$$= 1 - E[1+B]E\left[\frac{1}{1+A}\right]. \tag{62}$$

Thus, the maximum occurs at $\lambda = 1$ if and only if

$$E[1+B]E\left[\frac{1}{1+A}\right] \leq 1. \tag{63}$$

We can upper bound $E[1/(1+A)]$ using methods suggested in [22, Lemma 1]. Because the third derivative of $1/(1+A)$ is negative, the maximum of $E[1/(1+\tilde{A})]$ over all random variables $\tilde{A}$ with $\tilde{A} \geq 0$, mean $E[A]$, and second moment $E[A^2]$ is achieved by a two mass point distribution

$$p_{\tilde{A}}(x) = \begin{cases} 1 - \dfrac{(E[A])^2}{E[A^2]}, & x = 0 \\ \dfrac{(E[A])^2}{E[A^2]}, & x = \dfrac{E[A^2]}{E[A]}. \end{cases} \tag{64}$$

Therefore

$$E\left[\frac{1}{1+A}\right] \leq \frac{(E[A])^2}{E[A^2]} \frac{1}{1 + \dfrac{E[A^2]}{E[A]}} + \left(1 - \frac{(E[A])^2}{E[A^2]}\right) \tag{65}$$

$$= \frac{E[A] + E[A^2] - (E[A])^2}{E[A] + E[A^2]}. \tag{66}$$

Using this upper bound in (63) yields the sufficient condition of the lemma.

REFERENCES

[1] F. Adachi, M. T. Feeney, A. G. Williamson, and J. D. Parsons, "Crosscorrelation between the envelopes of 900 MHz signals received at a mobile radio base station site," *Proc. Inst. Elect. Eng.*, vol. 133, pt. F, no. 6, pp. 506–512, 1986.

[2] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, May 1989.

[3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antenna," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[4] S. Geman, "A limit theorem for the norm of random matrices," *Ann. Prob.*, vol. 8, pp. 252–261, Apr. 1980.

[5] D. Gerlach and A. Paulraj, "Base station transmitting antenna arrays for multipath environments," *Signal Processing*, vol. 54, no. 1, pp. 59–73, Oct. 1996.

[6] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston: Kluwer, 1992.

[7] T. Hattori and K. Hirade, "Multitransmitter digital signal transmission by using offset frequency strategy in a land-mobile telephone system," *IEEE Trans. Veh. Technol.*, vol. VT-27, pp. 231–238, Nov. 1978.

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

[9] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques.* New York: Wiley, 1974.

[10] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: Wiley, 1995, vol. 2.

[11] W. C. Y. Lee, "Effects of correlation between two mobile radio base-station antennas," *IEEE Trans. Commun.*, vol. COMM-21, pp. 1214–1224, Nov. 1973.

[12] M. J. Lopez and G. W. Wornell, "Transmitter antenna array broadcasting with side information," in *Proc. Int. Symp. Information Theory*, Aug. 1998.

[13] R. J. Muirhead, *Aspects of Multivariate Statistical Theory.* New York: Wiley, 1982.

[14] A. Narula, M. D. Trott, and G. W. Wornell, "Information-theoretic analysis of multiple-antenna transmission diversity," in *Proc. Int. Symp. Information Theory and Applications*, Sept. 1996.

[15] ——, "Information-theoretic analysis of multiple-antenna transmission diversity," *IEEE Trans. Inform. Theory*, to be published.

[16] A. Narula, "Information-theoretic analysis of multiple-antenna transmission diversity," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, May 1997.

[17] A. Narula and M. D. Trott, "Multiple-antenna transmission with partial side information," in *Proc. Int. Symp. Information Theory*, July 1996.

[18] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1293–1302, July 1993.

[19] J. Salz and A. D. Wyner, "On data transmission over cross coupled multi-input, multi-output linear channels with applications to mobile radio," AT&T Bell Labs Tech. Memo., May 1990.

[20] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency division duplex (FDD) transmission systems using transmitter antenna diversity," *Int. J. Wireless Inform. Networks*, vol. 1, no. 1, pp. 49–60, Jan. 1994.

