Space–Time Precoding for Mean and Covariance Feedback: Application to Wideband OFDM

Gwen Barriac and Upamanyu Madhow, *Fellow, IEEE*

***Abstract*—We consider optimization of the capacity of a multi-input single-output wideband cellular "downlink," in which the base station has estimates of the statistics of the spatial channel. Our main focus is on orthogonal frequency-division multiplexed (OFDM) systems, although some of our results apply to single-carrier systems, as well. Prior work has shown that estimates of the channel spatial covariance can be obtained without overhead for both frequency-division duplex (FDD) and time-division duplex (TDD) systems by suitably averaging uplink measurements. In this paper, we investigate the benefits of supplementing this "free" *covariance feedback* with *mean feedback,* where the latter refers to estimates of the spatial channel realization in each subcarrier. Mean feedback can be obtained using reciprocity for TDD systems, and requires explicit feedback for FDD systems. We first devise strategies for using both covariance and mean feedback, mainly restricting attention to beamforming, which is optimal or near-optimal for many outdoor channels with narrow spatial spread. Second, since mean feedback degrades rapidly with feedback delay for mobile channels, we develop quantitative rules of thumb regarding the accuracy required for the mean feedback to be a useful supplement to the already available, and robust, covariance feedback. Our results validate the following intuition: the accuracy requirements for mean feedback to be useful are more relaxed for channels with larger spatial spread, or for a larger number of transmit elements.**

***Index Terms*—Fading channels, information rates, multiple-input multiple-output (MIMO) systems.**

I. Introduction

We consider downlink communication in a wideband orthogonal frequency-division multiplexed (OFDM) cellular network, in which the base station (BS) transmitter has an antenna array, while the mobile receiver has one antenna. Higher throughput is essential for the success of fourth-generation (4G) and future wireless systems, and the use of feedback is known to dramatically increase capacity in multiantenna systems [1]–[5]. It has been shown [6] that for both time-division duplexed (TDD) and frequency-division duplexed (FDD) systems, robust feedback regarding the second-order statistics of the spatial channel can be obtained for "free" by averaging across frequency. In this paper, we provide analytical rules of thumb quantifying the additional benefits of using first-order channel statistics, given that second-order statistics are already available. Key issues that impact system design are the effect of channel time variations and feedback delay, and the cost of obtaining feedback.

It can be shown [7], [8] that the space–time channels seen by different subcarriers are identically distributed random vectors which decorrelate across frequency. This, in turn, leads to the observation in [6] and [9] that the covariance of the space–time channel for any subcarrier is the same. This *statistical reciprocity* enables the BS to obtain the spatial covariance matrix for use in downlink transmission by averaging over its uplink measurements. Such *implicit* feedback regarding the covariance matrix can be used for space–time transmit precoding on the downlink. Covariance feedback is particularly effective when the power-angle profile (PAP) of the channel seen by the BS is relatively narrow, as is the case for many outdoor systems in which the BS is far enough, and at high enough elevation, that signals reaching a given mobile leave the BS in a narrow spatial cone. In such cases, often one or two channel eigenmodes are dominant, and the BS can employ transmit beamforming along these eigenmodes to improve downlink performance while reducing complexity.

Covariance feedback becomes less effective as the number of dominant eigenmodes increases. This might occur, for example, when BSs are located at lower altitudes, not far enough removed from urban clutter. However, *first*-order feedback regarding the space–time channel in each subcarrier, if available, could still be used to improve performance. We term such feedback *mean* feedback, and provide estimates of the performance gains due to mean and covariance feedback, relative to the performance with covariance feedback alone. We also investigate the required accuracy of the mean feedback in order for the gains to be significant.

Mean feedback could be obtained implicitly using (classical) reciprocity for TDD systems. For FDD systems, mean feedback must be obtained by explicitly feeding back measurements at the mobile receiver to the BS. Since covariance information is averaged across subcarriers, the transmit strategy with covariance feedback alone is the same across subcarriers. In contrast, a precoding strategy that accounts for mean feedback (which varies across subcarriers) must be implemented on a per-subcarrier basis. Mean feedback is also more sensitive to feedback delay than covariance feedback: the time variation in the space–time channel realization for a specific subcarrier is much faster than

Paper approved by X. Wang, the Editor for Wireless Spread Spectrum of the IEEE Communications Society. Manuscript received May 21, 2004; revised February 5, 2005. This work was supported in part by Motorola, with matching funds from the University of California Industry-University Cooperative Research program, and in part by the National Science Foundation under Grants EIA 00-80134 and ANI 02-20118 (ITR). This paper was presented in part at the International Symposium on Information Theory, Chicago, IL, June 2004.

G. Barriac was with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, Santa Barbara, CA, 93106 USA. She is now with Qualcomm Inc., San Diego, CA 92121 USA (e-mail: gbarriac@qualcomm.com).

U. Madhow is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, Santa Barbara, CA, 93106 USA (e-mail: madhow@ece.ucsb.edu).

Digital Object Identifier 10.1109/TCOMM.2005.861665

the variation of the second-order statistics averaged across subcarriers. Keeping the preceding issues in mind, the goal of this paper is to provide a design framework for evaluating whether the additional overhead and complexity associated with mean feedback are justified.

While our interest is in obtaining design prescriptions for wideband systems, we can examine the fundamental tradeoffs regarding mean and covariance feedback by considering their effect on the ergodic capacity for a single subcarrier, which sees a narrowband flat-fading space–time channel. The parameters of the narrowband model are related to the array manifold at the BS, and the channel PAP. For the most part, we restrict attention to beamforming as a transmit strategy, since the channels we consider have relatively small angular spreads, where beamforming gives optimal or close to optimal results. In addition, beamforming is desirable, because it greatly reduces the receiver complexity.

A. Summary of Results

Information-theoretic results on ergodic capacity are derived. These, together with the use of analytical models for outdoor OFDM systems, are then employed to obtain system design tradeoffs, as follows.

1) We find an upper bound on ergodic capacity when the transmit strategy is restricted to beamforming and the BS has access to both mean and covariance feedback. We then show that beamforming in the direction that maximizes receive signal-to-noise ratio (SNR), denoted $\boldsymbol{\nu}_0$, gives results that are close to this upper bound. Since the optimal beamforming direction would be too computationally complex to compute in real time, the practical solution for maximizing the expected spectral efficiency is thus to send along $\boldsymbol{\nu}_0$. Note that the ergodic capacity when the transmit strategy is restriced to beamforming is a lower bound to capacity when the transmitter has no such restrictions.
2) For an OFDM system, we provide a simple model in which the quality of the available mean feedback is approximately characterized using a single parameter ρ, whose value is related to the goodness of minimum mean-squared error (MMSE) estimation of the channel in the desired time-frequency bin as a function of the available feedback. Restricting attention to beamforming as the transmit strategy, we find an approximate threshold on ρ for mean feedback to provide appreciable gains over the use of covariance feedback alone. The threshold depends on the number of transmit antennas and the channel eigenvalues. As the number of antennas increases, or the PAP becomes more spread out, the necessary accuracy decreases (i.e., mean feedback becomes more valuable).
3) Even when the transmitter is not restricted to beamforming, the general trends delineated in 2) still hold, i.e., mean feedback is more valuable for systems with larger N_T and wider PAPs.

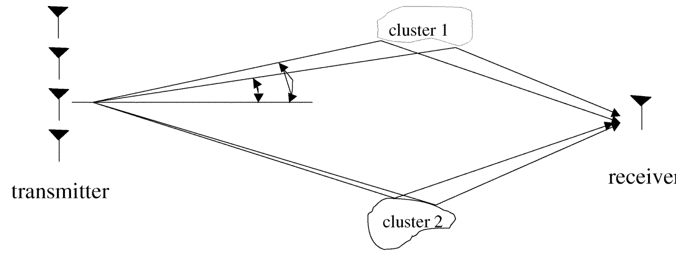

Fig. 1. Channel model setup. Specular rays from the BS to the mobile are grouped together in one or more clusters. The BS may have multiple antennas, while the receiver is restricted to a single antenna element.

B. Relation to Previous Work

The use of feedback in multiple-antenna systems has received considerable attention in the literature. Visotsky and Madhow [3] find the optimal transmit directions when the mobile has a single antenna and the BS has either mean feedback *or* covariance feedback. Jafar *et al.* [4] extend these results to systems where the mobiles have multiple antennas. Covariance feedback has been considered as a means of improving performance in [5], [10], and [11], and mean feedback has been used in [2] and [12], to name but a few. Downlink beamforming using covariance information has also been extensively studied in the context of multiuser systems, where the aim is to maximize the performance of a desired user while minimizing interference to other users [13]–[15]. The problem of maximizing performance while simultaneously minimizing interference is also considered in [16], except the BS is assumed to know the channel response to the user of interest. To the best of our knowledge, this is the first paper to consider optimizing capacity when the BS can separately obtain both mean and covariance feedback, and to consider how accurate mean feedback need be in order to be useful when covariance feedback is already available.

C. Paper Organization

The channel model is described in Section II. Optimization of beamforming capacity, given both mean and covariance feedback, is considered in Section III. Section IV provides a measure of how accurate mean feedback must be, in order to be useful. Section V concludes the paper with a discussion of how the analytical framework in the paper can be applied to some practical scenarios in order to determine the utility of mean feedback.

II. System Model

The space–time precoding strategies we consider can be understood in terms of a per-subcarrier mathematical model, described in Section II-A. This mathematical model follows from a physical model for a wideband OFDM system, described in Section II-B.

A. Mathematical Model for One Subcarrier

Consider a BS with N_T antennas sending to a mobile with a single antenna, as shown in Fig. 1. For a particular time-frequency bin of interest in an OFDM system, the received signal can be written as

$$y = \mathbf{h}^H \mathbf{s} + n \qquad (1)$$

where $\mathbf{s}$ is the $N_T \times 1$ transmitted symbol vector, $\mathbf{h}$ is the $N_T \times 1$ channel response, and n is circularly symmetric complex Gaussian noise, denoted $n \sim CN(0, \sigma^2)$. (Note that σ^2 is the total noise variance, not the noise variance per dimension.)

Model for Covariance Feedback Alone: When there is no feedback regarding the mean of the channel, $\mathbf{h}$ is modeled as a zero-mean complex Gaussian vector with covariance $\mathbf{C}$

$$\mathbf{h} \sim CN(0, \mathbf{C}). \tag{2}$$

Since the channel spatial covariance $\mathbf{C}$ can be accurately estimated by averaging uplink measurements [6], [9], it is assumed that the BS knows $\mathbf{C}$.

Note that $\mathbf{C}$ can be written

$$\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \tag{3}$$

where $\mathbf{U}$ is the eigenvector matrix $[\mathbf{u}_1 \ldots \mathbf{u}_{N_T}]$, and $\mathbf{\Lambda}$ is the diagonal channel eigenvalue matrix, where the eigenvalues $\{\lambda_i\}$ are arranged in decreasing order.

Model for Mean and Covariance Feedback: In addition to knowing $\mathbf{C}$, the BS also has some direct feedback $\mathbf{x}$ regarding the channel response $\mathbf{h}$. As discussed in Section II-B, $\mathbf{x}$ might be a vector comprised of measurements of channels in other time-frequency bins. Assuming that $\mathbf{x}$ is jointly complex Gaussian with the channel in the time-frequency bin of interest, the conditional distribution of $\mathbf{h}$ is given by

$$\mathbf{h} \sim CN(\mathbf{m}, \mathbf{K}) \tag{4}$$

where $\mathbf{m}$ is the MMSE estimate of $\mathbf{h}$ given $\mathbf{x}$, and $\mathbf{K}$ is the corresponding error covariance. Note that for consistency with the unconditional distribution of $\mathbf{h}$ in (2), $E[\mathbf{m}] = 0$ and $E[\mathbf{m}\mathbf{m}^\mathbf{H} + \mathbf{K}] = \mathbf{C}$.

The relation of $\mathbf{m}$ and $\mathbf{K}$ to the physical channel model is discussed in detail in Section II-B, where it is shown that as long as the feedback has sufficiently large signal-to-noise ratio (SNR), the conditional distribution (4) of the channel, given the feedback, can be approximated as follows:

$$\mathbf{h} \sim CN(\rho\mathbf{f}, (1-\rho^2)\mathbf{C}). \tag{5}$$

In other words, we can write

$$\begin{aligned} \mathbf{m} &= \rho\mathbf{f} \\ \mathbf{K} &= (1-\rho^2)\mathbf{C} \end{aligned} \tag{6}$$

where $\mathbf{f} \sim CN(0, \mathbf{C})$ and is an auxiliary random vector related to the MMSE estimate of $\mathbf{h}$ given $\mathbf{x}$, and ρ is a normalized correlation that measures the goodness of this MMSE estimate.

B. Model for a Wideband System

In this section, we connect the per-subcarrier models (4) and (5) to an OFDM system model. The system we consider is as in Fig. 1. The BS is assumed to be far away from the mobile, and at high enough altitude that there is little to no local scattering around it. Thus, both uplink and downlink signals for a given mobile are restricted to a fairly narrow spatial cone, from the viewpoint of the BS antenna array. In contrast, the mobile is assumed to be in a rich scattering environment.

A common approach to statistically modeling wideband space–time channels is to generate a large number of specular rays at random, so as to conform to a measured power-delay profile (PDP) and PAP, as well as measured delay and angle distributions. In [8], we developed an alternative vector tap delay line (TDL) model with complex Gaussian taps which is more amenable to analysis. The TDL model arises from classical arguments: paths which are spaced less than $1/W$ apart (where W is the bandwidth), cannot be distinguished, and hence, sum together to form taps. Since the paths are independent, with uniformly distributed phases, their sum is complex Gaussian (as long as there are enough paths for the central limit theorem to be effective). The channel model can be made consistent with measured PDP and PAP by appropriately choosing the distributions of the vector taps. A typical channel impulse response is then given by

$$\mathbf{p}_W(t,\tau) = \sum_{l=0}^{\infty} A_l \mathbf{v}_l(t)\delta\left(\tau - \frac{1}{W}\right) \tag{7}$$

where $\delta(t)$ denotes the Dirac delta, and where the tap weights A_l are proportional to the square root of the PDP, and are normalized so the channel has unit power. The vectors $\mathbf{v}_l(t)$ are independent zero-mean complex normal random vectors

$$\mathbf{v}_l(t) \sim CN(0, \mathbf{C}) \quad \forall l \quad \forall t \tag{8}$$

where $\mathbf{C}$ is the expected value of the outer product of the antenna array response, where the expectation is taken over the PAP, $P_\Omega(\cdot)$

$$\mathbf{C} = E[\mathbf{a}(\Omega')\mathbf{a}(\Omega')^H] = \int_{-\pi}^{\pi} \mathbf{a}(\Omega')\mathbf{a}(\Omega')^H P_\Omega(\Omega')d\Omega'. \tag{9}$$

Throughout this paper, we use for illustration an exponential PDP and a linear BS array with equally spaced antennas. The BS array response as a function of the angle of departure is given as follows:

$$\begin{aligned} \mathbf{a}(\Omega) &= [a_1 \ldots a_l \ldots a_{N_T}]^T \\ a_l(\Omega) &= e^{j(l-1)2\pi(d/\lambda)\sin(\Omega)}, \quad l = 1, \ldots, N_T \end{aligned} \tag{10}$$

where d is the interelement spacing, and λ the carrier wavelength.

Now, consider an OFDM system with N_T transmit antennas, one receive antenna, and N frequency bins. One OFDM symbol transmitted at time t_k consists of N symbol vectors $\mathbf{s}(f_i, t_k), i = 1, \ldots, N$ of length N_T, each transmitted at a different frequency. We can write

$$y(f_i, t_k) = \mathbf{h}(f_i, t_k)^H \mathbf{s}(f_i, t_k) + n(f_i, t_k) \tag{11}$$

where f_i denotes frequency i and t_k denotes time k. In other words, $y(f_i, t_k)$ is the received data vector for the ith tone of the kth OFDM symbol, $\mathbf{h}(f_i, t_k)$ is the $N_T \times 1$ channel frequency response at the ith tone at the time of the kth OFDM symbol, and $n(f_i, t_k)$ is additive white Gaussian noise (AWGN) satisfying $E[n(f_i, t_k)n(f_j, t_q)^H] = \sigma^2\delta_{ij}\delta_{kq}$, where δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ if $i = j$; $\delta_{ij} = 0$, otherwise).

Approximating the $\mathbf{v}_l(t)$ as constant over one OFDM symbol period, and denoting the start of the kth symbol period by t_k, the channel frequency response for symbol k at tone i is

$$\mathbf{h}(f_i, t_k) = \sum_{l=0}^{\infty} A_l \mathbf{v}_l(t_k) e^{-j2\pi f_i 1/W}. \tag{12}$$

It follows that the $\{\mathbf{h}(f_i, t_k)\}$ are well-modeled as identically distributed $N_T \times 1$ proper complex Gaussian random vectors with zero mean and covariance matrix $\mathbf{C}$

$$\mathbf{h}(f_i, t_k) \sim CN(0, \mathbf{C}) \quad \forall i \quad \forall k. \tag{13}$$

We can now justify our assumption that the BS knows the channel covariance $\mathbf{C}$. Because the channel responses are identically distributed, if the BS measures the channel from uplink measurements spaced further apart than the coherence bandwidth, it can reconstruct an accurate estimate of $\mathbf{C}$ [6], [9].

In the context of Section II-A, if the BS wishes to estimate the channel at frequency f_j and time t_q, we have

$$\mathbf{h} \equiv \mathbf{h}(f_j, t_q). \tag{14}$$

If the BS has no mean feedback, then (14) and (13) lead to the channel model given in (2).

We now show how the channel decorrelates in frequency and time; it is this decorrelation that fundamentally limits the efficacy of mean feedback. For wide-sense stationary (WSS) taps $\{\mathbf{v}_l\}$, we have

$$E[\mathbf{v}_l(t_1)\mathbf{v}_l^H(t_2)] = \kappa_t(\Delta_t)\mathbf{C} \tag{15}$$

where $\Delta_t \equiv t_2 - t_1$, and $\kappa_t(\Delta_t)$ is the correlation coefficient. Due to the mobile's rich scattering environment, in our numerical results, we will employ Clarke's Rayleigh fading model [17] for the channel taps $\{\mathbf{v}_l\}$ in (7). Thus, $\kappa_t(\Delta_t) \equiv J_0(2\pi f_D \Delta_t)$, where $J_0(\cdot)$ is a zeroth-order Bessel function of the first kind, and f_d is the maximum Doppler frequency.