[21] J. S. Silverstein, "The smallest eigenvalue of a large dimensional Wishart matrix," *Ann. Prob.*, vol. 13, pp. 1364–1368, Nov. 1985.

[22] S. Shamai (Shitz) and A. Lapidoth, "Bounds on the capacity of a spectrally constrained Poisson channel," *IEEE Trans. Inform. Theory*, vol. 39, pp. 19–29, Jan. 1993.

[23] G. W. Stewart, "The efficient generation of random orthogonal matrices with an application to condition estimators," *SIAM J. Num. Anal.*, vol. 17, no. 3, pp. 403–409, June 1980.

[24] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication—I: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[25] J. H. Winters, "Smart antennas for wireless systems," *IEEE Personal Commun.*, pp. 23–27, Feb. 1998.

[26] A. Wittneben, "Basestation modulation diversity for digital SIMULCAST," in *IEEE Vehicular Technology Conf.*, St. Louis, MO, May 1991, pp. 848–853.

[27] G. W. Wornell and M. D. Trott, "Efficient signal processing techniques for exploiting transmit antenna diversity on fading channels," *IEEE Trans. Signal Processing*, vol. 45, pp. 191–205, Jan. 1997.

[28] G. W. Wornell, "Linear diversity techniques for fading channels," in *Wireless Communications: Signal Processing Perspectives*, H. V. Poor and G. W. Wornell, Eds. Upper Saddle River, NJ: Prentice-Hall, 1998.

Aradhana Narula received the B.S.E. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1990 and the S.M. and Ph.D. degrees also in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 1992 and 1997, respectively.

While attending MIT, she was a Research Assistant at the MIT Lincoln Laboratory. She is currently in the Research and Advanced Development Group at Motorola Information Systems Group, Mansfield, MA. Her research interests include practical and theoretical aspects of wireless and wireline communication systems.

Michael J. Lopez received the B.S. and M.S.E. degrees in electrical and computer engineering from Johns Hopkins University, Baltimore, MD, in 1995.

During 1995–1996, he was with Digital Equipment Corporation, Shrewsbury, MA. Since 1996, he has been at the Massachusetts Institute of Technology, where he is a doctoral candidate in the Department of Electrical Engineering and Computer Science. His interests include communications and signal processing.

Since 1995, Mr. Lopez has held a National Science Foundation Graduate Research Fellowship.

Mitchell D. Trott received the B.S. and M.S. degrees in systems engineering from Case Western Reserve University, Cleveland, OH, in 1987 and 1988, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1992.

He was an Associate Professor in the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology from 1992 until 1998. He is now Director of Research at ArrayComm, Inc. His research interests include multiuser communication, information theory, and algebraic system theory.

Dr. Trott received an NSF Young Investigator Award in 1994 and the KDD Career Development Chair in 1995.

Gregory W. Wornell (S'83–M'85) received the B.A.Sc. degree (with honors) from the University of British Columbia, Canada, and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, all in electrical engineering, in 1985, 1987 and 1991, respectively.

Since 1991, he has been on the faculty of the Department of Electrical Engineering and Computer Science at MIT, where he is currently a Cecil and Ida Green Career Development Associate Professor. During the 1992–1993 academic year, he was on leave at AT&T Bell Laboratories, Murray Hill, NJ. His research interests include signal processing and wireless communications, and he is the author of the monograph *Signal Processing with Fractals: A Wavelet-Based Approach* and a coeditor of the volume *Wireless Communications: Signal Processing Perspectives*. He is also a consultant to industry and an inventor on several patents.

Dr. Wornell is currently an Associate Editor for the communications area for the IEEE SIGNAL PROCESSING LETTERS and serves on the Communications Technical Committee of the Signal Processing Society. Among the awards he has received for teaching and research are the MIT Goodwin Medal, the ITT Career Development Chair at MIT, an NSF Faculty Early Career Development Award, an ONR Young Investigator Award, and the MIT Junior Bose Award.