Using (12) and (15), we have that

$$E[\mathbf{h}(f_i, t_k)\mathbf{h}(f_i, t_q)^H] = \kappa_t(\Delta_t)\mathbf{C} \tag{16}$$

where $\Delta_t = t_q - t_k$. On account of the WSS across frequency bins which follows from our channel model, we can also write

$$E[\mathbf{h}(f_i, t_k)\mathbf{h}(f_j, t_k)^H] = \kappa_f(\Delta_f)\mathbf{C} \tag{17}$$

where $\Delta_f \equiv f_i - f_j$ and $\kappa_f(\Delta_f)$ depends on the PDP. For our running example of an exponential PDP, $\kappa_f(\Delta_f) = 1/(1 + j2\pi\tau_{\rm rms}\Delta_f)$. Using (15) and the fact that the $\{\mathbf{v}_l\}$ are independent, it follows that the correlations in frequency and time are separable, i.e., we can write

$$E[\mathbf{h}(f_i, t_k)\mathbf{h}(f_j, t_q)^H]) = \kappa_f(\Delta_f)\kappa_t(\Delta_t)\mathbf{C}. \tag{18}$$

Example (Noiseless Channel Measurement in One Time-Frequency Bin): We now consider a simple example that illustrates how the conditional mean and covariance of the desired channel, given mean feedback, relate to the preceding wideband channel model. Suppose that the BS has an exact channel measurement $\mathbf{h}(f_i, t_k)$, with which to estimate the channel $\mathbf{h}(f_j, t_q)$ in the desired time-frequency bin. In the notation of Section II-A

$$\mathbf{h} = \mathbf{h}(f_j, t_q), \quad \mathbf{x} = \mathbf{h}(f_i, t_k).$$

Letting $\hat{\mathbf{h}}$ denote the MMSE estimate of $\mathbf{h}$ given $\mathbf{x}$, given by

$$\hat{\mathbf{h}} = E[\mathbf{h}|\mathbf{x}] \tag{19}$$
$$= \mathbf{K_{hx}}\mathbf{K_{xx}}^{-1}\mathbf{x} \tag{20}$$

where

$$\mathbf{K_{hx}} = E[\mathbf{h}\mathbf{x}^H] \tag{21}$$
$$\mathbf{K_{xx}} = E[\mathbf{x}\mathbf{x}^H]. \tag{22}$$

In this case, we have

$$\mathbf{K_{hx}} = \rho\mathbf{C} \tag{23}$$
$$\mathbf{K_{xx}} = \mathbf{C} \tag{24}$$

where $\rho = \kappa(t_q - t_k, f_i - f_j)$, and thus

$$\hat{\mathbf{h}} = \rho\mathbf{x}. \tag{25}$$

Since $\mathbf{x}$ is complex Gaussian with zero mean, the MMSE estimate $\hat{\mathbf{h}}$ is also zero-mean complex Gaussian. Using (20), it can be seen that its distribution is given as follows:

$$\hat{\mathbf{h}} \sim CN(0, \mathbf{K_{hx}}\mathbf{K_{xx}}^{-1}\mathbf{K_{xh}}) \tag{26}$$
$$= CN(0, \rho^2\mathbf{C}). \tag{27}$$

The distribution of $\mathbf{h}$ conditioned on the feedback is then

$$\mathbf{h} \sim CN(\hat{\mathbf{h}}, (1-\rho^2)\mathbf{C}) \tag{28}$$

where we have used the fact that $\mathbf{h} \sim CN(0, \mathbf{C})$. Hence, referring back to (4), we have, in this case

$$\mathbf{m} = \hat{\mathbf{h}} = \rho\mathbf{x} \tag{29}$$
$$\mathbf{K} = (1-\rho^2)\mathbf{C}. \tag{30}$$

Since $\mathbf{x} \sim CN(0, \mathbf{C})$, this corresponds to the model (6) in Section II-A with $\mathbf{f} \equiv \mathbf{x}$. Note that the accuracy of mean feedback is summarized by a single parameter ρ, thus simplifying design considerations such as whether or not to use mean feedback. In the following proposition, we state that this feature holds more generally for noiseless channel measurements from several time-frequency bins, and also holds approximately for noisy channel measurements at high enough SNR.

Proposition 1: Let $\mathbf{h} = \mathbf{h}(f_j, t_q)$ denote the channel the BS wishes to estimate. If the feedback $\mathbf{x}$ consists of (possibly noisy) channel realizations in a set of time-frequency bins, then the channel $\mathbf{h}$ is conditionally circularly Gaussian

$$\mathbf{h} \sim CN(\mathbf{m}, \mathbf{K}) \tag{31}$$

where $\mathbf{m}$ is the MMSE estimate of $\mathbf{h}$ and $\mathbf{K}$ is the error covariance for the MMSE estimate. For "noiseless" feedback

$$\mathbf{m} = \rho\mathbf{f} \sim CN(0, \rho^2\mathbf{C}) \tag{32}$$
$$\mathbf{K} = (1-\rho^2)\mathbf{C} \tag{33}$$

where ρ is a parameter describing the goodness of the *noiseless* feedback. For "noisy" feedback (i.e., the channel measurements available to the BS are corrupted by complex WGN with variance σ_f^2 (the subscript f stands for feedback)

$$\mathbf{m} \sim CN(0, \rho^2\mathbf{C} - \sigma_f^2\mathbf{Q}) \quad (34)$$

$$\mathbf{K} = (1-\rho^2)\mathbf{C} + \sigma_f^2\mathbf{Q} \quad (35)$$

where $\mathbf{Q}$ is as defined in (84) in the proof, and ρ is, as before, a parameter measuring the goodness of the MMSE estimate for the corresponding noiseless feedback scenario.

Proof: See the Appendix.

Remark: From the proposition, we see that the conditional channel distribution can be written approximately as

$$\mathbf{h} \sim CN(\rho\mathbf{f}, (1-\rho^2)\mathbf{C}) \quad (36)$$

when the SNR on the channel measurements available to the BS is high enough. This noiseless feedback model also provides an upper bound on the performance obtained using mean feedback; thus, if mean feedback is not useful under the noiseless model, then it should certainly not be used when the feedback is noisy. In our discussions in Section IV, we restrict attention to (36), since it enables us to quantify the value of mean feedback under rather general assumptions using the single parameter ρ.

III. Achieving Beamforming Capacity

We now investigate how to maximize the expected spectral efficiency when both mean and covariance feedback are available, and the transmit strategy is restricted to beamforming. We consider beamforming for two main reasons. First, the channels we consider have narrow enough angular spreads (typical of outdoor cellular systems) such that, even with covariance feedback alone, beamforming gives optimal or close to optimal results [18]–[20]. (We consider an antenna spacing of half the carrier wavelength; with this spacing and the channel models of interest, beamforming performs well.) Second, beamforming greatly reduces the receiver complexity.

In [20], it is shown that for systems with covariance feedback, for a given covariance matrix, beamforming is always optimal up to a certain SNR, at which point it becomes optimal to transmit in multiple directions. Using the approach in [20], it can be shown that if the main eigenmode has 80% of the channel power, then the SNR must be higher than 20 dB for nonbeamforming strategies to be optimal. Similarly, if the main eigenvector has 70% of the power, then the SNR must higher than 10 dB. Furthermore, in [19], it is shown that the maximum capacity gain achievable over beamforming when transmitting in more than one direction is bounded by 0.577 nats/s/Hz, and that even for covariance matrices with multiple dominant eigenmodes operating in high-SNR regimes, transmitting in multiple directions gives minimal gains over beamforming. Hence, in realistic outdoor scenarios where complexity is an issue and angular spreads are narrow, it makes sense to restrict the transmitter to beamforming strategies.

For a given beamforming direction $\boldsymbol{\nu}$ (normalized to unit norm), the ergodic capacity of a narrowband channel is

$$C_{bf} = \max_{\boldsymbol{\nu}} E_{\mathbf{h}}[\log(1 + P\mathbf{h}^H\boldsymbol{\nu}\boldsymbol{\nu}^H\mathbf{h})] \quad (37)$$

$$= \max_{\boldsymbol{\nu}} E_{\mathbf{h}}[\log(1 + P|\mathbf{h}^H\boldsymbol{\nu}|^2)] \quad (38)$$

where the expectation is taken over $\mathbf{h}$, as denoted by the subscript $\mathbf{h}$ in $E_{\mathbf{h}}$, and $\mathbf{h}$ is distributed as given in (4). (We sometimes use a subscript for the expectation operator where it may be unclear as to what the expectation operator refers. We also use the subscript bf to differentiate beamforming capacity from standard capacity. We have also set the noise power to 1 without loss of generality.) (38) can be written as

$$C_{bf} = \max_{\boldsymbol{\nu}} I(\boldsymbol{\nu}) \quad (39)$$

where $I(\boldsymbol{\nu})$ is the expected mutual information when beamforming along the direction $\boldsymbol{\nu}$, i.e.,

$$I(\boldsymbol{\nu}) = E_{\mathbf{h}}[\log(1 + P|\mathbf{h}^H\boldsymbol{\nu}|^2)]. \quad (40)$$

The expectation is taken over the channel $\mathbf{h}$, given a fixed value for the feedback, (i.e., $\mathbf{m}$ is fixed). Letting $z \equiv \mathbf{h}^H\boldsymbol{\nu}$, we can also write the average mutual information (given the feedback) as a function of the complex Gaussian random variable z (note that $|z|$ is a Rician random variable)

$$\tilde{I}(z) = E_z[\log(1 + P|z|^2)]. \quad (41)$$

A. Max SNR Beamformer

Because the maximization in (39) to locate the optimal $\boldsymbol{\nu}$, $\boldsymbol{\nu}_{\text{opt}}$, is prohibitively complex for real-time calculations, we propose that, as a practical way of achieving near-optimal mutual information, the BS sends along the direction $\boldsymbol{\nu}_0$, which maximizes the receive SNR. (This strategy is commonly used [1], [13], [16], as receive SNR is often the desired performance metric.) The vector $\boldsymbol{\nu}_0$ is easily determined, and we show later that the expected mutual information $I(\boldsymbol{\nu}_0)$ is close to the maximum beamforming capacity C_{bf} (39). Writing $\boldsymbol{\nu}_0$ as

$$\boldsymbol{\nu}_0 = \arg\max_{\boldsymbol{\nu}} E_{\mathbf{h}}[P\boldsymbol{\nu}^H\mathbf{h}\mathbf{h}^H\boldsymbol{\nu}] \quad (42)$$

$$= \arg\max_{\boldsymbol{\nu}} P\boldsymbol{\nu}^H[\mathbf{m}\mathbf{m}^H + \mathbf{K}]\boldsymbol{\nu} \quad (43)$$

it is clear that $\boldsymbol{\nu}_0$ is the dominant eigenvector of $\mathbf{M} \equiv \mathbf{m}\mathbf{m}^{\mathbf{H}} + \mathbf{K}$, which can readily be computed by the BS.

B. An Upper Bound on Beamforming Capacity

To show that $I(\boldsymbol{\nu}_0)$ is close to C_{bf}, we find an upper bound on capacity, and show that $I(\boldsymbol{\nu}_0)$ is close to this bound. In order to do so, we need the following lemma.

Lemma 1: $\tilde{I}(z)$, as defined in (41), is an increasing function of both m_z and σ_z, where

$$m_z = |E[z]| \quad (44)$$

$$\sigma_z^2 = \frac{1}{2}var[z]. \quad (45)$$

Proof: See the Appendix.

We now use *Lemma 1* to find an upper bound on capacity.

Theorem 1: Consider a narrowband channel distributed as in (4)

$$\mathbf{h} \sim CN(\mathbf{m}, \mathbf{K}) \quad (46)$$

and let $z^\star$ be a complex Gaussian random variable with

$$E[z^\star] = \|\mathbf{m}\| \quad (47)$$

$$var[z^\star] = k_1 \quad (48)$$

where k_1 is the largest eigenvalue of $\mathbf{K}$. Then an upper bound for the beamforming capacity is given by

$$C_{bf} \leq \tilde{I}(z^\star). \tag{49}$$

Proof: From (40), we can write $I(\boldsymbol{\nu}) = E[\log(1 + P|z(\boldsymbol{\nu})|^2)]$, where $z(\boldsymbol{\nu}) = \mathbf{h}^H\boldsymbol{\nu}$, and is characterized by

$$E[z(\boldsymbol{\nu})] = E[\mathbf{h}^H\boldsymbol{\nu}] = \mathbf{m}^H\boldsymbol{\nu} \tag{50}$$

$$\mathrm{var}[z(\boldsymbol{\nu})] = \boldsymbol{\nu}^H\mathbf{K}\boldsymbol{\nu}. \tag{51}$$

If $E[z(\boldsymbol{\nu})] \leq \tilde{m}_z$ and $\mathrm{var}[z(\boldsymbol{\nu})] \leq \tilde{v}_z$, then by *Lemma 1*

$$I(\boldsymbol{\nu}) \leq \tilde{I}(z^\star) \tag{52}$$

where $E[z^\star] = \tilde{m}_z$ and $\mathrm{var}[z^\star] = \tilde{v}_z$. Unless $\mathbf{m} = \mathbf{0}$, $\mathbf{K} = \mathbf{0}$, or $\mathbf{K}$ is white, the inequality in (52) is strict, because it is not possible to maximize the mean and variance of $z(\boldsymbol{\nu})$ simultaneously. It now remains to find $\tilde{m}_z$ and $\tilde{v}_z$. From (50), it can be seen that $E[z(\boldsymbol{\nu})]$ is maximized when $\boldsymbol{\nu}$ lies along the direction of $\mathbf{m}$, giving $\tilde{m}_z = \|\mathbf{m}\|$. From (51), $\mathrm{var}[z(\boldsymbol{\nu})]$ is maximized when $\boldsymbol{\nu}$ is equal to the dominant eigenvector of $\mathbf{K}$, and hence, $\tilde{v}_z = k_1$. Taking the maximum of both sides of (52), and using (39), we have the desired result. □.

C. Beamforming Capacity for Extreme Cases

We can recover prior results in the literature [3], [18], [21] regarding beamforming capacity when the BS has *only* mean feedback, or *only* covariance feedback, as corollaries of *Theorem 1*. For completeness, these results are presented below.

Corollary 1: When the BS has covariance feedback, but no mean feedback, (i.e., $\mathbf{m} = \mathbf{0}$), then capacity (C_{bf}) is achieved when the BS sends along the dominant eigenvector of $\mathbf{K}$, which we denote $\mathbf{v}_1$ (i.e., when $\boldsymbol{\nu} = \mathbf{v}_1$).

Moreover

$$I(\mathbf{v_1}) = I(\boldsymbol{\nu}_0) = \tilde{I}(z^\star) = C_{bf}. \tag{53}$$

In other words, the beamforming direction that maximizes receive SNR also maximizes capacity.

Corollary 2: When there is only mean feedback (i.e., $\mathbf{K} = \mathbf{0}$, or $\mathbf{K}$ is white), then capacity is achieved when sending along the direction of the feedback

$$I\left(\frac{\mathbf{m}}{\|\mathbf{m}\|}\right) = I(\boldsymbol{\nu}_0) = \tilde{I}(z^\star) = C_{bf}. \tag{54}$$

Again, the beamforming direction that maximizes receive SNR also maximizes capacity.

D. Numerical Results

In this section, we provide numerical results showing that maximum SNR beamforming along $\boldsymbol{\nu}_0$ is close to optimal in the information-theoretic sense. For notational simplicity, for the rest of the paper, we use the "noiseless" channel model given in (6), which results when the BS has exact channel measurements at various time-frequency bins with which to estimate the channel of interest. Results using the "noisy" channel model have been shown to be very similar, provided the noise variance is small.

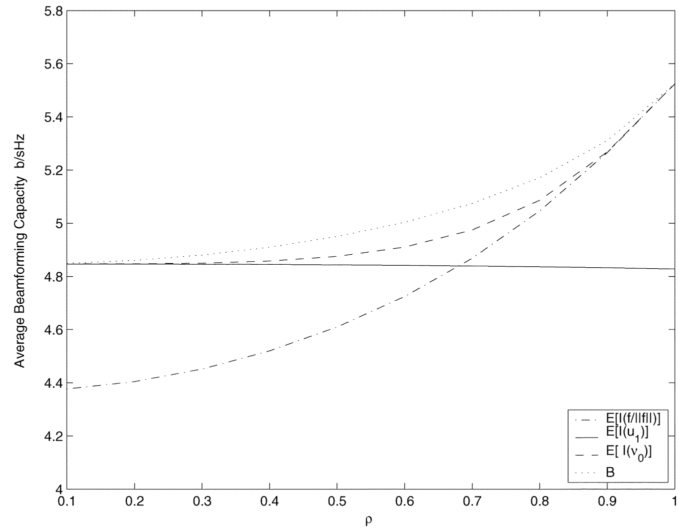

Fig. 2. Upper bound to average beamforming capacity, B, versus ρ, along with $\bar{I}(\mathbf{u}_1)$, $\bar{I}(\mathbf{f}/\|\mathbf{f}\|)$ and $\bar{I}(\nu_0)$. $N_T = 6$, the PAP is Laplacian with zero mean and angular spread 10°, and $P = 10$.

To recapitulate, the conditional channel distribution given feedback (6) is given by $\mathbf{h} \sim CN(\rho\mathbf{f}, (1-\rho^2)\mathbf{C})$, where $\mathbf{f}$ can be thought of as the feedback and has distribution $\mathbf{f} \sim CN(0, \mathbf{C})$, and ρ measures the accuracy of the feedback. Note that for this model, sending in the direction of the mean $\mathbf{m} \equiv \rho\mathbf{f}$ is equivalent to sending in the direction of $\mathbf{f}$, and $\boldsymbol{\nu}_0$ is the main eigenvector of $\mathbf{M} \equiv \rho^2\mathbf{f}\mathbf{f}^{\mathbf{H}} + (\mathbf{1} - \rho^{\mathbf{2}})\mathbf{C}$. Note also that the dominant eigendirection of $\mathbf{K}$ is $\mathbf{u}_1$, the dominant eigendirection of $\mathbf{C}$ (i.e., $\mathbf{v}_1 = \mathbf{u}_1$).

Let us define

$$\bar{I}(\boldsymbol{\nu}) \equiv E_{\mathbf{f}}[I(\boldsymbol{\nu})] \tag{55}$$

$$B \equiv E_{\mathbf{f}}[\tilde{I}(z^\star)] \tag{56}$$

where the expectation is taken over the feedback $\mathbf{f}$. Note that $I(\boldsymbol{\nu})$ depends on $\mathbf{f}$ not only because $\boldsymbol{\nu}$ may be a function of $\mathbf{f}$, but also because the channel distribution, given feedback, depends on $\mathbf{f}$, as shown above. Note also that $z^\star$ is a function of $\mathbf{f}$ because its mean depends on $\mathbf{f}$. $\bar{I}(\boldsymbol{\nu})$ is the expected mutual information when beamforming along direction $\boldsymbol{\nu}$. By construction, B is an upper bound on the average capacity.

As an example, we consider a system with $N_T = 6$ transmit-antenna elements at half-wavelength spacing. The channel has a single cluster, with a Laplacian PAP of zero mean and angular spread 10°. (We measure angular spread as $2\Omega_{\text{spread}}$ where $\Omega^2_{\text{spread}} = \mathrm{var}[\Omega]/2$. For this PAP and antenna spacing, the resulting covariance $\mathbf{C}$ is such that beamforming is the optimal transmit strategy, even if only covariance feedback is available. This can be verified using the sufficient and necessary conditions for beamforming given in [18]. Fig. 2 shows that $\bar{I}(\boldsymbol{\nu}_0)$ is close to the upper bound B, regardless of the correlation ρ between the mean feedback and the actual channel. Also displayed are $\bar{I}(\mathbf{u}_1)$, which is the average beamforming capacity with covariance feedback alone, and $\bar{I}(\mathbf{f}/\|\mathbf{f}\|)$, which is the average beamforming capacity with mean feedback alone (see *Corollaries 1* and *2*). As expected, when ρ is close to 1, $\bar{I}(\mathbf{f}/||\mathbf{f}||)$ approaches capacity, and as ρ gets close to 0, $\bar{I}(\mathbf{u}_1)$ approaches capacity.

Extensive computations, not reported here due to lack of space, show that the mutual information with maximum SNR beamforming is close to the upper bound of beamforming capacity for all choices of N_T, SNR, and PAP that we considered. Since $\boldsymbol{\nu}_0$ can be easily computed, given mean and covariance feedback, in contrast to the complexity of computing the optimal beamformer, maximum SNR beamforming would probably be the best practical strategy for exploiting mean and covariance feedback together. However, obtaining mean feedback can be costly in terms of resource consumption, and obtaining *accurate* feedback may even be infeasible. Since covariance feedback is implicitly available to the BS, it is desirable to know how good the mean feedback has to be in order to see significant performance benefits over using covariance feedback alone. This issue is addressed in the next section.

IV. When to Use Mean Feedback

In this section, we find an approximate measure of the accuracy needed for mean feedback to give appreciable gains over covariance feedback, in terms of the number of transmit antennas and the channel eigenvalues, given that beamforming is the transmit strategy. As the number of antennas increases, or the PAP becomes more spread out, the necessary accuracy decreases, meaning that mean feedback becomes more valuable as N_T increases and the channel becomes less spatially correlated. This remains true even when the transmitter is not restricted to beamforming.

The accuracy of the mean feedback can be gauged by its correlation with the true channel response. For some value ρ_0 of the correlation coefficient ρ, the average beamforming capacity with mean feedback alone equals the average beamforming capacity with covariance feedback alone. When $\rho > \rho_0$, i.e., the mean feedback is more accurate, using mean feedback gives better results than using covariance feedback, while the opposite is true when $\rho < \rho_0$. More precisely, when $\rho = \rho_0$, $\bar{I}(\mathbf{u}_1) = \bar{I}(\mathbf{f}/||\mathbf{f}||)$. In Fig. 2, it can be seen that $\rho_0 \approx 0.68$. We say that the mean feedback is accurate enough to be useful as long as $\rho \geq \rho_0$, and offer the following intuitively pleasing estimate of ρ_0, denoted by $\tilde{\rho}_0$. At $\rho = \tilde{\rho}_0$, the receive SNR is the same whether the BS beamforms along the direction of the dominant channel eigenmode, or along the direction of the channel mean. Mathematically, this is expressed as

$$E_{\mathbf{h},\mathbf{f}}[|z(\mathbf{u}_1)|^2|\rho = \tilde{\rho}_0)] = E_{\mathbf{h},\mathbf{f}}\left[\left|z\left(\frac{\mathbf{f}}{||\mathbf{f}||}\right)\right|^2 |\rho = \tilde{\rho}_0)\right] \quad (57)$$

where $z(\boldsymbol{\nu}) = \mathbf{h}^H\boldsymbol{\nu}$. Note that the above expressions depend on ρ, since the distribution of $\mathbf{h}$ depends on ρ. *Proposition 2* shows how $\tilde{\rho}_0$ can be calculated from N_T and the eigenvalues of $\mathbf{C}$.

Proposition 2: When the mean feedback and the channel have correlation coefficient $\rho = \tilde{\rho}_0$, the receive SNR is the same whether the BS beamforms along the direction of the dominant channel eigenmode, or along the direction of the channel mean. We can approximate $\tilde{\rho}_0$ in terms of the number of transmit antennas and the channel eigenvalues, as follows:

$$\tilde{\rho}_0 = \frac{\sqrt{\lambda_1 - A}}{\sqrt{N_T - A}} \quad (58)$$

where $A \equiv \sum_{i=1}^{N_T} \lambda_i^2/N_T$ and the $\{\lambda_i\}$ are the eigenvalues of $\mathbf{C}$, sorted in decreasing order (3).

Proof: We first find $E_{\mathbf{h},\mathbf{f}}[|z(\mathbf{u}_1)|^2|\rho]$ and $E_{\mathbf{h},\mathbf{f}}[|z(\mathbf{f}/||\mathbf{f}||)|^2|\rho]$ in terms of ρ, N_T, and $\{\lambda_i\}$. Given feedback $\mathbf{f}$ with correlation ρ, the channel can be written as

$$\mathbf{h} = \rho\mathbf{f} + \sqrt{1-\rho^2}\boldsymbol{\epsilon} \quad (59)$$

where $\boldsymbol{\epsilon} \sim CN(0, \mathbf{C})$. Expanding $\mathbf{f}$ and $\boldsymbol{\epsilon}$ in the basis of the eigenvectors of $\mathbf{C}$, $\mathbf{u}_1 \ldots \mathbf{u}_{N_T}$, we have

$$\mathbf{h} = \rho\alpha\sum_{i=1}^{N_T} a_i\mathbf{u}_i + \sqrt{1-\rho^2}\sum_{i=1}^{N_T} b_i\mathbf{u}_i \quad (60)$$

where $\alpha = ||\mathbf{f}||$ and $\sum_{i=1}^{N_T} a_i\mathbf{u}_i = \mathbf{f}/||\mathbf{f}||$. The $\{a_i\}$ and $\{b_i\}$ are independent, zero-mean complex Gaussian random variables with $E[|a_i|^2] = \lambda_i/N_T$ and $E[|b_i|^2] = \lambda_i$. Thus

$$z\left(\frac{\mathbf{f}}{||\mathbf{f}||}\right) = \mathbf{h}^H\frac{\mathbf{f}}{||\mathbf{f}||} \quad (61)$$

$$= \rho\alpha + \sqrt{1-\rho^2}\sum_{i=1}^{N_T} b_i^* a_i \quad (62)$$

and

$$\left|z\left(\frac{\mathbf{f}}{||\mathbf{f}||}\right)\right|^2 = \rho^2\alpha^2 + (1-\rho^2)\left|\sum_{i=1}^{N_T} b_i^* a_i\right|^2 + \rho\alpha\sqrt{1-\rho^2}\sum_{i=1}^{N_T} b_i^* a_i + \rho\alpha\sqrt{1-\rho^2}\sum_{i=1}^{N_T} b_i a_i^*. \quad (63)$$

Taking expectations with the feedback $\mathbf{f}$ fixed

$$E\left[\left|z\left(\frac{\mathbf{f}}{||\mathbf{f}||}\right)\right|^2 |\rho, \alpha, \{a_i\}\right] = \rho^2\alpha^2 + (1-\rho^2)\sum_{i=1}^{N_T} |a_i|^2\lambda_i \quad (64)$$

where we have used that the $\{a_i\}$, $\{b_i\}$, are independent with zero mean. Taking expectations over the feedback, we can write

$$E\left[\left|z\left(\frac{\mathbf{f}}{||\mathbf{f}||}\right)\right|^2 |\rho\right] = \rho^2 N_T + \frac{(1-\rho^2)}{N_T}\sum_{i=1}^{N_T} \lambda_i^2. \quad (65)$$

Similarly

$$E[|z(\mathbf{u}_1)|^2|\rho] = \lambda_1. \quad (66)$$

Equating the right-hand sides (RHS) of (65) and (66), we have the desired result. □

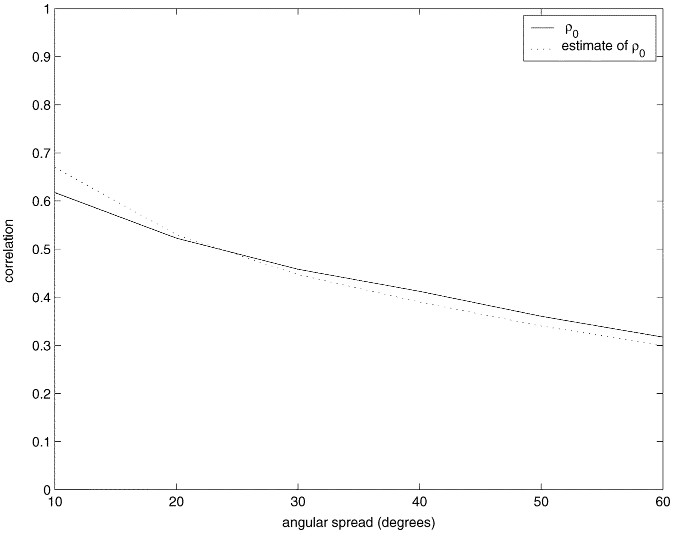

Fig. 3. Solid and dashed lines show ρ_0 and $\tilde{\rho}_0$, respectively, versus the angular spread of the PAP for a Laplacian profile centered at 0°. (The angular spread corresponds to $2\Omega_{\text{spread}}$). $\tilde{\rho}_0$ is calculated using the formula given in *Proposition 2*. $N_T = 6$ and the SNR $= 10$ dB.

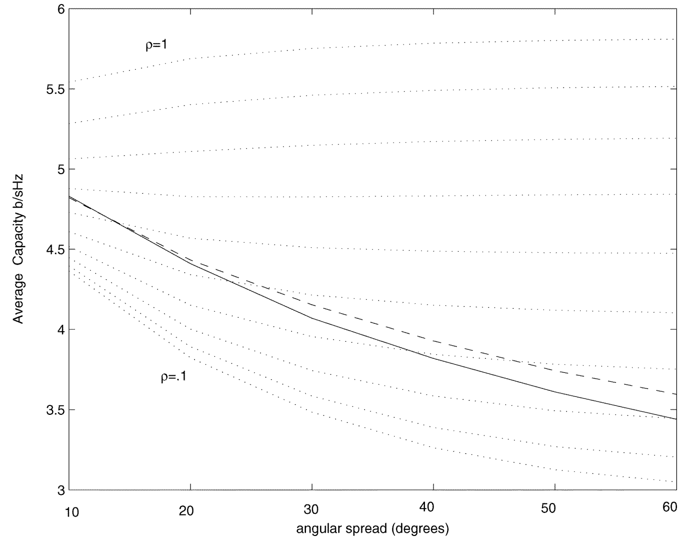

Fig. 4. Dotted lines show $\bar{I}(\mathbf{f}/\|\mathbf{f}\|)$, the average beamforming capacity when only mean feedback is available, versus the angular spread of the PAP for values of ρ between 0.1 and 1 in increments of 0.1. (The curves move up for increasing ρ.) The solid line is $\bar{I}(\mathbf{u}_1)$, the average beamforming capacity when only covariance feedback is available. Also shown is the average capacity when there is only covariance feedback, but the transmitter is not restricted to beamforming (the dashed line). $N_T = 6$ and the SNR $= 10$ dB.

Fig. 3 shows $\tilde{\rho}_0$, as defined in *Proposition 2*, as a function of the angular spread of the PAP ($2\Omega_{\text{spread}}$), for our running example of a one-cluster channel with a Laplacian PAP centered at 0°. As before, $N_T = 6$, and $P = 10$. Since $\tilde{\rho}_0$ is decreasing, the feedback need be less accurate for larger angular spreads. This is intuitive, since covariance feedback becomes less valuable as the channel correlations decrease. Also plotted is ρ_0 (the correlation at which $\bar{I}(\mathbf{u}_1) = \bar{I}(\mathbf{f}/\|\mathbf{f}\|)$), as obtained via simulation. As long as N_T is not too small, $\tilde{\rho}_0$ is a good approximation for ρ_0.

A clearer picture of how mean feedback compares to covariance feedback can be seen in Fig. 4. (The parameters are the same as in Fig. 3.) The expected beamforming capacity with mean feedback alone, $\bar{I}(\mathbf{f}/\|\mathbf{f}\|)$, is plotted versus angular spread for various values of ρ between 0.1 and 1, together with $\bar{I}(\mathbf{u}_1)$, the expected beamforming capacity with covariance feedback alone (plotted as a solid line). Having mean feedback with correlation coefficient 0.7 gives negligible gain over having only covariance feedback when the angular spread is around 10°, but delivers significant gains over covariance feedback (almost 1.5 b/s/Hz) when the angular spread is around 60°. This is expected, since the larger the angular spread, the less useful the covariance information. The dashed line in Fig. 4 shows the capacity when the BS only uses covariance information, but is not restricted to beamforming. Comparing this with the capacity attained when the BS uses only covariance information, but *is* restricted to beamforming, it can be seen that the added complexity produces minimal improvement. (If the mobile were equipped with multiple antennas, or if there were multiple clusters, using a more complex transmission strategy than beamforming would be much more beneficial [6].) The main conclusion is that for larger angular spreads, obtaining mean feedback may be worthwhile, even though covariance information is implicitly available.

Besides being a function of the $\{\lambda_i\}$, and hence the PAP, $\tilde{\rho}_0$ is also a function of N_T. Thus, the accuracy required for mean feedback to be more effective than covariance feedback depends on the number of transmit antennas. Since $\tilde{\rho}_0$ is a decreasing function of N_T, the requirements on the accuracy of mean feedback get relaxed as the number of transmit antennas increases. This can be seen when comparing Fig. 4 with Fig. 5. Both simulate the same one-cluster channel, but the number of transmit antennas is six in Fig. 4 and four in Fig. 5. Whether the BS uses beamforming, or the optimal transmit strategy, mean feedback gives better results than covariance feedback for lower values of ρ when the number of antennas is greater. For instance, at an angular spread of 10°, mean feedback is helpful for $\rho \geq 0.7$ when

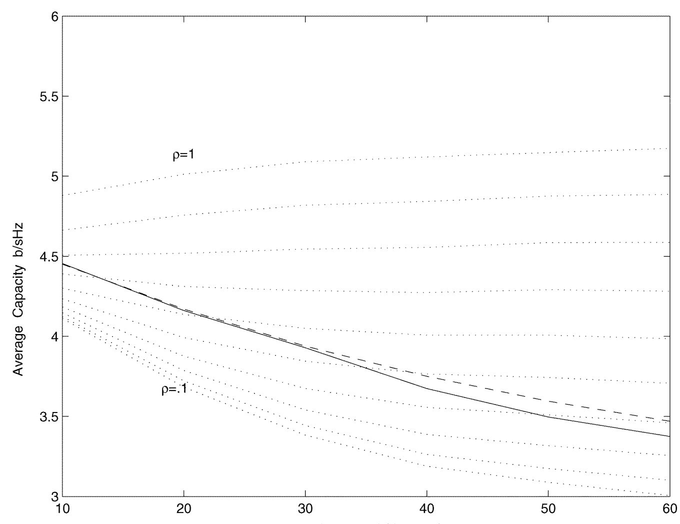

Fig. 5. $\bar{I}(\mathbf{f}/\|\mathbf{f}\|)$ versus angular spread for values of ρ between 0.1 and 1 in increments of 0.1 (the dotted lines). The solid line is $\bar{I}(\mathbf{u}_1)$, the average beamforming capacity when only covariance feedback is available. Also shown is the average capacity when there is only covariance feedback, but the transmitter is not restricted to beamforming (the dashed line). $N_T = 4$ and the SNR $= 10$ dB.

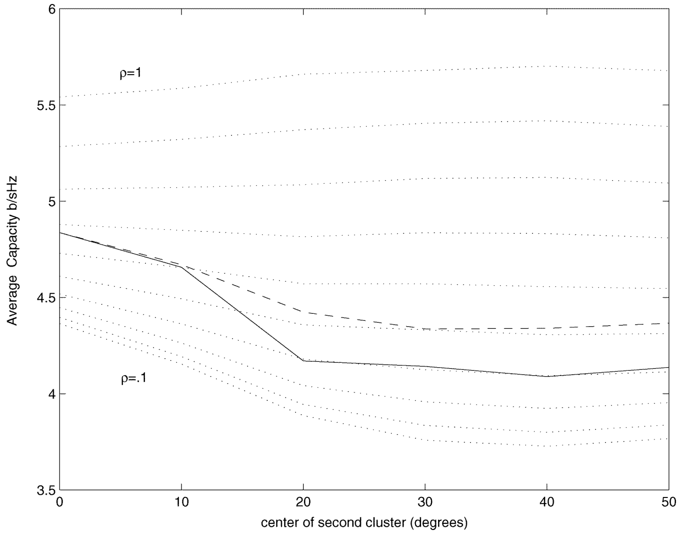

Fig. 6. Dotted lines show $\bar{I}(\mathbf{f}/\|\mathbf{f}\|)$ versus $\bar{\Omega}$ for values of ρ between 0.1 and 1 in increments of 0.1. The solid line is $\bar{I}(\mathbf{u}_1)$, and the dashed line is the average capacity when there is only covariance feedback, but the transmitter is not restricted to beamforming. $N_T = 6$ and the SNR $= 10$ dB.

there are six antennas, whereas it is helpful for $\rho \geq 0.8$ when there are four antennas.

So far, we have only considered one cluster channels. Fig. 6 shows results for a channel with two equipowered clusters, each with an angular spread of 10°. The number of transmit antennas is six. The center of the first cluster's PAP is fixed at 0°, while the center of the second cluster's PAP is varied. $\bar{I}(\mathbf{f}/\|\mathbf{f}\|)$ is plotted versus the center angle of the second cluster for various values of ρ between 0.1 and 1, together with $\bar{I}(\mathbf{u}_1)$, the expected beamforming capacity with covariance feedback alone, plotted as a solid line. The capacity when only covariance feedback is used, but the transmission strategy is optimal, is shown in the dashed line. In this particular scenario, since both clusters have narrow spreads and contribute equally to the PAP, capacity using only covariance information is achieved by sending along two eigendirections. It can be seen that mean feedback is more useful (in comparison with covariance feedback) when there are two clusters than when there is a single cluster, even when the optimal transmit strategy is used. (The single-cluster case corresponds to the second cluster being centered at 0°.) Note, however, that in this example, even when there are two clusters, we must have correlation $\rho \geq 0.7$ to offer improvements on the order of 0.5 b/s/Hz. In view of the overhead that may be required to obtain such accurate estimates, it is probably preferable to use only covariance information in this situation.

V. Discussion

We have provided a simple analytical framework for evaluating the gains of mean feedback in systems where covariance feedback is already available. The simplicity of this framework results from our use of a noiseless feedback model which provides an upper bound on performance with mean feedback, while capturing the fundamental problem with mean feedback: decorrelation of channel realizations across time and frequency. We now consider some examples which illustrate when mean feedback is of practical use in such systems. These examples

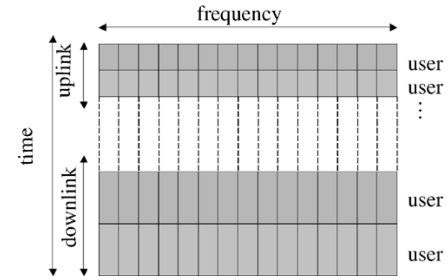

Fig. 7. TDD system with TDMA on the uplink and TDM on the downlink.

are similar to those in [1], where the author investigates mean feedback in the *absence* of covariance feedback. We assume an OFDM framework with 1024 subcarriers spaced 25 kHz apart, and centered at 1 GHz. We also assume the channel has a single cluster with an exponential PDP having a root mean square (rms) value of 0.5 μs.

A TDD system with time-division multiple access (TDMA)/time-division multiplexing (TDM) on the uplink/downlink is shown in Fig. 7. Each user sends to the BS using the entire frequency band for a certain amount of time, and subsequently the BS takes turns sending to the mobiles over the whole band. The BS can measure the channel on the uplink, and use these measurements to estimate the downlink channel. The longest a user will have to wait until it hears back from the BS is approximately the number of users in the system multiplied by the time the BS sends to each user. For a rate of 20 Mb/s and 10 packet payloads of 10 000 bits each, the time the BS sends to each mobile is approximately 5 ms. If there are 10 users, this means the total delay is around 50 ms.

Now, consider that the BS uses channel measurements from 10 frequency bins to estimate the channel of a particular bin. The frequencies used are as close as possible to the frequency of interest, but the measurements have a time delay of 50 ms, due to the nature of TDD systems. For a mobile traveling at vehicular speeds, the maximum Doppler frequency is on the order of 100 Hz, and hence, the correlation between the mean feedback and the desired channel is 0.1 (we calculate time decorrelation using Clarke's model, and frequency decorrelation using an exponential PDP, as discussed in Section II-B). Clearly, it is not worthwhile to try and estimate the first-order channel statistics in this case. For pedestrian mobile speeds, and a corresponding Doppler spread of 3 Hz, the correlation is 0.79. If the BS has six antennas and an SNR of 10 dB, then it can be seen from Fig. 4 that using mean feedback instead of covariance feedback gives gains ranging from 0.1 to 1.5 b/s/Hz, depending on the PAP of the channel. For a narrow angular spread of 10°, the cost of the mean feedback probably does not justify the benefits, but for larger angular spreads, the improvement is significant.

While our overall conclusion is that covariance feedback alone is very effective on typical outdoor mobile channels, the preceding results also illustrate that mean feedback can further enhance capacity for slow-moving users, especially if the channel has a large angular spread. Mean feedback is particularly attractive for high-bandwidth applications (e.g., file downloads or gaming) involving users who are sitting in one place. For FDD systems, the use of mean feedback is more onerous in terms of implementation, standardization, and overhead, since it must be sent back explicitly. It may be attractive in FDD systems, therefore, to selectively request mean feedback from slow-moving users (which the BS should be able to identify using Doppler estimation).

APPENDIX

A. Proof of Proposition 1

Let the BS have knowledge of M different channel estimates, $\{\tilde{\mathbf{h}}_{i_1k_1}, \tilde{\mathbf{h}}_{i_2k_2}, \ldots, \tilde{\mathbf{h}}_{i_Mk_M}\}$ from different time and frequency bins, with which to estimate the channel at frequency j and time q, denoted $\mathbf{h}_{jq} \equiv \mathbf{h}$. (Note that $\mathbf{h} \sim CN(0, \mathbf{C})$, and that we have switched to subscript notation for the time and frequency indexes for convenience.) Let $\mathbf{x}$ denote the $N_TM \times 1$ stacked vector of these realizations

$$\mathbf{x} = [\tilde{\mathbf{h}}_{i_1k_1}^T, \tilde{\mathbf{h}}_{i_2k_2}^T, \ldots, \tilde{\mathbf{h}}_{i_Mk_M}^T]^T. \tag{67}$$

We assume the estimates are equal to the true channel realization plus some AWGN. In other words

$$\tilde{\mathbf{h}}_{i_jk_j} = \frac{1}{\gamma}(\mathbf{h}_{i_jk_j} + \tilde{e}) \quad \forall j \tag{68}$$

where $\tilde{e}$ is AWGN with covariance $\sigma_f^2\mathbf{I}$ and $\gamma = \sqrt{1+\sigma_f^2}$ is a normalization constant. The MMSE estimate of $\mathbf{h}$, denoted $\hat{\mathbf{h}}$, is given by

$$\begin{aligned}\hat{\mathbf{h}} &= E[\mathbf{h}_{jq}|\mathbf{x}] & (69)\\ &= \mathbf{K_{hx}}\mathbf{K_{xx}}^{-1}\mathbf{x} & (70)\end{aligned}$$

where $\mathbf{K_{xx}}$ and $\mathbf{K_{xh}}$ are as defined in (22) and (21), respectively.

Since the $\{\mathbf{h}_{i_mk_m}\}$ are jointly complex Gaussian with zero mean, then $\hat{\mathbf{h}}$ is also zero-mean complex Gaussian. Using (70), it can be seen that its distribution is given as follows:

$$\hat{\mathbf{h}} \sim CN(0, \mathbf{K_{hx}}\mathbf{K_{xx}}^{-1}\mathbf{K_{xh}}). \tag{71}$$

From standard estimation theory, the conditional distribution of $\mathbf{h}$ is

$$\mathbf{h} \sim CN(\hat{\mathbf{h}}, \mathbf{C} - \mathbf{K_{hx}}\mathbf{K_{xx}}^{-1}\mathbf{K_{xh}}). \tag{72}$$

Equations (71) and (72) can be massaged into a more meaningful form as follows. Writing out $\mathbf{K_{hx}}$, we have

$$\mathbf{K_{hx}} = \frac{1}{\gamma}[\kappa_1\mathbf{C} \;\; \kappa_2\mathbf{C} \ldots\ldots \kappa_M\mathbf{C}] \tag{73}$$

where the $\{\kappa_i\}$ are correlation coefficients. This can be written more compactly as

$$\mathbf{K_{hx}} = \frac{1}{\gamma}\boldsymbol{\varsigma} \otimes \mathbf{C} \tag{74}$$

where $\boldsymbol{\varsigma}$ is a $1 \times M$ vector of correlation coefficients $\{\kappa_i\}$, and $\otimes$ denotes Kronecker product. Similarly, we can write

$$\mathbf{K_{xx}} = \frac{1}{\gamma^2}(\boldsymbol{\Upsilon} \otimes \mathbf{C} + \sigma_f^2\mathbf{I}) \tag{75}$$

where $\boldsymbol{\Upsilon}$ is an $M \times M$ matrix of correlation coefficients with ones along the diagonal, and $\mathbf{I}$ is the $MN_T \times MN_T$ identity matrix.

Noiseless Feedback: Let us for the moment assume that $\sigma_f^2 = 0$ ($\gamma = 1$). In this case

$$\mathbf{K_{xx}} = \boldsymbol{\Upsilon} \otimes \mathbf{C} \tag{76}$$

and

$$\mathbf{K_{xx}}^{-1} = \boldsymbol{\Upsilon}^{-1} \otimes \mathbf{C}^{-1}. \tag{77}$$

Hence, we have that

$$\mathbf{K_{hx}}\mathbf{K_{xx}}^{-1}\mathbf{K_{xh}} = \boldsymbol{\varsigma}\boldsymbol{\Upsilon}^{-1}\boldsymbol{\varsigma}^H\mathbf{C}. \tag{78}$$

Letting $\rho^2 = \boldsymbol{\varsigma}\boldsymbol{\Upsilon}^{-1}\boldsymbol{\varsigma}^H$, we now have that

$$\hat{\mathbf{h}} \sim CN(0, \rho^2\mathbf{C}) \tag{79}$$

and

$$\mathbf{h} \sim CN(\hat{\mathbf{h}}, (1-\rho^2)\mathbf{C}) \tag{80}$$

where the value of ρ indicates the accuracy of the feedback. ($\rho = 1$ would indicate perfect feedback.) Hence, we can write $\mathbf{m} = \rho\mathbf{f}$, where $\mathbf{f} \sim CN(0, \mathbf{C})$ and $\mathbf{K} = (1-\rho^2)\mathbf{C}$, as desired.

Noisy Feedback: Let us now consider $\sigma_f^2 \neq 0$. Then $\mathbf{K_{xx}}$ is as given in (75). The error term (the second term) in (75) results in the covariance of $\hat{\mathbf{h}}$, (79) having an extra term which is proportional to σ_f^2. This can be seen as follows. We first write $\mathbf{K_{xx}}^{-1}$ as follows, so as to separate out the terms due to noise:

$$\mathbf{K_{xx}}^{-1} = \gamma^2\boldsymbol{\Upsilon}^{-1} \otimes \mathbf{C}^{-1} - \gamma^2\sigma_f^2\boldsymbol{\Upsilon}^{-1} \otimes \mathbf{C}^{-1}(\boldsymbol{\Upsilon} \otimes \mathbf{C} + \sigma_f^2\mathbf{I})^{-1}. \tag{81}$$

Substituting (81) and (74) in (71) and (72), we get that

$$\hat{\mathbf{h}} \sim CN(0, \rho^2\mathbf{C} - \sigma_f^2\mathbf{Q}) \tag{82}$$

and

$$\mathbf{h} \sim CN(\hat{\mathbf{h}}, (1-\rho^2)\mathbf{C} + \sigma_f^2\mathbf{Q}) \tag{83}$$

where

$$\mathbf{Q} \equiv (\boldsymbol{\varsigma} \otimes \mathbf{C})\left(\boldsymbol{\Upsilon}^{-1} \otimes \mathbf{C}^{-1}(\boldsymbol{\Upsilon} \otimes \mathbf{C} + \sigma_f^2\mathbf{I})^{-1}\right)(\boldsymbol{\varsigma} \otimes \mathbf{C})^H. \tag{84}$$

Hence, $\mathbf{m} \sim CN(0, \rho^2\mathbf{C} - \sigma_f^2\mathbf{Q})$ and $\mathbf{K} = (1-\rho^2)\mathbf{C} + \sigma_f^2\mathbf{Q})$ as desired. Note that, to first order, we can approximate $\sigma_f^2\mathbf{Q}$ as follows:

$$\begin{aligned}\sigma_f^2\mathbf{Q} &\approx \sigma_f^2(\boldsymbol{\varsigma}\otimes\mathbf{C})(\boldsymbol{\Upsilon}^{-1}\otimes\mathbf{C}^{-1})(\boldsymbol{\Upsilon}^{-1}\otimes\mathbf{C}^{-1})(\boldsymbol{\varsigma}\otimes\mathbf{C})^H & (85)\\ &= \sigma_f^2\boldsymbol{\varsigma}\boldsymbol{\Upsilon}^{-2}\boldsymbol{\varsigma}^H\mathbf{I}. & (86)\end{aligned}$$

Thus, for small σ_f^2, $\sigma_f^2\mathbf{Q}$ is small (this also relies on the structure of $\boldsymbol{\varsigma}$ and $\boldsymbol{\Upsilon}$), and thus (82) and (83) are well approximated by (79) and (80). □

B. Proof of Lemma 1

The complex Gaussian random variable z can be written as

$$z \sim m_z + \sigma_z w_1 + j\sigma_z w_2 \tag{87}$$

where w_1 and w_2 are independent and identically distributed (i.i.d.) standard Gaussian random variables. We now show that $\partial\tilde{I}(z)/\partial m_z \geq 0$. (WLOG we set $P = 1$.) Substituting the RHS of (87) into (41), and noting that the expectation is now over w_1 and w_2, we have that

$$\frac{\partial\tilde{I}(z)}{\partial m_z} = E\left[\frac{2(m_z+\sigma_z w_1)}{1+(m_z+\sigma_z w_1)^2+(\sigma_z w_2)^2}\right] \geq 0. \tag{88}$$

To see why the inequality holds, note that w_1 is symmetrically distributed around 0 and $m_z \geq 0$. Fix the value of w_2, and consider a negative value of w_1 that makes $m_z + \sigma_z w_1 < 0$; for instance, let $w_1 = -a_1$ make $m_z + \sigma_z w_1 = -x$. (Of course, $a_1, x > 0$.) Then, there exists an $a_2 < a_1$ ($a_2 > 0$) such that when $w_1 = a_2$, $m_z + \sigma_z w_1 = x$. The value inside the expectation in (88) is equal to some y ($y > 0$) when $w_1 = a_2$, and $-y$ when $w_1 = -a_1$. However, since $a_2 < a_1$, the positive value is more probable (a_2 is more probable since w_1 is Gaussian with zero mean), and therefore, the expected value in (88) over w_1 with fixed w_2 is greater than or equal to zero. Since this holds for all values of w_2, we have the desired inequality.

We show that $\partial\tilde{I}(z)/\partial\sigma_z \geq 0$ in a similar manner

$$\frac{\partial\tilde{I}(z)}{\partial\sigma_z} = E\left[\frac{2w_1(m_z+\sigma_z w_1)+2\sigma_z w_2^2}{1+(m_z+\sigma_z w_1)^2+(\sigma_z w_2)^2}\right]. \tag{89}$$

The denominator of (89) can be written as

$$1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2\left(1+\frac{2\sigma_z w_1 m_z}{1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2}\right). \tag{90}$$

Letting

$$s \equiv \frac{2\sigma_z w_1 m_z}{1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2}$$

and using the inequality

$$\frac{1}{1+s} \geq 1-s$$

we have from (89) and (90) that

$$\begin{aligned}\frac{\partial\tilde{I}(z)}{\partial\sigma_z} &\geq E\left[\frac{2w_1(m_z+\sigma_z w_1)+2\sigma_z w_2^2}{1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2}(1-s)\right] \qquad (91)\\ &= E\left[\frac{2w_1m_z+2\sigma_z w_1^2+2\sigma_z w_2^2}{1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2}\right]\\ &\quad - E\left[\frac{4w_1^2m_z^2\sigma_z+4\sigma_z^2w_1^3m_z+4\sigma_z^2w_2^2w_1m_z}{(1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2)^2}\right] \qquad (92)\\ &= E\left[\frac{2\sigma_z w_1^2+2\sigma_z w_2^2}{1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2}\right]\\ &\quad - E\left[\frac{4\sigma_z w_1^2m_z^2}{(1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2)^2}\right] \qquad (93)\\ &= E\left[\frac{2\sigma_z w_1^2+2\sigma_z w_2^2+2(\sigma_z w_1^2+\sigma_z w_2^2)(\sigma_z^2w_1^2+\sigma_z^2w_2^2)}{(1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2)^2}\right]\\ &\quad - E\left[\frac{-2(\sigma_z w_1^2+\sigma_z w_2^2)m_z^2+4\sigma_z w_1^2m_z^2}{(1+m_z^2+\sigma_z^2w_1^2+\sigma_z^2w_2^2)^2}\right] \qquad (94)\end{aligned}$$

where we have again used that $w1$, $w2$ are symmetrically distributed around 0. The second term in (94) is equal to 0, since w_1 and w_2 are i.i.d., and the first term is positive, hence

$$\frac{\partial\tilde{I}(z)}{\partial\sigma_z} \geq 0. \tag{95}$$

□

REFERENCES

[1] J. Cavers, "Single-user and multiuser adaptive maximal ratio transmission for Rayleigh channels," *IEEE Trans. Veh. Technol.*, vol. 49, no. 6, pp. 2043–2050, Nov. 2002.

[2] A. Narula, M. Lopez, D. Trott, and G. Wornell, "Efficient use of side information in multiple antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 10, pp. 1423–1435, Oct. 1998.

[3] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect channel feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 9, pp. 2632–2639, Sep. 2001.

[4] S. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Int. Conf. Commun.*, vol. 7, 2001, pp. 2266–2270.

[5] F. Vook, T. Thomas, and X. Zhuang, "Transmit diversity and transmit adaptive arrays for broadband mobile OFDM systems," *Wireless Commun. Netw.*, vol. 1, pp. 44–49, Mar. 2003.

[6] G. Barriac and U. Madhow, "Space-time communication for OFDM with implicit channel feedback," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3111–3129, Dec. 2004.

[7] ——, "Characterizing outage capacity for space-time communication over wideband wireless channels," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 2, 2002, pp. 1513–1517.

[8] ——, "Characterizing the outage capacity for space–time communication over wideband channels," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2198–2208, Dec. 2004.

[9] ——, "Wideband space-time communication with implicit channel feedback," in *Proc. 7th Int. Symp. Signal Process. Appl.*, vol. 1, 2003, pp. 225–228.

[10] G. Raleigh and V. Jones, "Adaptive antenna transmission for frequency duplex digital wireless communication," in *Proc. Int. Conf. Commun.*, vol. 6, Jun. 1997, pp. 641–646.

[11] Y. Liang and F. Chin, "Downlink channel covariance matrix estimation and its applications in wireless DS-CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 2, pp. 222–232, Feb. 2001.

[12] S. Zhou and G. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2599–2613, Oct. 2002.

[13] D. Morgan, "Downlink adaptive array algorithms for cellular mobile communications," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 476–488, Mar. 2003.

[14] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[15] C. Farsakh and J. Nossek, "Spatial covariance based downlink beamforming in an SDMA mobile radio system," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1497–1506, Nov. 1998.

[16] P. Rashid-Farrokhi, K. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 10, pp. 1437–1449, Oct. 1998.

[17] W. C. Jakes and D. C. Cox, Eds., *Microwave Mobile Communications*. New York: Wiley-IEEE Press, 1994.

[18] S. Jafar and A. Goldsmith, “On optimality of beamforming for multiple antenna systems with imperfect feedback,” in *Proc. IEEE Int. Symp. Inf. Theory*, 2001, p. 321.
[19] E. Jorswieck and H. Boche, “Optimal transmission with imperfect channel state information at the transmit antenna array,” *Wireless Pers. Commun.*, vol. 27, no. 1, pp. 33–56, 2003.
[20] H. Boche and E. Jorswieck, “On the optimality range of beamforming for MIMO systems with covariance fedback,” *IEICE Trans. Fundam. Electron., Commun., Comput. Sci.*, vol. E85-A, pp. 2521–2528, Nov. 2002.
[21] S. Simon and A. Moustakas, “Optimizing MIMO antenna systems with channel covariance feedback,” *IEEE J. Sel. Areas Commun.*, vol. 21, no. 4, pp. 406–417, Apr. 2003.

Gwen Barriac received the B.S. degree in electrical engineering from Princeton University, Princeton, NJ, in 1999, and the M.S. and Ph.D. degrees in electrical engineering from the University of California at Santa Barbara in 2004.

She is currently with Qualcomm, San Diego, CA.

Upamanyu Madhow (S'86–M'90–SM'96–F'05) received the bachelor's degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1985, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, in 1987 and 1990, respectively.

From 1990 to 1991, he was a Visiting Assistant Professor at the University of Illinois. From 1991 to 1994, he was a Research Scientist with Bell Communications Research, Morristown, NJ. From 1994 to 1999, he was on the faculty of the Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign. Since December 1999, he has been with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, where he is currently a Professor. His research interests are in communication systems and networking, with current emphasis on wireless communication, sensor networks, and data hiding.

Dr. Madhow is a recipient of the NSF CAREER award. He has served as Associate Editor for Spread Spectrum for the IEEE TRANSACTIONS ON COMMUNICATIONS, and as Associate Editor for Detection and Estimation for the IEEE TRANSACTIONS ON INFORMATION THEORY.